Computer Vision, Graphics, and Pattern Recognition Group
Department of Mathematics and Computer Science
University of Mannheim
D-68131 Mannheim, Germany

Reihe Informatik
4/1999

Optic Flow Calculations with Nonlinear Smoothness Terms Extended into the Temporal Domain

Joachim Weickert, Christoph Schnörr

Technical Report 4/1999
Computer Science Series
March 1999

Optic Flow Calculations with Nonlinear Smoothness Terms Extended into the Temporal Domain

Joachim Weickert and Christoph Schnörr
Computer Vision, Graphics and Pattern Recognition Group
Department of Mathematics and Computer Science
University of Mannheim
D-68131 Mannheim, Germany
{Joachim.Weickert, Christoph.Schnoerr}@ti.uni-mannheim.de

Abstract

Non–quadratic variational regularization is a well known and powerful approach for the discontinuity–preserving computation of optical flow. In the present paper, we introduce an extension of nonlinear spatial smoothness terms to nonlinear spatio-temporal and flow-driven regularization. To assess the performance of our approach, the implementation is purely based on the corresponding reaction–diffusion system and dispenses with any pre–smoothing typically used for canceling out noise and estimating partial derivatives. Results for real–world scenes show that our spatio–temporal approach (i) improves optical flow fields significantly, (ii) smoothes out background noise efficiently, and (iii) enhances true motion boundaries. The computational costs required are only twice as high as with a pure 2D spatial approach.

1 Introduction

Robust motion estimation is of central importance in computer vision. Motion is linked to the notion of optic flow, the distribution of apparent velocities of movement of brightness pattern in an image. Numerous methods for calculating optic flow have been proposed in the last two decades. A survey of the state-of-the-art can be found in a paper by Mitiche and Bouthemy [14], and for a performance evaluation of some of the most popular algorithms we refer to Barron *et al.* [1].

Bertero *et al.* [2] pointed out that, depending on its formulation, optic flow calculations may be ill-conditioned or even ill-posed. It is therefore common to use implicit or explicit smoothing steps in order to stabilize or regularize the process.

Implicit smoothing steps appear for instance in the robust calculation of image derivatives, where it is very common to use spatial or temporal smoothing (averaging over several frames). It is not rare that these steps are only described as algorithmic details, but indeed they are often very crucial for the quality of the algorithm.

Thus, it would be consequent to make the role of smoothing more explicit by incorporating it already in a continuous problem formulation. This way has been pioneered by Horn and Schunck [11] and improved by Nagel and Enkelmann [16] and many others. Approaches of this type calculate optic flow as the minimizer of an energy functional, which consists of a data term and a smoothness term. The data term involves optic flow constraints such as the assumption that corresponding pixels in different frames should reveal the same grey value.

The smoothness term usually requires that the optic flow field should vary smoothly in space. Such a term may be modified in an *image-driven* way in order to suppress smoothing across image boundaries; see e.g. [16]. More recently also *flow-driven* modifications have been proposed which reduce smoothing across flow discontinuities [3, 5, 6, 13, 20, 23]. These nonlinear methods have already led to rather good results in spite of the fact that the smoothness term imposed only *spatial* smoothness of the flow field.

The goal of this paper is to indroduce an extension of spatial flow-driven smoothness terms into *spatio-temporal flow-driven* regularizations. Such an extension is natural and leads to equations which are hardly more complicated than in the pure spatial case. Our experiments on real-world sequences, however, indicate that this approach leads to significantly robuster results, and that it is no longer necessary to use additional implicit smoothing operations. Our research is in line with a growing tendency to consider a larger number of frames or the complete image sequence in order to improve motion analysis; see e.g. Jähne [12]

and the references therein.

To the best of our knowledge, temporal flow-driven smoothness terms have not been used in the computer vision community before. The only extension of a smoothness constraint into the temporal domain we are aware of is a proposal by Nagel [15]. He suggested a spatio-temporal smoothness constraint which was image-driven. No experiments have been presented there.

