Numerical Justification for Multiresolution Optical Flow Computation

Nils Papenberg Andrés Bruhn Thomas Brox
Joachim Weickert
Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science
Saarland University, Building 27, 66041 Saarbrücken, Germany
{papenberg, bruhn, brox, weickert}@mia.uni-saarland.de

Abstract

We introduce a numerical scheme for the minimisation of an energy functional for computing optical flow. This functional combines a brightness constancy assumption, and a discontinuity-preserving spatio-temporal smoothness constraint. In order to allow for large displacements, linearisations in the data term is strictly avoided. The presented numerical scheme is based on two nested fixed point iterations. By proving that this scheme implements a coarse-to-fine warping strategy, we give a theoretical foundation for warping which has been used on a mainly experimental basis so far.

1 Introduction

In the last two decades the quality of optical flow estimation methods has increased dramatically. Starting from the original approaches of Horn and Schunck [11] as well as Lucas and Kanade [15], many new concepts have been developed for dealing with shortcomings of previous models. In order to handle discontinuities in the flow field, the quadratic regulariser in the Horn and Schunck model was replaced by smoothness constraints that permit piecewise smooth results [1, 9, 20, 22, 23]. Some of these ideas are close in spirit to methods for joint motion estimation and motion segmentation [10, 18], and to optical flow methods motivated from robust statistics where outliers are penalised less severely [6, 7]. Spatio-temporal approaches have ameliorated the results by using the information of an additional dimension [19, 6, 24, 10].

Since image sequences are often undersampled in time direction, large displacements are common. In this case nonlinearised models [20, 2] as well as coarse-to-fine strategies [3, 7, 17] have been experimentally demonstrated to be highly useful. Unfortunally – apart from a very nice paper by Lefébure and Cohen [14] – not many results are available that provide a theoretical foundation for this experimentally successful coarse-to-fine warping strategy. The goal of this paper is to close this gap.

To this end we consider a variational method with a nonlinearised data term. The minimizer of this energy functional is approximated by a specific numerical method. This scheme provides a novel foundation for the coarse-to-fine warping that is commonly used in image sequence analysis. This has two important effects: Firstly, it becomes possible to integrate the warping technique, which was so far only algorithmically motivated, into a variational framework. Secondly, it shows a theoretically sound way of how image correspondence problems can be solved with an efficient multi-resolution technique.

The experimental evaluation shows that our method yields excellent results. Compared to those in the literature, their accuracy is always higher.

Paper organisation. In the next section, our variational model is described, first by discussing all model assumptions, and then in form of an energy based formulation. Section 3 derives a minimisation scheme for this energy. The theoretical foundation of warping methods as a numerical approximation step is given in Section 4. An experimental evaluation is presented in Section 5, followed by a brief summary in Section 6.

2 The Variational Model

Before deriving a variational formulation for our optical flow method, we give an intuitive idea of which constraints in our view should be included in such a model.

- **Grey value constancy assumption.**
  Since the beginning of optical flow estimation, it has been assumed that the grey value of a pixel is not changed by the displacement.

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

  Here $I : \Omega \subset \mathbb{R}^3 \to \mathbb{R}$ denotes a rectangular image sequence, and $\mathbf{w} := (u, v, 1)^\top$ is the searched displacement vector between an image at time t and another

image at time $t+1$. The linearised version of the grey value constancy assumption yields the famous optical flow constraint [11]

$$I_x u + I_y v + I_t = 0 \tag{2}$$

where subscripts denote partial derivatives. However, this linearisation is only valid under the assumption that the image changes linearly along the displacement, which is in general not the case, especially for large displacements. Therefore, our model will use the original, non-linearised grey value constancy assumption (1).

