A Multigrid Approach for Hierarchical Motion Estimation

E. Mémin

Valoria/Université de Bretagne Sud
BP 1104, 56014 Vannes, France
and
IRISA, 35042 Rennes Cedex, France
memin@irisa.fr

P. Pérez

IRISA/INRIA
35042 Rennes Cedex, France

perez@irisa.fr

Abstract

This paper focuses on the estimation of the apparent motion field between two consecutive frames in an image sequence. The approach developed here is a trade-off between methods based on global parameterized flow models and local dense optic flow estimators.

The method relies on an adaptive multigrid minimization approach. In addition to accelerated convergence toward good estimates, it allows to mix different parameterizations of the estimate relative to adaptive partitions of the image.

The performances of the resulting algorithms are demonstrated in the difficult context of a non-convex energy. Experimental results on real world Meteosat sequences are presented.

1 Introduction

Energy-based models constitute a powerful way to cope with low-level image processing problems. These methods are issued either from continuous formalism such as PDEs or from discrete modelization such as Markov random fields. When the energy is convex one can seek the unique minimum with deterministic descent minimization algorithms. However, when it comes to designing accurate and robust methods, able for instance to handle and precisely locate discontinuities, more suitable energies are non-quadratic with numerous local minima. The optical flow estimation is not an exception to this rule [4, 7, 13, 15]. Even within an incremental *multiresolution* formulation of the problem – which is commonly thought as yielding a smooth version of the original problem – one still has to face to difficult sequence of optimization problems.

To cope efficiently with such global energy minimizations, we propose here a family of deterministic optimization algorithms. These *multigrid* algorithms are extensions of the method proposed in [13]. They allow to properly combine a multigrid minimization strategy with the multiresolution framework commonly used in motion analysis. The key idea consists in the minimization of the energy function through an appropriate hierarchy of subspaces of the whole configuration space. Each subspace is defined as a set of configurations constrained to be piecewise parametric relative to a certain partition of the image.

These multigrid minimization algorithms turn out to provide an accelerated convergence toward improved estimates. Furthermore, associated to a robust model of dense optic flow estimation, the multigrid framework provides efficient motion estimators allowing to combine different parameterizations of the velocity field through an adaptive partition of the image grid. A compromise solution between local dense methods involving smoothness constraints [4, 7, 13, 15] and global parametric approaches assuming low order polynomial representation of the velocity field [1, 3, 5] is thus introduced. The resulting approach takes benefit both from the robustness and the richness of global parametric flow descriptions and from the flexibility of the local smoothness constraint associated to dense estimators. Contrary to standard parametric approach assuming independent motion models on the regions of the segmentation, the parametric models can here interact together through local smoothness constraint at patches borders. As a consequence, the estimation of the image partition is a less critical issue in our case.

2 Robust incremental estimation

Let us first present the robust model for the optical estimation [4, 13]. We consider here an incremental estimation of the flow field captured both by the optic-flow multiresolution setup and by the multigrid strategy *used at each resolution level* [13]. The multiresolution framework, based on a pyramidal decomposition of the image data and allowing to estimate large velocities, is standard and won't be emphasized herein. In the following, we shall assume to work at a

given resolution of this structure. However, one has to keep in mind that the expressions and computations are meant to be reproduced at each resolution level according to a coarse-to-fine strategy.

Let us thus assume that a rough estimate $\boldsymbol{w} = \{\boldsymbol{w}_s, s \in S\}$ of the unknown velocity field is available (e.g., from an estimation at lower resolution or grid level), on the pixel lattice S. Let $f(t) = \{f(s,t), s \in S\}$ be the luminance function at time t. Under the constancy brightness assumption from time t to $t+1$, a small *increment field* $\mathsf{d}\boldsymbol{w} \in \Omega \subset (\mathbb{R} \times \mathbb{R})^S$ can be estimated by minimizing the functional $H \triangleq H_1 + \alpha H_2$, with [4, 13]:

$$H_1 \triangleq \sum_{s \in S} \rho_1[\underbrace{\boldsymbol{\nabla} f(s+\boldsymbol{w}_s)^T \mathsf{d}\boldsymbol{w}_s + f_t(s, \boldsymbol{w}_s)}_{\triangleq \mathrm{ofc}(s, \boldsymbol{w}_s, \mathsf{d}\boldsymbol{w}_s)}], \quad (1)$$