Our paper is organized as follows. In Section 2 we review optic flow approaches with spatial smoothness terms, and Section 3 describes our novel method using a spatio-temporal smoothness constraint. A simple numerical algorithm is sketched in Section 4, and Section 5 illustrates the potential of our approach by applying it to real-world image sequences. The paper is concluded with a summary in Section 6.

2 Spatial Smoothness Terms

Let us denote an image sequence by some function $f(x, y, z)$ where (x,y) denotes the location within some rectangular image domain Ω and z is the time. In its simplest case, optic flow calculations based on energy functionals determine the optic flow vector $(u, v)^T$ based on two assumptions:

1. Corresponding features are supposed to maintain their intensity over time. This leads to the *optic flow constraint (OFC) equation*

$$f_x u + f_y v + f_z = 0, \tag{1}$$

where the subscripts denote partial derivatives. Numerous generalizations to multiple constraint equations and/or different "conserved quantities" (replacing intensity) exist; see e.g. [8, 21].
Evidently such a single equation (1) is not sufficient to determine the two unknown functions u and v uniquely. In order to obtain a unique flow field, a second constraint is needed.

2. Such a second constraint may impose that the flow field should vary smoothly in space. This can be expressed by requiring that

$$\int_\Omega \Psi\left(|\nabla u|^2 + |\nabla v|^2\right) dx\, dy \tag{2}$$

should be close to 0, where $\Psi : \mathbb{R} \to \mathbb{R}$ is an increasing differentiable function and $\nabla := (\partial_x, \partial_y)^T$ denotes the 2D nabla operator. This assumption is called *smoothness constraint.* In the sequel we shall assume that $\Psi(s^2)$ is convex in s. Examples for Ψ will be presented at the end of this section.

In order to satisfy both the optic flow and smoothness constaint as good as possible, they are assembled into an energy functional to be minimized:

$$\begin{aligned} E(u,v) &:= \int_\Omega \Big((f_x u + f_y v + f_z)^2 \\ &+ \alpha\, \Psi(|\nabla u|^2 + |\nabla v|^2) \Big)\, dx\, dy \end{aligned} \tag{3}$$

where the *regularization parameter* $\alpha > 0$ specifies the weight of the second summand *(smoothness term)* relative to the first one *(data term).* Larger values for α lead to smoother flow fields.

It is a classic result from the calculus of variations [7] that a solution (u, v) minimizing an energy functional of type

$$E(u,v) := \int_\Omega F(x, y, u, v, \nabla u, \nabla v)\, dx\, dy \tag{4}$$

satisfies necessarily the so-called Euler equations

$$\partial_x F_{u_x} + \partial_y F_{u_y} - F_u = 0, \tag{5}$$
$$\partial_x F_{v_x} + \partial_y F_{v_y} - F_v = 0 \tag{6}$$

with reflecting boundary conditions:

$$\partial_n u = 0 \quad \text{on } \partial\Omega, \tag{7}$$
$$\partial_n v = 0 \quad \text{on } \partial\Omega. \tag{8}$$

Hereby, $\partial\Omega$ denotes the boundary of the image domain with normal vector n.

Applying this relation it is easily seen that minimization of (3) corresponds to solving

$$\begin{aligned} 0 &= \nabla\cdot \left(\Psi'\left(|\nabla u|^2 + |\nabla v|^2\right) \nabla u\right) \\ &\quad - \tfrac{1}{\alpha} f_x (f_x u + f_y v + f_z), \end{aligned} \tag{9}$$
$$\begin{aligned} 0 &= \nabla\cdot \left(\Psi'\left(|\nabla u|^2 + |\nabla v|^2\right) \nabla v\right) \\ &\quad - \tfrac{1}{\alpha} f_y (f_x u + f_y v + f_z), \end{aligned} \tag{10}$$

