Dense Estimation and Object-Based Segmentation of the Optical Flow with Robust Techniques

Etienne Mémin and Patrick Pérez

Abstract—In this paper, we address the issue of recovering and segmenting the apparent velocity field in sequences of images. As for motion estimation, we minimize an objective function involving two robust terms. The first one cautiously captures the *optical flow constraint*, while the second (*a priori*) term incorporates a *discontinuity-preserving* smoothness constraint. To cope with the nonconvex minimization problem thus defined, we design an efficient deterministic *multigrid* procedure. It converges fast toward estimates of good quality, while revealing the large discontinuity structures of flow fields. We then propose an extension of the model by attaching to it a flexible *object-based segmentation device* based on deformable closed curves (different families of curve equipped with different kinds of prior can be easily supported). Experimental results on synthetic and natural sequences are presented, including an analysis of sensitivity to parameter tuning.

Index Terms—Closed segmenting curve, incremental multiresolution, motion segmentation, multigrid nonconvex minimization, optical flow, robust estimators.

I. INTRODUCTION

MANY TASKS in computer vision and image analysis can be expressed as global optimization problems. The general issue is to find the global minimum of a cost function (or *energy*) involving the data and the "hidden" variables of interest to be extracted from the data. Usually, a first part of the energy expresses the interaction between the unknown variables and the data, while the second one captures some kind of prior knowledge about the researched information. This latter ingredient is often a mere regularization term that only encodes a weak prior, but whose essential role is to remedy the ill-posed nature of the problem at hand (it guarantees, to some extent, the existence and the uniqueness of a consistent solution which continuously depends on the data). To keep the energy-based model tractable, the cost function usually decomposes as a sum of local interaction functions associated with a neighborhood system [19], [29].

Within this energy minimization framework, we address here the particular problem of optical flow estimation, and its possible association with some kind of optical flow-based segmentation. The motion estimator that we present belongs to the class of differential methods which make use of the so-called optical flow constraint (OFC). Unfortunately, this modeling is very sensitive to noise and behaves badly when spatial or temporal variations of the luminance function are too large. Another source of concern arises from the smoothness regularizing prior, which is usually associated to these models. The frontiers that demarcate the different apparent motions coexisting within the same scene are ignored by a blind smoothing. This deficiency results in a bad estimation nearby these border lines. A great deal of studies has been dedicated to this specific problem of discontinuity-preserving regularization in computing optical flow (and in computer vision in general): within Markovian framework, binary edge variables similar to Geman and Geman's line processes [19] have thus been introduced (see, for instance, [22], [28] and [41]); within anisotropic diffusion framework, nonlinear Euler–Lagrange PDE's have been devised along the same philosophy [11], [14], [15], [32], [35].

Adopting a more global viewpoint, Black has pointed out in [3] that the different problems we have just evoked can all be seen as "deviations" from a model (either the data model or the smoothing prior model). Though different in nature, they can hopefully be located and treated within a unified framework: the one offered by the *robust statistics* whose original aim is the estimation of models in presence of many deviating occurrences among the data [25].

Nevertheless, the introduction of robust estimators in energy-based image applications leads most of the time to a global nonlinear minimization in presence of numerous local minima. Even within a *multiresolution* formulation of the problem (which is almost inescapable in case of long range motions to be estimated), one has to deal with a sequence of global optimization problems which remain tricky.

To avoid the use of greedy stochastic algorithms, some authors [3], [29] have proposed to get benefit from the "scale" parameter involved in standard robust estimators. A proper and progressive tuning of this parameter allows to define a minimization strategy similar in spirit to continuation methods such as the "graduated nonconvexity" algorithm proposed by Black and Zisserman [10].

For the same purpose (i.e., definition of an efficient deterministic algorithm to deal with the global optimization problem at hand), we propose here to extend the *multigrid* relaxation method presented in [23]. This method consists in the minimization of the energy function through a hierarchy of nested subspaces of the whole configuration space. These subspaces correspond to configurations constrained to be piecewise constant over smaller and smaller pixel subsets.

This is equivalent to handling velocity fields which lie on reduced grids, whereas observations are still viewed at the current resolution.

It is well known that the estimation of the optical flow in the one hand, and the segmentation of the images with respect to the apparent motion in the other hand, are two important issues of motion analysis which should help each other to provide better results as well as richer information [12], [38]. Following this philosophy, we propose a connection of the optical flow estimator under concern with an object-based motion segmentation. This is addressed within an *augmented cost function*, which involves one or several segmenting closed curves.

From the point of view of the optical flow estimation, this curve is used to hopefully drive and structure the apparition of spatial discontinuities, thus improving the quality of the estimation. Conversely, the segmenting curve, which provides interesting information of its own, is driven by the spatial discontinuities of the velocity field under estimation. Without much pain, the efficient multigrid algorithm can be adapted to this joint model.

The paper is organized as follows. In Section II, we start from the standard gradient-based optical flow estimation, to design an energy-based multiresolution model for a robust and discontinuity-preserving estimation of the apparent motion. The global optimization issue is then addressed in Section III. A deterministic multigrid algorithm involving iteratively reweighted least squares estimation is proposed. In Section IV, we explore a way of coupling the motion estimation process with a simultaneous segmentation of the current optical flow. We especially show how a segmenting closed curve can interact with the discontinuity "indicators" to extract moving entities while improving the optical flow estimation nearby. In Section V, experimental results of optical flow estimation and object-based motion segmentation on synthetic and real-world sequences are reported and discussed.

II. ROBUST INCREMENTAL OPTICAL FLOW ESTIMATION

Let $\boldsymbol{w} = \{\boldsymbol{w}_s, s \in S\} \in \Omega$ and $f(t) = \{f(s, t), s \in S\}$ be, respectively, the unknown bidimensional velocity field at time t and the luminance function at time t, both defined on the rectangular pixel lattice S. The configuration space to explore, Ω, is a bounded subset of $(\mathbb{R} \times \mathbb{R})^S$.

Assuming that the luminance of a given "physical point" does not change much between times t and $t + 1$, and that the velocity field is reasonably smooth, one often addresses the optical flow recovery problem by minimizing an objective function of the following type [28]:

$$U(\boldsymbol{w}; f) \triangleq \sum_{s \in S} [f(s + \boldsymbol{w}_s, t + 1) - f(s, t)]^2 + \alpha \sum_{\langle s, r \rangle \in \mathcal{C}} \|\boldsymbol{w}_s - \boldsymbol{w}_r\|^2 \qquad (1)$$

where $\mathcal{C}$ is the set of neighboring site pairs (with respect to the first- or second-order neighborhood system ν), and $\alpha > 0$ is a parameter controlling the balance between the smoothness constraint and the global adequacy to the brightness constancy

assumption. The first term of this energy, which depends on what the data exactly are, is highly nonconvex. To cope with this problem, it is usual (as in Gauss–Newton minimization procedures) to linearize this term, provided that a good estimate $\boldsymbol{w}$ is, somehow, already available. If so, an *incremental refinement* $\mathrm{d}\boldsymbol{w}$ is searched [2] by minimizing

$$U(\boldsymbol{w} + \mathrm{d}\boldsymbol{w}; f) \approx \sum_{s \in S} [\nabla f(s + \boldsymbol{w}_s, t + 1)^T \, \mathrm{d}\boldsymbol{w}_s + f_t(s, \boldsymbol{w}_s)]^2 + \alpha \sum_{\langle s, r \rangle \in \mathcal{C}} \|(\boldsymbol{w}_s + \mathrm{d}\boldsymbol{w}_s) - (\boldsymbol{w}_r + \mathrm{d}\boldsymbol{w}_r)\|^2$$

where $\nabla f = [f_x \quad f_y]^T$ stands for the spatial gradient of luminance f, and $f_t(s, \boldsymbol{w}_s) \triangleq f(s + \boldsymbol{w}_s, t + 1) - f(s, t)$ is the displaced frame difference. The first term of this approximate expression now measures the deviation from the well-known OFC, here displaced according to $\boldsymbol{w}$.

The weaknesses of this differential modeling, as follows, are well known.

1) The first-order expansion used is valid for some s if the displacement $\mathrm{d}\boldsymbol{w}_s$ falls into the domain of "linearity" of the luminance $f(\bullet, t + 1)$ around $s + \boldsymbol{w}_s$, this region being smaller as the spatial gradient gets larger in the direction of $\mathrm{d}\boldsymbol{w}_s$. As a consequence, large contribution to this data term can be expected for large increments $\mathrm{d}\boldsymbol{w}_s$ and at sharp edges.

2) The underlying assumption of brightness constancy is already very likely to be violated in cases of occlusions, transparency, specular reflection, change of illumination, etc.

3) Most of "real" velocity fields are, at most, piecewise smooth: they usually exhibit motion discontinuities that risk to be ignored and smoothed out by the quadratic prior.

To efficiently cope with the large deviations both from the data model and from the prior model, one can replace the quadratic penalties by *robust* penalty functions [3], [5] in the approximated energy. Thus, we now consider the new energy function $H \triangleq H_1 + \alpha H_2$, with

$$H_1(\mathrm{d}\boldsymbol{w}; f, \boldsymbol{w}) \triangleq \sum_{s \in S} \rho_1[\nabla f(s + \boldsymbol{w}_s, t + 1)^T \, \mathrm{d}\boldsymbol{w}_s + f_t(s, \boldsymbol{w}_s)], \qquad (2)$$

$$H_2(\mathrm{d}\boldsymbol{w}; \boldsymbol{w}) \triangleq \sum_{\langle s, r \rangle \in \mathcal{C}} \rho_2[\|(\boldsymbol{w}_s + \mathrm{d}\boldsymbol{w}_s) - (\boldsymbol{w}_r + \mathrm{d}\boldsymbol{w}_r)\|] \qquad (3)$$

where functions ρ_1 and ρ_2 are standard robust estimators.[1]

The estimation of the "guess" $\boldsymbol{w}$ is usually addressed by embedding this refinement approach in a *multiresolution coarse-to-fine* procedure [2], [3], [16], [22]. To this end, for each instant t of the sequence, a pyramid of images $f^k(t) = \{f^k(s, t), s \in S^k\}, k = 0 \cdots N$, is derived from the original frame $f(t)$ by successive Gaussian smoothings and regular

---

[1] See [8] for a catalogue of such robust objective functions in computer vision. Most of them are bounded, and therefore *nonconvex*.
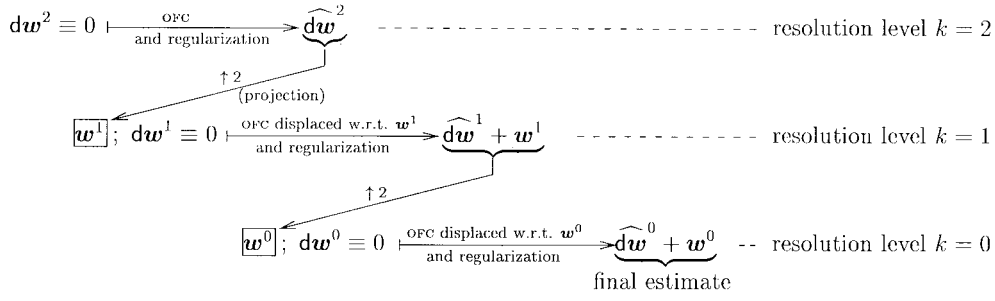
Fig. 1. Synopsis of the multiresolution incremental optical flow estimation on three levels ($N = 2$).
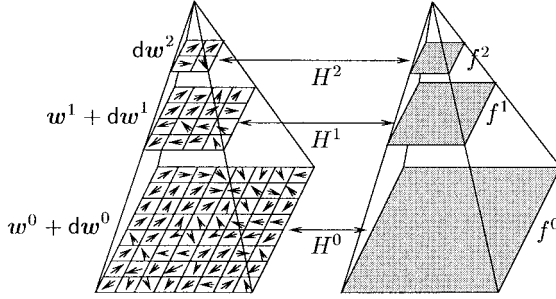
Fig. 2. Energetic structure of the multiresolution model ($N = 2$).

subsamplings by a factor of two in each direction.[2] The created pyramidal structure then allows to incrementally estimate the velocity field, by using *ad hoc* "reductions" of energy H: Low resolution components are estimated at coarsest level where the domain of validity of the linearized data model is hopefully larger (due to the joint reduction of spatial gradients by smoothing, and of motion magnitude by subsampling). This crude estimate is then refined step by step: at resolution k, an increment velocity field $\mathrm{d}\boldsymbol{w}^k$ in $\Omega^k \subset (\mathbb{R} \times \mathbb{R})^{S^k}$ is estimated around the "projection" (by duplication or bilinear interpolation) $\boldsymbol{w}^k$ of the final estimate at previous resolution level $k + 1$ (see Fig. 1).