- **Smoothness assumption.**
  So far, the model estimates the displacement of a pixel only locally without taking any interaction between neighbouring pixels into account. Therefore, it runs into problems as soon as the gradient vanishes somewhere, or if only the flow in normal direction to the gradient can be estimated (*aperture problem*). Furthermore, one would expect some outliers in the estimates. Hence, it is useful to introduce as a further assumption the smoothness of the flow field. This smoothness constraint can either be applied solely to the spatial domain, if there are only two frames available, or to the spatio-temporal domain, if the displacements in a sequence of images are wanted. As the optimal displacement field will have discontinuities at the boundaries of objects in the scene, it is sensible to generalise the smoothness assumption by demanding a *piecewise smooth* flow field.

- **Multiscale approach.**
  In the case of displacements that are larger than one pixel per frame, the cost functional in a variational formulation must be expected to be multi-modal, i.e. a minimisation algorithm could easily be trapped in a local minimum. In order to find the global minimum, it can be useful to apply multiscale ideas: One starts with solving a coarse, smoothed version of the problem, which may have a unique minimum, hopefully close to the global minimum of the original problem, and uses the result as an initialisation for solving a refined version of the problem. Instead of smoothing the problem, i.e. the image sequence, it is more efficient to downsample the images respecting the sampling theorem, so the model ends up in a multiresolution strategy.

With this description, it is straightforward to derive an energy functional that penalises deviations from these model assumptions. Let $\mathbf{x} := (x, y, t)^\top$ and $\mathbf{w} := (u, v, 1)^\top$. Then the global deviations from the grey value constancy assumption and the gradient constancy assumption are measured by the energy

$$E_{Data}(u, v) = \int_\Omega \left(|I(\mathbf{x} + \mathbf{w}) - I(\mathbf{x})|^2 \right) \mathbf{dx}. \tag{3}$$

Since with quadratic penalisers, outliers get too much influence on the estimation, an increasing concave function $\Psi(s^2)$ is applied, leading to a robust energy [7, 16]:

$$E_{Data}(u, v) = \int_\Omega \Psi\left(|I(\mathbf{x} + \mathbf{w}) - I(\mathbf{x})|^2 \right) \mathbf{dx}. \tag{4}$$

We use the function $\Psi(s^2) = \sqrt{s^2 + \epsilon^2}$ which results in (modified) L^1 minimisation. Due to the small positive constant ϵ, $\Psi(s)$ is still convex which offers advantages in the minimisation process. Moreover, this choice of Ψ does not introduce any additional parameters, since the small numerical parameter ϵ can be set to a fixed value, say 0.001.
Finally, a smoothness term has to describe the model assumption of a piecewise smooth flow field. This is achieved by penalising the total variation of the flow field [21, 8], which can be expressed as

$$E_{Smooth}(u, v) = \int_\Omega \Psi\left(|\nabla_3 u|^2 + |\nabla_3 v|^2 \right) \mathbf{dx}. \tag{5}$$

with the same function for Ψ as above. The spatio-temporal gradient $\nabla_3 := (\partial_x, \partial_y, \partial_t)^\top$ indicates that a spatio-temporal smoothness assumption is involved. For applications with only two images available it is replaced by the spatial gradient.
The total energy is the weighted sum between the data term and the smoothness term

$$E(u, v) = E_{Data} + \alpha E_{Smooth} \tag{6}$$

with some regularisation parameter $\alpha > 0$. Now the goal is to find the functions u and v that minimise this energy.

3 Minimisation

3.1 Euler–Lagrange Equations

Since $E(u, v)$ is highly nonlinear, the minimisation is not trivial. For better readability we define the following abbreviations:

$$\begin{aligned} I_x &:= \partial_x I(\mathbf{x} + \mathbf{w}), \\ I_y &:= \partial_y I(\mathbf{x} + \mathbf{w}), \\ I_z &:= I(\mathbf{x} + \mathbf{w}) - I(\mathbf{x}). \end{aligned} \tag{7}$$

According to the calculus of variations, a minimiser of (6) must fulfill the Euler-Lagrange equations

$$\begin{aligned} 0 &= \Psi'(I_z^2) \cdot (I_x I_z) \\ &\quad -\alpha \operatorname{div} \left(\Psi'(|\nabla_3 u|^2 + |\nabla_3 v|^2) \nabla_3 u \right), \\ 0 &= \Psi'(I_z^2) \cdot (I_y I_z) \\ &\quad -\alpha \operatorname{div} \left(\Psi'(|\nabla_3 u|^2 + |\nabla_3 v|^2) \nabla_3 v \right) \end{aligned}$$

with reflecting boundary conditions.