where Ψ' is the derivative of Ψ with respect to its argument, and $\nabla\cdot$ denotes the 2D divergence operator, i.e. $\nabla\cdot \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) := \partial_x a + \partial_y b$. The preceding equations can be recovered as the steady-state of the diffusion–reaction processes

$$\begin{aligned} u_t &= \nabla\cdot \left(\Psi'\left(|\nabla u|^2 + |\nabla v|^2\right) \nabla u\right) \\ &\quad - \tfrac{1}{\alpha} f_x (f_x u + f_y v + f_z), \end{aligned} \tag{11}$$
$$\begin{aligned} v_t &= \nabla\cdot \left(\Psi'\left(|\nabla u|^2 + |\nabla v|^2\right) \nabla v\right) \\ &\quad - \tfrac{1}{\alpha} f_y (f_x u + f_y v + f_z). \end{aligned} \tag{12}$$

The diffusion time t is an artificial evolution parameter which should not be mixed up with the time z

of the image sequence $f(x, y, z)$. For $t \to \infty$, the solution (u, v) gives the minimum of $E(u, v)$. It is unique if $\Psi(s^2)$ is convex in s. We may also regard the preceding diffusion–reaction system as a gradient descent method for minimizing (3).

The diffusivity in both equations is given by $\Psi'\left(|\nabla u|^2+|\nabla v|^2\right)$. It steers the activity of the smoothing process: diffusion is strong at locations where the diffusivity is large, and smoothing is reduced at places where the diffusivity is small. We shall now consider some examples which demonstrate how the choice of Ψ influences the smoothing process.

1. Horn and Schunck [11] considered the linear case $\Psi(s^2) = s^2$. This corresponds to the constant diffusivity $\Psi'(s^2) = 1$. Therefore, the smoothing activity of the Horn and Schunck method does not depend on the flow variation $s^2 = |\nabla u|^2 + |\nabla v|^2$. As a consequence, the flow is also smoothed across motion boundaries. This explains a well-known drawback of this method: a blurry flow field which is ignorant of the true motion boundaries.

2. Many modifications have been proposed to alleviate this problem. Nagel and Enkelmann [16] for instance reduce diffusion across image boundaries with large $|\nabla f|$. Thus, their method considers an *image-driven* smoothness term for the flow field. In many cases this modification outperforms the Horn and Schunck approach. In specific situations, however, image discontinuities may not coincide with flow discontinuities: strongly textured objects, for example, may have numerous texture edges which are not motion boundaries. In such situations an image-driven smoothness term would lead to an oversegmentation and it would be desirable to replace it by one which respects flow discontinuities instead of image discontinuities.

3. A *flow-driven* smoothness term can be constructed by using a nonlinear smoothness potential $\Psi(s^2)$ with a decreasing diffusivity $\Psi'(s^2)$. This ensures that the smoothing is reduced at locations where the flow magnitude is small. Such methods have been considered by Schnörr [20] and Weickert [23]. In [23] the potential

$$\Psi(s^2) = \lambda^2 \sqrt{1 + s^2/\lambda^2} \qquad (\lambda > 0). \tag{13}$$

is used, leading to the diffusivity [4]

$$\Psi'(s^2) = \frac{1}{\sqrt{1 + s^2/\lambda^2}}. \tag{14}$$

We observe that λ can be regarded as a *contrast parameter*: If the flow variation $s^2 = |\nabla u|^2+|\nabla v|^2$ is large compared to λ^2, then the diffusivity is close to 0, and for $s^2 << \lambda^2$ the diffusivity tends to 1. Choosing a very small value for λ relates this method to total variation regularization, a powerful denoising technique permitting discontinuous solutions [18].