This is performed by minimizing the reduced energy function $H^k \triangleq H_1^k + \alpha H_2^k$, with

$$H_1^k(\mathrm{d}\boldsymbol{w}^k; f^k, \boldsymbol{w}^k)$$
$$\triangleq \sum_{s \in S^k} \rho_1[\nabla f^k(s + \boldsymbol{w}_s^k, t+1)^T \mathrm{d}\boldsymbol{w}_s^k + f_t^k(s, \boldsymbol{w}_s^k)] \quad (4)$$

$$H_2^k(\mathrm{d}\boldsymbol{w}^k; \boldsymbol{w}^k)$$
$$\triangleq \sum_{\langle s, r \rangle \in \mathcal{C}^k} \rho_2[\|(\boldsymbol{w}_s^k + \mathrm{d}\boldsymbol{w}_s^k) - (\boldsymbol{w}_r^k + \mathrm{d}\boldsymbol{w}_r^k)\|] \quad (5)$$

where $\mathcal{C}^k$ is the set of neighboring site pairs (w.r.t. neighborhood system ν) lying on grid S^k. Fig. 2 shows this multiresolution energy setup.

From a minimization point of view, the complete multiresolution procedure can be viewed as a Gauss–Newton-type minimization of the initial nonconvex energy U defined by (1). However, with the introduction of the different pyramidal ingredients (smoothing/sampling of data, duplication/reduction

---

[2]The resolution superscript k will span from 0 for the finest resolution, to N for the coarsest one.

of energy expressions, top-down propagation of estimates), along with the degree of nonlinearity of U depending on the data, no convergence guaranty of the whole procedure seems to be within reach. However, this standard and pragmatic multiresolution approach is of particular interest when the scene at hand exhibits *large displacements* at some locations in the image.

We now return to the robust aspect of the model. Roughly speaking, robust objective functions are continuous even functions, increasing on $\mathbb{R}^+$, and which penalize large "residual" values less drastically than quadratic functions do. This is usually achieved by letting their derivative, called *influence function* in robust statistics [25], have finite limit at infinity (usually zero). In the model currently discussed, this characteristic makes them robust to data model outliers (for ρ_1) or to spatial discontinuities of the apparent motion field (for ρ_2).

To give an insight into robust estimation as well as a practical way of handling it, it is fruitful transforming its use in terms of dual optimization problem involving *auxiliary variables* [3], [8], [13], [18]. To this end, we use the following reformulation result (see [8] or [13] for a complete account): Let ρ be a real-valued continuously differentiable even function such that

1) ρ is increasing on $\mathbb{R}^+$;
2) $\phi(v) \triangleq \rho(\sqrt{v})$ is *strictly concave* on $\mathbb{R}^+$;
3) $\lim_{v \to +\infty} \phi'(v) = 0$;
4) $\tau \triangleq \lim_{v \to 0^+} \phi'(v) < +\infty$.

There exits a function ψ, continuously differentiable on $(0, 1]$, such that

$$\forall u \in \mathbb{R}^+, \qquad \rho(u) = \min_{z \in (0, 1]} \tau z u^2 + \psi(z). \quad (6)$$

This means that the graph of ρ is the inferior envelope of a family of parabolas continuously indexed by $z \in (0, 1]$. The minimum in (6) is given in closed form by [8] and [13]:

$$\arg\min_{z \in (0, 1]} \tau z u^2 + \psi(z) = \frac{\rho'(u)}{2\tau u} = \sigma^2 \phi'(u^2) \quad (7)$$

where parameter $\sigma^2 \triangleq \tau^{-1}$ is analog to a variance [$\rho(u) \sim_0 u^2/\sigma^2$]. Function ψ is obtained as $\psi(z) \triangleq \phi \circ (\phi')^{-1}(\tau z) - \tau z (\phi')^{-1}(\tau z)$.

With such a cost function ρ, one can thus replace the multidimensional minimization in x of some $\sum_i \rho[g_i(x)]$ by the minimization in $(x, \{z_i\})$ of $\sum_i [\tau z_i g_i(x)^2 + \psi(z_i)]$ since both sums have the same global minimum in x. The extra

(a)   (b)
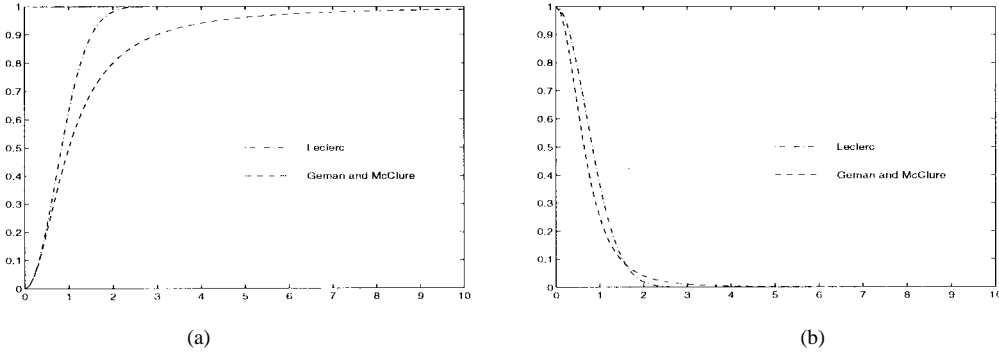
Fig. 3.   (a) Leclerc's estimator $\rho(u) = 1 - \exp[-(u^2/\sigma^2)]$ and Geman–McClure's one $\rho(u) = u^2/(\sigma^2 + u^2)$ for $\sigma = 1$. (b) Associated optimal weight functions $\hat{z}(u) = \exp[-(u^2/\sigma^2)]$ and $\hat{z}(u) = \sigma^4/(\sigma^2 + u^2)^2$, for $\sigma = 1$.

variables z_is act as *adaptative weights* continuously lying in $(0, 1]$. Note that in practice, each function g_i only depends on a few components of x (one or two in our problem, as we shall see).

The minimization of the new compound function is usually lead alternatively with respect to x and to the z_i,s, as follows.

• The z_is being frozen, the minimization in x becomes

$$\arg \min_x \sum_i \left[\tau z_i g_i(x)^2 + \psi(z_i) \right]$$

$$= \arg \min_x \sum_i \tau z_i g_i(x)^2 = \hat{x}(z).$$

If g_is are affine forms, one has to face a standard weighted least squares problem equivalent to the resolution of a (sparse) linear system.

• x being frozen, the *simultaneous* minimization in $\{z_i\}$ yields

$$\hat{z}_i(x) = \frac{\rho'[g_i(x)]}{2\tau g_i(x)} = \sigma^2 \phi'[g_i(x)^2] \qquad (8)$$

according to (7). The assumptions about ρ mean that ϕ' strictly decreases from $\mathbb{R}^+$ into $(0, \tau]$: when the ith "residual" $|g_i(x)|$ gets larger, the corresponding optimal weight $\hat{z}_i$ gets smaller and smaller, providing the robustness of the estimator.

In case g_is are affine, the whole alternate procedure constitutes an *iteratively reweighted least squares* estimation [24]. Fig. 3 shows the two robust estimators we used in our experiments and their associated optimal weight functions [20], [29].

Apart from the practical advantage in case of affine residuals, transforming a robust estimator-based energy into its auxiliary variable formulation offers appealing modeling flexibility: the dual compound model can be extended by adding interactions among auxiliary variables (e.g., to capture some "geometric" *a priori* knowledge, either local or more global, on the discontinuity configurations). The extended model presented in Section IV relies on these variables to couple the estimation of motion with an object-based motion segmentation.

In our case the weights are of two natures: (a) *data outliers weights* (related to the dual formulation of H_1^k), and (b) *discontinuity weights* lying on the dual grid of S^k (provided
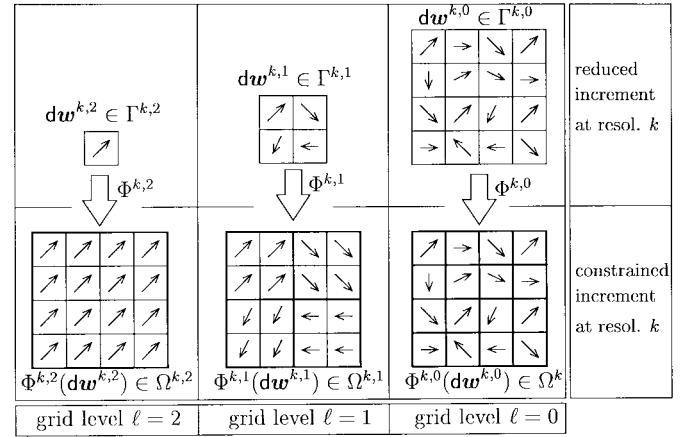

Fig. 4.   Constrained increment fields (at resolution level k, from $\Omega^{k,2}$, $\Omega^{k,1}$, Ω^k), and reduced increment fields associated to them through $[\Phi^{k,\ell}]^{-1}$, $\ell = 2, 1, 0$.

by the dual formulation of H_2^k). The first set of weights, denoted by $\delta^k \triangleq \{\delta_s^k, s \in S^k\}$, allows to attenuate the effect of data for which the OFC is violated. The second one, denoted by $\beta^k \triangleq \{\beta_{sr}^k, \langle s, r \rangle \in \mathcal{C}^k\}$, prevents from oversmoothing in locations obviously exhibiting significant velocity discontinuities. The estimation at resolution level k is now expressed as the global minimization of $\mathcal{H}^k \triangleq \mathcal{H}_1^k + \alpha \mathcal{H}_2^k$, with

$$\mathcal{H}_1^k(\mathrm{d}\boldsymbol{w}^k, \delta^k; f^k, \boldsymbol{w}^k)$$
$$\triangleq \sum_{s \in S^k} \{ \tau_1 \delta_s^k [\nabla f^k(s + \boldsymbol{w}_s^k, t+1)^T \, \mathrm{d}\boldsymbol{w}_s^k$$
$$+ f_t^k(s, \boldsymbol{w}_s^k)]^2 + \psi_1(\delta_s^k) \} \qquad (9)$$

$$\mathcal{H}_2^k(\mathrm{d}\boldsymbol{w}^k, \beta^k; \boldsymbol{w}^k)$$
$$\triangleq \sum_{\langle s, r \rangle \in \mathcal{C}^k} [\tau_2 \beta_{sr}^k \|(\boldsymbol{w}_s^k + \mathrm{d}\boldsymbol{w}_s^k) - (\boldsymbol{w}_r^k + \mathrm{d}\boldsymbol{w}_r^k)\|^2$$
$$+ \psi_2(\beta_{sr}^k)]. \qquad (10)$$

However, the underlying energy function H^k being often nonconvex with respect to the unknown variables of interest (it will be the case in our experiments), we still have to deal
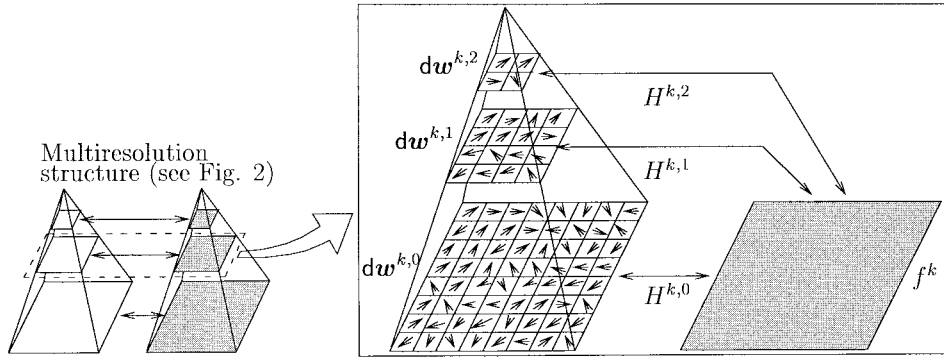
Fig. 5. Multigrid structure ($L = 2$) for relaxation at resolution level k.

with a tough optimization problem, despite the reformulation. In particular, the alternate minimization procedure is not guaranteed to reach the global minimum, even though each step is an exact minimization (but with respect to only a subset of variables). Actually, only a "local" minimum depending on the initialization is reached. For this reason, we designed an extension of the multigrid method proposed in [23], which hopefully converges fast without getting stuck in high local minima.

III. MULTIGRID DETERMINISTIC OPTIMIZATION

To efficiently cope with the global optimization problem at resolution k, we design a hierarchical "constrained" exploration of the configuration space Ω^k: the optimization is led through a sequence of nested configuration subspaces $\Omega^{k,L} \subset \Omega^{k,L-1} \subset \cdots \subset \Omega^{k,1} \subset \Omega^{k,0} = \Omega^k$, where $\Omega^{k,\ell}$ is the set of increment fields which are piecewise constant according to a $2^\ell \times 2^\ell$-block partition of grid S^k. Denote $\mathcal{B}^{k,\ell} \triangleq \{\mathcal{B}_n^{k,\ell}, n = 1 \cdots N_{k,\ell}\}$ this partition, the number of blocks being $N_{k,\ell} = |S^k|/4^\ell = |S|/4^{\ell+k}$. Each constrained field of $\Omega^{k,\ell}$ is equivalent to a *reduced increment field* $\mathrm{d}\boldsymbol{w}^{k,\ell}$ lying on the grid $S^{k,\ell} \triangleq \{1, \cdots, N_{k,\ell}\}$ associated with $\mathcal{B}^{k,\ell}$. Let $\Gamma^{k,\ell}$ be the set of such reduced fields and let $\Phi^{k,\ell}$ be the one-to-one mapping from $\Gamma^{k,\ell}$ into $\Omega^{k,\ell}$ (see Fig. 4).