3.2 Numerical Approximation

The preceding Euler-Lagrange equations are nonlinear in their argument $\mathbf{w} = (u, v, 1)^\top$. A first step towards a linear system of equations, which can be solved with common numerical methods, is the use of fixed point iterations on $\mathbf{w}$. In order to implement a multiscale approach, necessary to better approximate the global optimum of the energy, these fixed point iterations are combined with a downsampling strategy. Instead of the standard downsampling factor of 0.5 on each level, it is proposed here to use an *arbitrary* factor $\eta \in (0, 1)$, what allows smoother transitions from one scale to the next[1]. Moreover, the full pyramid of images is used, starting with the smallest possible image at the coarsest grid. Let $\mathbf{w}^k = (u^k, v^k, 1)^\top$, $k = 0, 1, \ldots$, with the initialisation $\mathbf{w}^0 = (0, 0, 1)^\top$ at the coarsest grid. Further, let I_*^k be the abbreviations defined in (7) but with the iteration variable $\mathbf{w}^k$ instead of $\mathbf{w}$. Then $\mathbf{w}^{k+1}$ will be the solution of

$$\begin{aligned} 0 &= \Psi'((I_z^{k+1})^2) \cdot (I_x^k I_z^{k+1}) \\ &\quad -\alpha \operatorname{div} \left(\Psi'(|\nabla_3 u^{k+1}|^2 + |\nabla_3 v^{k+1}|^2)\nabla_3 u^{k+1}\right) \\ 0 &= \Psi'((I_z^{k+1})^2) \cdot (I_y^k I_z^{k+1}) \\ &\quad -\alpha \operatorname{div} \left(\Psi'(|\nabla_3 u^{k+1}|^2 + |\nabla_3 v^{k+1}|^2)\nabla_3 v^{k+1}\right). \end{aligned} \quad (8)$$

As soon as a fixed point in $\mathbf{w}^k$ is reached, we change to the next finer scale and use this solution as initialisation for the fixed point iteration on this scale.

Notice that we have a fully implicit scheme for the smoothness term and a semi-implicit scheme for the data term. Implicit schemes are used to yield higher stability and faster convergence. However, this new system is still nonlinear because of the nonlinear function Ψ' and the symbols I_*^{k+1}. In order to remove the nonlinearity in I_z^{k+1}, a first order Taylor expansion is used:

$$I_z^{k+1} \approx I_z^k + I_x^k du^k + I_y^k dv^k,$$

where $u^{k+1} = u^k + du^k$ and $v^{k+1} = v^k + dv^k$. So we split the unknowns u^{k+1}, v^{k+1} in the solutions of the previous iteration step u^k, v^k and unknown increments du^k, dv^k. For better readability let

$$\begin{aligned} (\Psi')^k_{Data} &:= \Psi'\left((I_z^k + I_x^k du^k + I_y^k dv^k)^2\right) \\ (\Psi')^k_{Smooth} &:= \Psi'(|\nabla_3(u^k + du^k)|^2 + |\nabla_3(v^k + dv^k)|^2), \end{aligned} \quad (9)$$

where $(\Psi')^k_{Data}$ can be interpreted as a robustness factor in the data term, and $(\Psi')^k_{Smooth}$ as a diffusivity in the smoothness term. With this the first equation in system (8) can be written as

$$\begin{aligned} 0 &= (\Psi')^k_{Data} \cdot \left(I_x^k \left(I_z^k + I_x^k du^k + I_y^k dv^k\right)\right) \\ &\quad - \alpha \operatorname{div} \left((\Psi')^k_{Smooth} \nabla_3(u^k + du^k)\right), \end{aligned} \quad (10)$$

[1] Since the grid size in both x- and y-direction is reduced by η, the image size in fact shrinks with a factor η^2 at each scale.

and the second equation can be expressed in a similar way. This is still a nonlinear system of equations for a fixed k, but now in the unknown increments du^k, dv^k. As the only remaining nonlinearity is due to Ψ', and Ψ has been chosen to be a convex function, the remaining optimisation problem is a convex problem, i.e. there exists a unique minimum solution.

