Universität des Saarlandes

Fachrichtung 6.1 – Mathematik

Preprint Nr. 82

Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optic Flow Methods

Joachim Weickert, Andrés Bruhn and Christoph Schnörr

Saarbrücken 2003

Fachrichtung 6.1 – Mathematik
Universität des Saarlandes

Preprint No. 82
submitted: April 24, 2003

Lucas/Kanade Meets Horn/Schunck: Combining Local and Global Optic Flow Methods

Joachim Weickert
Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science
Saarland University
Building 27.1
66041 Saarbrücken
Germany
weickert@mia.uni-saarland.de

Andrés Bruhn
Mathematical Image Analysis Group
Faculty of Mathematics and Computer Science
Saarland University
Building 27.1
66041 Saarbrücken
Germany
bruhn@mia.uni-saarland.de

Christoph Schnörr
Computer Vision, Graphics and Pattern Recognition Group
Faculty of Mathematics and Computer Science
University of Mannheim
68131 Mannheim
Germany
schnoerr@uni-mannheim.de

Edited by
FR 6.1 – Mathematik
Universität des Saarlandes
Postfach 15 11 50
66041 Saarbrücken
Germany

Fax: + 49 681 302 4443
e-Mail: preprint@math.uni-sb.de
WWW: http://www.math.uni-sb.de/

Abstract

Differential methods belong to the most widely used techniques for optic flow computation in image sequences. They can be classified into local methods such as the Lucas–Kanade technique or Bigün's structure tensor method, and into global methods such as the Horn/Schunck approach and its extensions. Often local methods are more robust under noise, while global techniques yield dense flow fields. The goal of this paper is to contribute to a better understanding and the design of differential methods in four ways: (i) We juxtapose the role of smoothing/regularisation processes that are required in local and global differential methods for optic flow computation. (ii) This discussion motivates us to describe and evaluate a novel method that combines important advantages of local and global approaches: It yields dense flow fields that are robust against noise. (iii) Spatiotemproal and nonlinear extensions to this hybrid method are presented. (iv) We propose a simple confidence measure for optic flow methods that minimise energy functionals. It allows to sparsify a dense flow field gradually, depending on the reliability required for the resulting flow. Comparisons with experiments from the literature demonstrate the favourable performance of the proposed methods and the confidence measure.

AMS 2000 Subject Classification: 68T45, 49K20, 65K10, 35J60, 65N04
Key Words: computer vision, optic flow, differential techniques, variational methods, structure tensor, partial differential equations, confidence measures, performance evaluation.

Contents

1 Introduction

Ill-posedness is a problem that is present many image processing and computer vision techniques: Edge detection, for example, requires the computation of image derivatives. This problem is ill-posed in the sense of Hadamard[1], as small perturbations in the signal may create large fluctuations in its derivatives [62]. Another example consists of optic flow computation, where the ill-posedness manifests itself in the nonuniqueness due to the aperture problem [7]: The data allow to compute only the optic flow component normal to image edges. Both types of ill-posedness problems appear jointly in so-called *differential methods* for optic flow recovery, where optic flow estimation is based on computing spatial and temporal image derivatives. These techniques can be classified into *local* methods that may optimise some local energy-like expression, and *global* strategies which attempt to minimise a global energy functional. Examples of the first category include the Lucas–Kanade method [36, 37] and the structure tensor approach of Bigün *et al.* [8, 9], while the second category is represented by the classic method of Horn and Schunck [28] and its numerous discontinuity-preserving variants [1, 3, 10, 16, 26, 34, 41, 44, 47, 50, 51, 58]. Differential methods are rather popular: Together with phase-based methods such as [23] they belong to the techniques with the best performance [6, 24]. Local methods may offer relatively high robustness under noise, but do not give dense flow fields. Global methods, on the other hand, yield flow fields with 100 % density, but are experimentally known to be more sensitive to noise [6, 24].
A typical way to overcome the ill-posedness problems of differential optic flow methods consists of the use of smoothing techniques and smoothness assumptions: It is common to smooth the image sequence prior to differentiation in order to remove noise and to stabilise the differentiation process. Local techniques use spatial constancy assumptions on the optic flow field in the case of the Lucas–Kanade method, and spatiotemporal constancy for the Bigün method. Global approaches, on the other hand, supplement the

[1]A problem is called *well-posed* in the sense of Hadamard, if it has a unique solution that depends continuously on the data. If one of these conditions is violated, it is called *ill-posed*.

optic flow constraint with a regularising smoothness term. Surprisingly, the actual role and the difference between these smoothing strategies, however, has hardly been addressed in the literature so far. In a first step of this paper we juxtapose the role of the different smoothing steps of these methods. We shall see that each smoothing process offers certain advantages that cannot be found in other cases. Consequently, it would be desirable to combine the different smoothing effects of local and global methods in order to design novel approaches that combine the high robustness of local methods with the high density of global techniques. One of the goals of the present paper is to propose and analyse such an embedding of local methods into global approaches. This results in a technique that is robust under noise and gives flow fields with 100 % density. Hence, there is no need for a postprocessing step where sparse data have to be interpolated.

On the other hand, it has sometimes been criticised that there is no reliable confidence measure that allows to sparsify the result of a dense flow field such that the remaining flow is more reliable [6]. In this way it would be possible to compare the real quality of dense methods with the characteristics of local, nondense approaches. In our paper we shall present such a measure. It is simple and applicable to the entire class of energy minimising global optic flow techniques. Our experimental evaluation will show that this confidence measure can give excellent results.

Our paper is organised as follows. In Section 2 we discuss the role of the different smoothing processes that are involved in local and global optic flow approaches. Based on these results we propose two *combined local-global (CLG) methods* in Section 3, one with spatial, the other one with spatiotemporal smoothing. In Section 4 nonlinear variants of the CLG method are presented. Our numerical algorithm is described in Section 5. In Section 6, we introduce a novel confidence measure for all global optic flow methods that use energy functionals. Section 7 is devoted to performance evaluations of the CLG methods and the confidence measure. A summary and an outlook to future work is given in Section 8. In the Appendix, we show how the CLG principle has to be modified if one wants to replace the Lucas–Kanade method by the structure tensor method of Bigün et al. [8, 9].

Related Work. In spite of the fact that there exists a very large number number of publications on motion analysis (see e.g. [39, 53] for reviews), there has been remarkably little work devoted to the integration of local and global optic flow methods. Schnörr [49] sketched a framework for supplementing global energy functionals with multiple equations that provide local data constraints. He suggested to use the output of Gaussian filters shifted in frequency space [23] or local methods incorporating second-order derivatives

[55, 56], but did not consider methods of Lucas–Kanade or Bigün type.
Our proposed technique differs from the majority of global regularisation methods by the fact that we also use spatiotemporal regularisers instead of spatial ones. Other work with spatiotemporal regularisers includes publications by Murray and Buxton [40], Nagel [42], Black and Anandan [10], Elad and Feuer [18], and Weickert and Schnörr [59].
While the noise sensitivity of local differential methods has been studied intensively in recent years [5, 22, 30, 33, 45, 52], the noise sensitivity of global differential methods has been analysed to a significantly smaller extent. In this context, Galvin *et al.* [24] have compared a number of classical methods where small amounts of Gaussian noise had been added. Their conclusion was similar to the findings of Barron *et al.* [6]: the global approach of Horn and Schunck is more sensitive to noise than the local Lucas–Kanade method.
A preliminary shorter version of the present paper has been presented at a conference [14].

2 Role of the Smoothing Processes

In this section we discuss the role of smoothing techniques in differential optic flow methods. For simplicity we focus on spatial smoothing. All spatial smoothing strategies can easily be extended into the temporal domain. This will usually lead to improved results [59].
Let us consider some image sequence $g(x, y, t)$, where (x, y) denotes the location within a rectangular image domain Ω, and $t \in [0, T]$ denotes time. It is common to smooth the image sequence prior to differentiation [6, 33], e.g. by convolving each frame with some Gaussian $K_\sigma(x, y)$ of standard deviation σ:

$$f(x, y, t) \ := \ (K_\sigma * g)(x, y, t), \tag{1}$$

The low-pass effect of Gaussian convolution removes noise and other destabilising high frequencies. In a subsequent optic flow method, we may thus call σ the *noise scale.*
Many differential methods for optic flow are based on the assumption that the grey values of image objects in subsequent frames do not change over time:

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

where the displacement field $(u, v)^\top(x, y, t)$ is called *optic flow.* For small displacements, we may perform a first order Taylor expansion yielding the *optic flow constraint*

$$f_x u + f_y v + f_t = 0, \tag{3}$$

where subscripts denote partial derivatives. Evidently, this single equation is not sufficient to uniquely compute the two unknowns u and v *(aperture problem)*: For nonvanishing image gradients, it is only possible to determine the flow component parallel to $\nabla f := (f_x, f_y)^\top$, i.e. normal to image edges. This so-called *normal flow* is given by

$$w_n = -\frac{f_t}{|\nabla f|}. \tag{4}$$

Figure 1(a) depicts one frame from the famous Hamburg taxi sequence[2]. We have added Gaussian noise, and in Figure 1(b)–(d) we illustrate the impact of presmoothing the image data on the normal flow. While some moderate presmoothing improves the results, great care should be taken not to apply too much presmoothing, since this would severely destroy important image structure.