$$H_2 \triangleq \sum_{<s,r> \in \mathcal{C}} \rho_2 \left[\|(\boldsymbol{w}_s + \mathsf{d}\boldsymbol{w}_s) - (\boldsymbol{w}_r + \mathsf{d}\boldsymbol{w}_r)\|\right], \quad (2)$$

where $\alpha > 0$, $\mathcal{C}$ is the set of neighboring site pairs lying on grid S equipped with some neighborhood system ν, $\boldsymbol{\nabla} f$ stands for the spatial gradient of $f(t+1)$ and $f_t(s, \boldsymbol{w}_s) \triangleq f(s + \boldsymbol{w}_s, t+1) - f(s,t)$ is the displaced frame difference. Functions ρ_1 and ρ_2 are standard *robust M-estimators* (with hyper-parameters σ_1 and σ_2). They penalize the deviations both from the data model (i.e., the well known *optical flow constraint*) and both from the first order smoothing prior. Unlike the quadratic norm, these robust norm functions have a saturating property (finite limit or assymptotic limit at infinity) in presence of large residuals.

Under some simple conditions (mainly concavity of $\phi(v) \triangleq \rho(\sqrt{v})$, see [6, 8] for a complete account), any multidimensional minimization problem of the form "find $\arg\min_x \sum_i \rho(x)$" can be turned into a "semi"-quadratic minimization problem "find $\arg\min_{x,z} \sum_i [m z_i x^2 + \psi(z_i)]$" involving *auxiliary variables* (or *weights*) z_is continuously lying in $(0,1]$. ψ is a continuously differentiable function, depending on ρ, and $m \triangleq \lim_{v \to 0+} \phi'(v)$. The new minimization is then usualy lead alternatively with respect to x and to the z_is. The expression being quadratic w.r.t. x the corresponding minimization is lead through a standard *weighted least squares* minimization. In turn x being frozen, the best weights are given by the following closed form [6, 8]:

$$\hat{z}_i(x) = \frac{\rho'[x]}{2mx} = \frac{1}{m}\phi'[x^2]. \quad (3)$$

The overall alternate procedure constitutes an *iteratively reweighted least squares* estimation.

Following this reformulation result, the estimation is now expressed as the global minimization in $(\mathsf{d}\boldsymbol{w}, \delta, \beta)$ of $\mathcal{H} \triangleq \mathcal{H}_1 + \alpha \mathcal{H}_2$ where $\mathcal{H}_1$ and $\mathcal{H}_2$ are respectively:

$$\sum_{s \in S} \left[m_1 \delta_s \left[\mathrm{ofc}(s, \boldsymbol{w}_s, \mathsf{d}\boldsymbol{w}_s)^2\right] + \psi_1(\delta_s)\right] \text{ and} \quad (4)$$

$$\sum_{<s,r> \in \mathcal{C}} \left[m_2 \beta_{sr} \|(\boldsymbol{w}_s + \mathsf{d}\boldsymbol{w}_s) - (\boldsymbol{w}_r + \mathsf{d}\boldsymbol{w}_r)\|^2 + \psi_2(\beta_{sr})\right]. \quad (5)$$

The two sets of weights $\delta = \{\delta_s, s \in S\}$ and $\beta = \{\beta_{sr}, <s,r> \in \mathcal{C}\}$ are respectively the weights related to the data-model term (H_1) and the discontinuity weights related to the smoothing term (H_2). The first one allows to attenuate the confidence on data where the OFC is violated whereas the second set, lying on the dual edge grid of S, permits to introduce discontinuities between sites exhibiting significant differences of motion.

Despite leading to a semi-quadratic energy $\mathcal{H}$, the underlying energy function H is ususally non-convex. We still have to deal with a tough optimization problem.

3 Multigrid minimization

To efficiently cope with our global optimization problem, we design a hierarchical "constrained" exploration of the configuration space Ω: the optimization is lead through a sequence of constrained configuration subspaces of increasing sizes

$$\dim(\Omega^L) < \dim(\Omega^{L-1}) < \cdots < \dim(\Omega^0), \quad (6)$$

with $\Omega^0 = \Omega$ and where Ω^ℓ is the set of increment fields which are constrained to be piecewise parametric according to a partition of grid S. Let $\mathcal{B}^\ell \triangleq \{\mathcal{B}^\ell_n,\ n = 1, \ldots, N_\ell\}$ denotes this partition and S^ℓ the vertices of the associated connectivity graph. Each patch of the incremental field is defined such as:

$$\forall n \in S^\ell,\ \forall s \in \mathcal{B}^\ell_n, \quad \mathsf{d}\boldsymbol{w}_s = \Phi^\ell_n(\boldsymbol{\Theta}^\ell_n, s), \quad (7)$$

where $\boldsymbol{\Theta}^\ell_n$ is a parameter vector. The whole increment is then expressed as $\mathsf{d}\boldsymbol{w} = \Phi^\ell(\boldsymbol{\Theta}^\ell)$, with $\boldsymbol{\Theta}^\ell = \{\boldsymbol{\Theta}^\ell_n, n \in S^\ell\}$ lying in the reduced parameter space Γ^ℓ. The full-rank function Φ^ℓ is the *interpolation operator* between the reduced subspace Γ^ℓ and the full original configuration space Ω. It is a one-to-one mapping from Γ^ℓ into $\Omega^\ell = \mathrm{Im}\Phi^\ell$.

3.1 Different parameterizations and their mixing

In this work we focus on linear forms of the interpolation operator:

$$\forall n \in S^\ell,\ \forall s \in \mathcal{B}^\ell_n,\ \mathsf{d}\boldsymbol{w}_s = P_n(s)\boldsymbol{\Theta}^\ell_n, \quad (8)$$

where $P_n(s)$ is a 2 by p_n matrix. The corresponding parameter spaces are in these cases $\Gamma^\ell = \mathbb{R}^{p_1} \times \cdots \times \mathbb{R}^{p_{N_\ell}}$. The standard parametric models used in motion analysis correspond to p_n=2, 4, 6 or 8 [3]. In this work we will consider three different parameterizations of the configuration subspaces (in the following x_s and y_s are the coordinates of site s).

Blockwise constant model

Here, the configurations of Ω^ℓ are constrained to be piecewise constant over each patch of the partition $\mathcal{B}^\ell$:

$$\begin{aligned} \forall n \in S^\ell,\ \forall s \in \mathcal{B}_n^\ell,\ P_n(s) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \triangleq \mathbb{I}_2, \\ \text{i.e., } \mathtt{d}\boldsymbol{w}_s = \boldsymbol{\Theta}_n^\ell = \begin{bmatrix} \mathtt{d}u_n^\ell & \mathtt{d}v_n^\ell \end{bmatrix}^T. \end{aligned} \tag{9}$$

This kind of constraint has been extensively used in the regularization-free context of conventional block-matching techniques for video coding.

Piecewise simplified affine model

In this second model the projection between the subspaces Γ^ℓ and the configuration space Ω^ℓ is described by a four parameters transformation:

$$\begin{aligned} &\forall n \in S^\ell,\ \forall s \in \mathcal{B}_n^\ell, \quad P_n(s) = \begin{bmatrix} 1 & 0 & x_s & y_s \\ 0 & 1 & y_s & -x_s \end{bmatrix} \\ &\text{and } \boldsymbol{\Theta}_n^\ell = \begin{bmatrix} tx_n^\ell & ty_n^\ell & div_n^\ell & curl_n^\ell \end{bmatrix}^T. \end{aligned} \tag{10}$$

This model conjectures that the partition patches correspond to the projection of 3D planar facets parallel to the image plane with apparent motion restricted to translation, divergence and rotation. It is known as the simplified affine model.

Piecewise affine model

Usually a more complete affine model is preferred. Here the 3D planar patches of surface are not anymore supposed to be parallel to the image plane and their motions are confined to rotations around the optical axis and to translation in the facet's plan. The constraint is described by a six parameters vector:

$$\begin{aligned} &\forall n \in S^\ell, \forall s \in \mathcal{B}_n^\ell,\ P_n(s) = \begin{bmatrix} 1 & x_s & y_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_s & y_s \end{bmatrix} \\ &\text{and } \boldsymbol{\Theta}_n^\ell = \begin{bmatrix} a_n^\ell & b_n^\ell & c_n^\ell & d_n^\ell & e_n^\ell & f_n^\ell \end{bmatrix}^T. \end{aligned} \tag{11}$$

The affine model has been widely used for motion-based segmentation [5] or for the estimation of dominant motion over the entire image [4, 10]. It is considered as a good trade-off between model complexity and model efficiency [5]. In segmentation applications, the affine model defines the profile of the flow inside each segment of the associated image partition. The accuracy of the field estimation depends obviously on the quality of the segmentation. If the considered regions are too large then a single affine model may badly represents the motion of the associated 3D surface. The planarity hypotheses are indeed very likely to be violated in this case. On the opposite, since usually no "continuity" between patches is maintained on the velocity field, smaller regions may lead to inaccurate motion estimation (occlusion or uniform areas for example). Recently, a blockwise smooth parametric model involving inter-block regularization has been considered [11]. In this work, the field is constrained to be locally affine on a block partition and a regularization term in the parameter space is added in order to enforce smoothness across patches. However, the regularization prior considered does not allow to easily support different parameterizations of the flow field. Furthermore, this approach is not defined hierarchically and thus implies to consider mixtures of an unknown number of models when the block involves several motions or a complex motion profile.