Constrained optimization in $\Omega^{k,\ell}$ is then equivalent to the minimization of the new energy function

$$\mathcal{H}^{k,\ell}(\mathrm{d}\boldsymbol{w}^{k,\ell}, \delta^k, \beta^k; f^k, \boldsymbol{w}^k)$$
$$\triangleq \mathcal{H}^k[\Phi^{k,\ell}(\mathrm{d}\boldsymbol{w}^{k,\ell}), \delta^k, \beta^k; f^k, \boldsymbol{w}^k]. \quad (11)$$

At each resolution, we now have a cascade of optimization problems of reduced complexity:

$$\underset{\mathrm{d}\boldsymbol{w}^{k,\ell}, \delta^k, \beta^k}{\arg \min} \ \mathcal{H}^{k,\ell}(\mathrm{d}\boldsymbol{w}^{k,\ell}, \delta^k, \beta^k; f^k, \boldsymbol{w}^k), \ell = L \cdots 0 \quad (12)$$

where $\mathrm{d}\boldsymbol{w}^{k,\ell} \in \Gamma^{k,\ell}$ lies on the reduced grid $S^{k,\ell}$, while weights and data remain attached to S^k, whatever the grid level ℓ (see Fig. 5, where $H^{k,\ell} = \min_{\delta^k, \beta^k} \mathcal{H}^{k,\ell}$). Each of these problems is processed in terms of iteratively reweighted least squares within a multigrid coarse-to-fine strategy: the final estimate at level $\ell+1$ has a natural image at level ℓ (through $[\Phi^{k,\ell}]^{-1} \circ \Phi^{k,\ell+1}$), which is used as an initial configuration for the deterministic relaxation algorithm at that level. This

procedure is repeated until the finest level $\ell = 0$ is reached (see [23] for further details). Note that the definition of the reduced energy by (11) along with the successive inclusion of the configuration subspaces, ensures that the energy $\mathcal{H}^k$ keeps *decreasing* as the multigrid iterations proceed. Experiments reported in Section V will show that the multigrid approach is of great benefit both in terms of convergence speed and quality of final estimates.

A. Multigrid Energy Derivation

We now go into deeper details about the new multigrid function $\mathcal{H}^{k,\ell}$, which is obviously composed of two terms similar to those of $\mathcal{H}^k$: $\mathcal{H}^{k,\ell} = \mathcal{H}_1^{k,\ell} + \alpha \mathcal{H}_2^{k,\ell}$. For sake of readability, we will omit the resolution superscript k in all expressions throughout the remainder of this section. *All computations will be meant to concern resolution level k.*

1) Data Model Adequation Term: For any $n \in S^l$, denote $s_1, \cdots, s_{4^\ell}$ the sites of block $\mathcal{B}_n^\ell$, and define the following blockwise expressions:

$$\boldsymbol{\delta}_n^\ell \triangleq [\delta_{s_1} \cdots \delta_{s_{4^\ell}}]^T, \ \Psi_1^\ell(\boldsymbol{\delta}_n^\ell) \triangleq \sum_{s \in \mathcal{B}_n^\ell} \psi_1(\delta_s),$$

$$\Delta_n^\ell \triangleq \mathrm{diag}(\delta_{s_1}, \cdots, \delta_{s_{4^\ell}}),$$

$$\boldsymbol{f}_t^\ell(n, \boldsymbol{w}) \triangleq [f_t(s_1, \boldsymbol{w}_{s_1}) \cdots f_t(s_{4^\ell}, \boldsymbol{w}_{s_{4^\ell}})]^T,$$

$$\boldsymbol{f}_\bullet^\ell(n, \boldsymbol{w}) \triangleq [f_\bullet(s_1 + \boldsymbol{w}_{s_1}, t+1) \cdots$$
$$f_\bullet(s_{4^\ell} + \boldsymbol{w}_{s_{4^\ell}}, t+1)]^T, \quad \text{for } \bullet = x \text{ or } y,$$

and

$$\nabla f^\ell(n, \boldsymbol{w}) \triangleq [\boldsymbol{f}_x^\ell(n, \boldsymbol{w}) \boldsymbol{f}_y^\ell(n, \boldsymbol{w})].$$

Also, we will denote $\langle X|Y \rangle_n \triangleq X^T \Delta_n^\ell Y$ for any two 4^ℓ-row matrices or vectors, and $\|X\|_n^2 \triangleq \langle X|X \rangle_n$ for any 4^ℓ-component column vector. It is then easy to get the following compact expression:

$$\boxed{\begin{aligned} &\mathcal{H}_1^\ell(\mathrm{d}\boldsymbol{w}^\ell, \delta; f, \boldsymbol{w}) \\ &= \sum_{n \in S^\ell} [\tau_1 \|\nabla f^\ell(n, \boldsymbol{w}) \mathrm{d}\boldsymbol{w}_n^\ell + \boldsymbol{f}_t^\ell(n, \boldsymbol{w})\|_n^2 + \Psi_1^\ell(\boldsymbol{\delta}_n^\ell)] \end{aligned}}$$

$$(13)$$

which is *very similar* to the one of the "parent" energy $\mathcal{H}_1$ (9). For each block, one gets a blockwise optical flow constraint expression involving aggregated observations.

2) Smoothing Term: Let $\mathcal{C}_n^\ell \triangleq \{\langle s, r \rangle \in \mathcal{C}: \langle s, r \rangle \subset \mathcal{B}_n^\ell\}$ be the set of neighboring site pairs included in block $\mathcal{B}_n^\ell$ and $\mathcal{C}_{nm}^\ell \triangleq \{\langle s, r \rangle \in \mathcal{C}: s \in \mathcal{B}_n^\ell, r \in \mathcal{B}_m^\ell\}$ the set of neighboring site pairs straddling blocks $\mathcal{B}_n^\ell$ and $\mathcal{B}_m^\ell$. These sets $\{\mathcal{C}_n^\ell\}$ and $\{\mathcal{C}_{nm}^\ell\}$ form a partition of $\mathcal{C}$ and reduced grid S^ℓ turns out to be equipped with the *same neighborhood system* as S (i.e., first- or second-order neighborhood system). The corresponding set of neighboring pairs will be denoted by $\mathcal{C}^\ell$. The smoothing term of $\mathcal{H}^\ell$ is

$$
\mathcal{H}_2^\ell(\mathrm{d}\boldsymbol{w}^\ell, \beta; \boldsymbol{w})
$$
$$
= \tau_2 \left[\sum_{n \in S^\ell} \sum_{\langle s, r \rangle \in \mathcal{C}_n^\ell} \beta_{sr} \|\boldsymbol{w}_s - \boldsymbol{w}_r\|^2 + \sum_{\langle n, m \rangle \in \mathcal{C}^\ell} \sum_{\langle s, r \rangle \in \mathcal{C}_{nm}^\ell} \right.
$$
$$
\left. \cdot \, \beta_{sr} \|(\boldsymbol{w}_s + \mathrm{d}\boldsymbol{w}_n^\ell) - (\boldsymbol{w}_r + \mathrm{d}\boldsymbol{w}_m^\ell)\|^2 \right]
$$
$$
+ \sum_{\langle s, r \rangle \in \mathcal{C}} \psi_2(\beta_{sr}). \tag{14}
$$

Noting that

$$
\tau_2 \left[\sum_{n \in S^\ell} \sum_{\langle s, r \rangle \in \mathcal{C}_n^\ell} \beta_{sr} \|\boldsymbol{w}_s - \boldsymbol{w}_r\|^2 \right.
$$
$$
\left. + \sum_{\langle n, m \rangle \in \mathcal{C}^\ell} \sum_{\langle s, r \rangle \in \mathcal{C}_{nm}^\ell} \beta_{sr} \|\boldsymbol{w}_s - \boldsymbol{w}_r\|^2 \right]
$$
$$
+ \sum_{\langle s, r \rangle \in \mathcal{C}} \psi_2(\beta_{sr})
$$
$$
= \sum_{\langle s, r \rangle \in \mathcal{C}} [\tau_2 \beta_{sr} \|\boldsymbol{w}_s - \boldsymbol{w}_r\|^2 + \psi_2(\beta_{sr})]
$$
$$
= \mathcal{H}_2(\boldsymbol{0}, \beta; \boldsymbol{w})
$$

one gets the following reduced prior energy:

$$
\boxed{
\begin{aligned}
\mathcal{H}_2^\ell&(\mathrm{d}\boldsymbol{w}^\ell, \beta; \boldsymbol{w}) \\
&= \mathcal{H}_2(\boldsymbol{0}, \beta; \boldsymbol{w}) + \tau_2 \sum_{\langle n, m \rangle \in \mathcal{C}^\ell} [\beta_{nm}^\ell \|\mathrm{d}\boldsymbol{w}_n^\ell - \mathrm{d}\boldsymbol{w}_m^\ell\|^2 \\
&\quad + 2(\mathrm{d}\boldsymbol{w}_n^\ell - \mathrm{d}\boldsymbol{w}_m^\ell)^T \overline{\Delta \boldsymbol{w}}_{nm}^\ell]
\end{aligned}
}
$$
$$
\tag{15}
$$

with $\beta_{nm}^\ell \triangleq \sum_{\langle s, r \rangle \in \mathcal{C}_{nm}^\ell} \beta_{sr}$ and $\overline{\Delta \boldsymbol{w}}_{nm}^\ell \triangleq \sum_{\langle s, r \rangle \in \mathcal{C}_{nm}^\ell} \beta_{sr}(\boldsymbol{w}_s - \boldsymbol{w}_r)$.

B. Energy Minimization

The current reduced increment estimate $\mathrm{d}\boldsymbol{w}^\ell$ being fixed, we know that the optimal weight values are directly accessible. According to (8) in combination with energy definitions (4) and (5), these values are

$$
\hat{\delta}_s = \sigma_1^2 \phi_1' \{[\nabla f(s + \boldsymbol{w}_s, t + 1)^T \mathrm{d}\boldsymbol{w}_n^\ell + f_t(s, \boldsymbol{w}_s)]^2\},
$$
$$
\forall s \in \mathcal{B}_n^\ell, \tag{16}
$$
$$
\hat{\beta}_{sr} = \begin{cases} \sigma_2^2 \phi_2'[\|(\boldsymbol{w}_s + \mathrm{d}\boldsymbol{w}_n^\ell) \\ \quad - (\boldsymbol{w}_r + \mathrm{d}\boldsymbol{w}_m^\ell)\|^2], & \forall \langle s, r \rangle \in \mathcal{C}_{nm}^\ell, \\ \sigma_2^2 \phi_2'[\|\boldsymbol{w}_s - \boldsymbol{w}_r\|^2], & \forall \langle s, r \rangle \in \mathcal{C}_n^\ell. \end{cases} \tag{17}
$$

According to (17), the discontinuity weights β_{sr} located in between two neighboring blocks of $\mathcal{B}^\ell$ (i.e., $\langle s, r \rangle \in \mathcal{C}_{nm}^\ell$ for some $\langle n, m \rangle \in \mathcal{C}^\ell$) are the only ones to be iteratively updated as $\mathrm{d}\boldsymbol{w}^\ell$ evolves. The others only depend on $\boldsymbol{w}$ which is fixed along the whole multigrid procedure at level ℓ. Therefore, they can be computed right away at the first iteration of the current level. As soon as the values of all weights are computed and frozen, the energy function $\mathcal{H}^\ell(\mathrm{d}\boldsymbol{w}^\ell, \delta, \beta; f, \boldsymbol{w})$ is quadratic with respect to $\mathrm{d}\boldsymbol{w}^\ell$. Its minimization is equivalent to the resolution of a linear system whose solution is searched with an iterative Gauss–Seidel scheme. All sites of S^ℓ are repeatedly visited, until convergence. If n is the current site of S^ℓ, the reduced increment vector $\mathrm{d}\boldsymbol{w}_n^\ell$ is updated according to (18), shown at the bottom of the page, with

$$
\overline{\boldsymbol{w}}_n^\ell \triangleq \frac{\displaystyle\sum_{m \in \nu(n)} (\beta_{nm}^\ell \mathrm{d}\boldsymbol{w}_m^\ell - \overline{\Delta \boldsymbol{w}}_{nm}^\ell)}{\displaystyle\sum_{m \in \nu(n)} \beta_{nm}^\ell},
$$
$$
\gamma \triangleq \frac{\alpha \tau_2}{\tau_1} \sum_{m \in \nu(n)} \beta_{nm}^\ell,
$$
$$
A \triangleq \nabla f^{\ell T} \Delta_n^\ell \nabla f^\ell = \begin{bmatrix} \|\boldsymbol{f}_x^\ell\|_n^2 & \langle \boldsymbol{f}_x^\ell | \boldsymbol{f}_y^\ell \rangle_n \\ \langle \boldsymbol{f}_x^\ell | \boldsymbol{f}_y^\ell \rangle_n & \|\boldsymbol{f}_y^\ell\|_n^2 \end{bmatrix}
$$

and "det A," "trace A," "com A" stand, respectively, for the determinant, the trace and the cofactor matrix of A. Note that in the above expressions, $\boldsymbol{f}_x^\ell$, $\boldsymbol{f}_y^\ell$, $\boldsymbol{f}_t^\ell$ vectors, and ∇f^ℓ matrices as well, are displayed without $(n, \boldsymbol{w})$ for the sake of concision.