In order to cope with the aperture problem, Lucas and Kanade [36, 37] proposed to assume that the unknown optic flow vector is constant within some neighbourhood of size ρ. In this case it is possible to determine the two *constants* u and v at some location (x, y, t) from a weighted least square fit by minimising the function

$$E_{LK}(u,v) \;:=\; K_\rho * \left((f_x u + f_y v + f_t)^2\right). \tag{5}$$

Here the standard deviation ρ of the Gaussian serves as an *integration scale* over which the main contribution of the least square fit is computed.

A minimum (u, v) of E_{LK} satisfies $\partial_u E_{LK} = 0$ and $\partial_v E_{LK} = 0$. This gives the linear system

$$\begin{pmatrix} K_\rho * (f_x^2) & K_\rho * (f_x f_y) \\ K_\rho * (f_x f_y) & K_\rho * (f_y^2) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} -K_\rho * (f_x f_t) \\ -K_\rho * (f_y f_t) \end{pmatrix} \tag{6}$$

which can be solved provided that its system matrix is invertible. This is not the case in flat regions where the image gradient vanishes. In some other regions, the smaller eigenvalue of the system matrix may be close to 0, such that the aperture problem remains present and the data do not allow a reliable determination of the full optic flow. All this results in nondense flow fields. They constitute the most severe drawback of local gradient methods: Since many computer vision applications require dense flow estimates, subsequent interpolation steps are required. On the other hand, one may use the smaller eigenvalue of the system matrix as a confidence measure that

[2]The taxi sequence is available from `ftp://csd.uwo.ca` under the directory `pub/vision`

characterises the reliability of the estimate. Experiments by Barron *et al.* [6] indicated that this performs better than the trace-based confidence measure in [52].

Figures 1(e),(f) show the influence of the integration scale ρ on the final result. In these images we have displayed the entire flow field regardless of its local reliability. We can see that in each case, the flow field has typical structures of order ρ. In particular, a sufficiently large value for ρ is very successful in rendering the Lucas–Kanade method robust under noise.

In order to end up with dense flow estimates one may embed the optic flow constraint into a regularisation framework. Horn and Schunck [28] have pioneered this class of global differential methods. They determine the unknown *functions* $u(x, y, t)$ and $v(x, y, t)$ as the minimisers of the global energy functional

$$E_{HS}(u, v) \;=\; \int_{\Omega} \left((f_x u + f_y v + f_t)^2 + \alpha \left(|\nabla u|^2 + |\nabla v|^2 \right) \right) dx\, dy \qquad (7)$$

where the smoothness weight $\alpha > 0$ serves as *regularisation parameter:* Larger values for α result in a stronger penalisation of large flow gradients and lead to smoother flow fields.

Minimising this convex functional comes down to solving its corresponding Euler–Lagrange equations [17, 19]. They are given by

$$0 \;=\; \Delta u - \tfrac{1}{\alpha}\,(f_x^2 u + f_x f_y v + f_x f_t), \qquad (8)$$

$$0 \;=\; \Delta v - \tfrac{1}{\alpha}\,(f_x f_y u + f_y^2 v + f_y f_t). \qquad (9)$$

with reflecting boundary conditions. Δ denotes the spatial Laplace operator:

$$\Delta := \partial_{xx} + \partial_{yy}. \qquad (10)$$

The solution of these diffusion–reaction equations is not only unique [48], it also benefits from the *filling-in effect:* At locations with $|\nabla f| \approx 0$, no reliable local flow estimate is possible, but the regulariser $|\nabla u|^2 + |\nabla v|^2$ fills in information from the neighbourhood. This results in dense flow fields and makes subsequent interpolation steps obsolete. This is a clear advantage over local methods.

It has, however, been criticised that for such global differential methods, no good confidence measures are available that would help to determine locations where the computations are more reliable than elsewhere [6]. It has also been observed that they may be more sensitive to noise than local differential methods [6, 24].

An explanation for this behaviour can be given as follows. Noise results in high image gradients. They serve as weights in the data term of the regularisation functional (7). Since the smoothness term has a constant weight

α, smoothness is relatively less important at locations with high image gradients than elsewhere. As a consequence, *flow fields are less regularised at noisy image structures.* This sensitivity under noise is therefore nothing else but a side-effect of the desired filling-in effect. Figures 1(g),(h) illustrate this behaviour. Figure 1(g) shows that the flow field does not reveal a uniform scale: It lives on a fine scale at high gradient image structures, and the scale may become very large when the image gradient tends to zero. Increasing the regularisation parameter α will finally also smooth the flow field at noisy structures, but at this stage, it might already be too blurred in flatter image regions (Figure 1(h)).

3 A Combined Local–Global Method

We have seen that both local and global differential methods have complementary advantages and shortcomings. Hence it would be interesting to construct a hybrid technique that constitutes the best of two worlds: It should combine the robustness of local methods with the density of global approaches. This shall be done next. We start with spatial formulations before we extend the approach to the spatiotemporal domain.

3.1 Spatial Approach

In order to design a *combined local–global (CLG) method*, let us first reformulate the previous approaches. Using the notations

$$
\begin{aligned}
w &:= (u, v, 1)^\top, && (11)\\
|\nabla w|^2 &:= |\nabla u|^2 + |\nabla v|^2, && (12)\\
\nabla_3 f &:= (f_x, f_y, f_t)^\top, && (13)\\
J_\rho(\nabla_3 f) &:= K_\rho * (\nabla_3 f\, \nabla_3 f^\top) && (14)
\end{aligned}
$$

it becomes evident that the Lucas–Kanade method minimises the quadratic form

$$E_{LK}(w) = w^\top J_\rho(\nabla_3 f)\, w, \tag{15}$$

while the Horn–Schunck technique minimises the functional

$$E_{HS}(w) = \int_\Omega \left(w^\top J_0(\nabla_3 f)\, w + \alpha |\nabla w|^2 \right) dx\, dy. \tag{16}$$

This terminology suggests a natural way to extend the Horn–Schunck functional to the desired CLG functional. We simply replace the matrix $J_0(\nabla_3 f)$

