Optical Flow Estimation with Consistent Spatio-temporal Coherence Models

Javier Sánchez, Agustín Salgado and Nelson Monzón
Centro de Tecnologías de la Imagen (CTIM), Departamento de Informática y Sistemas, University of Las Palmas de Gran Canaria, Las Palmas de Gran Canaria, Spain

Keywords: Optical Flow, Variational Methods, PDE, Temporal Coherence.

Abstract: In this work we propose a new variational model for the consistent estimation of motion fields. The aim of this work is to develop appropriate spatio-temporal coherence models. In this sense, we propose two main contributions: a nonlinear flow constancy assumption, similar in spirit to the nonlinear brightness constancy assumption, which conveniently relates flow fields at different time instants; and a nonlinear temporal regularization scheme, which complements the spatial regularization and can cope with piecewise continuous motion fields. These contributions pose a congruent variational model since all the energy terms, except the spatial regularization, are based on nonlinear warpings of the flow field. This model is more general than its spatial counterpart, provides more accurate solutions and preserves the continuity of optical flows in time. In the experimental results, we show that the method attains better results and, in particular, it considerably improves the accuracy in the presence of large displacements.

1 INTRODUCTION

The estimation of motion fields is a key problem in computer vision. It serves as a basis for many applications, such us stereoscopic vision and 3D scene reconstruction, medical image analysis, structure from motion, object tracking and others. If we are given a video sequence, and we want to find the motion of the objects in the images, our method should provide a solution that is consistent through the sequence. In this work, we address the problem of temporal coherence in optical flow methods. The aim is to devise new methods that allow finding continuous flow fields in time.

Optical flow methods can be further improved if temporal information is properly managed (Weickert and Schnörr, 2001). In this work, the authors propose a method that is a straight extrapolation of the spatial coherence model to the temporal dimension, based on a continuous spatio-temporal regularization scheme. More recently, some authors have generalized the use of the flow temporal derivative. Typically, the temporal information is coupled with the spatial gradient in the form of a non-quadratic 3D smoothing operator. However, in (Sánchez et al., 2012), the authors analyze the behavior of a continuous temporal regularizer and show several experiments where it fails.

Black (Black, 1994) uses robust functionals to deal with outliers and introduces a *temporal continuity* strategy to account for the temporal coherence of the sequence. This temporal continuity is based on a prediction step and an attachment of the flow to the predicted value. It warps the flow field to estimate its value in the following frame. This is interesting, because the warping allows finding the correct flow correspondences. More recently, there has been several works dealing with temporal coherence in different ways: for instance, in (Sun et al., 2010) the temporal consistency is established reasoning on the segmentation on layers.

We propose several contributions: on the one hand, we introduce a nonlinear flow constancy assumption that fits with the nonlinear data assumption; on the other hand, we propose a novel nonlinear flow regularization scheme that can deal with non-continuous optical flows. Another contribution is a new anistropic diffusion operator based on the Nagel-Enkelmann operator. This new operator allows respecting the object boundaries during the diffusion process, at the same time that it avoids oversegmentation in texture regions.

The former contribution was motivated by the results presented in (Salgado and Sánchez, 2006). The experimental results showed that the use of a nonlin-

Sánchez J., Salgado A. and Monzón N..
Optical Flow Estimation with Consistent Spatio-temporal Coherence Models.
DOI: 10.5220/0004199903660369
In *Proceedings of the International Conference on Computer Vision Theory and Applications* (VISAPP-2013), pages 366-369
ISBN: 978-989-8565-48-8

ear temporal formulation of the flow field provided very good results. That was the first time that such a nonlinear flow assumption was introduced. For the second contribution, we introduce a non-continuous flow regularization scheme at the PDE level. This is a pure regularization approach that replaces the traditional continuous temporal smoothing.