4. Other flow-driven smoothness terms from the literature [5, 6, 13] replace the smoothness potential $\Psi(|\nabla u|^2+|\nabla v|^2)$ by $\Psi(|\nabla u|^2) + \Psi(|\nabla v|^2)$. This leads to two diffusion–reaction equations where the joint diffusivity $\Psi'(|\nabla u|^2+|\nabla v|^2)$ is replaced by $\Psi'(|\nabla u|^2)$ and $\Psi'(|\nabla v|^2)$, respectively. Hence, the coupling between the two equations becomes weaker and flow discontinuities may be formed at different locations for u and v. Problems may also arise from the fact that in this case the energy functional is not necessarily rotationally invariant.

3 Spatio-Temporal Smoothness Terms

Using the knowledge from the previous section it is straightforward to extend the smoothness constraint into the temporal domain. Instead of calculating the optic flow (u, v) as the minimizer of the two-dimensional integral (3) for each time frame z, we now minimize a single three-dimensional integral whose solution is the optic frow for *all* frames $z \in [0, T]$:

$$E(u, v) := \int_{\Omega \times [0,T]} \Big((f_x u + f_y v + f_z)^2 + \alpha\, \Psi(|\nabla_3 u|^2 + |\nabla_3 v|^2) \Big)\, dx\, dy\, dz \tag{15}$$

where $\nabla_3 := (\partial_x, \partial_y, \partial_z)^T$ denotes the spatio-temporal nabla operator. In the same way as in Section 2 we derive that the minimizer can be recovered as the solution of the diffusion–reaction system

$$\begin{aligned} u_t &= \nabla_3 \cdot \left(\Psi'\left(|\nabla_3 u|^2+|\nabla_3 v|^2\right) \nabla_3 u\right) \\ &\quad -\tfrac{1}{\alpha} f_x (f_x u + f_y v + f_z), \qquad (16) \\ v_t &= \nabla_3 \cdot \left(\Psi'\left(|\nabla_3 u|^2+|\nabla_3 v|^2\right) \nabla_3 v\right) \\ &\quad -\tfrac{1}{\alpha} f_y (f_x u + f_y v + f_z). \qquad (17) \end{aligned}$$

for $t \to \infty$. In the present paper we study this process for the nonlinear potential given in (13).

The diffusion part in (16)–(17) is closely related to nonlinear diffusion filters for regularizing three-dimensional vector-valued images. Such methods have first been applied by Gerig *et al.* [9] in the context of medical imaging. The latter approach, however, uses

diffusivities from [17] which may create ill-posed processes. This cannot happen in our case, where convex smoothness terms in the energy functional create well-posed diffusion–reaction processes [19]. For an extensive discussion of nonlinear diffusion filtering, we refer to [10, 22].

4 Numerical Aspects

We approximate the 2-D diffusion–reaction system (11)–(12) and its 3-D counterpart (16)–(17) by finite differences. Derivatives in x, y and z are approximated by central differences, and for the discretization in t direction we use a slightly modified explicit (Euler forward) scheme. It should be stressed that no (!) additional presmoothing or postprocessing is applied.

We used the time step size $\tau = 1/4$ in the 2-D case and $\tau = 1/6$ in the 3-D case. The iterations were stopped when the Euclidean norm of the relative residue was below 0.001.

It should be noted that the 3-D scheme uses only about two times the computational effort of a corresponding 2-D scheme when applied to all subsequent frame pairs of an image sequence. The main difference is an increased memory effort, since, in the 3-D case, the whole sequence is processed simultaneously. For the typical test sequences in computer vision, this does not lead to problems when conventional workstations are used.

5 Experiments

In this section we illustrate the behaviour of our method by applying it to two real-image sequences shown in Fig. 1 and Fig. 3. The second one is available via anonymous ftp from `ftp://csd.uwo.ca` under the directory `pub/vision`.

We used 16 frames for the first sequence, 20 frames for the second one, and compared pure 2D processing (eqns. (11)–(12)) with 3D processing (eqns. (16)–(17)) using the same parameter values. For better visibility, vectors w with $|w| < 0.2$ pixels have not been drawn.

Comments and a description for each result are given in the figure captions below.