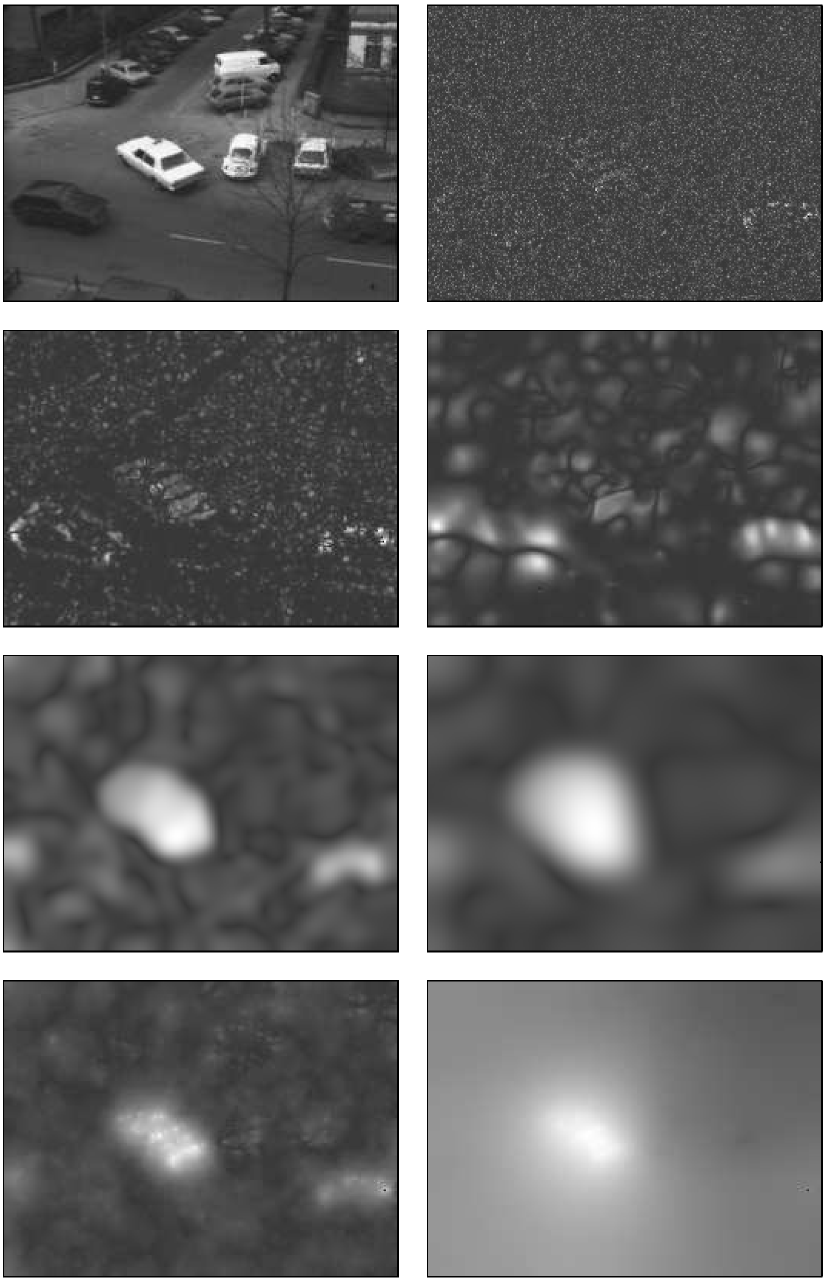

Figure 1: *From left to right, and from top to bottom: (a)* Frame 10 of the Hamburg taxi sequence, where Gaussian noise with standard deviation $\sigma_n = 10$ has been added. The white taxi turns around the corner, the left car drives to the right, and the right van moves to the left. *(b)* Normal flow magnitude without presmoothing. *(c)* Normal flow magnitude, presmoothing with $\sigma = 1$. *(d)* Ditto, presmoothing with $\sigma = 5$. *(e)* Lucas-Kanade method with $\sigma = 0$, $\rho = 7.5$. *(f)* Ditto, $\sigma = 0$, $\rho = 15$. *(g)* Optic flow magnitude with the Horn-Schunck approach, $\sigma = 0$, $\alpha = 10^5$. *(h)* Ditto, $\sigma = 0$, $\alpha = 10^6$.

by the structure tensor $J_\rho(\nabla_3 f)$ with some integration scale $\rho > 0$. Thus, we propose to minimise the functional

$$E_{CLG}(w) = \int_\Omega \left(w^\top J_\rho(\nabla_3 f)\, w + \alpha |\nabla w|^2 \right) \, dx\, dy. \tag{17}$$

Its minimising flow field (u, v) satisfies the Euler–Lagrange equations

$$0 \;=\; \Delta u - \tfrac{1}{\alpha} \left(K_\rho * (f_x^2)\, u + K_\rho * (f_x f_y)\, v + K_\rho * (f_x f_t) \right), \tag{18}$$

$$0 \;=\; \Delta v - \tfrac{1}{\alpha} \left(K_\rho * (f_x f_y)\, u + K_\rho * (f_y^2)\, v + K_\rho * (f_y f_t) \right). \tag{19}$$

It should be noted that these equations are hardly more complicated than the original Horn–Schunck equations (8)–(9). All one has to do is to evaluate the terms containing image data at a nonvanishing integration scale. The basic structure with respect to the unknown functions $u(x, y, t)$ and $v(x, y, t)$ is identical. It is therefore not surprising that the well-posedness proof for the Horn–Schunck method that was presented in [48] can also be extended to this case.

3.2 Spatiotemporal Approach

The previous approaches used only *spatial* smoothness operators. Rapid advances in computer technology, however, makes it now possible to consider also spatiotemporal smoothness operators. Formal extensions in this direction are straightforward. In general, one may expect that spatiotemporal formulations give better results than spatial ones because of the additional denoising properties along the temporal direction.

A spatiotemporal variant of the Lucas–Kanade approach simply replaces convolution with 2-D Gaussians by spatiotemporal convolution with 3-D Gaussians. This still leads to a 2×2 linear system of equations for the two unknowns u and v.

Spatiotemporal versions of the Horn-Schunck method have been considered by Elad and Feuer [18], while discontinuity preserving global methods with spatiotemporal regularisers have been proposed in different formulations in [10, 40, 42, 59].

Combining the temporal extended variant of both the Lucas–Kanade and the Horn–Schunck method we obtain a spatiotemporal version of our CLG functional given by

$$E_{CLG3}(w) = \int\limits_{\Omega \times [0,T]} \left(w^\top J_\rho(\nabla_3 f)\, w + \alpha |\nabla_3 w|^2 \right) \, dx\, dy\, dt \tag{20}$$

where convolutions with Gaussians are now to be understood in a spatiotemporal way, and

$$|\nabla_3 w|^2 \ := \ |\nabla_3 u|^2 + |\nabla_3 v|^2. \tag{21}$$

In general, the spatiotemporal Gaussians may have different standard deviations in space and time. Let us denote by J_{nm} the component (n, m) of the structure tensor $J_\rho(\nabla_3 f)$. Then the Euler–Lagrange equations for (20) are given by

$$\Delta_3 u - \tfrac{1}{\alpha}\,(J_{11} u + J_{12} v + J_{13}) \ = \ 0, \tag{22}$$

$$\Delta_3 v - \tfrac{1}{\alpha}\,(J_{12} u + J_{22} v + J_{23}) \ = \ 0. \tag{23}$$

One should note that they have the same structure as (18)–(19), apart from the fact that spatiotemporal Gaussian convolution is used, and that the spatial Laplacean Δ is replaced by the spatiotemporal Laplacean

$$\Delta_3 := \partial_{xx} + \partial_{yy} + \partial_{tt}. \tag{24}$$

The spatiotemporal Lucas–Kanade method is similar to the approach of Bigün *et al.* [8, 9]. In the Appendix we show how the latter method can be embedded in a global energy functional.

4 Nonquadratic Approach

So far the underlying Lucas–Kanade and Horn–Schunck approaches are *linear* methods that are based on *quadratic* optimisation. It is not very difficult to replace them by *nonquadratic* optimisation problems that lead to *nonlinear* methods. From a statistical viewpoint this can be regarded as applying methods from robust statistics where outliers are penalised less severely than in quadratic approaches [25, 29]. In general, nonlinear methods give better results at locations with flow discontinuities. Robust variants of the Lucas–Kanade method have been investigated by Black and Anandan [11] and by Yacoob and Davies [60], respectively, while a survey of the numerous convex discontinuity-preserving regularisers for global optic flow methods is presented in [58].

In order to render our approach more robust against outliers in both the data and the smoothness term we propose the minimisation of the following functional:

$$E_{CLG3-N}(w) = \int\limits_{\Omega\times[0,T]} \left(\psi_1\left(w^\top J_\rho(\nabla_3 f)\, w\right) + \alpha\, \psi_2\left(|\nabla_3 w|^2\right)\right) dx\, dy\, dt \tag{25}$$

where $\psi_1(s^2)$ and $\psi_2(s^2)$ are nonquadratic penalisers. Encouraging experiments with related continuous energy functionals have been performed by Hinterberger *et al.* [27]. In order to guarantee well–posedness for the remaining problem, we focus on penalisers that are convex in s. In particular, we use a function that has been proposed by Charbonnier *et al.* [15]:

$$\psi_i(s^2) \;=\; 2\beta_i^2 \sqrt{1+\frac{s^2}{\beta_i^2}}, \qquad i \in 1,2 \tag{26}$$

where β_1 and β_2 are scaling parameters. Under some technical requirements, the choice of convex penalisers ensures a unique solution of the minimisation problem and allows to construct simple globally convergent algorithms. The Euler–Lagrange equations of the energy functional (25) are given by

$$\begin{aligned} 0 \;&=\; \operatorname{div}\left(\psi_2'\left(|\nabla_3 w|^2\right)\,\nabla_3 u\right) \\ &\quad-\; \tfrac{1}{\alpha}\,\psi_1'\left(w^\top J_\rho(\nabla_3 f)\,w\right)\left(J_{11}\,u + J_{12}\,v + J_{13}\right), \qquad (27) \\ 0 \;&=\; \operatorname{div}\left(\psi_2'\left(|\nabla_3 w|^2\right)\,\nabla_3 v\right) \\ &\quad-\; \tfrac{1}{\alpha}\,\psi_1'\left(w^\top J_\rho(\nabla_3 f)\,w\right)\left(J_{21}\,v + J_{22}\,u + J_{23}\right). \qquad (28) \end{aligned}$$