3.2 Multigrid energy derivation

Let us now see more precisely how these different parametric models may be embedded within an unified hierarchical minimization scheme over constrained configuration subspaces Ω^ℓ. Recalling that our purpose is to build a hierarchy of such subspaces the partitions have to satisfy the following property:

$$\forall n \in S^\ell, \exists\ !\ \bar{n} \in S^{\ell+1} :\ \mathcal{B}_n^\ell \subset \mathcal{B}_{\bar{n}}^{\ell+1}, \tag{12}$$

which expresses that $\mathcal{B}^\ell$ corresponds to a subdivision of elements of $\mathcal{B}^{\ell+1}$. This induces a natural tree structure for which $\bar{n}$ is the parent of n.

The constrained optimization in $\Omega^\ell = \mathrm{Im}\Phi^\ell$ is obviously equivalent to the minimization of the new energy function:

$$\mathcal{H}^\ell(\boldsymbol{\Theta}^\ell, \delta, \beta; f, \boldsymbol{w}) \triangleq \mathcal{H}(\Phi^\ell(\boldsymbol{\Theta}^\ell), \delta, \beta; f, \boldsymbol{w}), \tag{13}$$

defined over Γ^ℓ, whereas the weights, the data, and the field to be refined, remain the same (i.e., defined on the original grid S). From this family of energy functions, we are now able to define our minimization scheme as a cascade (from $\ell = L$ to $\ell = 0$) of optimization problems of reduced complexity:

$$(\hat{\boldsymbol{\Theta}}^\ell, \hat{\delta}, \hat{\beta}) = \underset{\boldsymbol{\Theta}^\ell, \delta, \beta}{\arg\min} \mathcal{H}^\ell(\boldsymbol{\Theta}^\ell, \delta, \beta; f, \boldsymbol{w}^\ell), \ell = L, \ldots, 0. \tag{14}$$

The field $\boldsymbol{\Theta}^\ell \in \Gamma^\ell$ lies on the reduced grid S^ℓ whereas the weights and the main velocity components, $\boldsymbol{w}^\ell \triangleq \boldsymbol{w}^{\ell+1} + \Phi^{\ell+1}(\hat{\boldsymbol{\Theta}}^{\ell+1})$, are attached to S, whatever the grid level ℓ. The initial field $\boldsymbol{w}^L$ comes from an estimation at a coarser resolution or from a given initialization.

Each of these successive minimizations is processed in terms of iteratively reweighted least squares and a multigrid coarse-to-fine strategy is developed: the increment field estimated at level $\ell+1$, $\Phi^{\ell+1}(\widehat{\boldsymbol{\Theta}}^{\ell+1})$ is added to $\boldsymbol{w}^{\ell+1}$ to form the new main component of the velocity field $\boldsymbol{w}^{\ell}$, $f(t)$ is warped accordingly for the computation of the spatial and temporal derivatives and a new increment is then estimated at level ℓ. This procedure is repeated until the finest level $\ell = 0$ is reached.

The new multigrid energy functions[1] $\mathcal{H}^{\ell}$ may be easily derived from the original one (4-5). They are also composed of two terms $\mathcal{H}^{\ell} = \mathcal{H}_1^{\ell} + \alpha\mathcal{H}_2^{\ell}$:

$$\mathcal{H}_1^{\ell} = \sum_{n\in S^{\ell}} \sum_{s\in\mathcal{B}_n^{\ell}} m_1 \delta_s \Big[\mathrm{ofc}(s, \boldsymbol{w}_s^{\ell}, P_n(s)\boldsymbol{\Theta}_n^{\ell})\Big]^2 + \psi_1(\delta_s) \quad (15)$$