6 Conclusion and Further Work

We have presented an nonlinear *spatio–temporal* regularization approach for the computation of piecewise smooth optic flow. Our results show a significant improvement over pure 2D processing at low additional computational costs.

Our future work will include a critical performance evaluation (cf. [1]) and research on highly efficient numerical methods for the nonlinear variational regularization proposed in this paper.

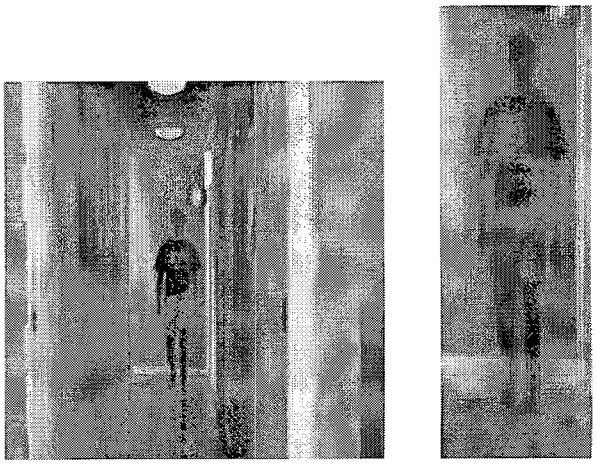

Figure 1: The hallway sequence. A person (see the section right) is moving towards the camera.

Acknowledgements

We thank Ole Fogh Olsen and Mads Nielsen (Department of Computer Science, University of Copenhagen) for providing the hallway sequence.

References

[1] J.L. Barron, D.J. Fleet, S.S. Beauchemin, *Performance of optical flow techniques*, Int. J. Comput. Vision, Vol. 12, 43–77, 1994.

[2] M. Bertero, T.A. Poggio, V. Torre, *Ill-posed problems in early vision*, Proc. IEEE, Vol. 76, 869–889, 1988.

[3] M.J. Black, P. Anandan, *The robust estimation of multiple motions: Parametric and piecewise–smooth flow fields*, Comp. Vision and Image Underst., Vol. 63, 75–104, 1996.

[4] P. Charbonnier, L. Blanc–Féraud, G. Aubert, M. Barlaud, *Two deterministic half-quadratic regularization algorithms for computed imaging*, Proc. IEEE Int. Conf. Image Processing (ICIP–94, Austin, Nov. 13–16, 1994), Vol. 2, IEEE Computer Society Press, Los Alamitos, 168–172, 1994.

[5] I. Cohen, *Nonlinear variational method for optical flow computation*, Proc. Eighth Scandinavian Conf. on Image Analysis (SCIA '93, Tromsø, May 25–28, 1993), Vol. 1, 523–530, 1993.

[6] R. Deriche, P. Kornprobst, G. Aubert, *Optical-flow estimation while preserving its discontinuities: A variational approach*, Proc. Second Asian Conf. Computer Vision (ACCV '95, Singapore, December 5–8, 1995), Vol. 2, 290–295, 1995.