5 Algorithmic Realisation

Let us now discuss a suitable algorithm for the CLG method (18)–(19) and its spatiotemporal variant. To this end we consider the unknown functions $u(x,y,t)$ and $v(x,y,t)$ on a rectangular pixel grid of size h, and we denote by u_i the approximation to u at some pixel i with $i = 1,...,N$. Gaussian convolution is realised in the spatial / spatiotemporal domain by discrete convolution with a truncated and renormalised Gaussian, where the truncation took place at 3 times the standard deviation. Symmetry and separability has been exploited in order to speed up these discrete convolutions. Spatial derivatives of the image data have been approximated using a sixth-order approximation with the stencil $(-1, 9, -45, 0, 45, -9, 1)/(60h)$, while temporal derivatives are approximated with a simple two-point stencil. Let us denote by J_{nmi} the component (n,m) of the structure tensor $J_\rho(\nabla f)$ in some pixel i. Furthermore, let $\mathcal{N}(i)$ denote the set of (4 in 2-D, 6 in 3-D) neighbours of pixel i. Then a finite difference approximation to the Euler–Lagrange equations (18)–(19) is given by

$$0 = \sum_{j \in \mathcal{N}(i)} \frac{u_i - u_j}{h^2} - \frac{1}{\alpha} \left(J_{11i}\, u_i + J_{12i}\, v_i + J_{13i} \right), \qquad (29)$$

$$0 = \sum_{j \in \mathcal{N}(i)} \frac{v_i - v_j}{h^2} - \frac{1}{\alpha} \left(J_{21i}\, u_i + J_{22i}\, v_i + J_{23i} \right) \qquad (30)$$

for $i = 1,...,N$. This sparse linear system of equations may be solved iteratively. The *successive overrelaxation (SOR) method* [61] is a good compromise between simplicity and efficiency. If the upper index denotes the iteration step, the SOR method can be written as

$$u_i^{k+1} = (1-\omega)\, u_i^k + \omega \frac{\sum\limits_{j \in \mathcal{N}^-(i)} u_j^{k+1} + \sum\limits_{j \in \mathcal{N}^+(i)} u_j^k - \frac{h^2}{\alpha} \left(J_{12i}\, v_i^k + J_{13i} \right)}{|\mathcal{N}(i)| + \frac{h^2}{\alpha} J_{11i}}, \qquad (31)$$

$$v_i^{k+1} = (1-\omega)\, v_i^k + \omega \frac{\sum\limits_{j \in \mathcal{N}^-(i)} v_j^{k+1} + \sum\limits_{j \in \mathcal{N}^+(i)} v_j^k - \frac{h^2}{\alpha} \left(J_{21i}\, u_i^{k+1} + J_{23i} \right)}{|\mathcal{N}(i)| + \frac{h^2}{\alpha} J_{22i}} \qquad (32)$$

where

$$\mathcal{N}^-(i) := \{ j \in \mathcal{N}(i) \mid j < i \}, \qquad (33)$$

$$\mathcal{N}^+(i) := \{ j \in \mathcal{N}(i) \mid j > i \} \qquad (34)$$

and $|\mathcal{N}(i)|$ denotes the number of neighbours of pixel i that belong to the image domain. The *relaxation parameter* $\omega \in (0, 2)$ has a strong influence on the convergence speed. For $\omega = 1$ one obtains the well-known *Gauß–Seidel method*. We usually use values for ω between 1.9 and 1.99. This numerically inexpensive overrelaxation step results in a speed-up by one order of magnitude compared with the Gauß–Seidel approach. We initialised the flow components for the first iteration by 0. The specific choice of the initialisation is not critical since the method is globally convergent.

It should be noted that the iteration scheme does not require many computations per step, since one may compute expressions of type $\frac{h^2}{\alpha} J_{nmi}$ before entering the iteration loop. Moreover, any practical implementation requires only a single vector of size N for storing each of the two flow components u and v: Since the components are updated sequentially, there is no need for two vectors for the iteration levels k and $k+1$.

Our CLG method has been implemented in ANSI C. For computing the optic flow between two image frames of size 316×252 on a 1533 GHz *Athlon* PC, one iteration takes 4 CPU milliseconds. In the 3-D case using 15 frames of

the same sequence, one iteration takes 70 CPU milliseconds. For our performance evaluations in Section 7, we used 1000 iterations in the 2-D case and 200 iterations in the 3-D case. Since the iterative process converges fast in the beginning and slows down afterwards, one may get perceptually similar solutions already after significantly less iterations. The memory requirement was 5.9 MB in the 2-D example, and 63 MB in the 3-D case.
In the nonlinear case, the discretisation of the Euler-Lagrange equations (27)-(28) is straightforward. For the sake of clarity let us denote the derivatives of the penalising functions ψ_1 and ψ_2 at some pixel i by $\psi'_{1i} := \psi'_1 \left(w_i^\top J_{\rho i}\, w_i \right)$ respectively $\psi'_{2i} := \psi'_2 \left(|\nabla_3 w_i|^2 \right)$. Then the obtained nonlinear system of equations reads

$$0 = \sum_{j \in \mathcal{N}(i)} \frac{\psi'_{2i} + \psi'_{2j}}{2} \, \frac{u_i - u_j}{h^2} - \frac{\psi'_{1i}}{\alpha} \left(J_{11i}\, u_i + J_{12i}\, v_i + J_{13i} \right), \quad (35)$$

$$0 = \sum_{j \in \mathcal{N}(i)} \frac{\psi'_{2i} + \psi'_{2j}}{2} \, \frac{v_i - v_j}{h^2} - \frac{\psi'_{1i}}{\alpha} \left(J_{21i}\, u_i + J_{22i}\, v_i + J_{23i} \right) \quad (36)$$

for $i = 1,...,N$. One should keep in mind that the nonlinearity results from the dependency of ψ'_{1i} and ψ'_{2i} on u_i and v_i. Therfore it is not surprising that, when choosing the quadratic penalisers $\psi_1(s^2) = \psi_2(s^2) = s^2$, the nonlinear case comes down to the linear one, since $\psi'_1(s^2) = \psi'_2(s^2) = 1$.

6 A Confidence Measure for Energy-Based Methods

While global optic flow methods typically yield dense flow fields, it is clear that the flow estimates cannot have the same reliability at all locations. Local methods, on the other hand, have natural confidence measures that help to avoid computing flow values at locations where there is not enough information for a reliable estimate. It would thus be interesting to find a confidence measure that allows to assess the reliability of a dense optic flow field. Barron *et al.* [6] have identified the absence of such good measure as one of the main drawbacks of energy-based global optic flow techniques: Simple heuristics such as using $|\nabla f|$ as a confidence measure did not work well.
In order to address this problem, we propose a measure that may be applied to any energy-based global differential method for computing the optic flow: Since the energy functional E penalises deviations from model assumptions by summing up the deviations E_i from all pixels i in the image domain, it appears natural to use E_i for assessing the local reliability of the computation.

Table 1: Average angular errors computed for the *office* sequence with varying standard deviations σ_n of Gaussian noise. 2-D implementations of the methods of Lucas/Kanade (LK), Horn/Schunck (HS) and the combined local-global approach (CLG) are compared.

σ_n	LK	HS	CLG
0	5.71°	4.36°	4.32°
10	6.79°	6.17°	5.89°
20	8.43°	8.30°	7.75°
40	11.47°	11.76°	10.73°

All we have to do is to consider the cumulative histogram of the contributions E_i with $i = 1,...,N$. As an approximation to the p per cent locations with the highest reliability, we look for the p per cent locations where the contribution E_i is lowest. There are very efficient algorithms available for this purpose; see e.g. [46, Section 8.5]. In the next section we shall observe that this simple criterion may work well over a large range of densities.

7 Experiments

7.1 Evaluation of the CLG Method

Figure 2 shows our first experiment. It depicts a zoom into a synthetic office scene where divergent motion is dominating. This test sequence is available from `www.cs.otago.ac.nz/research/vision/`. It has been created by Galvin *et al.* [24]. We have added Gaussian noise with zero mean and standard deviation $\sigma_n = 20$ to this sequence, and we used the 2-D CLG method (i.e. with spatial regularisation) for computing the flow field. Figure 2(d) shows that the recovered flow field is not very sensitive to Gaussian noise and that it coincides well with the ground truth flow field in Figure 2(c).
These qualitative results are confirmed by the quantitative evaluations in Table 1, where we compare the average angular errors of the Lucas–Kanade, Horn–Schunck, and the CLG method for different noise levels and optimised smoothing parameters σ, ρ, and α. We computed the angular error via

$$\arccos\left(\frac{u_c u_e + v_c v_e + 1}{\sqrt{(u_c^2 + v_c^2 + 1)(u_e^2 + v_e^2 + 1)}}\right) \tag{37}$$

where (u_c, v_c) denotes the correct flow, and (u_e, v_e) is the estimated flow (cf.