In order to remove the remaining nonlinearity in Ψ', a second, inner, fixed point iteration loop is applied. Let $du^{k,0} := 0$, $dv^{k,0} := 0$ be our initialisation and let $du^{k,l}, dv^{k,l}$ denote the iteration variables at some step l. Furthermore, let $(\Psi')^{k,l}_{Data}$ and $(\Psi')^{k,l}_{Smooth}$ denote the robustness factor and the diffusivity defined in (9) at iteration k, l. Then finally the *linear* system of equations in $du^{k,l+1}, dv^{k,l+1}$ reads

$$\begin{aligned} 0 &= (\Psi')^{k,l}_{Data} \cdot \left(I_x^k \left(I_z^k + I_x^k du^{k,l+1} + I_y^k dv^{k,l+1}\right)\right) \\ &\quad - \alpha \operatorname{div} \left((\Psi')^{k,l}_{Smooth} \nabla_3(u^k + du^{k,l+1})\right) \end{aligned} \quad (11)$$

for the first equation. Using standard discretisations for the derivatives, the resulting sparse linear system of equations can now be solved with common numerical methods, such as Gauss-Seidel or SOR iterations. Expressions of type $I(\mathbf{x} + \mathbf{w}^\mathbf{k})$ are computed by means of bilinear interpolation.

4 Relation to Warping Methods

Coarse-to-fine warping techniques are a frequently used tool for improving the performance of optic flow methods [3, 7, 18]. While they are often introduced on a purely experimental basis, we show in this section that they can be theoretically justified as a numerical approximation. Let us simplify the model by assuming solely spatial smoothness, as in [18]. Under these condition, (10) can be written as

$$\begin{aligned} -(\Psi')^k_{Data} I_z^k \nabla I^k &= (\Psi')^k_{Data} \nabla I^k (\nabla I^k)^\top \begin{pmatrix} du^k \\ dv^k \end{pmatrix} \\ &\quad -\alpha \begin{pmatrix} \operatorname{div}\left((\Psi')^k_{Smooth} \nabla(u^k + du^k)\right) \\ \operatorname{div}\left((\Psi')^k_{Smooth} \nabla(v^k + dv^k)\right) \end{pmatrix} \end{aligned}$$

For a fixed k, this system is equivalent to the Euler–Lagrange equations described in [18]. Also there, only the increments du and dv between the first image and the warped second image are estimated. The same increments appear in the outer fixed point iterations of our approach in order to resolve the nonlinearity of the grey value constancy assumption. *This shows that the warping technique implements the minimisation of a non-linearised constancy assumption by means of fixed point iterations on* $\mathbf{w}$.

In earlier approaches, the main motivation for warping has been the coarse-to-fine strategy. Due to solutions u and v computed on coarser grids, only an increment du and

dv had to be computed on the fine grid. Thus, the estimates used to have a magnitude of less than one pixel per frame, independent of the magnitude of the total displacement. This ability to deal with larger displacements proved to be a very important aspect in differential optical flow estimation.

A second strategy to deal with large displacements has been the usage of the non-linearised grey value constancy assumption [20, 2]. Here, large displacements are allowed from the beginning. However, the nonlinearity results in a multi-modal functional. In such a setting, the coarse-to-fine strategy is not only wanted, but even necessary to better approximate the global minimum. At the end, both strategies not only lead to similar results. In fact, as we have seen above, they are completely equivalent. As a consequence, the coarse-to-fine warping technique can be formulated as a single minimisation problem, and image registration techniques relying on non-linearised constancy assumptions get access to an efficient multiresolution method for minimising their energy functionals.