We have completely described a multigrid reweighted least squares minimization algorithm, which is here devoted to robust optical flow estimation. The use of such a multigrid iterative relaxation in the present context allows to build an optical flow estimator which is formally speaking similar to Horn and Schunck's estimator. In contrast, it is robust to the failures of the OFC-based model, and it is able to localize and preserve the discontinuities of the optical flow.

$$
\boxed{\widehat{\mathrm{d}\boldsymbol{w}}_n^\ell = \overline{\boldsymbol{w}}_n^\ell - \frac{\gamma \langle \nabla f^\ell | \nabla f^\ell \overline{\boldsymbol{w}}_n^\ell + \boldsymbol{f}_t^\ell \rangle_n + \det A \, \overline{\boldsymbol{w}}_n^\ell + \mathrm{com}\, A \langle \nabla f^\ell | \boldsymbol{f}_t^\ell \rangle_n}{\gamma(\gamma + \mathrm{trace}\, A) + \det A}}
$$
$$
\tag{18}
$$

IV. COUPLING WITH AN OBJECT-BASED MOTION SEGMENTATION

We now introduce an extension of the model to couple the motion estimation process with an object-based motion segmentation. By this, we mean that we aim at simultaneously extracting the silhouette of one dynamical entity.[3] Hereafter, the segmentation is defined as a single non-self-intersecting closed curve lying in the image plane. Some discretization scheme allows to associate with it a unique partition of the pixel set in terms of "interior," R, and "exterior," $R' = S - R$. In the following R will be called a *segment*. The precise definition and parameterization of this curve may be of different natures, depending on the problem at hand, on the amount of known or assumed prior knowledge, and on the affordable computational burden.

A parameterization of low dimension can be used to specify a strong geometric prior knowledge on the region shape. This is allowed by the *deformable template* framework proposed by Grenander [21]. Conversely, a "weak" regularizing prior can be captured with plane curves controlled by a large number of parameters. This situation is thoroughly addressed within the framework of *active contour* models and *snakes* [26]. Originally used in still image segmentation, these different closed curve models have recently appeared as promising tools to cope with object-based motion issues [9], [17], [27], [30], [33], [36], [37].

Our purpose here is neither to choose among the different deformable shape models, nor to propose a new one. We would rather like to show how such a curve-based segmentation may be connected to the optical flow estimator we propose. Therefore, in the coming section, we remain general by neither restricting ourselves to a specific family of possible curves, nor to a specific prior knowledge. As for the experiments, we will demonstrate the feasibility and the interest of the approach by using two simple types of object-based segmentation models that illustrate two extreme cases of parameterization.

A. Energy Design

The extension of the energy-based estimation model is obtained by adding two terms to the global energy function $\mathcal{H}$.[4] The first one, E_{prior}, captures the *a priori* knowledge about the segmenting curve. The second one, E_{interact}, specifies the mode of interaction between the segment and the rest of the estimation model (i.e., velocity field, weights, and data). Different ways of interaction may be considered: The segment can directly interact with w by "cutting" the regularization through its border, as a set of binary line processes; The segment can interact indirectly with the velocity field through discontinuity and/or data outlier weights; The segment can interact with the data to capture for instance the fact that the boundary of a moving region is very likely to exhibit large photometric discontinuities, while the interior should ex-

hibit, to some extent, a noticeable spatial/temporal brightness consistence.

In our first attempt to equip our complete estimation model with an object-based interacting segmentation device, we have chosen so far an interaction mode which leads to a simple energy formulation, and to easy computations. The segment will only interact with the estimation process through the discontinuity weights. The corresponding cost function component $E_{\text{interact}}(R, \beta)$ exhibits two terms: The first one is proportional to the mean value of β_{sr}s over $\mathcal{C}_{\partial R} \triangleq \{\langle s, r \rangle \in R \times R'\}$. It then i) favors low values (close to zero) of discontinuity weights along the border of the segment, and ii) drives the curve toward the more significant gathering of low-valued β_{sr}'s. The second one is proportional to the opposite mean value of β_{sr}s over $\mathcal{C}_R \triangleq \{\langle s, r \rangle \subset R\}$. Its role is i) to favor large values (close to unity) inside, and ii) to make the curve surround areas with uniform velocity. The global energy of the extended model is designed as follows:

$$\mathbb{H}(\mathrm{d}w, \delta, \beta, R; f, w) \triangleq \mathcal{H}(\mathrm{d}w, \delta, \beta; f, w) + E_{\text{prior}}(R) + E_{\text{interact}}(R, \beta), \quad (19)$$

where

$$E_{\text{interact}}(R, \beta) \triangleq \frac{\mu_1}{|\mathcal{C}_{\partial R}|} \sum_{\langle s, r \rangle \in \mathcal{C}_{\partial R}} \beta_{sr} - \frac{\mu_2}{|\mathcal{C}_R|} \sum_{\langle s, r \rangle \in \mathcal{C}_R} \beta_{sr} \quad (20)$$

with some positive parameters μ_1 and μ_2. As for the class of admissible segments, and associated prior energy, we have considered two extreme cases in our experiments.

The first case corresponds to a tight geometric constraint associated with a weak prior on the shape: the segmenting curve is a convex quadrilateral with prior energy $E_{\text{prior}}(R) \triangleq \lambda(|\mathcal{C}_{\partial R}|^2/|R|)$ (for some $\lambda > 0$) favoring "compact" shapes. For a given surface (i.e., $|R|$ fixed), the prior on the segment is all the lower that the shape of the segment is closed to a square. As concerns the size of the segment, this prior obviously does not depend on it: there is no *a priori* on the apparent size of the region. When updating the segment such defined, local deformations will be simply obtained by moving each vertex within a small window (e.g., 3×3) around its current location. As for global transformations, rotations, translations and scalings will be considered, which do not alter the prior energy.

This prior modeling of the segment is a very simple instance of deformable templates [21]. More sophisticated modelings could be used in this context, both in terms of admissible curves (class of parametric shapes, number of control parameters) and energy (e.g., prior on the angles of successive edges of the polygonal silhouette [21], [33]).

The second choice corresponds to a loose geometric constraint associated with a classical minimum length description (MDL) prior [29]: the segmenting curve is any non-self-intersecting closed curve whose cost is proportional to its length above a certain threshold, i.e., $E_{\text{prior}}(R) \triangleq \lambda \max\{|\mathcal{C}_{\partial R}|, \varepsilon\}$ for some $\lambda > 0$ and $\varepsilon > 0$. Contrary to the previous prior, this energy favors segment with short and

---

[3] The extension to a *fixed* number of such entities would be straightforward. The estimation of an *unknown* and *varying* number of regions is a tough problem that we do not address in the present context.

[4] Even though the superscript remains omitted, we still suppose in the coming developments that some resolution level k is concerned.

smooth border (the threshold is thus necessary in general not to have the segment shrinking to a point). Iterative local deformations of the segment will be obtained by "moving" border sites from R to R', and the other way around. As for global transformations, scalings will change the prior energy.

This prior can be viewed as a very simple instance of snake [26], with only a first-order smoothing. A more sophisticated snake-type prior could probably be substituted to it, involving a parameterized smooth closed curve with stretching and bending penalties (first- and second-order smoothing).

B. Energy Minimization

The alternate minimization spirit is still considered. Compared to the segmentation-free optical flow estimation of Section III-B, two extra features have to be described: i) the new computation rules for the optimal discontinuity weights, and ii) the segment updating given the current optical flow under estimation.

For given velocity field $\boldsymbol{w}+\mathrm{d}\boldsymbol{w}$ and segment R, the optimal weights $\hat{\beta}_{sr}$, $\langle s, r \rangle \in \mathcal{C}$, are

$$\hat{\beta}_{sr} = \arg\min_{\beta_{sr} \in (0,1]} \left\{ \beta_{sr} \left[\tau_2 \|\Delta\boldsymbol{w}_{sr}\|^2 + \frac{\mu'_1}{|\mathcal{C}_{\partial R}|} \right. \right.$$
$$\left. \left. \cdot \mathbf{1}_{\partial R}(s, r) - \frac{\mu'_2}{|\mathcal{C}_R|} \mathbf{1}_R(s, r) \right] + \psi_2(\beta_{sr}) \right\} \quad (21)$$

where $\Delta\boldsymbol{w}_{sr} \triangleq (\boldsymbol{w}_s + \mathrm{d}\boldsymbol{w}_s) - (\boldsymbol{w}_r + \mathrm{d}\boldsymbol{w}_r)$ as a notational convenience, $\mathbf{1}_R(s, r)$ [resp., $\mathbf{1}_{\partial R}(s, r)$] $= 1$ if $\langle s, r \rangle \in \mathcal{C}_R$ (resp., $\langle s, r \rangle \in \mathcal{C}_{\partial R}$), and zero otherwise, and $\mu'_i \triangleq \mu_i/\alpha$, $i = 1, 2$. One gets the new optimal weight computation rule (see Appendix A for the proof)

if $\langle s, r \rangle \subset R'$,
$$\hat{\beta}_{sr} = \sigma_2^2 \phi'_2[\|\Delta\boldsymbol{w}_{sr}\|^2]; \quad (22)$$
if $\langle s, r \rangle \in R \times R'$,
$$\hat{\beta}_{sr} = \sigma_2^2 \phi'_2 \left[\|\Delta\boldsymbol{w}_{sr}\|^2 + \frac{\mu'_1}{\tau_2 |\mathcal{C}_{\partial R}|} \right]; \quad (23)$$
if $\langle s, r \rangle \subset R$,
$$\hat{\beta}_{sr} = \sigma_2^2 \phi'_2 \left\{ \left[\|\Delta\boldsymbol{w}_{sr}\|^2 - \frac{\mu'_2}{\tau_2 |\mathcal{C}_R|} \right]^+ \right\} \quad (24)$$

where $[\bullet]^+ \triangleq \max\{0, \bullet\}$. Basically, the optimal weight computation rule is only changed along and within the segment, according to a simple shift of the argument of ϕ'_2 under the constraint that it remains positive. Along the border, this (positive) shift results in a decrease of the weights (i.e., smoothing reduction), which is all the more important that the border is short. Inside the region, the (bounded negative) shift results in an increase of the weights (i.e., smoothing accentuation), with a saturation at one. In this context, Leclerc's estimator $\rho(u) = 1 - \exp(-\tau u^2)$ yields a very convenient updating rule: the computation of the optimal weight turns out to be simply related to the *segmentation-free optimal weights* ($\hat{\beta}_{sr}^0 \triangleq \sigma_2^2 \phi'_2[\|\Delta\boldsymbol{w}_{sr}\|^2]$) through a single
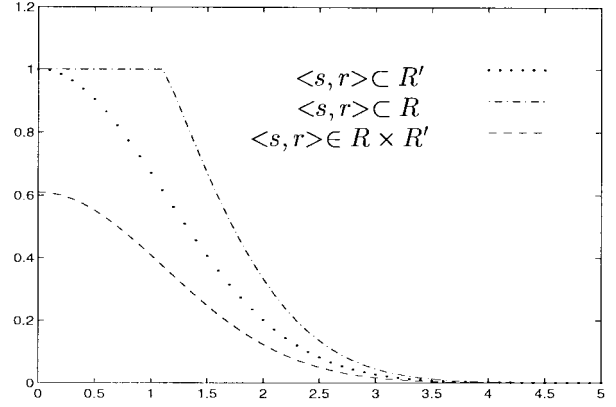

Fig. 6. Modified discontinuity weight functions for Leclerc's estimator ($\sigma_2^2 = 2.5$, $\mu'_2/|\mathcal{C}_R| = 0.5$, $\mu'_1/|\mathcal{C}_{\partial R}| = 0.5$), according to location of pair $\langle s, r \rangle$ w.r.t. segment R.

multiplication (see Fig. 6 for a plot example):

if $\langle s, r \rangle \subset R'$,
$$\hat{\beta}_{sr} = \hat{\beta}_{sr}^0; \quad (25)$$
if $\langle s, r \rangle \in R \times R'$,
$$\hat{\beta}_{sr} = \underbrace{\exp\left(-\frac{\mu'_1}{|\mathcal{C}_{\partial R}|}\right)}_{\triangleq \, K_{\partial R} < 1} \hat{\beta}_{sr}^0; \quad (26)$$
if $\langle s, r \rangle \subset R$,
$$\hat{\beta}_{sr} = \min\left\{ 1, \underbrace{\exp\left(\frac{\mu'_2}{|\mathcal{C}_R|}\right)}_{\triangleq \, K_R > 1} \hat{\beta}_{sr}^0 \right\}. \quad (27)$$

A multigrid version with piecewise constant velocity increments is readily derived.