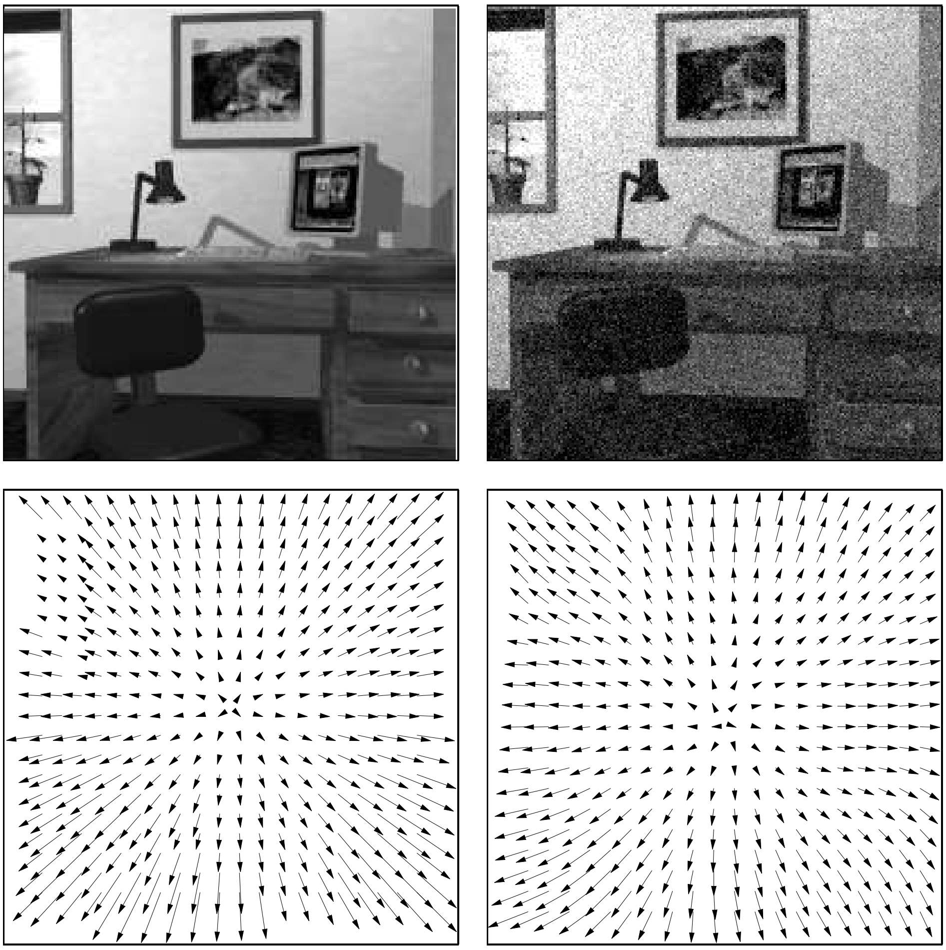

Figure 2: *(a) Top left:* Frame 10 of the synthetic *office* sequence. *(b) Top right:* Degraded by Gaussian noise with $\sigma_n = 20$. *(c) Bottom left:* Ground truth optic flow field. *(d) Bottom right:* Computed optic flow field using the 2-D CLG method for the noisy sequence.

also [6]).
Table 1 shows that for small noise levels, Horn–Schunck performs better than Lucas–Kanade. This indicates that the filling-in effect of the Horn–Schunck algorithm is very useful here. For higher noise levels, Lucas–Kanade becomes somewhat more robust than Horn–Schunck, since the former does not reduce smoothing at noisy structures. The CLG method appears to be able to pick up the best of two worlds: It may benefit from filling-in effects in flat regions without renouncing robustness against noise. Table 1 shows that this combined effect leads to results that may be better than both the Lucas–Kanade and the Horn–Schunck method.
Another example demonstrating the robustness of the 2-D CLG method under Gaussian noise is shown in Figure 3. It depicts the results for the synthetic Yosemite sequence with cloudy sky. This sequence, which is available from `ftp://csd.uwo.ca` under the directory `pub/vision`, combines divergent motion with the translational motion of the sky. It has been used by Barron *et al.* [6] for evaluating a number of optic flow algorithms. Also in this example we can observe that the flow computations using the CLG method do hardly suffer from severe degradations by Gaussian noise.
Let us now investigate the sensitivity of the CLG method with respect to parameter variations. This is done in Table 2 for the *Yosemite* sequence with and without clouds. The modified variant without cloudy sky is available from `http://www.cs.brown.edu/people/black/images.html`. We observe that the average angular error does hardly deteriorate when smoothness parameters are used that differ from their optimal settings by as much as a factor 2. Only the noise scale σ, that is responsible for the presmoothing of the original sequence, is slightly more sensitive. This stability under parameter variations may be regarded as another experimental confirmation of the well-posedness of the CLG approach. Moreover, this also indicates that the method performs sufficiently robust in practice even if non-optimised default parameter settings are used.
In Table 3 we have studied the effect of replacing spatial smoothing steps by spatiotemporal ones for both *Yosemite* sequences. As one may expect, both the quality of the optic flow estimates and their robustness under Gaussian noise improve when temporal coherence is taken into account.
In Section 4 we have presented nonlinear variants of our spatial and spatiotemporal CLG approaches. Our next experiment compares these nonlinear versions to their linear counterparts for a variety of sequences. Generally, flow discontinuities do not cover more than a few percent of the estimated flow field, so only moderate improvements should be expected. These considerations are confirmed by Table 4 where the computed average angular errors are listed. A qualitative example of such a nonlinear variant is given in figure 4.

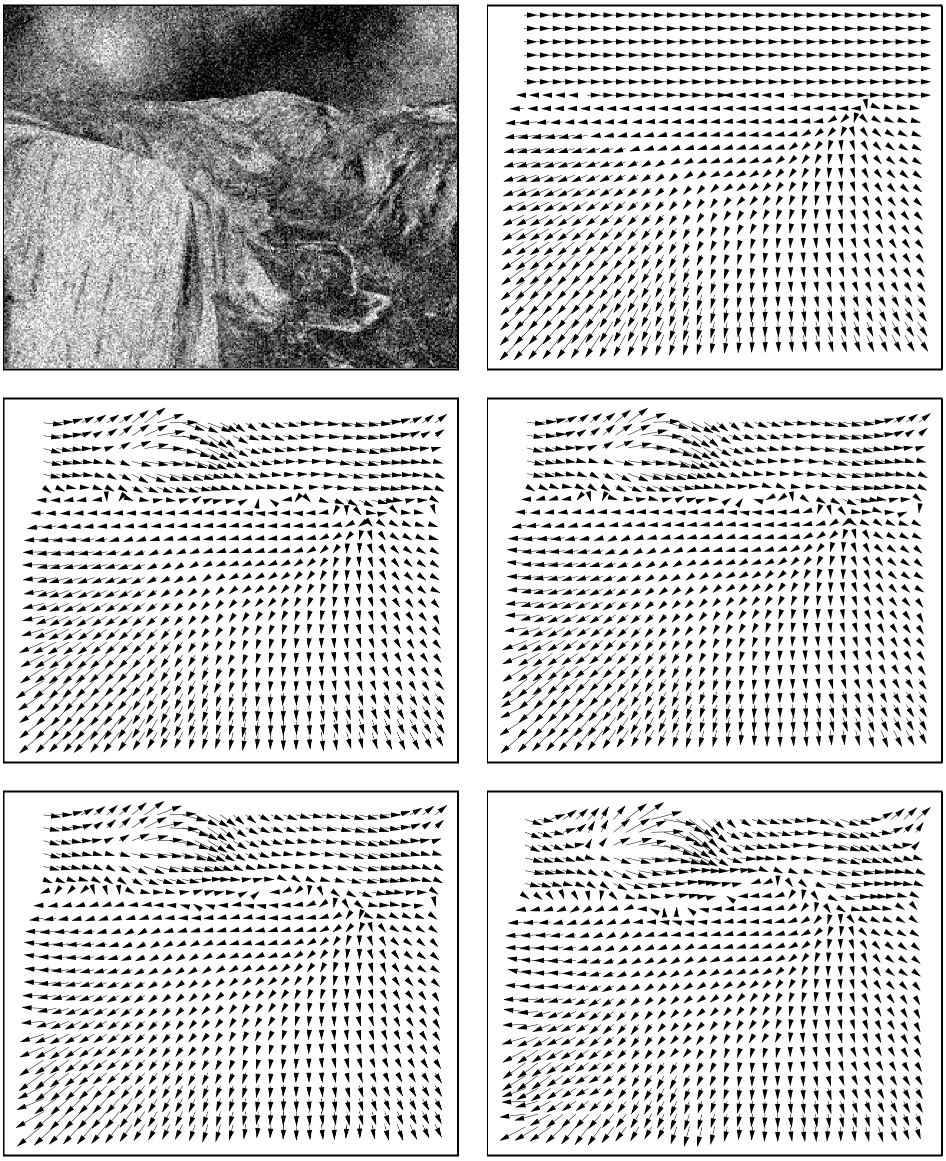

Figure 3: *(a) Top left:* Frame 8 of the *Yosemite* sequence severely degraded by Gaussian noise with $\sigma_n = 40$. *(b) Top right:* Ground truth flow field. *(c) Middle left:* Computed flow field for $\sigma_n = 0$. *(d) Middle right:* Ditto for $\sigma_n = 10$. *(e) Bottom left:* $\sigma_n = 20$. *(f) Bottom right:* $\sigma_n = 40$.

Table 2: Stability of the 2-D CLG method under variations of the smoothing parameters. Two of the three parameters have been set to their optimal value, while the other one may deviate from its optimum by a factor 2. The data refer to the *Yosemite* sequence with and without clouds degraded by Gaussian noise with $\sigma_n = 10$. AAE = average angular error.