5 Evaluation

For evaluation purposes, experiments with both synthetic and real-world image data were performed. The presented angular errors were computed via

$$\arccos\left(\frac{u_c u_e + v_c v_e + 1}{\sqrt{(u_c^2 + v_c^2 + 1)(u_e^2 + v_e^2 + 1)}}\right) \qquad (12)$$

where the subscripts c and e denote the correct resp. the estimated flow (cf. [5]).

Let us start our evaluation with the two variants of a famous sequence: the *Yosemite* sequence with and without cloudy sky. The original version of the sequence with cloudy sky was created by *Lynn Quam* and is available at `ftp://csd.uwo.ca` under the directory `pub/vision`. It depicts a flight through the Yosemite national park and combines divergent and translational motion. The version without clouds is available at `http://www.cs.brown.edu/people/black/images.html`.

Tab.1 shows a comparison of our results for both sequences to the best results from the literature. As one can see, our variational approach outperforms all other methods. The corresponding flow fields presented in Fig.1 give a qualitative impression of these raw numbers: They match the ground truth well.

In a second experiment we compare the results of our new linearisation-method to those which are based on the linearised modell using the optical flow constraint (2). Both modells include the same smoothness assumption so differences in the result are only ascribed to the later linerisation of the data term. The outcome listed in Table 2 shows that our methods leads to results which are about 35% better than those from the linearised case. So shifting the linearisation to the numerical approximation improves the accurancy of the algorithm even if there are no large displacements as in the Yosemite sequence.

Table 1: Comparison between the best results from the literature with 100 % density and our results for the *Yosemite* sequence with and without cloudy sky. AAE = average angular error. STD = standard deviation. 2D = spatial smoothness assumption. 3D = spatio-temporal smoothness assumption.

Yosemite with clouds

Technique	AAE	STD
Nagel [5]	10.22°	16.51°
Horn–Schunck, mod. [5]	9.78°	16.19°
Uras *et al.* [5]	8.94°	15.61°
Alvarez *et al.* [2]	5.53°	7.40°
Mémin–Pérez [17]	4.69°	6.89°
Our method (2D)	**4.63**°	**6.89**°
Our method (3D)	**3.94**°	**6.28**°

Yosemite without clouds

Technique	AAE	STD
Ju *et al.* [12]	2.16°	2.00°
Bab-Hadiashar–Suter [4]	2.05°	2.92°
Lai–Vemuri [13]	1.99°	1.41°
Our method (2D)	**1.72**°	**1.37**°
Mémin–Pérez [17]	1.58°	1.21°
Farnebäck [10]	1.14°	2.14°
Our method (3D)	**1.09**°	**1.13**°

For evaluating the performance of our method for real-world image data, the *Ettlinger Tor* traffic sequence by Nagel was used. This sequence consists of 50 frames of size 512×512. It is available at `http://i21www.ira.uka.de/image_sequences/`. In Fig. 2 the computed flow field and its magnitude are shown. Our estimation gives very realistic results, and the algorithm hardly suffers from interlacing artifacts that are present in all frames. Moreover, the flow boundaries are rather sharp and can be used directly for segmentation purposes by applying a simple thresholding step.

6 Conclusion

In this paper we have present a new numerical scheme for the minimisation of a continuous, rotationally invariant energy functional for optical flow computation based on two terms: a robust data term with a brightness constancy, com-