The minimization of the energy according to the segment is addressed in two successive ways. First, on the coarsest grid ($\ell = L$) of the coarsest resolution ($k = N$) where the complete procedure starts and where the dimensionality of the problem is drastically reduced, one can sweep efficiently over the set of possible segmentations. A crude estimate of the segment is thus obtained at low cost by a stochastic algorithm, with *no need of any manual initialization*. The location and shape of the segment is then refined deterministically through the following grids of resolution N, and through the finest grids ($\ell = 0$) of the following resolution levels $N - 1, \cdots, 0$. For both kinds of segment updating, the velocity field being fixed, an iterative scheme is used which considers at each step different possible segments R *along with the associated optimal weights* $\hat{\beta}$ provided by (22)–(24). The part of energy $E(R)$ actually concerned by the updating process reduces to the one involving the segment and/or the discontinuity weights

$$E(R) \triangleq E_{\text{prior}}(R) + \sum_{\langle s, r \rangle \in \mathcal{C}} \left\{ \hat{\beta}_{sr} \left[\tau_2 \|\Delta\boldsymbol{w}_{sr}\|^2 + \frac{\mu'_1}{|\mathcal{C}_{\partial R}|} \right. \right.$$
$$\left. \left. \cdot \mathbf{1}_{\partial R}(s, r) - \frac{\mu'_2}{|\mathcal{C}_R|} \mathbf{1}_R(s, r) \right] + \psi_2(\hat{\beta}_{sr}) \right\}.$$

TABLE I
COMPARATIVE RESULTS ON YOSEMITE

Technique	Average error	Standard deviation	Density	
Horn and Schunck (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%	
Robust multigrid	**5.38^o**	**7.73^o**	**100%**	
Szeliski and Coughlan [39]	2.45^o	3.05^o	100%	
Szeliski and Shum [40]	2.20^o	5.87^o	23.1%	
Black and Anandan [6]	4.46^o	4.21^o	100%	without sky
Black [4]	3.52^o	3.25^o	100%	
Black and Jepson [7]	2.29^o	2.25^o	100%	
Robust multigrid	**2.34^o**	**1.45^o**	**100%**	

In the first stage ($\ell = L$, $k = N$), in absence of any "initial guess," the set of possible segmentations has to be explored thoroughly. This is done by using a simulated annealing algorithm based on Metropolis dynamics. Given the current segment R at step n, a new segment Q is proposed by randomly applying a global deformation (rotation, translation, and scaling) to R [27], [33]. The new segment is accepted with probability $\exp\{-[E(Q) - E(R)]^+/T_n\}$, according to a geometric cooling $T_n = T_0 \times (0.99^n)$.

For the second type of updating, an initial guess is always available by properly interpolating the segment obtained at the previous grid level (if $k = N$) or resolution (if $k < N$). Given this initial guess, we simply seek a local minimizer of the energy nearby this guess, proceeding with small deformations. In the case of quadrilateral segment for instance, the four vertices are displaced in a 3×3 window around their current location, seeking for the largest energy decrease.

Before turning to experimental results, let us come back to the energy change computation. In case of Leclerc's estimator, using a first-order expansion for inner contour sites such that $\|\Delta w_{sr}\|^2 - \mu'_2/\tau_2|\mathcal{C}_R| < 0$, one gets an approximate energy variations $E(Q) - E(R) \approx \tilde{E}(Q) - \tilde{E}(R)$ with

$$\tilde{E}(R) \triangleq E_{\text{prior}}(R) + (1 - K_{\partial R}) \sum_{\langle s, r \rangle \in \mathcal{C}_{\partial R}} \hat{\beta}^0_{sr}$$
$$+ (1 - K_R) \sum_{\langle s, r \rangle \in \mathcal{C}_R} \hat{\beta}^0_{sr}$$

which is very easy to compute (see Appendix B for details).

V. EXPERIMENTAL RESULTS

In this section, we present results of optical flow estimation alone (Section V-A), and results of joint estimation/segmentation (Section V-B).

A. Optical Flow Estimation

The optical flow estimation model presented in the first part of this paper has been validated both on synthetic and real-world sequences. The first one, Yosemite [see Fig. 7(a) and (b)], is the most complex (though synthetic) sequence from the comparative study by Barron *et al.* [1] for which a "ground-truth" exists. The two other sequences are real. They exhibit far more motion discontinuities and OFC violations. Therefore, they are probably more adapted than Yosemite to demonstrate the nice characteristics and abilities of our method, except that there is no ground-truth for them. Fig. 9(a) shows a parking lot sequence that involves two moving cars. The camera pans the scene and the wind shakes the trees in the background. Calendar [Fig. 10(a)] is a TV sequence involving large displacements. It includes several different moving objects and a horizontal panning camera motion. The calendar translates vertically and the toy train pushes a rolling ball.

The choice of the two robust estimators ρ_1 and ρ_2 has been based on heuristic considerations arising from our experience. Since frequent and large deviations from the brightness constancy assumption are very likely to occur, a strongly saturating estimator seems to be well suited to the corresponding component of the energy function. We selected Leclerc's estimator [29] (see Fig. 3). As for the regularization, a softer saturation seems to provide a better behavior of the alternate minimization procedure. For that reason we chose Geman and McClure's estimator [20] (Fig. 3) to be embedded within the smoothness constraint.

The values of the different parameters for the three sequences are the following: $N = 1$, $L = 3$, $\alpha = 320$, $\sigma_1^2 = 6$, and $\sigma_2^2 = 0.7$ for Yosemite; $N = 0$ (the displacements are small), $L = 4$, $\alpha = 200$, $\sigma_1^2 = 6$, and $\sigma_2^2 = 0.2$ for parking lot; $N = 2$, $L = 4$, $\alpha = 200$, $\sigma_1^2 = 7$, and $\sigma_2^2 = 0.4$ for calendar.

Following [1], quantitative comparative results on Yosemite are provided for different algorithms. For each estimate, the deviation with respect to the "real" flow is measured at each pixel location by converting the two-dimensional (2-D) vectors into three-dimensional (3-D) unit vectors, and by computing the angle $\arccos[(u_s u_s^{\text{real}} + v_s v_s^{\text{real}} + 1)/(\sqrt{\|w_s\|^2 + 1}\sqrt{\|w_s^{\text{real}}\|^2 + 1})]$ between them.

TABLE II
INFLUENCE OF THE DIFFERENT INGREDIENTS (MULTIGRID MINIMIZATION, ROBUST PENALTY ON SMOOTHING TERM, ROBUST PENALTY ON DATA TERM) ON RESULTS FOR YOSEMITE SEQUENCE, WITH AND WITHOUT SKY

Technique	with sky	without sky	"equiv. sweeps"
quadratic, monogrid ($\beta \equiv 1, \delta \equiv 1, L = 0$)	10.33 ± 9.29^o	8.31 ± 7.08^o	261
quadratic, multigrid ($\beta \equiv 1, \delta \equiv 1, L = 3$)	6.37 ± 8.28^o	3.47 ± 2.51^o	183
quadratic on data, multigrid ($\delta \equiv 1, L = 3$)	5.88 ± 8.44^o	2.98 ± 2.29^o	190
robust multigrid ($L = 3$)	5.38 ± 7.73^o	2.34 ± 1.45^o	206
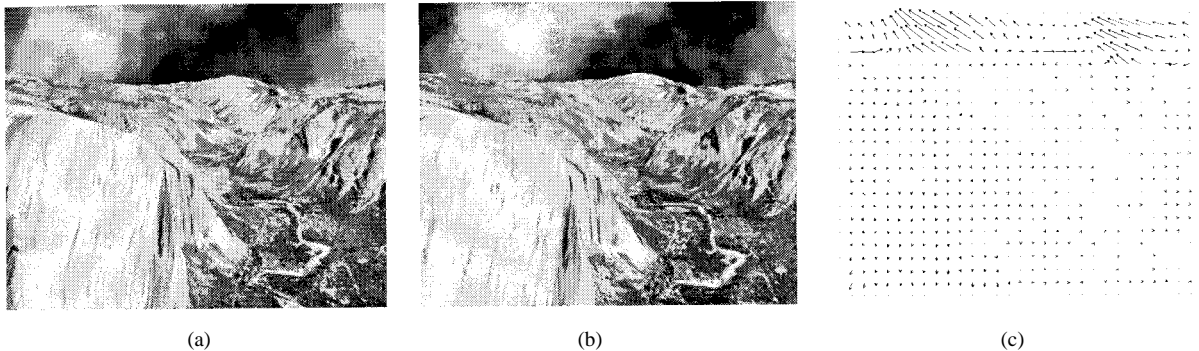
Fig. 7.   (a) and (b) Two frames of Yosemite and (c) difference ($\times$ 10) with the "true" flow.

For some methods reported in [1], estimates are only available at so-called "reliable" locations. The percentage of such locations is the "density" of the estimate. The errors are actually computed only at these locations. Table I lists the average and standard deviation of these angular discrepancies for different algorithms. The five top-lines recall results presented by Barron *et al.* (see references therein). They concern two different versions of Horn and Schunck's algorithms (whose model is the basis of our approach), the best full-density algorithm (Uras *et al.*) and the two algorithms yielding the best results, but with reduced densities (Lucas and Kanade, Fleet and Jepson). Other authors have provided similar comparisons, but on a subsequence where the sky was removed. As a matter of fact, this region is extremely tricky due to the complex luminance and shape evolution of the moving clouds before the sun. The second part of the table compares our method to those by Szeliski and Coughlan [39], Szeliski and Shum [40], Black and Anandan [6], Black [4], Black and Jepson [7], on this reduced sequence.

On the complete scene, our method provides a *dense* estimate almost as good as those obtained with the best (nondense) mentioned methods. In addition, the obtained standard deviation is the smallest one. On the subscene, the average error is slightly lower than the one obtained by Szeliski and Coughlan, and slightly higher than the one obtained by Black and Jepson, but with a standard deviation significantly reduced. The difference between our estimate and the "real" flow is displayed in Fig. 7(c) (subsampled and *magnified ten times*). It clearly appears that most of the discrepancies are gathered around the two moving clearings in the clouds, which the sunlight breaks through.

In order to quantitatively evaluate the influence of the different ingredients (multigrid minimization, robust penalty on smoothing term, robust penalty on data term) both on the quality of the final estimate and on the computational load, we provide corresponding results on Yosemite in Table II. The computational load is measured as a number of "monogrid/monoresolution equivalent sweeps." This is the ratio of the total number of single-site updates (whatever ℓ and k) over the number of sites in lattice S (i.e., one complete sweep at level k, ℓ corresponds to $1/4^{k+\ell}$ of such equivalent sweeps).

It appears that the multigrid minimization allows to significantly improve the global quality of results when put on top of the linear Horn and Schunck's model, while reducing the computational load.[5] The introduction of robust penalization first in the smoothing term, and then in the data term, provides further local improvements. The cost for these late improvements is kept reasonable by the multigrid speed-up. As a rough complexity comparison, our algorithm took around 500 s on a SunSparc 10 (174 s on a SunSparc 20), which is the same time as reported by Black and Anandan for their continuation minimization method. In Black and Jepson's approach, a similar load seems to be necessary to provide the *initial* dense optical flow estimate which is then refined through parametric segmentation. Note that a nice feature of the multigrid approach is that good results can be already obtained on intermediate grids (thus at reduced cost). It is the case on Yosemite where a slightly *lower* angular error 5.27 ± 7.88^0 was actually obtained at convergence for $k = 0$, $\ell = 1$, after a total CPU time of 88 s (compared to the 174 s of the complete procedure).

To demonstrate the low sensitiveness of our method to parameters, we run it on Yosemite for 1000 triples $(\alpha, \sigma_1^2, \sigma_2^2) \in [140, 320] \times [3, 12] \times [0.1, 1]$. In Fig. 8, we plot the corre-

---

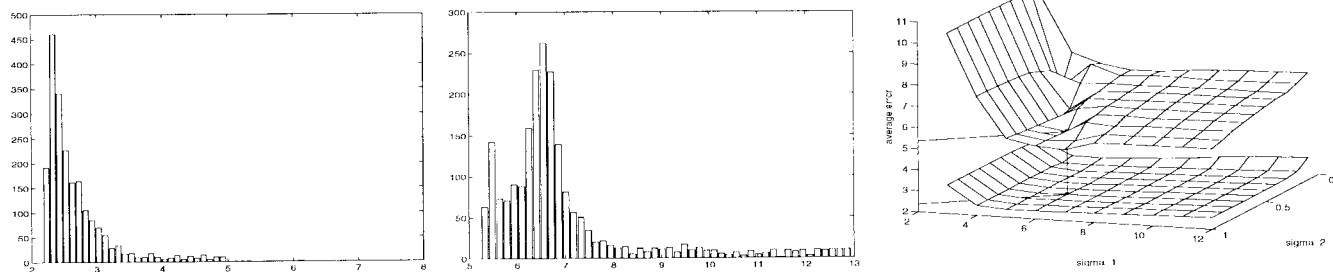[5]Similar conclusions were reported in [23] for other (nonconvex) energy-based motion analysis models.