In Section 2 we examine the new energy model and explain the novel temporal coherence strategy. The minimization of the energy model and some numerical details are explained in Section 3. In the experimental results – Section 4 – we test our method using a synthetic sequence. Finally the conclusions in Section 5.

2 NONLINEAR VARIATIONAL MODEL

If we have a set of images $I_j(\mathbf{x})$, with $j = 1,..,N$, N the number of frames and $\mathbf{x} = (x,y)$, the aim is to find a set of optical flow functions, $\{\mathbf{h}_i(\mathbf{x})\}$, with $i = 1,..,N-1$. We decompose our energy functional in two separate parts:

$$E\left(\{\mathbf{h}_i(\mathbf{x})\}\right) = E_S\left(\{\mathbf{h}_i(\mathbf{x})\}\right) + E_T\left(\{\mathbf{h}_i(\mathbf{x})\}\right). \quad (1)$$

The first term on the right, E_S, stands for the spatial energy model and the second term, E_T, is the energy model corresponding to the temporal coherence strategy. The spatial model reads as follows:

$$\begin{aligned} E_S =& \int \sum_{i=1}^{N-1} \Psi\left(\left(I_i(\mathbf{x}) - I_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))\right)^2\right) \mathbf{dx} \\ &+ \gamma \int \sum_{i=1}^{N-1} \Psi\left(\|\nabla I_i(\mathbf{x}) - \nabla I_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))\|^2\right) \mathbf{dx} \\ &+ \alpha \int \sum_{i=1}^{N-1} \Psi\left(\mathcal{N}(\nabla I_i, \nabla \mathbf{h}_i)\right) \mathbf{dx}, \end{aligned} \quad (2)$$

with $\Psi\left(s^2\right) = \sqrt{s^2+\varepsilon^2}$ (ε a prefixed small constant, e.g. 0.01). This kind of function mitigates the effect of outliers and behaves like TV regularization approaches when used in the smoothness term. The advantage of a TV smoothing scheme is that it preserves discontinuities of the flow. We use the anisotropic diffusion operator, $\mathcal{N}(\nabla I_i, \nabla \mathbf{h}_i) = trace\left(\nabla \mathbf{h}_i^T(\mathbf{x}) D(\nabla I_i) \nabla \mathbf{h}_i(\mathbf{x})\right)$, proposed in (Nagel and Enkelmann, 1986), which preserves discontinuities of the images in the flow field, $D(.)$ defined as:

$$D(\nabla I) = \frac{\nabla I^{\perp T} \nabla I^{\perp} + \lambda^2 \mathbf{Id}}{\|\nabla I\|^2 + 2\lambda^2},$$

with $\mathbf{Id}$ the identity matrix. λ determines the gradient value from which the anisotropy is activated. This parameter can be computed from the more intuitive *isotropic fraction*, $0 \leq s \leq 1$, introduced in (Álvarez et al., 2000).

For the temporal energy model, we follow the ideas presented in (Salgado and Sánchez, 2006). Given that an object in the sequence may undergo large displacements, we have to deal with information that is warped through the flows. In fact, given a flow $\mathbf{h}_i(\mathbf{x})$, at instant i, its corresponding flow in the following time instant is $\mathbf{h}_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))$. If $\mathbf{h}_i(\mathbf{x})$ is large, then the temporal derivative cannot be computed, but the previous correspondence still holds. Thus, one way to relate motion fields at different time instants is through the *flow constancy assumption* (FCA), $\mathbf{h}_i(\mathbf{x}) = \mathbf{h}_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))$. Therefore, the temporal coherence model, E_T, can be formulated as,

$$E_T = \beta \int \sum_{i=1}^{N-2} \Phi\left(\|\mathbf{h}_i(\mathbf{x}) - \mathbf{h}_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))\|^2\right) \mathbf{dx}, \quad (3)$$

with $\Phi\left(s^2\right) = e^{-\|\nabla I\|^{\kappa}} \sqrt{s^2+\varepsilon^2}$, with $\kappa = 0.8$ and $\varepsilon = 0.01$.