$$\mathcal{H}_2^{\ell} = \sum_{<n,m>\in\mathcal{C}^{\ell}} \sum_{<s,r>\in\mathcal{C}_{nm}^{\ell}} m_2\beta_{sr} \Big\| \boldsymbol{W}_s^{\ell}(n) - \boldsymbol{W}_r^{\ell}(m) \Big\|^2 + \psi_2(\beta_{sr}) + \sum_{n\in S^{\ell}} \sum_{<s,r>\in\mathcal{C}_n^{\ell}} m_3\beta_{sr} \Big\| \boldsymbol{W}_s^{\ell}(n) - \boldsymbol{W}_r^{\ell}(n) \Big\|^2 + \psi_2(\beta_{sr}), \quad (16)$$

with $\boldsymbol{W}_s^{\ell}(n) \triangleq \boldsymbol{w}_s^{\ell} + P_n(s)\hat{\boldsymbol{\Theta}}_n^{\ell}$ and where $\mathcal{C}_n^{\ell} \triangleq \{<s,r>\in\mathcal{C} : <s,r>\subset\mathcal{B}_n^{\ell}\}$ is the set of neighboring site pairs included in patch $\mathcal{B}_n^{\ell}$ and $\mathcal{C}_{nm}^{\ell} \triangleq \{<s,r>\in\mathcal{C} : s\in\mathcal{B}_n^{\ell}, r\in\mathcal{B}_m^{\ell}\}$ the set of neighboring site pairs straddling $\mathcal{B}_n^{\ell}$ and $\mathcal{B}_m^{\ell}$. All possible $\mathcal{C}_n^{\ell}$ and $\mathcal{C}_{nm}^{\ell}$ site sets form a partition of $\mathcal{C}$. Reduced site set S^{ℓ} turns out to be equipped with a new neighborhood system ν^{ℓ}. The corresponding neighboring pair set will be denoted by $\mathcal{C}^{\ell}$. In $\mathcal{H}_2$, we consider different hyper-parameter values, m_2 and m_3, depending on whether the site pairs are inside a patch or straddling its border. A low value of this hyper-parameter inside each patch (i.e., a reduced regularization) will allow us to favor a larger class of affine motion increments inside each block since all the variables are then almost decorrelated. The last term of $\mathcal{H}_2$ can thus be made negligible.

3.3 Energy minimization

The current reduced parameter estimate $\boldsymbol{\Theta}^{\ell}$ being fixed, we know that the optimal weight values are directly accessible. According to (3), in combination with energy definitions (15-16), these values are:

$$\forall s\in\mathcal{B}_n^{\ell}, \delta_s = \frac{1}{m_1}\phi_1'\left[\Big(\mathrm{ofc}(s, \boldsymbol{w}_s^{\ell}, P_n(s)\boldsymbol{\Theta}_n^{\ell})\Big)^2\right], \quad (17)$$

$$\forall <s,r>\in\mathcal{C}_{nm}^{\ell}, \beta_{sr} = \frac{1}{m_2}\phi_2'\left[\Big\| \boldsymbol{W}_s^{\ell}(n) - \boldsymbol{W}_r^{\ell}(m) \Big\|^2\right], \quad (18)$$

$$\forall <s,r>\in\mathcal{C}_n^{\ell}, \beta_{sr} = \frac{1}{m_3}\phi_2'\left[\Big\| \boldsymbol{W}_s^{\ell}(n) - \boldsymbol{W}_r^{\ell}(n) \Big\|^2\right]. \quad (19)$$

As soon as the values of weights are computed and frozen, the energy function $\mathcal{H}^{\ell}(\boldsymbol{\Theta}^{\ell}, \delta, \beta; f, \boldsymbol{w}^{\ell})$ is quadratic with respect to $\boldsymbol{\Theta}^{\ell}$. Its minimization is equivalent to the resolution of a linear system whose solution is searched with an iterative Gauss-Seidel scheme. Each update is obtained by solving a linear equation in $\boldsymbol{\Theta}_n^{\ell}$ for the current block $\mathcal{B}_n^{\ell}$. See [14] for detailed derivations.

3.4 Adaptive grids

The different parametric models previously presented do not imply the same computational load. Furthermore, they are different in terms of locality and accuracy. For example the six parameter affine model allows to describe the global motion profile of large areas whereas the constant model is accurate only for small patches. In order to better fit the content of the image, it seems interesting to mix different levels of parameterization along with an adapted partionning. Such adaptive grids are considered in [16] where a blockwise parametric motion model (a spline based description) is locally refined on smaller blocks if the mean square error between frame (t) and the frame $(t+1)$ backward registered exceeds a certain threshold. However this approach does not make use of regularization. A different approach based on finite elements method considers a threshold on the normal flow field (which are directly accessible from the OFC) [7]. In this case the computation of the adaptive grid is done *a priori*. Our robust incremental motion model permits to directly have access to such information through the data auxiliary variables [2] and to use it on line. We consider an adaptive grid structure relying on the following subdivision criterion: at convergence on the current grid level, a patch is split in four parts if the standard deviation of the data outliers on that patch is greater than a certain threshold.