Fig. 8. Average angular error histograms for 1000 parameter triples $(\alpha, \sigma_1^2, \sigma_2^2) \in [140, 320] \times [3, 12] \times [0.1, 1]$: without the sky (left) and with the sky (right); Average angular error for 100 parameter triples $(\sigma_1, \sigma_2) \in [3, 12] \times [0.1, 1]$, with $\alpha = 320$: without sky (lower surface) and with the sky (upper surface). In both cases, the minimal error is obtained for $\sigma_1^2 = 6$ and $\sigma_2^2 = 0.7$.
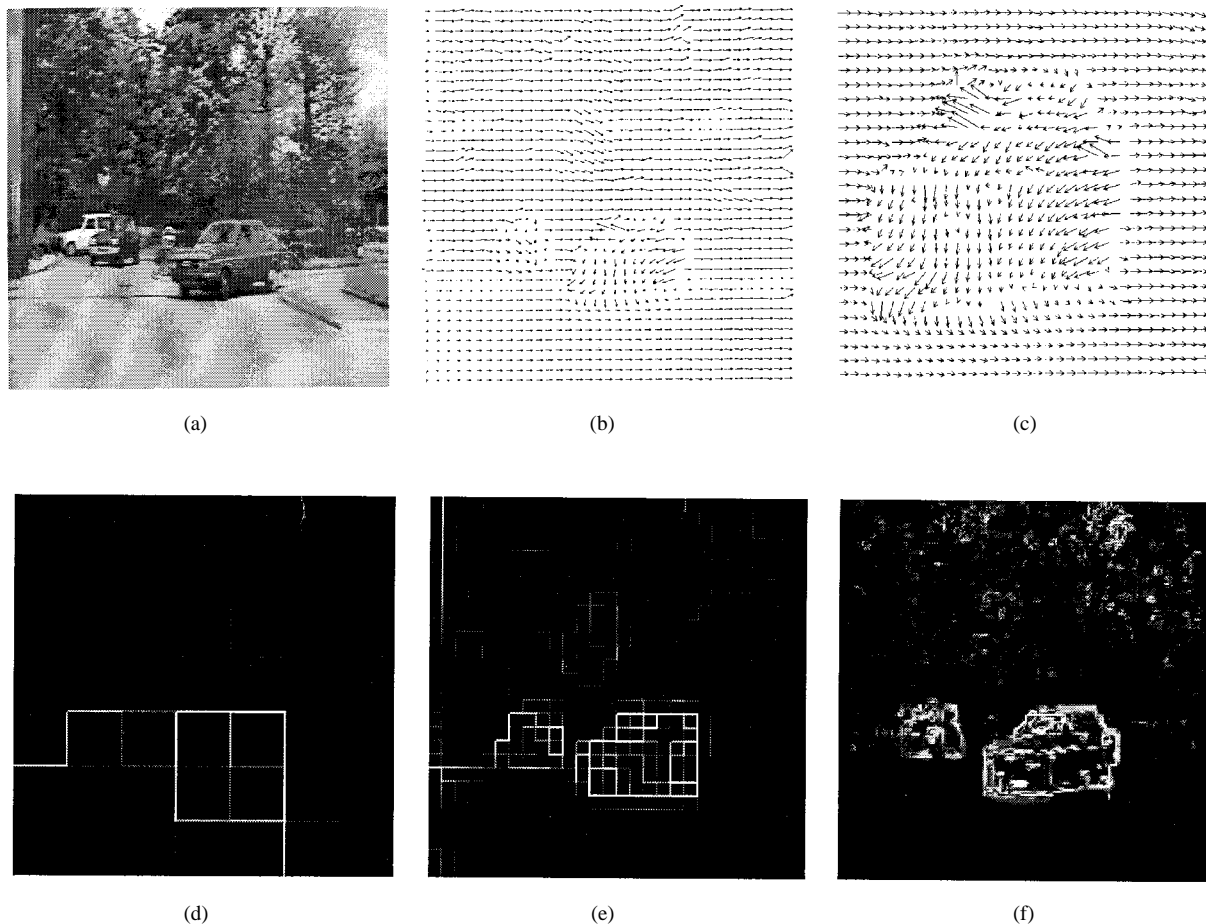

Fig. 9. Results on parking lot: (a) one frame, (b) flow estimate, (c) zoom on the foreground car, and (d)–(f) discontinuity weights at resolution level $k = 0$, on grid levels $\ell = 4, 2, 0$.

sponding histograms of average angular errors for the complete scene and of the subscene. In both cases, a high robustness is exhibited. This is also noticeable in the plot of the average error versus (σ_1^2, σ_2^2) for fixed $\alpha = 320$.

Figs. 9 and 10 present final flow estimates and discontinuity weights at different resolutions and/or grid levels for parking lot and calendar. The discontinuity weights are displayed on 256 grey levels on sites of the dual lattice: white points represent zero weights (i.e., maximal discontinuity) and black points are for unity weights (i.e., no discontinuity).

Visually, the recovered optical flow fields seem of good quality. The location of spatial discontinuities fits with a good accuracy the boundaries of the different motion regions (e.g., the first car in parking lot, the train in calendar). This allows the estimation in quite uniform regions not to be contaminated by nearby flows of large amplitude belonging to other motion regions. For instance, the estimation of the left-to-right apparent motion induced by the camera panning within the street region beneath the two cars would be dramatically influenced in case of plain quadratic estimator, resulting in an almost zero
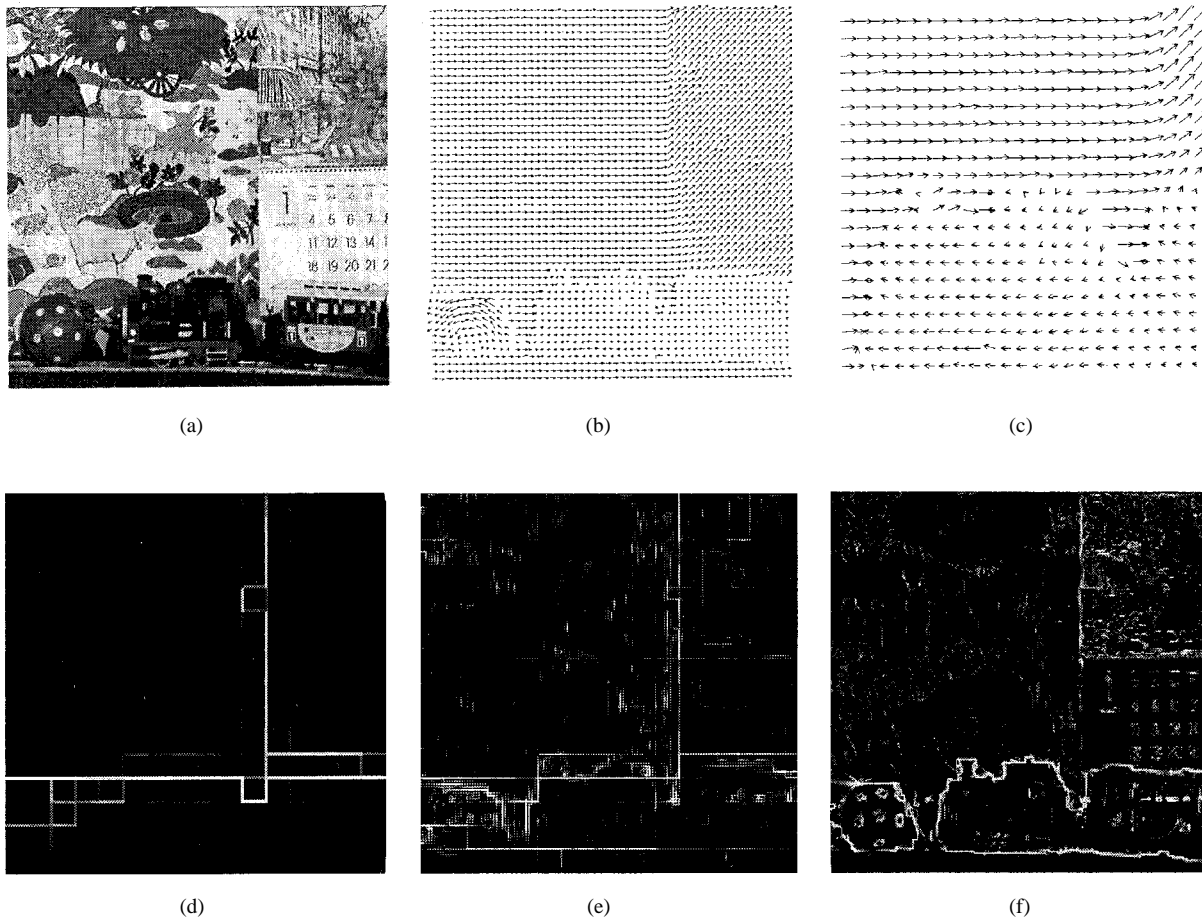
Fig. 10.   Results on calendar: (a) one frame, (b) flow estimate, (c) zoom on the train, and (d)–(f) discontinuity weights in $\Omega^{2,3}$, $\Omega^{1,4}$, and Ω.
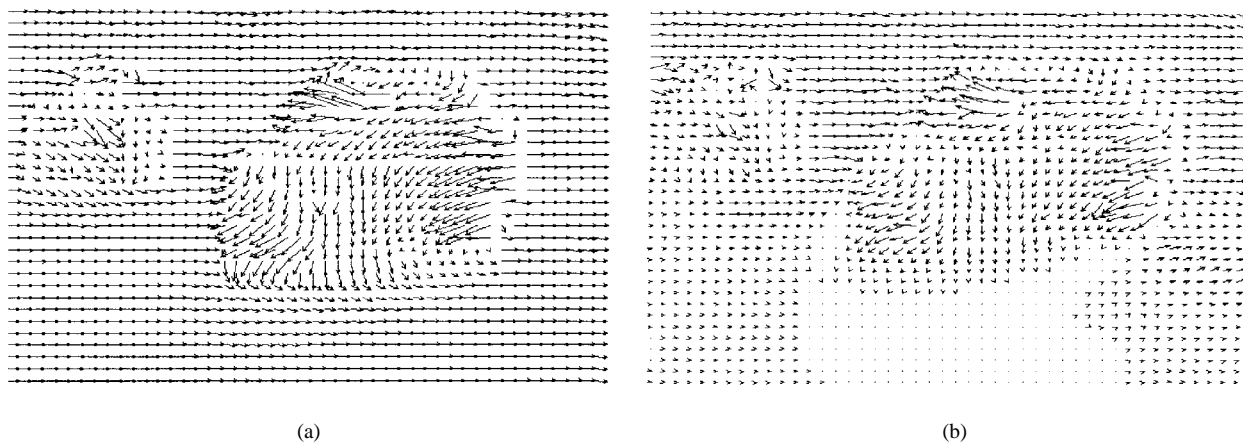

Fig. 11.   Comparison of estimates on parking lot. (a) Estimation with the proposed model. (b) Estimation with a plain quadratic model.

estimate (see Fig. 11 for a comparison). One can also notice that the motion estimation remains consistent in tough parts of the dynamic scenes at hand. For example, the motion of the calendar is quite well estimated despite highly textured portions exhibiting periodical patterns (like the drawing of houses). Another interesting example is found on the left part of the front car windshield in parking lot. The motion of a patch of specular reflection does not disturb the car velocity

estimation in the neighborhood. From the final velocity field estimate, it appears as an independent moving patch that goes in a direction opposite to the one of the car motion.