This term is congruent with the brightness and gradient constancy terms. In the presence of large displacements, this temporal model is coherent with the spatial formulation and relates values at the correct positions. Note that when object displacements are very small, this term can be seen as an approximation of the temporal derivative of the flow, which has shown to be effective in a continuous setting (e.g., (Weickert and Schnörr, 2001) or (Papenberg et al., 2006)).

3 MINIMIZING THE ENERGY MODEL

In this section we derive the Euler-Lagrange equations of (2) and (3). Then, we introduce a nonlinear regularization scheme at the PDE, which closely resembles a continuous temporal smoothing approach.

The Euler-Lagrange equations for the spatial energy model (2) are:

$$\begin{aligned} \mathbf{0} =& \Psi'\left(\left(I_i(\mathbf{x}) - I_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))\right)^2\right) \\ & \cdot \left(I_i(\mathbf{x}) - I_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))\right) \cdot \nabla I_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x})) \\ & + \gamma\, \Psi'\left(\|\nabla I_i(\mathbf{x}) - \nabla I_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))\|^2\right) \\ & \cdot \left(\nabla I_i(\mathbf{x}) - \nabla I_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))\right) \cdot \mathcal{H} I_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x})) \\ & + \alpha \operatorname{div}\left(\Psi'\left(\mathcal{N}(\nabla I_i, \nabla \mathbf{h}_i)\right) \cdot D\left(\nabla I_i\right) \cdot \nabla \mathbf{h}_i\right), \end{aligned} \quad (4)$$

where $\mathcal{H}I_{i+1}$ is the Hessian matrix. The temporal energy model (3) yields the following Euler-Lagrange equations:

$$\begin{aligned}\mathbf{0} =& \beta\, \Phi'\left(\|\mathbf{h}_i(\mathbf{x}) - \mathbf{h}_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))\|^2\right) \\ &\cdot \Big((\mathbf{h}_i(\mathbf{x}) - \mathbf{h}_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x})))^T \\ &\cdot \left(\mathbf{Id} - \nabla\mathbf{h}^T_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))\right)\Big) \\ &+ \beta\, \Phi'\left(\|\mathbf{h}_i(\mathbf{x}) - \mathbf{h}_{i-1}(\mathbf{x}+\mathbf{h}^*_{i-1}(\mathbf{x}))\|^2\right) \\ &\cdot \left((\mathbf{h}_i(\mathbf{x}) - \mathbf{h}_{i-1}(\mathbf{x}+\mathbf{h}^*_{i-1}(\mathbf{x}))) \cdot |\mathcal{J}(\mathbf{x})|\right), \qquad (5)\end{aligned}$$

where $|\mathcal{J}(\mathbf{x})|$ stands for the absolute value of the Jacobian matrix, with $\mathcal{J}(\mathbf{x}) = \left(1+u^*_{i-1,x}\right)\left(1+v^*_{i-1,y}\right) - u^*_{i-1,y}v^*_{i-1,x}$. $\mathbf{h}^*_{i-1} = \left(u^*_{i-1}, v^*_{i-1}\right)^T$ is the backward flow from frame I_i to I_{i-1}.

In order to derive $(\mathbf{h}_i(\mathbf{x}) - \mathbf{h}_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x})))$ with respect to $\mathbf{h}_{i+1}(\mathbf{x})$, we can use the change of variables $\mathbf{z} = \mathbf{x}+\mathbf{h}_{i-1}(\mathbf{x})$. This change allows us to remove the nonlinearity inside the flow. The backward flow, $\mathbf{h}^*_{i-1}$, naturally appears due to this change of variables.

We use a gradient descent approach to find the solution of the above PDE. The nonlinear terms, e.g. $I_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))$, are linearized using first order Taylor expansions. In the temporal coherent framework, we use Dirichlet boundary conditions for the last and first frames, whereas Neumann boundary conditions are used in the spatial domain. We use a standard coarse-to-fine strategy to deal with large displacements, based on a pyramidal structure. The system of equations is sparse, so it can be efficiently solved by means of the Gauss-Seidel or SOR method in each scale.