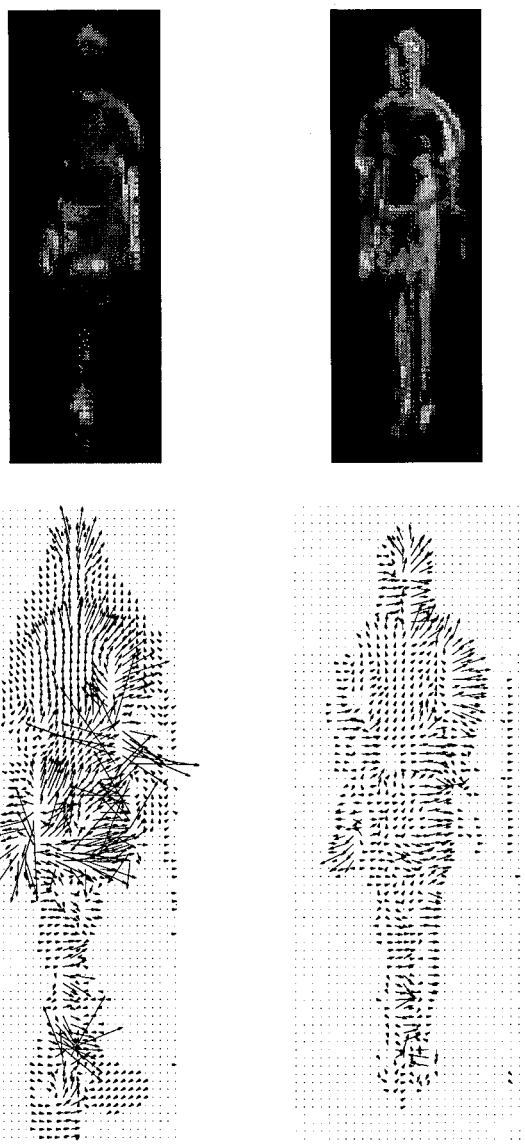

Figure 2: Results computed using the section shown in Figure 1, right. **Top:** The magnitudes of velocity vectors shown as grey value plot. **Top left:** The result using 2D processing (eqns. (11)–(12)). **Top right:** The result using 3D processing (eqns. (16)–(17)). Note how outliers dominate the image on the left such that other regions get scaled down. The right image shows that the proposed extension of adaptive smoothing to the temporal axis gives a much more coherent and complete result. **Bottom:** The velocity vector fields corresponding to the images above (subsampled by factor 2 for better visibility).

Figure 3: The well–known Hamburg taxi scene.

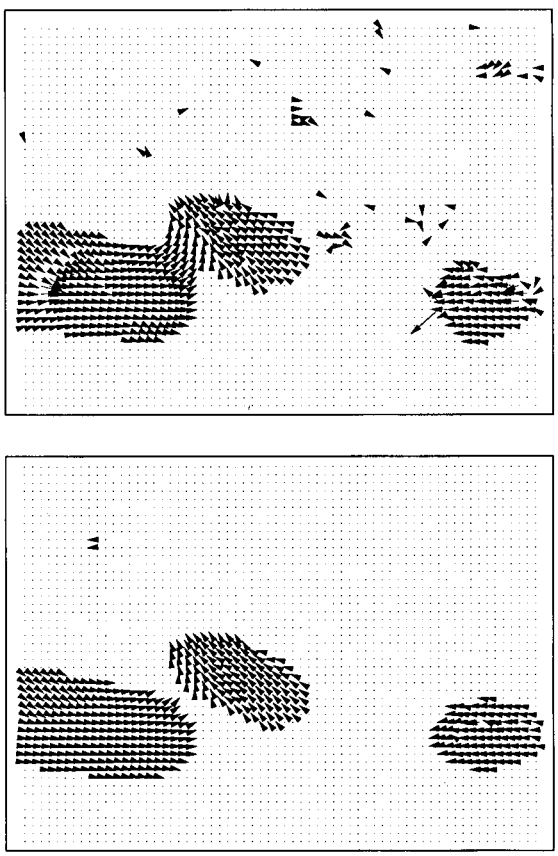

Figure 4: Results computed using the taxi scene (subsampled by factor 4 for better visibility). **Top:** 2D processing. **Bottom:** 3D processing. A comparison clearly shows that the proposed extension of adaptive smoothing to the temporal axis (i) improves the vector fields significantly, (ii) smoothes out background noise, and (iii) enhances true motion boundaries.