Yosemite with clouds				Yosemite without clouds			
α	ρ	σ	AAE	α	ρ	σ	AAE
475	4.550	1.770	9.31°	1000	4.550	1.950	4.57°
633	"	"	9.23°	1666	"	"	4.44°
950	"	"	9.18°	2000	"	"	4.43°
1425	"	"	9.24°	3000	"	"	4.54°
1900	"	"	9.37°	4000	"	"	4.79°
950	2.275	1.770	9.25°	2000	2.275	1.950	4.46°
"	3.033	"	9.21°	"	3.033	"	4.44°
"	4.550	"	9.18°	"	4.550	"	4.43°
"	6.825	"	9.24°	"	6.825	"	4.49°
"	9.100	"	9.39°	"	9.100	"	4.62°
950	4.550	0.885	13.65°	2000	4.550	0.975	7.48°
"	"	1.180	10.58°	"	"	1.300	5.39°
"	"	1.770	9.18°	"	"	1.950	4.43°
"	"	2.655	10.24°	"	"	2.975	5.60°
"	"	3.540	12.30°	"	"	3.800	7.09°

Table 3: Results for the 2-D and 3-D CLG method using the *Yosemite* sequence with and without cloudy sky. Gaussian noise with varying standard deviations σ_n was added, and the average angular errors and their standard deviations were computed.

Yosemite with clouds

σ_n	2-D CLG	3-D CLG
0	$7.14° \pm 9.28°$	$6.18° \pm 9.19°$
10	$9.19° \pm 9.62°$	$7.25° \pm 9.39°$
20	$10.17° \pm 10.50°$	$8.62° \pm 9.97°$
40	$15.82° \pm 11.53°$	$11.21° \pm 11.19°$

Yosemite without clouds

σ_n	2-D CLG	3-D CLG
0	$2.64° \pm 2.27°$	$1.79° \pm 2.34°$
10	$4.45° \pm 2.94°$	$2.53° \pm 2.75°$
20	$6.93° \pm 4.31°$	$3.47° \pm 3.37°$
40	$11.30° \pm 7.41°$	$5.34° \pm 3.81°$

Table 4: Results for linear and nonlinear variants of our CLG method using various sequences. Average angular error and standard deviation have been computed.

Sequence	2-D linear	2-D nonlin.	3-D linear	3-D nonlin.
Yosemite (clouds)	7.14°	6.03°	6.18°	5.18°
Yosemite (no clouds)	2.64°	2.31°	1.79°	1.46°
Office	4.33°	4.13°	3.60°	3.24°
Marble	5.30°	5.14°	2.06°	1.70°

The *Marble* sequence that was used for this purpose can be downloaded at the following internet address: `http://i21www.ira.uk.de/image_sequences`. It is easy to see that the flow discontinuities are much better preserved using the nonlinear variant.
A comparison with other methods from the literature that yield dense flow fields is shown in Table 5. We observe that both the linear 2-D and the 3-D version of our method performs favourably compared to other techniques that do not use multiscale focusing strategies. Since the *Yosemite* sequence contains also large displacements up to 5 pixels per frame, it is not surprising that recent multiscale differential methods are in a somewhat more advantageous position here. However, using nonlinear methods, we are even able to perform better than those techniques.
Nevertheless, it seems possible that the CLG method can also be improved further by embedding it in a suitable multiscale focusing framework. Since our CLG techniques are based on *convex* energy functionals, one would have to correct the image sequence by the estimated coarse scale motion field before going to a finer scale then. If this is not done, the algorithm would remain globally convergent, and multiscale strategies would only speed up the convergence, but the final result would not change.

7.2 Evaluation of the Confidence Measure

Let us now evaluate the quality of our energy-based confidence measure. To this end we have depicted in Figure 5(a) the 20 % quantile of locations where the 3-D CLG method has lowest contributions to the energy. A comparison with Figure 5(b) – which displays the result of a theoretical confidence measure that would be optimal with respect to the average angular error – demonstrates that the energy-based confidence method leads to a fairly realistic thinning of flow fields. In particular, we observe that this confidence

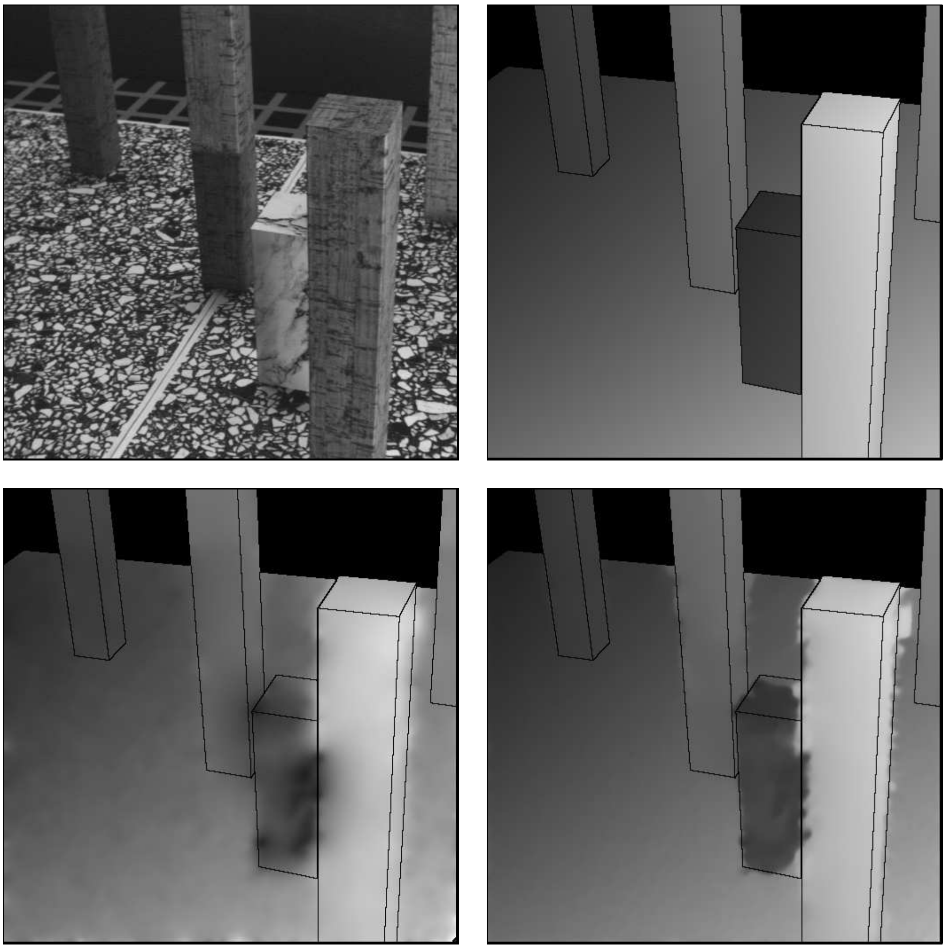

Figure 4: *(a) Top left:* Frame 10 of the *Marble* sequence. *(b) Top right:* Ground truth magnitude. *(c) Bottom left:* Computed flow field magnitude using linear 2-D CLG method. *(d) Bottom right:* Ditto for the nonlinear 3-D CLG variant.

Table 5: Comparison between the results from the literature with 100 % density and our results. All data refer to the *Yosemite* sequence with cloudy sky. Multiscale means that some focusing strategy using linear scale-space or pyramids has been applied. AAE = average angular error.

Technique	multiscale	AAE
Horn/Schunck, original [6]	no	31.69°
Singh, step 1 [6]	no	15.28°
Anandan [6]	no	13.36°
Singh, step 2 [6]	no	10.44°
Nagel [6]	no	10.22°
Horn/Schunck, modified [6]	no	9.78°
Uras *et al.*, unthresholded [6]	no	8.94°
2-D CLG linear	no	7.09°
3-D CLG linear	no	6.18°
2-D CLG nonlinear	no	6.03°
Alvarez *et al.* [2]	yes	5.53°
Mémin/Pérez [38]	yes	5.38°
3-D CLG nonlinear	no	5.18°

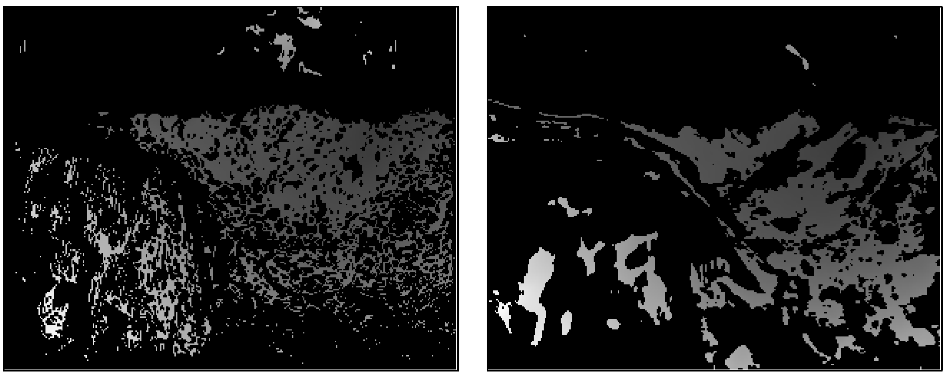

Figure 5: Confidence criterion for the *Yosemite* sequence. *(a) Left:* Locations with the lowest contributions to the energy (20 % quantile). The non-black grey values depict the optic flow magnitude. *(b) Right:* Locations where the angular error is lowest (20 % quantile).