4 Experimental results

In this section we present results obtained both on synthetic and real-world sequences. The first sequence, Yosemite, is the most complex (though synthetic) sequence from the comparative study by Barron *et al.* [2] for which a "ground-truth" exists.

The second test sequence named Depression (Fig. 1a) is a meteorological sequence involving large displacements. It includes a trough of low pressure and some moving clouds driven by different motions.

In our experiments we have tested different motion estimators arising from our approach. First of all, although our multigrid scheme is defined for any parti-

[1] We use here the dual form of M-estimators. This could as well be described in their initial formulation.

[2] More precisely, they account for the linearized errors between the frame $(t+1)$ backward registered and frame (t).

tion of the image, for sake of simplicity we will consider here only square block partitions. The multigrid framework has been evaluated for the three kinds of constraint previously described. This yields different estimation models which have been run both on regular grids and adaptive grids. In the first case, the multigrid minimization is associated to a regular and complete subdivision scheme leading to consider a single parametric class at every grid level. Whereas, in the second case we have to deal with irregular grids issued from the adaptive subdivision strategy described in section 3.4 and where the type of parameterization on each patch may depend on its size.

The different estimation models are denoted $\mathcal{M}_{\mathbf{6}}$, $\mathcal{M}_{\mathbf{4}}$, $\mathcal{M}_{\mathbf{2}}$, $\mathcal{M}_{\mathbf{6,4}}$, $\mathcal{M}_{\mathbf{6,2}}$ and $\mathcal{M}_{\mathbf{6,4,2}}$. The first three concern respectively the 6-parameter affine model ($\mathcal{M}_{\mathbf{6}}$), the simplified affine model ($\mathcal{M}_{\mathbf{4}}$) and the constant model ($\mathcal{M}_{\mathbf{2}}$). The others are multi-parametric versions where several parameterizations are successively imposed (regular case) or mixed (adaptive case). $\mathcal{M}_{\mathbf{6,4}}$ associates the 6-parameter affine model and simplified affine model whereas $\mathcal{M}_{\mathbf{6,2}}$ involves constant model and affine model and finally $\mathcal{M}_{\mathbf{6,4,2}}$ includes the three models. The finest grid level (resp. the finest block size in case of adaptive grids) actually considered in the single parametric estimation model depends on the parameterization used. It is $\ell_0 = 3$ (8×8 blocks) for $\mathcal{M}_{\mathbf{6}}$, $\ell_0 = 2$ (4×4 blocks) for $\mathcal{M}_{\mathbf{4}}$, and $\ell_0 = 0$ for $\mathcal{M}_{\mathbf{2}}$. By contrast, the current parameterization to consider in the multi-parametric versions depends on the grid level (resp. the block size). We imposed to have the 6-parameter affine model in $\mathcal{M}_{\mathbf{6,4}}$, $\mathcal{M}_{\mathbf{6,2}}$ and $\mathcal{M}_{\mathbf{6,4,2}}$ if $\ell > 2$ (resp. $|\mathcal{B}_n^\ell| > 4 \times 4$). For $\mathcal{M}_{\mathbf{6,4}}$ and $\mathcal{M}_{\mathbf{6,4,2}}$ we used the simplified affine model if $\ell = 2$ (this level being the finest grid level for $\mathcal{M}_{\mathbf{6,4}}$). The constant model is involved if ℓ=1 or 0 for $\mathcal{M}_{\mathbf{6,4,2}}$ and if ℓ=2, 1 or 0 for $\mathcal{M}_{\mathbf{6,2}}$

We must outline that the different algorithms were run with the same set of hyper-parameters (see [14] for explicite values). Based on heuristic considerations [14], we choose Leclerc's estimator [12] $\rho_1(u) \triangleq 1 - \exp(-u^2/\sigma_1^2)$ for the data-model term. For the smoothing term we selected a "softer" saturating estimator (i.e., a slower decreasing rate of the estimator's derivative) [9]: $\rho_2(u) \triangleq \frac{u^2}{u^2+\sigma_2}$.