We believe that the good quality of these results is partly due to the multigrid structure of the estimator. It produces good intermediate coarse flows. Besides, it allows to extract long structures of discontinuities. This is particularly noticeable in calendar [see Fig. 10(d) and (e)].
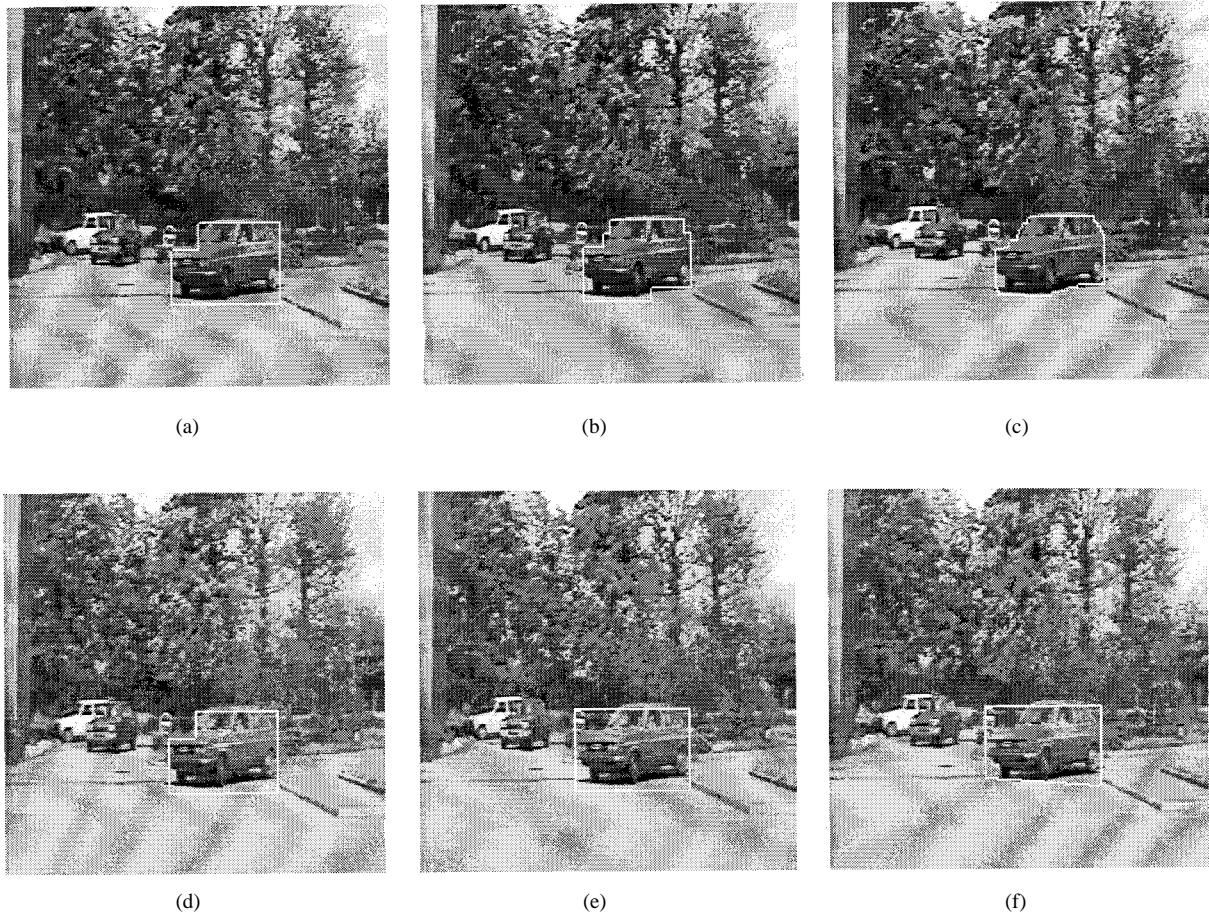
Fig. 12. Single object motion segmentation on parking lot: (a)–(c) with an *unconstrained closed curve* at grid levels 4, 2, and 0, and (d)–(f) with a *convex quadrilateral* at grid levels 4, 2, and 0.

B. Coupling with the Segmentation

We now report experiments of the augmented model (with a single segmenting closed curve) on two real traffic sequences. The first one is parking lot, which has already been introduced. The second one, called "street," is shot from a car driving in a stream of moving vehicles (see Fig. 13). The motion of the camera results in a divergent motion component in the image.

As for the choice of the two robust functions, we explained in Section V-A why we chose Geman and McClure's function in the smoothing term. However, as pointed out in Section IV-B, Leclerc's estimator, due to its exponential nature, yields simpler computations as far as the smoothing energy term is concerned in the estimation/segmentation model. Therefore, in place of Geman and McClure's estimator, we made this convenient choice for ρ_2 in experiments about the augmented model.

Two "extreme" cases of *a priori* on the curve have been considered: in the first case, the curve is constrained to be a convex quadrilateral along with a prior favoring compact shapes; in the second case, the curve is only constrained to be closed and nonself-intersecting with an MDL-type prior. Except for the number of resolutions ($N = 0$ for parking lot and $N = 2$ for street), the values of the parameters were set the same for both sequences, namely: $L = 4$, $\alpha = 200$, $\sigma_1^2 = 8$,

$\sigma_2^2 = 0.1$ for the quadrilateral and 0.3 for the unconstrained curve, $\mu_1' = 0.5$, and $\mu_2' = 0.05$ for the quadrilateral and 0.5 for the unconstrained curve.

Similar results of good quality are obtained with both kinds of curve, on both sequences. In each case (see Figs. 12 and 13) the region exhibiting the largest motion discontinuities along its border (the car in the foreground in parking lot, the car entering the image plane in the foreground, and moving away from the camera in street) are correctly picked (even with the MDL-type prior which favors short borders).

A low dimension parameterization (as with the quadrilaterals) yields computations of lower cost, while allowing to capture a strong geometric knowledge (if available) on the shapes of the moving entities. At the same time, this can become a drawback if the prior knowledge is not sufficient or the restricted family of curves is not well suited to the scene content [compare (c) and (f) in Figs. 12 and 13]. In our case, quadrilaterals are not able to completely fit the complex shape of moving vehicles. As a consequence, the relative importance of the border-based energy risks to shrink: the region-based term has to be reduced by tuning μ_2' to a very low value in order to keep a proper balance. Otherwise, the template would be mainly driven by regionwise flow uniformity, risking to get stuck around spurious locations such as parts of still background in case of nonmoving camera. These problems can

Fig. 13. Single object motion segmentation on street: (a)–(c) with an *unconstrained closed curve*, in $\Omega^{2,2}$, Ω^2, and Ω, and (d)–(f) with a *convex quadrilateral* in $\Omega^{2,2}$, Ω^2, and Ω.

hopefully be circumvented by using less constrained curves, but at increased cost.

VI. CONCLUSION

In this paper, we have presented a multiresolution/multigrid framework for optical flow estimation and object-based motion segmentation.

The estimation 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 data model deviations. The minimization is processed through a multigrid algorithm which consists in imposing successively weaker and weaker constraints on the searched estimates. Applied to a dual formulation of the original energy function, this method leads to a multigrid iteratively reweighted least squares minimization which is efficient in term of convergence rate and in term of quality of the produced estimates. It is worth noting that this is a *general purpose* multigrid approach which can be easily applied to most of image analysis objective functions. Besides, it allows to define efficient and original parallel relaxation algorithms [31], [34].

In the motion estimation context, one of the nicest features of the approach is that it gives access to a consistent and meaningful sequence of finer and finer configurations at each single level of data resolution. The coarser grain estimates, which are far easier to compute due to their reduced dimensionality, reveal large discontinuity structures of the apparent motion field. This kind of significant reduced estimates cannot be produced by continuation-type minimization methods since the earliest estimates they provide are, to some extent, "smoothed versions" of their final estimates [3], [10], [29].

We get benefit from this compact and structured information by introducing a closed curve-based device. It allows to recover with improved accuracy the location of spatial discontinuities and to naturally handle *edge grouping* to get an object-based motion segmentation. The proposed model can support any kind of parameterized or nonparameterized family of curves, equipped with any prior energy function.

We thus contribute here to the efforts done to define global approaches for a joint and cooperative handling of two important interleaved issues of motion analysis. As far as our object-based motion segmentation is concerned, it is worth noting that neither knowing the motion of the camera (if any),

nor initializing manually the curve are required. This work has now to be extended to deal with complete motion-based segmentation (i.e., *partition* of the whole image plane into a variable number of regions with different motions). As a fact, there is a straightforward way to extend the interaction energy term E_{interact} to this case, with the MDL-type prior on the segmentation becoming a standard Potts model [12], [29]. However, algorithmic issues related to the choice of the initial segmentation, and of the update strategy (in terms of deformation, merge, split, etc.) need attention in this case. Another direction for such an extension would consist in joining a region-based *parametric* flow interacting with the dense flow under estimation, as proposed in [36].

The interaction mechanism we designed for closed curve, also suggests to use it with *open* curves. In this case, one would not segment the images, but this would be useful for preserving and refining through grid and resolution levels, the precious information captured by the (independent) discontinuity weights on smaller grids.

APPENDIX A
MINIMIZATION OF $\mathbb{H}$ WITH RESPECT TO β_{sr}

Let $\langle s, r \rangle \in \mathcal{C}$ be a pair of neighboring sites. The part of $\mathbb{H}$ which actually depends on β_{sr} is

$$\underbrace{\tau_2 \beta_{sr} \left[\|\Delta \boldsymbol{w}_{sr}\|^2 + \frac{\mu_1'}{\tau_2 |\mathcal{C}_{\partial R}|} \mathbf{1}_{\partial R}(s, r) - \frac{\mu_2'}{\tau_2 |\mathcal{C}_R|} \mathbf{1}_R(s, r) \right]}_{\triangleq A_{sr}(\boldsymbol{w}, R)}$$
$$+ \psi_2(\beta_{sr}). \tag{28}$$

If $A_{sr}(\boldsymbol{w}, R) < 0$, which means that $\langle s, r \rangle \subset R$ and $\|\Delta \boldsymbol{w}_{sr}\|^2 < \mu_2'/\tau_2|\mathcal{C}_R|$, then $\mathbb{H}$ is a decreasing function of β_{sr}, since ψ_2 is decreasing [indeed, it is easy to derive $\psi_2'(z) = -\tau_2 (\phi_2')^{-1}(\tau_2 z) < 0$]. Its minimizer in $(0, 1]$ is then $\hat{\beta}_{sr} = 1$ [and the corresponding "energy contribution" reduces to $\tau_2 A_{sr}(\boldsymbol{w}, R)$ since $\psi_2(1) = 0$] . In other cases $[A_{sr}(\boldsymbol{w}, R) \geq 0]$, the minimizer zeros the partial derivative

$$\frac{\partial \mathbb{H}}{\partial \beta_{sr}} (\hat{\beta}_{sr}) = 0 \Leftrightarrow \tau_2 A_{sr}(\boldsymbol{w}, R) + \psi_2'(\hat{\beta}_{sr})$$
$$= 0 \Leftrightarrow \hat{\beta}_{sr} = \sigma_2^2 \phi_2'[A_{sr}(\boldsymbol{w}, R)]. \tag{29}$$

Merging both cases provides the following optimal weight computation rule:

- if the site pair is outside R ($\langle s, r \rangle \subset R'$): $\hat{\beta}_{sr} = \sigma_2^2 \phi_2'[\|\Delta \boldsymbol{w}_{sr}\|^2]$;

- if the site pair straddles the "border" of R ($\langle s, r \rangle \in R \times R'$): $\hat{\beta}_{sr} = \sigma_2^2 \phi_2'[\|\Delta \boldsymbol{w}_{sr}\|^2 + \mu_1'/\tau_2|\mathcal{C}_{\partial R}|]$;
- if the site pair is inside R ($\langle s, r \rangle \subset R$)

$$\hat{\beta}_{sr} = \sigma_2^2 \phi_2' \left(\left[\|\Delta \boldsymbol{w}_{sr}\|^2 - \frac{\mu_2'}{\tau_2 |\mathcal{C}_R|} \right]^+ \right)$$
$$= \begin{cases} \sigma_2^2 \phi_2' \left[\|\Delta \boldsymbol{w}_{sr}\|^2 - \dfrac{\mu_2'}{\tau_2 |\mathcal{C}_R|} \right] & \text{if } \|\Delta \boldsymbol{w}_{sr}\|^2 \geq \dfrac{\mu_2'}{\tau_2 |\mathcal{C}_R|}, \\ 1 & \text{otherwise.} \end{cases}$$

Also, notice that in case $A_{sr}(\boldsymbol{w}, R) \geq 0$, it comes from (29) that

$$\tau_2 \hat{\beta}_{sr} A_{sr}(\boldsymbol{w}, R) + \psi_2(\hat{\beta}_{sr}) = -\hat{\beta}_{sr} \psi_2'(\hat{\beta}_{sr}) + \psi_2(\hat{\beta}_{sr})$$
$$= \tau_2 \hat{\beta}_{sr} (\phi_2')^{-1}(\tau_2 \hat{\beta}_{sr}) + \psi_2(\hat{\beta}_{sr})$$
$$= \phi_2 \circ (\phi_2')^{-1}(\tau_2 \hat{\beta}_{sr})$$

since $\psi_2(z) = \phi_2 \circ (\phi_2')^{-1}(\tau_2 z) - \tau_2 z (\phi_2')^{-1}(\tau_2 z)$ by definition. If $\phi_2(v) = 1 - \exp\{-\tau_2 v\}$ (Leclerc's estimator), the above energetic contribution becomes

$$\tau_2 \hat{\beta}_{sr} A_{sr}(\boldsymbol{w}, R) + \psi_2(\hat{\beta}_{sr}) = 1 - \hat{\beta}_{sr}. \tag{30}$$