Table 6: Comparison between the "nondense" results from Barron *et al.* [6] and our results for the *Yosemite* sequence with cloudy sky. AAE = average angular error. CLG = average angular error of the 3-D CLG method with the same density. The sparse flow field has been created using our energy-based confidence criterion. The table shows that using this criterion clearly outperforms all results in the evaluation of Barron *et al.*

Technique	Density	AAE	CLG
Singh, step 2, $\lambda_1 \leq 0.1$	97.7 %	10.03°	6.02°
Heeger, level 0	64.2 %	22.82°	2.89°
Weber/Malik [57]	64.2 %	4.31°	2.89°
Horn/Schunck, original, $\|\nabla f\| \geq 5$	59.6 %	25.33°	2.61°
Heeger, combined	44.8 %	15.93°	1.98°
Lucas/Kanade, $\lambda_2 \geq 1.0$	35.1 %	4.28°	1.62°
Fleet/Jepson, $\tau = 2.5$	34.1 %	4.63°	1.59°
Horn/Schunck, modified, $\|\nabla f\| \geq 5$	32.9 %	5.59°	1.55°
Nagel, $\|\nabla f\| \geq 5$	32.9 %	6.06°	1.55°
Fleet/Jepson, $\tau = 1.25$	30.6 %	5.28°	1.48°
Heeger, level 1	15.2 %	9.87°	1.13°
Uras *et al.*, $\det(H) \geq 1$	14.7 %	7.55°	1.11°
Singh, step 1, $\lambda_1 \leq 6.5$	11.3 %	12.01°	1.05°
Waxman *et al.*, $\sigma_f = 2.0$	7.4 %	20.05°	0.94°
Heeger, level 2	2.4 %	12.93°	0.76°

criterion is very successful in removing the cloudy sky regions. These locations are well-known to create large angular errors in many optic flow methods [6]. A number of authors have thus only used the modified *Yosemite* sequence without cloudy sky, or they have neglected the flow values from the sky region for their evaluations [4, 11, 12, 20, 21, 31, 32, 35, 54]. As we have seen one may get significantly lower angular errors than for the full sequence with cloudy sky.

A quantitative evaluation of our confidence measure is given in Table 6. Here we have used the energy-based confidence measure to sparsify the dense flow field such that the reduced density coincides with densities of well-known optic flow methods. Most of them have been evaluated by Barron *et al.* [6]. We observe that the sparsified 3-D CLG method performs very favourably: It has a far lower angular error than all corresponding methods with the same density. In several cases there is an order of magnitude between these approaches. At a flow density of 2.4 %, an average angular error of 0.76 ° is reached. To our knowledge, these are the best values that have been obtained for this sequence in the entire literature. It should be noted that these results have been computed from an image sequence that suffers from quantisation errors since its grey values have been stored in 8-bit precision only.

In Table 6 we also observe that the angular error decreases *monotonically* under sparsification over the entire range from 100 % down to 2.4 %. This in turn indicates in interesting finding that may seem counterintuitive at first glance: *Regions in which the filling-in effect dominates give particularly small angular errors.* In such flat regions, the data term vanishes such that a smoothly extended flow field may yield only a small local contribution to the energy functional. If there were large angular errors in regions with such low energy contributions, our confidence measure would not work well for low densities. This also confirms the observation that $|\nabla f|$ is not necessarily a good confidence measure [6]: Areas with large gradients may represent noise or occlusions, where reliable flow information is difficult to obtain. The filling-in effect, however, may create more reliable information in flat regions by averaging less reliable information that comes from all the surrounding high-gradient regions. The success of our confidence measure also confirms our previous findings that it is beneficial to supplement local methods with a global regulariser.

8 Summary and Conclusions

In this paper we have analysed the smoothing effects in local and global differential methods for optic flow computation. As a prototypes of local methods we used the least-square fit of Lucas and Kanade [36, 37], while the Horn and Schunck approach [28] was our representative for a global method. We saw that the smoothing steps in each of these methods serve different purposes and have different advantages and shortcomings. As a consequence, we proposed a combined local-global (CLG) approach that incorporates the advantages of both paradigms: It is highly robust under Gaussian noise while giving dense flow fields. In order to improve the performance of our method further we considered spatiotemporal variants and the use of nonlinear penalising functions that are well-known from robust statistics. Experiments have shown that the CLG method is not very sensitive under parameter variations. This method may serve as an example of how one can supplement local methods with a regulariser such that dense flow fields are obtained. As is shown in the Appendix, it can also be extended to the embedding of Bigün's structure tensor method into a global energy functional.

We have also proposed a simple confidence measure that allows to sparsify

the dense flow fields of energy-based global methods, such that the most reliable local estimates can easily be found. This enables fair comparisons of the quality of local and global approaches. Our evaluations have shown that the proposed confidence measure may give excellent results over a large range of densities. Last but not least, its success has triggered a surprising finding: For global energy-based optic flow methods, flat regions in which the filling-in effect dominates may offer particularly reliable flow estimates. This explains the common observation that the image gradient magnitude is not a good confidence measure for global variational methods.
While we have already taken efforts to use efficient numerical methods, we have certainly not reached the end of the road yet. Therefore, we are currently investigating multigrid implementations of our technique and we are studying parallelisation possibilities on low latency networks.

Acknowledgements

Our optic flow research is partly funded by the *Deutsche Forschungsgemeinschaft (DFG)* under the project SCHN 457/4-1. 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] G. Aubert, R. Deriche, and P. Kornprobst. Computing optical flow via variational techniques. *SIAM Journal on Applied Mathematics*, 60(1):156–182, 1999.

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

[5] A. Bainbridge–Smith and R. G. Lane. Determining optical flow using a differential method. *Image and Vision Computing*, 15(1):11–22, Jan. 1997.

[6] 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.

[7] M. Bertero, T. A. Poggio, and V. Torre. Ill-posed problems in early vision. *Proceedings of the IEEE*, 76(8):869–889, Aug. 1988.

[8] J. Bigün and G. H. Granlund. Optical flow based on the inertia matrix in the frequency domain. In *Proc. SSAB Symposium on Picture Processing*, Lund, Sweden, Mar. 1988.

[9] J. Bigün, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8):775–790, Aug. 1991.

[10] 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.

[11] 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.

[12] M. J. Black and A. Jepson. Estimating optical flow in segmented images using variable-order parametric models with local deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):972–986, Oct. 1996.

[13] T. Brox and J. Weickert. Nonlinear matrix diffusion for optic flow estimation. In L. Van Gool, editor, *Pattern Recognition*, volume 2449 of *Lecture Notes in Computer Science*, pages 446–453. Springer, Berlin, 2002.

[14] A. Bruhn, J. Weickert, and C. Schnörr. Combining the advantages of local and global optic flow methods. In L. Van Gool, editor, *Pattern Recognition*, volume 2449 of *Lecture Notes in Computer Science*, pages 454–462. Springer, Berlin, 2002.

[15] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud. Two deterministic half-quadratic regularization algorithms for computed imaging. In *Proc. 1994 IEEE International Conference on Image Processing*,

volume 2, pages 168–172, Austin, TX, Nov. 1994. IEEE Computer Society Press.

[16] 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.

[17] R. Courant and D. Hilbert. *Methods of Mathematical Physics*, volume 1. Interscience, New York, 1953.

[18] M. Elad and A. Feuer. Recursive optical flow estimation – adaptive filtering approach. *Journal of Visual Communication and Image Representation*, 9(2):119–138, June 1998.

[19] L. E. Elsgolc. *Calculus of Variations.* Pergamon, Oxford, 1961.

[20] G. Farnebäck. Fast and accurate motion estimation using orientation tensors and parametric motion models. In *Proc. 15th International Conference on Pattern Recognition*, volume 1, pages 135–139, Barcelona, Spain, Sept. 2000.

[21] 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.

[22] C. Fermüller, D. Shulman, and Y. Aloimonos. The statistics of optical flow. *Computer Vision and Image Understanding*, 82(1):1–32, Apr. 2001.

[23] D. J. Fleet and A. D. Jepson. Computation of component image velocity from local phase information. *International Journal of Computer Vision*, 5(1):77–104, Aug. 1990.

[24] B. Galvin, B. McCane, K. Novins, D. Mason, and S. Mills. Recovering motion fields: an analysis of eight optical flow algorithms. In *Proc. 1998 British Machine Vision Conference*, Southampton, England, Sept. 1998.