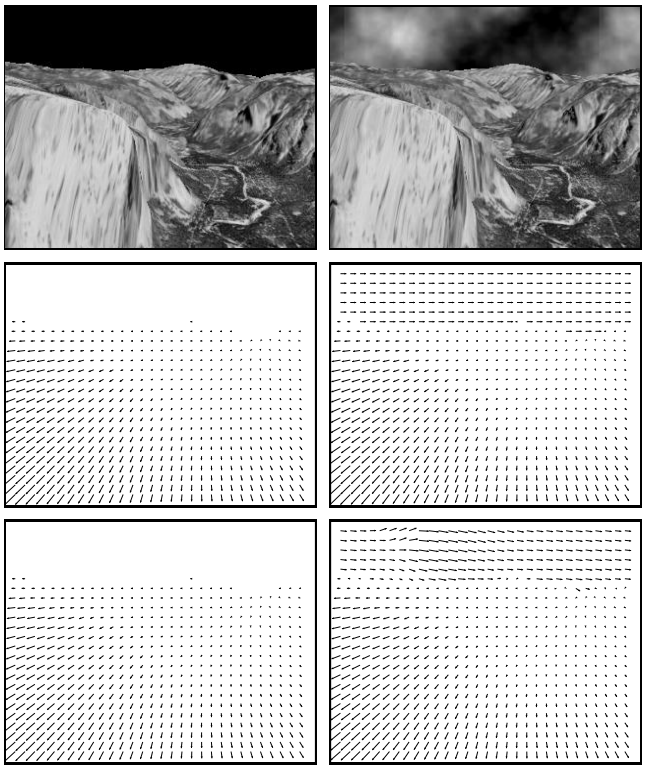

Figure 1: *(a) Top left:* Frame 8 of the *Yosemite* sequence without clouds. *(b) Top right:* Corresponding frame of the sequence *with* clouds. *(c) Middle left:* Ground truth without clouds. *(d) Middle right:* Ground truth *with* clouds. *(e) Bottom left:* Computed flow field by our 3D method for the sequence without clouds. *(f) Bottom right:* Ditto for the sequence *with* clouds.

bined with a discontinuity-preserving spatio-temporal TV regulariser. It should be stressed that we have avoided any linearisations in the data term in order to allow also for large displacements. We have shown that their combination outperforms all methods from the literature so far. One of the main reasons for this performance is the use of an energy functional with *non-linearised* data term and our strategy to consequently postpone all linearisations to the *numerical* scheme: *While linearisations in the model immediately compromise the overall performance of the system, linearisations in the numerical scheme can help to improve convergence to the global minimum.* The important result in our paper is the proof that *the widely-used warping can be theoretically justified as a numerical approximation strategy that does not influence the continuous model*.
As further work we want to transfer this numerical strategy to non-linearised data terms which are based on derivatives of the image sequence function. Our goal is to enhance the estimation of the flow field in areas where the grey value constancy assumption is disturbed like the area of the clouds in the Yosemite sequence.

Table 2: Comparison between our numerical linearisation-method and algorithms using the optical flow constraint as a linearisation of the model. AAE = average anngular error. STD = standard deviation. 2D = spatial smoothness assumption. 3D = spatio-temporal smoothness assumption.

Yosemite with clouds

	our method		ofc-method	
Technique	AAE	STD	AAE	STD
2D	4.63°	6.89°	6.12°	8.49°
3D	3.94°	6.28°	5.43°	8.24°

Yosemite without clouds

	our method		ofc-method	
Technique	AAE	STD	AAE	STD
2D	1.72°	1.37°	2.40°	1.94°
3D	1.09°	1.13°	1.57°	1.48°

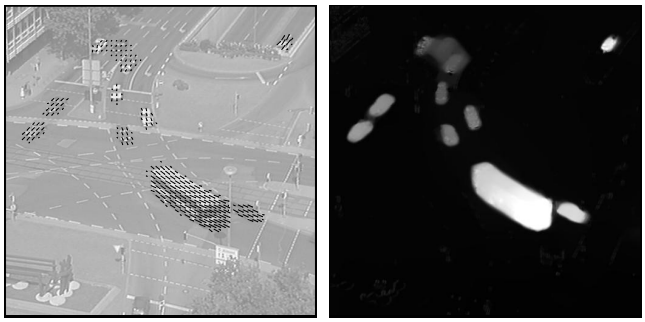

Figure 2: *(a) Left:* Computed flow field between frame 5 and 6 of the *Ettlinger Tor* traffic sequence. *(b) Right:* Computed magnitude of the optical flow field.