APPENDIX B
COMPUTATION OF $\Delta E(R)$ AND ITS APPROXIMATION FOR LECLERC'S ESTIMATOR

Suppose $\boldsymbol{w}$ is fixed. Let R be some segment, and $\hat{\beta}$ the optimal set of discontinuity weights associated to it according to the computations previously derived. The part of energy concerned by the segment updating process is

$$E(R) = E_{\text{prior}}(R) + \sum_{\langle s, r \rangle \in \mathcal{C}} \tau_2 \hat{\beta}_{sr} A_{sr}(\boldsymbol{w}, R) + \psi_2(\hat{\beta}_{sr})$$
$$= E_{\text{prior}}(R) + \sum_{\langle s, r \rangle : A_{sr}(\boldsymbol{w}, R) \geq 0} (1 - \hat{\beta}_{sr})$$
$$+ \sum_{\langle s, r \rangle : A_{sr}(\boldsymbol{w}, R) < 0} \tau_2 A_{sr}(\boldsymbol{w}, R),$$

according to Appendix A. $A_{sr}(\boldsymbol{w}, R) < 0$ possibly holds only for $\langle s, r \rangle \subset R$, where $A_{sr}(\boldsymbol{w}, R) = \|\Delta \boldsymbol{w}_{sr}\|^2 - (\mu_2'/\tau_2|\mathcal{C}_R|)\mathbf{1}_R(s, r)$.

$$E(R) \approx E_{\text{prior}}(R) - \sum_{\langle s, r \rangle \subset R'} \hat{\beta}_{sr}^0 - K_{\partial R} \sum_{\langle s, r \rangle \in R' \times R} \hat{\beta}_{sr}^0 - K_R \sum_{\langle s, r \rangle \subset R} \hat{\beta}_{sr}^0 + |\mathcal{C}|$$
$$= E_{\text{prior}}(R) + \underbrace{(1 - K_{\partial R}) \sum_{\langle s, r \rangle \in R' \times R} \hat{\beta}_{sr}^0 + (1 - K_R) \sum_{\langle s, r \rangle \subset R} \hat{\beta}_{sr}^0 + |\mathcal{C}| - \sum_{\langle s, r \rangle} \hat{\beta}_{sr}^0}_{\triangleq \tilde{E}(R)} \tag{31}$$

Using (25)–(27), we have

$$
\begin{aligned}
E(R) = E_{\text{prior}}(R) &+ \sum_{\langle s,r \rangle \subset R'} (1 - \hat{\beta}_{sr}^0) \\
&+ \sum_{\langle s,r \rangle \in R' \times R} (1 - K_{\partial R}\hat{\beta}_{sr}^0) \\
&+ \sum_{\langle s,r \rangle \subset R: \, \|\Delta\boldsymbol{w}_{sr}\|^2 \geq \mu_2'/\tau_2|\mathcal{C}_R|} (1 - K_R\hat{\beta}_{sr}^0) \\
&+ \sum_{\langle s,r \rangle \subset R: \, \|\Delta\boldsymbol{w}_{sr}\|^2 < \mu_2'/\tau_2|\mathcal{C}_R|} \\
&\cdot \left(\tau_2 \|\Delta\boldsymbol{w}_{sr}\|^2 - \frac{\mu_2'}{|\mathcal{C}_R|} \right).
\end{aligned}
$$

In the latter term of the right-hand side, if $\tau_2\|\Delta\boldsymbol{w}_{sr}\|^2 - \mu_2'/|\mathcal{C}_R| \approx 0$, then a first order Taylor expansion yields[6]

$$
\begin{aligned}
\tau_2 \|\Delta\boldsymbol{w}_{sr}\|^2 - \frac{\mu_2'}{|\mathcal{C}_R|} &\approx 1 - \exp\left\{ -\tau_2\|\Delta\boldsymbol{w}_{sr}\|^2 + \frac{\mu_2'}{|\mathcal{C}_R|} \right\} \\
&= 1 - K_R\hat{\beta}_{sr}^0.
\end{aligned}
$$

Then energy approximation (31), shown at the bottom of the previous page, is given.

The velocity field $\boldsymbol{w}$ being fixed (and therefore such is $\hat{\beta}^0$), the energy comparison of two segments R and Q is then driven by $E(Q) - E(R) \approx \tilde{E}(Q) - \tilde{E}(R)$, since $\sum_{\langle s,r \rangle} \hat{\beta}_{sr}^0$ does not depend on the segment by definition of the $\hat{\beta}_{sr}^0$s. The approximate expression is very easy to compute. Especially, if R and Q have the same shape, the only terms changing from $\tilde{E}(R)$ to $\tilde{E}(Q)$ are $\sum_{\langle s,r \rangle \subset R} \hat{\beta}_{sr}^0$ and $\sum_{\langle s,r \rangle \in R' \times R} \hat{\beta}_{sr}^0$.

ACKNOWLEDGMENT

REFERENCES

[1] J. Barron, D. Fleet, and S. Beauchemin, "Performance of optical flow techniques," *Int. J. Comput. Vis.,* vol. 12, pp. 43–77, 1994.
[2] J. Bergen, P. Anandan, K. Hanna, and R. Hingorani, "Hierarchical model-based motion estimation," in *Proc. Europ. Conf. Computer Vision,* G. Sandini, Ed. Berlin, Germany: Springer-Verlag, 1992, pp. 237–252.
[3] M. Black, "Robust incremental optical flow," Ph.D. dissertation, Yale Univ., New Haven, CT, 1992.
[4] _____, "Recursive nonlinear estimation of discontinuous flow fields," in *Proc. Europ. Conf. Computer Vision,* Stockholm, Sweden, 1994, pp. 138–145.
[5] M. Black and P. Anandan, "Robust incremental optical flow," in *Proc. Conf. Computer Vision and Pattern Recognition,* J. O. Eklundh, Ed. Berlin, Germany: Springer-Verlag, 1992, pp. 296–302.
[6] _____, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," *Comput. Vis. Image Understand.,* vol. 63, pp. 75–104, 1996.
[7] M. Black and P. Jepson, "Estimating optical flow in segmented images using variable-order parametric models with local deformations," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 18, pp. 972–986, 1996.
[8] M. Black and A. Rangarajan, "On the unification of line processes, outlier rejection, and robust statistics with applications in early vision," *Int. J. Comput. Vis.,* vol. 19, pp. 75–104, 1996.
[9] A. Blake, R. Curwen, and A. Zisserman, "A framework for spatiotemporal control in the tracking of visual contours," *Int. J. Comput. Vis.,* vol. 11, pp. 127–145, 1993.
[10] A. Blake and A. Zisserman, *Visual Reconstruction.* Cambridge, MA: MIT Press, 1987.
[11] L. Blanc-Féraud, M. Barlaud, and T. Gaidon, "Motion estimation involving discontinuities in a multiresolution scheme," *Opt. Eng.,* vol. 32, pp. 1475–1482, July 1993.
[12] P. Bouthemy and E. Francois, "Motion segmentation and qualitative dynamic scene analysis from an image sequence," *Int. J. Comput. Vis.,* vol. 10, pp. 157–182, 1993.
[13] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud, "Deterministic edge-preserving regularization in computed imaging," *IEEE Trans. Image Processing,* vol. 6, pp. 298–311, 1997.
[14] I. Cohen and I. Herlin, "Optical flow and phase portrait methods for environmental satellite image sequences," in *Proc. Europ. Conf. Computer Vision,* Cambridge, U.K., Apr. 1996, pp. 141–150.
[15] R. Deriche, P. Kornprobst, and G. Aubert, "Optical flow estimation while preserving its discontinuities: A variational approach," in *Proc. Asian Conf. Computer Vision,* Singapore, Dec. 1995, vol. 1, pp. 290–295.
[16] W. Enkelmann, "Investigation of multigrid algorithms for the estimation of optical flow fields in image sequences," *Comput. Vis. Graph. Image Process.,* vol. 43, pp. 150–177, 1988.
[17] K. Fujimura, N. Yokoya, and K. Yamamoto, "Motion tracking of deformable objects by active contour models using multiscale dynamic programming," *J. Vis. Commun. Image Represent.,* vol. 4, pp. 382–391, 1993.
[18] D. Geman and G. Reynolds, "Constrained restoration and the recovery of discontinuities," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 14, no. 3, pp. 367–383, 1992.
[19] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 6, pp. 721–741, 1984.
[20] S. Geman and D. McClure, "Statistical methods for tomographic image reconstruction," in *Bull. ISI, Proc. 46th Session Int. Statistical Institute,* vol. 52, 1987.
[21] U. Grenander, *General Pattern Theory.* Oxford, U.K.: Oxford Univ. Press, 1993.
[22] F. Heitz and P. Bouthemy, "Multimodal estimation of discontinuous optical flow using Markov random fields," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 15, pp. 1217–1232, 1993.
[23] F. Heitz, P. Pérez, and P. Bouthemy, "Multiscale minimization of global energy functions in some visual recovery problems," *CVGIP: Image Understand.,* vol. 59, pp. 125–134, 1994.
[24] P. Holland and R. Welsch, "Robust regression using iteratively reweighted least-squares," *Commun. Stat. Theoret. Meth.,* vol. A6, pp. 813–827, 1977.
[25] P. Huber, *Robust Statistics.* New York: Wiley, 1981.
[26] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.,* vol. 1, pp. 321–331, 1988.
[27] C. Kervrann and F. Heitz, "A hierarchical statistical framework for the segmentation of deformable objects in image sequences," in *Proc. Computer Vision and Pattern Recognition,* Seattle, WA, June 1994, pp. 724–728.
[28] J. Konrad and E. Dubois, "Bayesian estimation of motion vector fields," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 14, pp. 910–927, 1992.
[29] Y. Leclerc, "Constructing simple stable descriptions for image partitioning," *Int. J. Comput. Vis.,* vol. 3, pp. 73–102, 1989.
[30] F. Leymarie and M. Levine, "Tracking deformable objects in the plane using an active contour model," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 15, pp. 617–634, 1993.
[31] E. Mémin, F. Heitz, and F. Charot, "Efficient parallel nonlinear multigrid relaxation algorithms for low-level vision applications," *J. Parallel Distr. Comput.,* vol. 29, pp. 96–103, 1995.
[32] P. Nesi, "Variational approach to optical flow estimation managing discontinuities," *Image Vis. Comput.,* vol. 11, pp. 419–439, 1993.
[33] P. Pérez and B. Gidas, "Motion segmentation and tracking using deformable templates," in *Proc. Int. Conf. Image Processing,* Austin, TX, Nov. 1994.
[34] P. Piscaglia *et al.*, "Parallelized robust multiresolution motion estimation," in *Proc. Int. Conf. Image Processing,* Lausanne, Switzerland, Sept. 1996.
[35] P. Proesmans, L. Van Gool, E. Pauwels, and A. Oosterlinck, "Determination of optical flow and its discontinuities using nonlinear diffusion," in *Proc. Europ. Conf. Computer Vision,* Stockholm, Sweden, 1994, vol. 2, pp. 295–304.
[36] C. Schnörr, "Computation of discontinuous optical flow by domain decomposition and shape optimization," *Int. J. Comput. Vis.,* vol. 8, no. 2, pp. 153–165, 1992.

---

[6] In our experiments, this quantity turned out to be of order 10^{-3} or less.

[37] C. Schnörr and W. Peckard, "Motion-based identification of deformable templates," in *Proc. Int. Conf. Computer Analysis of Images and Patterns,* V. Hlavac and R. Sara, Eds. Prague, Czech Republic, Sept. 1995.

[38] C. Stiller, "Object-based estimation of dense motion fields," *IEEE Trans. Image Processing,* vol. 6, pp. 234–250, 1997.

[39] R. Szeliski and J. Coughlan, "Spline-based image registration," Tech. Rep. CRL 94/1, Digital Equipment Corp., Apr. 1994.

[40] R. Szeliski and H.-Y. Shum, "Motion estimation with quadtree splines," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 18, pp. 1199–1210, 1996.

[41] J. Zhang and G. Hanauer, "The application of mean field theory to image motion estimation," *IEEE Trans. Image Processing,* vol. 4, pp. 19–33, 1995.

Patrick Pérez was born in 1968. He graduated from Ecole Centrale Paris, France, in 1990, and received the Ph.D. degree in signal processing and telecommunications from the University of Rennes, France, in 1993.

He spent one year as an INRIA post-doctoral fellow at the Department of Applied Mathematics, Brown University, Providence, RI. He now holds a full-time research position at the INRIA Center, Rennes. His research interests include statistical and/or hierarchical models for large inverse problems in image analysis, with a particular focus on motion analysis issues (detection, segmentation, and estimation of the apparent motion, tracking, analysis of elastic and fluid motion, etc.).

Etienne Mémin was born in 1965. He received the Ph.D. degree in computer science from the University of Rennes, France, in 1993.

He is currently an Assistant Professor at the University of Bretagne Sud, France. His research interests include computer vision, statistical models for image (sequence) analysis, and parallel algorithms for computer vision.