Following [2], quantitative comparative results on Yosemite are provided for different algorithms. For each estimate, angular deviations with respect to the real flow are computed at "reliable" locations (the percentage of such locations is the "density" of the estimate). In our case the proposed method yields always a full density. Table 1 lists, for all the different algorithms proposed here, the mean angular error ($\bar{\mu}$) and the associated standard deviation (σ). The CPU times, measured on a SUN ULTRA SPARC, are also reported in there. The table 2 recalls some of the results presented in [2]. [3]

Regular grids				Adaptive grids			
model	$\bar{\mu}$	σ	cpu	model	$\bar{\mu}$	σ	cpu
$\mathcal{M}_{\mathbf{6}}$	5.07^o	7.20^o	34s	$\mathcal{M}_{\mathbf{6}}$	5.09^o	7.20^o	28s
$\mathcal{M}_{\mathbf{4}}$	5.56^o	9.20^o	72s	$\mathcal{M}_{\mathbf{4}}$	6.57^o	9.20^o	58s
$\mathcal{M}_{\mathbf{2}}$	4.97^o	7.64^o	64s	$\mathcal{M}_{\mathbf{2}}$	6.53^o	7.64^o	42s
$\mathcal{M}_{\mathbf{6,4}}$	5.01^o	7.23^o	58s	$\mathcal{M}_{\mathbf{6,4}}$	5.17^o	7.62^o	44s
$\mathcal{M}_{\mathbf{6,2}}$	4.69^o	6.89^o	78s	$\mathcal{M}_{\mathbf{6,2}}$	5.25^o	7.87^o	72s
$\mathcal{M}_{\mathbf{6,4,2}}$	4.75^o	6.89^o	92s	$\mathcal{M}_{\mathbf{6,4,2}}$	5.31^o	7.86^o	73s

Table 1: *Results on* Yosemite

Technique	$\bar{\mu}$	σ	dens.
H. and S. (original)	31.69^o	31.18^o	100%
Horn and Schunck (modified)	9.78^o	16.19^o	100%
Uras *et al.*	8.94^o	15.61^o	100%
Lucas and Kanade	4.28^o	11.41^o	35.1%
Fleet and Jepson	4.63^o	13.42^o	34.1%

Table 2: *Comparative results on* Yosemite

On Yosemite sequence our methods associated to regular grids provide a *dense* estimate almost as good as those obtained with the best (non-dense) mentioned methods. Besides, except in the case of the simplified affine model, the standard deviation are significantly lowered. The best results are obtained for multiple parameterization models on regular grids. This is particularly noticeable when the three parametric models are associated or when the 6-parameter affine model is coupled with the constant model.

The $\mathcal{M}_{\mathbf{6}}$ model on adaptive grids gives the lowest computation times. Compared to the other models on irregular grids, it also yields the best results. In that case in particular, multi-parametric models do not improved the results. At that point, let us remark that the comparison test developed by Barron *et al.* does not take really into account the ability of a method to detect more or less accurately the spatial discontinuities of the velocity field. It measures only an average deviation between the real flow field and the estimated one. As one would expect, the result of $\mathcal{M}_{\mathbf{4}}$ and $\mathcal{M}_{\mathbf{2}}$ on adaptive grids are less good. These models have to be used with a regular subdivision strategy.

To complete our comparisons, we show results obtained on a real world sequence. Figure 1 presents for the atmospheric image Depression, the final velocity fields estimated by two different multigrid estimators. The first flow (Fig. 1c) is the final flow obtained by $\mathcal{M}_{\mathbf{2}}$ on regular grids. The second one (Fig. 1d) has

[3] Results on a cropped sequence where the sky is removed are also sometimes presented [4, 11, 13].

been produced by $\mathcal{M}_6$ on adaptive grid structure (the corresponding final grid is shown in figure 1b). The two flows are displayed the same way, namely subsampled by 6 and magnified by four.

We can notice that with the constant constraint, the flow is drastically under-estimated compared to the one produced with the affine constraint. Besides, the blockwise constant multigrid yields in this case an over-smoothing of the solution. As a consequence on Depression, local features of interest such as the depression center in left upper corner of the image are concealed. This is not the case with affine modeling where the depression center is visible and may be easily identified in an automatic way.