[25] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel. *Robust Statistics: The Approach Based on Influence Functions.* MIT Press, Cambridge, MA, 1986.

[26] F. Heitz and P. Bouthemy. Multimodal estimation of discontinuous optical flow using Markov random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(12):1217–1232, Dec. 1993.

[27] W. Hinterberger, O. Scherzer, C. Schnörr, and J. Weickert. Analysis of optical flow models in the framework of calculus of variations. *Numerical Functional Analysis and Optimization*, 23(1/2):69–89, May 2002.

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

[29] P. J. Huber. *Robust Statistics.* Wiley, New York, 1981.

[30] B. Jähne. *Digitale Bildverarbeitung.* Springer, Berlin, 2001.

[31] S. Ju, M. Black, and A. Jepson. Skin and bones: multi-layer, 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.

[32] J. Karlholm. *Local Signal Models for Image Sequence Analysis.* PhD thesis, Linköping University, Sweden, 1998. Dissertation No. 536.

[33] J. K. Kearney, W. B. Thompson, and D. L. Boley. Optical flow estimation: an error analysis of gradient-based methods with local optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(2):229–244, Mar. 1987.

[34] A. Kumar, A. R. Tannenbaum, and G. J. Balas. Optic flow: a curve evolution approach. *IEEE Transactions on Image Processing*, 5(4):598–610, Apr. 1996.

[35] 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.

[36] 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.

[37] B. D. Lucas. *Generalized Image Matching by the Method of Differences.* PhD thesis, School of Computer Science, Carnegie–Mellon University, Pittsburgh, PA, 1984.

[38] 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.

[39] A. Mitiche and P. Bouthemy. Computation and analysis of image motion: a synopsis of current problems and methods. *International Journal of Computer Vision*, 19(1):29–55, July 1996.

[40] D. W. Murray and B. F. Buxton. Scene segmentation from visual motion using global optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(2):220–228, Mar. 1987.

[41] H.-H. Nagel. Constraints for the estimation of displacement vector fields from image sequences. In *Proc. Eighth International Joint Conference on Artificial Intelligence*, volume 2, pages 945–951, Karlsruhe, West Germany, August 1983.

[42] 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.

[43] H.-H. Nagel and A. Gehrke. Spatiotemporally adaptive estimation and segmentation of OF-fields. In H. Burkhardt and B. Neumann, editors, *Computer Vision – ECCV '98*, volume 1407 of *Lecture Notes in Computer Science*, pages 86–102. Springer, Berlin, 1998.

[44] P. Nesi. Variational approach to optical flow estimation managing discontinuities. *Image and Vision Computing*, 11(7):419–439, Sept. 1993.

[45] N. Ohta. Uncertainty models of the gradient constraint for optical flow computation. *IEICE Transactions on Information and Systems*, E79-D(7):958–962, July 1996.

[46] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C.* Cambridge University Press, Cambridge, UK, second edition, 1992.

[47] M. Proesmans, L. Van Gool, E. Pauwels, and A. Oosterlinck. Determination of optical flow and its discontinuities using non-linear diffusion. In J.-O. Eklundh, editor, *Computer Vision – ECCV '94*, volume 801 of *Lecture Notes in Computer Science*, pages 295–304. Springer, Berlin, 1994.

[48] C. Schnörr. Determining optical flow for irregular domains by minimizing quadratic functionals of a certain class. *International Journal of Computer Vision*, 6(1):25–38, Apr. 1991.

[49] C. Schnörr. On functionals with greyvalue-controlled smoothness terms for determining optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15:1074–1079, 1993.

[50] 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.

[51] D. Shulman and J. Hervé. Regularization of discontinuous flow fields. In *Proc. Workshop on Visual Motion*, pages 81–90, Irvine, CA, Mar. 1989. IEEE Computer Society Press.

[52] E. P. Simoncelli, E. H. Adelson, and D. J. Heeger. Probability distributions of optical flow. In *Proc. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 310–315, Maui, HI, June 1991. IEEE Computer Society Press.

[53] C. Stiller and J. Konrad. Estimating motion in image sequences. *IEEE Signal Processing Magazine*, 16:70–91, 1999.

[54] R. Szeliski and J. Coughlan. Hierarchical spline-based image registration. In *Proc. 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 194–201, Seattle, WA, June 1994. IEEE Computer Society Press.

[55] O. Tretiak and L. Pastor. Velocity estimation from image sequences with second order differential operators. In *Proc. Seventh International Conference on Pattern Recognition*, pages 16–19, Montreal, Canada, July 1984.

[56] S. Uras, F. Girosi, A. Verri, and V. Torre. A computational approach to motion perception. *Biological Cybernetics*, 60:79–87, 1988.

[57] J. Weber and J. Malik. Robust computation of optical flow in a multi-scale differential framework. *International Journal of Computer Vision*, 14:67–81, 1995.

[58] 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.

[59] 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.

[60] Y. Yacoob and L. S. Davis. Temporal multi-scale models for flow and acceleration. *International Journal of Computer Vision*, 32(2):1–17, Sept. 1999.

[61] D. M. Young. *Iterative Solution of Large Linear Systems.* Academic Press, New York, 1971.

[62] A. L. Yuille and T. A. Poggio. Scaling theorems for zero crossings. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(1):15–25, Jan. 1986.

Appendix: Extension to the Structure Tensor Method

The CLG approach can be extended in a straightforward way to the embedding of basically any local differential method into a global energy functional. Let us illustrate this general principle by focusing on another popular local method: the structure tensor approach of Bigün *et al.* [8, 9].
In Subsection 3.2 we have used a spatiotemporal variant of the Lucas–Kanade technique for the temporal extension of our CLG functional. This method is closely related to the approach of Bigün *et al.* [8, 9]. While Lucas and Kanade make use of a least square fit to overcome the aperture problem, Bigün *et al.* follow a slightly different strategy: They minimise the quadratic form

$$E_{BG}(\tilde{w}) \;=\; \tilde{w}^\top J_\rho(\nabla_3 f)\,\tilde{w} \tag{38}$$

where $\tilde{w} := (\tilde{u}, \tilde{v}, r)^\top$, and the normalisation constraint

$$\tilde{w}^\top \tilde{w} = 1. \tag{39}$$

has to be fulfilled. This is achieved by searching for the eigenvector $\tilde{w}$ that corresponds to the smallest eigenvalue of the structure tensor $J_\rho(\nabla_3 f)$. Normalising its third component to 1 yields $u = \frac{\tilde{u}}{r}$ and $v = \frac{\tilde{v}}{r}$ as the first two components.
In order to combine the local approach of Bigün *et al.* with some global differential optic flow technique, the method can be reformulated as an unconstraint minimisation of the local energy

$$E_{BG}(w) \;=\; \frac{w^\top J_\rho(\nabla_3 f)\, w}{|w|^2}. \tag{40}$$

This reformulation allows a comparison of (15) and (40), which shows that both types of least square fits differ only by the normalisation factor $\frac{1}{|w|^2}$.

The CLG functional obtained by the embedding of Bigün's method in the spatiotemporal variant of Horn and Schunck has the following structure:

$$E_{CLG3-B}(w) = \int\limits_{\Omega\times[0,T]} \left(\frac{w^\top J_\rho(\nabla_3 f)\, w}{|w|^2} + \alpha |\nabla_3 w|^2 \right) dx\, dy\, dt. \tag{41}$$

Its corresponding Euler–Lagrange equations are given by

$$\begin{aligned} 0 &= \Delta_3 u - \frac{1}{\alpha(u^2+v^2+1)^2} \Big(J_{11}(uv^2+u) + J_{12}(v^3-u^2v+v) \\ &+ J_{13}(v^2-u^2+1) - J_{22}uv^2 - 2J_{23}uv - J_{33}u \Big), \end{aligned} \tag{42}$$

$$\begin{aligned} 0 &= \Delta_3 v - \frac{1}{\alpha(u^2+v^2+1)^2} \Big(-J_{11}u^2v + J_{12}(u^3-uv^2+u) \\ &- 2J_{13}uv + J_{22}(u^2v+v) + J_{23}(u^2-v^2+1) - J_{33}v \Big). \end{aligned} \tag{43}$$

These nonlinear equations are somewhat more complicated than their linear Lucas–Kanade counterpart (22)–(23).
In order to encourage discontinuity-preserving optic flow fields, one can also introduce nonquadratic penalisers into the functional (41). This yields

$$E_{CLG3-BN}(w) = \int\limits_{\Omega\times[0,T]} \left(\psi_2\left(\frac{w^\top J_\rho(\nabla_3 f)\, w}{|w|^2} \right) + \alpha \psi_1(|\nabla_3 w|^2) \right) dx\, dy\, dt. \tag{44}$$

Such a strategy may be regarded as an alternative to the discontinuity-preserving structure tensor methods in [13, 43]. It gives dense flow fields without additional postprocessing steps.