We introduce a nonlinear temporal smoothing scheme. Its formulation is intuitively derived from the second order temporal derivative of the flow field, $u_{tt} \approx u_{i,j,k+1} - 2u_{i,j,k} + u_{i,j,k-1}$. In the PDE, this second order derivative has a continuous temporal regularizing effect that is consistent if the flow field varies smoothly across the image sequence. We propose a new solution, which is similar in spirit to this numerical approximation, and is suitable for dealing with non-continuous displacements. This is a nonlinear formulation that puts into correspondence the correct flow values in different frames. It is not evident how to abstract this idea at the energy level in Equation (3). As before, we also use L^1 functions to turn the method more robust against outliers, in the following way:

$$\begin{aligned}T_S =& \delta\, \Phi'\left(\|\mathbf{h}_{i-1}(\mathbf{x}+\mathbf{h}^*_{i-1}(\mathbf{x})) - \mathbf{h}_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))\|^2\right) \\ &\cdot \left(\mathbf{h}_{i-1}(\mathbf{x}+\mathbf{h}^*_{i-1}(\mathbf{x})) - 2\mathbf{h}_i(\mathbf{x}) + \mathbf{h}_{i+1}(\mathbf{x}+\mathbf{h}_i(\mathbf{x}))\right) \qquad (6)\end{aligned}$$

This term provides a new scheme at the PDE level and has to be combined with the previous PDE equations (4) and (5). In the experiments, we show that this nonlinear smoothing provides very good results: it has a similar gain as in the continuous case, but it correctly handles large discontinuities in the motion field.

4 EXPERIMENTAL RESULTS

Next we examine the behavior of the temporal models introduced in equations (1) and (6). For this, we use a simple sequence of a square translating over a textured background. The square is moving 15 pixels per frame, while the background moves 3 pixels in the same direction. In the first row of Fig. 1, we show the third frame of the square sequence, its ground truth, and the best spatial solution found. In the second row, we show three temporal solutions: the first for the nonlinear temporal attachment defined in (3); the second, for the nonlinear temporal smoothnes approach defined in (6); and, finally, using both temporal terms. The color, in the motion field, represents the direction and, the intensity, its magnitude.

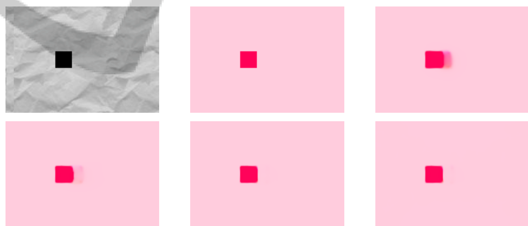

Figure 1: Square sequence. First row: one of the images of the Square sequence, the ground truth and the best spatial solution found. Second row: three temporal solutions with $\beta = 8$, $\delta = 25$ and ($\beta = 1, \delta = 25$), respectively.

The improvement of the temporal methods with respect to the spatial solution is important. As expected, the spatial method produces higher errors at the motion discontinuities and, more significantly, at the occlusions. Table 1 shows the average End-point (EPE) and Angular (AAE) errors for these results. The first temporal result, corresponding to the first image in the second row of Fig. 1, provides an important improvement on the EPE and, more noticeable, on the AAE. The improvement in accuracy is still more important if we use the nonlinear temporal smoothing scheme (Equation (6)) or a combination of both.

We observe that the nonlinear temporal smoothing scheme (6) behaves better than the temporal attachment, even at the motion boundaries. The graphics in Fig. 2 show the EPE for every frame on the

Table 1: EPE and AAE for the Square sequence.