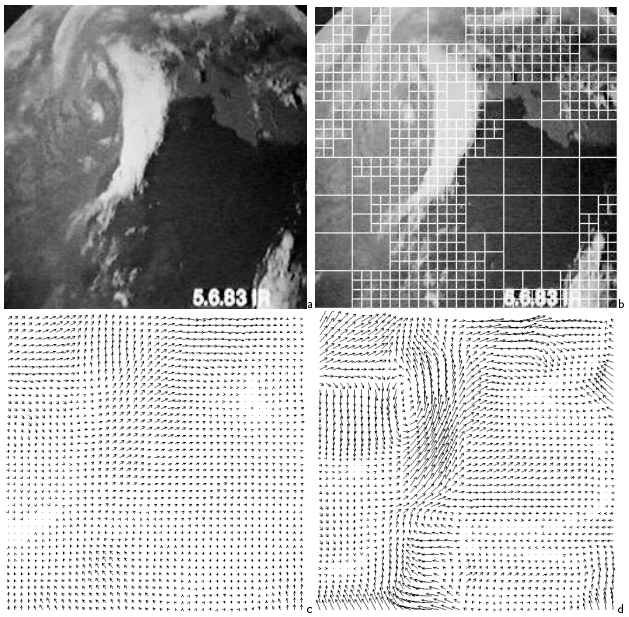

Figure 1: *Results on* Depression*: (***a***) one frame, (***b***) final adaptive grid, (***c***) flow estimate with the constant regular multigrid algorithm (***cpu: 105s***), (*d*) flow estimate with the affine adaptive multigrid algorithm (***cpu: 24s***).*

Finally, note the whole multiresolution/multigrid algorithm converges quickly, since only ten or so low cost iterations are required at each grid level.

5 Conclusion

In this paper, we have presented a comprehensive multigrid framework for incremental optical flow estimation. The problem is expressed as the global minimization of an energy function which involves robust estimators to avoid spatial over-smoothing and to attenuate the influence of large deviations from the OFC. The minimization is efficiently performed through a multigrid algorithm which consists in imposing successively weaker and weaker constraints on the searched estimates. The formulation of this framework is general and allows to mix different parameterization levels of the flow field. Furthermore, it yields a unified and coherent description of an hierarchical motion estimators family which gives good results on sequences involving fluid or rigid motions [13].

References

[1] G. Adiv. Determining three-dimensional motion and structure from optical flow generated by several moving objects. *PAMI*, 7:384–401, jul 1985.

[2] J. Barron, D. Fleet, and S. Beauchemin. Performance of optical flow techniques. *IJCV*, 12(1):43–77, 1994.

[3] J.R. Bergen, P. Anandan, K.J. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In *ECCV'92*, pages 237–252, May 1992.

[4] M. Black and P. Anandan. The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. *CVIU*, 63(1):75–104, 1996.

[5] P. Bouthemy and E. François. Motion segmentation and qualitative dynamic scene analysis from an image sequence. *IJCV*, 10(2):157–182, 1993.

[6] P. Charbonnier, L. Blanc-FÉRaud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Processing*, 3(1), 1996.

[7] I. Cohen and I. Herlin. Optical flow and phase portrait methods for environmental satellite image sequences. In *ECCV'96*, pages 141–150, April 1996.

[8] D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *PAMI*, 14(3):367–383, 1992.

[9] S. Geman and D. E. McClure. Statistical methods for tomographic image reconstruction. In *Bull. ISI, Proc. 46th Session Int. Statistical Institute*, volume 52, 1987.

[10] M. Irani, B. Rousso, and S. Peleg. Computing occluding and transparaent motions. *IJCV*, 12(1):5–16, 1994.

[11] X. Ju, M.J. Black, and A.D. Jepson. skin and bones: multilayer, locally affine, optical flow and regularization with transparency. In *CVPR'96*, pages 307–314, 1996.

[12] Y. Leclerc. Constructing simple stable descriptions for image partitioning. *IJCV*, 3:73–102, 1989.

[13] E. Mémin and P. Pérez. Robust discontinuity-preserving model for estimating optical flow. In *ICPR'96*, Vienna, Austria, August 1996.

[14] E. Mémin, P. Pérez, and D. Machecourt. Dense estimation and object-based segmentation of the optical flow with robust techniques. Technical Report 991, IRISA, March 1996 (**http://wwww.irisa.fr/**).

[15] P. Proesmans, L. van Gool, E. Pauwels, and A. Oosterlinck. Determination of optical flow and its discontinuities using non-linear diffusion. In *ECCV'94*, volume 2, pages 295–304, Stokholm, Sweden, 1994.

[16] R. Szeliski H.-Y. Shum. Motion estimation with quadtree splines. In *ICCV'95*, pages 757–763, 1995.