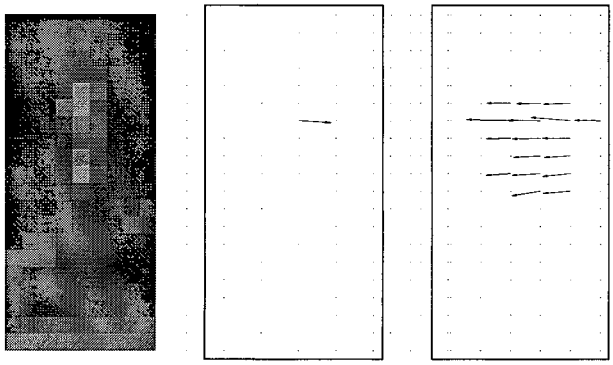

Figure 5: **Left:** Section showing the pedestrian of the upper left part of the taxi scene in Figure 3. **Middle:** The corresponding section from Figure 4, top. The vector fields have been scaled by factor 4 for better visibility. The 2D processing smoothes out the noisy local motion data (normal flow). **Right:** The corresponding section from Figure 4, bottom. The 3D processing computes a coherent flow field for the "rigid part" of the pedestrian. Regions with moving limbs are interpreted as noise at such a small spatial scale.

[7] L.E. Elsgolc, *Calculus of variations*, Pergamon, Oxford, 1961.

[8] D.J. Fleet, A.D. Jepson, *Computation of component image velocity from local phase information*, Int. J. Comput. Vision, Vol. 5, 77–104, 1990

[9] G. Gerig, O. Kübler, R. Kikinis, F.A. Jolesz, *Nonlinear anisotropic filtering of MRI data*, IEEE Trans. Medical Imaging, Vol. 11, 221–232, 1992.

[10] B.M. ter Haar Romeny (Ed.), *Geometry-driven diffusion in computer vision*, Kluwer, Dordrecht, 1994.

[11] B. Horn, B. Schunck, *Determining optical flow*, Artif. Intell., Vol. 17, 185–203, 1981.

[12] B. Jähne, *Spatio-temporal image processing*, Lecture Notes in Computer Science, Vol. 751, Springer, Berlin, 1993.

[13] A. Kumar, A.R. Tannenbaum, G.J. Balas, *Optic flow: a curve evolution approach*, IEEE Trans. Image Proc., Vol. 5, 598–610, 1996.

[14] A. Mitiche, P. Bouthemy, *Computation and analysis of image motion: a synopsis of current problems and methods*, Int. J. Comput. Vision, Vol. 19, 29–55, 1996.

[15] H.H. Nagel, *Extending the 'oriented smoothness constraint' into the temporal domain and the estimation of derivatives of optical flow*, O. Faugeras (Ed.), Computer vision – ECCV '90, Lecture Notes in Computer Science, Vol. 427, Springer, Berlin, 139–148, 1990.

[16] H.H. Nagel, W. Enkelmann, *An investigation of smoothness constraints for the estimation of displacement vector fields from images sequences*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 8, 565–593, 1986.

[17] P. Perona, J. Malik, *Scale space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 12, 629–639, 1990.

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

[19] C. Schnörr, *Unique reconstruction of piecewise smooth images by minimizing strictly convex non-quadratic functionals*, J. Math. Imag. Vision, Vol. 4, 189–198, 1994.

[20] C. Schnörr, *Segmentation of visual motion by minimizing convex non-quadratic functionals*, Proc. 12th Int. Conf. Pattern Recognition (ICPR 12, Jerusalem, Oct. 9–13, 1994), Vol. A, IEEE Computer Society Press, Los Alamitos, 661–663, 1994.

[21] J. Weber, J. Malik, *Robust computation of optical flow in a multi-scale differential framework*, Int. J. Comput. Vision, Vol. 14, 67–81, 1995.

[22] J. Weickert, *Anisotropic diffusion in image processing*, Teubner-Verlag, Stuttgart, 1998.

[23] J. Weickert, *On discontinuity-preserving optic flow*, S. Orphanoudakis, P. Trahanias, J. Crowley, N. Katevas (Eds.), Proc. Computer Vision and Mobile Robotics Workshop (CVMR '98, Santorini, Sept. 17–18, 1998), 115–122, 1998.