Method	EPE	AAE
Spatial	0.071	0.629^o
Temporal 1 ($\beta = 8$)	0.049	0.204^o
Temporal 2 ($\delta = 25$)	0.036	$\mathbf{0.134}^o$
Temporal 3 ($\beta = 1, \delta = 25$)	**0.035**	0.138^o

square sequence. Frame by frame, the optical flows are more accurate in the temporal methods. We also observe that the results are very stable, especially in the middle of the 'Temporal 2 (δ)' line. Reasonably, the frames at the beginning and end of the sequence present higher errors, due to the Dirichlet boundary conditions.

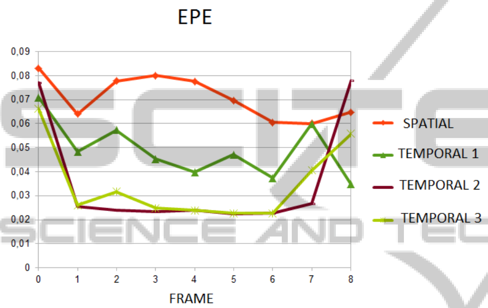

Figure 2: EPE in each optical flow of the Square sequence.

5 CONCLUSIONS

In this paper we have presented a new spatio-temporal coherence model for the consistent estimation of optical flows. We have focused on different nonlinear flow assumptions that are more confident in the estimation of motion fields than previous approaches. These nonlinear assumptions correctly fit with the standard nonlinear brightness and gradient constancy terms, can cope with general image sequences and provide better solutions. In particular, we have proposed two main contributions: on the one hand, we have introduced the nonlinear flow constancy assumption (FCA) in the energy model. This term relates flow fields at different time instants and is consistent with the rest of the energy terms. On the other hand, we have proposed a nonlinear temporal diffusion scheme at the PDE level, which produces continuous flows in time. We have seen that this new scheme is more general than using the continuous temporal regularization of the flow, with the advantage that it conveniently deals with continuous and non-continuous velocities. In fact, if the motion is very small, this term approximates a continuous temporal smoothing scheme. In the experimental results, we have shown that the method provides important accuracy improvements, specially in the presence of large displacements. The results are promising in both cases, although we observe a better performance for the nonlinear temporal smoothing scheme in general. Another interesting result of the temporal coherence schemes is that the background motion oscillations tend to disapper. These oscillations clearly appear in the spatial method, in regions where there is no apparent motion.

ACKNOWLEDGEMENTS

This work has been partly founded by the Spanish Ministry of Science and Innovation through the research project TIN2011-25488.

REFERENCES

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

Black, M. J. (1994). Recursive non-linear estimation of discontinuous flow fields. In *Proceedings of the third European conference on Computer vision (vol. 1)*, ECCV '94, pages 138–145, Secaucus, NJ, USA. Springer-Verlag New York, Inc.

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

Papenberg, N., Bruhn, A., Brox, T., Didas, S., and Weickert, J. (2006). Highly Accurate Optic Flow Computation with Theoretically Justified Warping. *International Journal of Computer Vision*, 67(2):141–158.

Salgado, A. and Sánchez, J. (2006). A temporal regularizer for large optical flow estimation. In *IEEE International Conference on Image Processing ICIP*, pages 1233–1236.

Sánchez, J., Monzón, N., and Salgado, A. (2012). Robust optical flow estimation. *IPOL: Image Processing Online*, Preprint:1–18.

Sun, D., Sudderth, E., and Black, M. J. (2010). Layered Image Motion with Explicit Occlusions, Temporal Consistency, and Depth Ordering. In Lafferty, J., Williams, C. K. I., Shawe-Taylor, J., Zemel, R., and Culotta, A., editors, *Advances in Neural Information Processing Systems 23*, volume 23, pages 2226–2234.

Weickert, J. and Schnörr, C. (2001). Variational Optic Flow Computation with a Spatio-Temporal Smoothness Constraint. *Journal of Mathematical Imaging and Vision*, 14(3):245–255.