Acknowledgements

Our optical flow research is partly funded by the Graduiertenkolleg *Leistungsgarantien für Computersysteme* and the *Deutsche Forschungsgemeinschaft (DFG)*. This is gratefully acknowledged.

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, Las Palmas de Gran Canaria, Spain, Sept. 1999.

[2] L. Alvarez, J. Weickert, and J. Sánchez. Reliable estimation of dense optical flow fields with large displacements. *International Journal of Computer Vision*, 39(1):41–56, Aug. 2000.

[3] P. Anandan. A computational framework and an algorithm for the measurement of visual motion. *International Journal of Computer Vision*, 2:283–310, 1989.

[4] A. Bab-Hadiashar and D. Suter. Robust optic flow computation. *International Journal of Computer Vision*, 29(1):59–77, Aug. 1998.

[5] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, Feb. 1994.

[6] M. J. Black and P. Anandan. Robust dynamic motion estimation over time. In *Proc. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 292–302, Maui, HI, June 1991. IEEE Computer Society Press.

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

[8] I. Cohen. Nonlinear variational method for optical flow computation. In *Proc. Eighth Scandinavian Conference on Image Analysis*, volume 1, pages 523–530, Tromsø, Norway, May 1993.

[9] R. Deriche, P. Kornprobst, and G. Aubert. Optical-flow estimation while preserving its discontinuities: a variational approach. In *Proc. Second Asian Conference on Computer Vision*, volume 2, pages 290–295, Singapore, Dec. 1995.

[10] G. Farnebäck. Very high accuracy velocity estimation using orientation tensors, parametric motion, and simultaneous segmentation of the motion field. In *Proc. Eighth International Conference on Computer Vision*, volume 1, pages 171–177, Vancouver, Canada, July 2001. IEEE Computer Society Press.

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

[12] S. Ju, M. Black, and A. Jepson. Skin and bones: multilayer, locally affine, optical flow and regularization with transparency. In *Proc. 1996 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 307–314, San Francisco, CA, June 1996. IEEE Computer Society Press.

[13] S.-H. Lai and B. C. Vemuri. Reliable and efficient computation of optical flow. *International Journal of Computer Vision*, 29(2):87–105, Oct. 1998.

[14] M. Lefébure and L. D. Cohen. Image registration, optical flow and local rigidity. *Journal of Mathematical Imaging and Vision*, 14(2):131–147, Mar. 2001.

[15] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. Seventh International Joint Conference on Artificial Intelligence*, pages 674–679, Vancouver, Canada, Aug. 1981.

[16] E. Mémin and P. Pérez. Dense estimation and object-based segmentation of the optical flow with robust techniques. *IEEE Transactions on Image Processing*, 7(5):703–719, May 1998.

[17] E. Mémin and P. Pérez. A multigrid approach for hierarchical motion estimation. In *Proc. Sixth International Conference on Computer Vision*, pages 933–938, Bombay, India, Jan. 1998. Narosa Publishing House.

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

[19] H.-H. Nagel. Extending the 'oriented smoothness constraint' into the temporal domain and the estimation of derivatives of optical flow. In O. Faugeras, editor, *Computer Vision – ECCV '90*, volume 427 of *Lecture Notes in Computer Science*, pages 139–148. Springer, Berlin, 1990.

[20] H.-H. Nagel and W. Enkelmann. 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, 1986.

[21] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[22] C. Schnörr. Segmentation of visual motion by minimizing convex non-quadratic functionals. In *Proc. Twelfth International Conference on Pattern Recognition*, volume A, pages 661–663, Jerusalem, Israel, Oct. 1994. IEEE Computer Society Press.

[23] J. Weickert and C. Schnörr. A theoretical framework for convex regularizers in PDE-based computation of image motion. *International Journal of Computer Vision*, 45(3):245–264, Dec. 2001.

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