I R I S A
INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTÈMES ALÉATOIRES

PUBLICATION
INTERNE
N° 1197

SEMI-PARAMETRIC ESTIMATION AND SEGMENTATION OF OPTIC FLOW

ETIENNE MÉMIN AND PATRICK PÉREZ

ISSN 1166-8687

IRISA
CAMPUS UNIVERSITAIRE DE BEAULIEU - 35042 RENNES CEDEX - FRANCE

IRISA
Tél. : (33) 02 99 84 71 00 – Fax : (33) 02 99 84 71 71
http://www.irisa.fr

Semi-parametric estimation and segmentation of optic flow

Etienne Mémin* and Patrick Pérez**

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Vista

Publication interne n ° 1197 — Août 1998 — 37 pages

Abstract: In this paper we present a comprehensive energy-based framework for the estimation and the segmentation of apparent motion. The robust cost functions and the associated hierarchical minimization techniques that we propose allow to mix efficiently non-parametric (dense) representations, local interacting parametric representations, and global non-interacting parametric representations related to a partition into regions. Experimental comparisons both on synthetic and real images allow to demonstrate the various advantages of the approach, depending on the type of sequence. It is thus demonstrated that good dense estimations of fluid motions can be achieved, as well as joint estimations-segmentations of 3D rigid motions.

Key-words: Optic flow, robust discontinuity-preserving estimation, motion-based segmentation, hierarchical non-linear minimization, dense and parametric representations.

(Résumé : tsvp)

* Université de Bretagne Sud, E-mail : memin@irisa.fr
** E-mail : perez@irisa.fr

Centre National de la Recherche Scientifique (UPRESSA 6074) Université de Rennes 1 – Insa de Rennes
Institut National de Recherche en Informatique et en Automatique – unité de recherche de Rennes

Estimation et segmentation semi-paramétrique du mouvement apparent

Résumé : Nous introduisons un cadre énergétique complet pour l'estimation et la segmentation du mouvement apparent. Les fonctions de coût robustes et les techniques hiérarchiques de minimisation que nous proposons offrent la possibilité de manipuler conjointement une représentation non-paramétrique dense, une représentation localement paramétrique avec interactions, et une représentation globale en régions paramétriques indépendantes. Les comparaisons expérimentales sur des données aussi bien synthétiques que réelles permettent de mettre en evidence les apports de l'approche, en fonction du type de séquence. Il est ainsi montré que d'intéressantes estimations denses peuvent être obtenues dans le cas de mouvements fluides complexes, ainsi que des estimations-segmentations dans le cas de scènes 3D plus structurées.

Mots clés : Mouvement apparent, estimation robuste, préservation des discontinuités, segmentation au sens du mouvement, minimisation non-convexe hiérarchique, représentation dense et paramétrique.

I. Introduction

Among early vision problems, the one dealing with the estimation and the segmentation of apparent motion from an image sequence is particularly intricate. It is a two-fold problem which lies at the heart of most issues related to the analysis of image sequences. It is thus a critical part of a number of computer vision applications such as motion detection in a scene, 3D motion and scene structure recovery, obstacle avoidance in robotics, etc. (see for example [23] for a recent review on motion analysis issues).

Stemming either from discrete Markovian framework or from deterministic continuous one, energy-based models seem very appealing to handle in a versatile way high-dimensional inverse problems, even when the variables are of very different natures and interact in a non-linear fashion. For motion analysis purposes, such models have been thoroughly investigated, usually based on the brightness conservation assumption. An infinitesimal expansion of this constraint leads to the well known *optical flow constraint equation* [18]. This equation links the projection of the unknown motion vector on the spatial gradient of the luminance function to the temporal partial derivative of the luminance function. The model is equipped with some *a priori* which is either *locally* captured by a smoothness term in the cost function [18], or more *globally* defined as a parametric representation of the unknown motion [1, 2, 4, 9].

These two types of priors have their own advantages and drawbacks. Opposite to local smoothing approach, parametric modelization relies on larger (and delimited) spatial supports of estimation. It is therefore more likely to be reliable and robust, as far as selected parameterization as well as the notion of motion regions make sense from physical point of view. Local non-parametric models are, in that sense, more general since they only capture smoothness assumptions on the desire solution. They are also independent of any partition of the image and local attributes of the motion field such as discontinuities may be easily handled [8].

Local non-parametric regularization is usually involved in *dense* motion estimation [8, 12, 14, 24] whereas parametric representations are more dedicated to motion-based segmentation where areas with cinematic meanings have to be extracted from the images [2, 4, 9, 26].

The limitations of each approach can be illustrated in the case of images involving fluid phenomena. In sequences of this nature, like those encountered in meteorology (atmospheric satellite images) or fluid mechanics (videos of wind tunnel or water tank experiments), it is very common to observe very low photometric contrasts. Although the optical flow constraint equation could easily be turned into a more "physical" *transport equation* [30, 33], it will always be plagued by the absence of consistent

photometric information to rely on. In that case, parametric approach, with its extended estimation supports, would seem more appropriate. However, the physics of the scene makes it delicate to apply: in fluid motions there is no real objects or motion regions with borders, and the involved motions can be more complex than those usually captured by standard parametric models.

In light of these preliminary remarks, our purpose is to propose methods that allow to mix both kinds of priors, thus offering a trade-off between local non-parametric smoothing and more global parametric representation. We are thus looking for what could be called a *semi-parametric* estimation framework. We actually present two different (and not exclusive) methods to reach that goal. The first one concerns a particular *constrained* minimization technique used with an energy-based dense motion estimation model. The second method deals with an energy-based model for the joint estimation/segmentation of the apparent motion. In both approaches (whose cost functions are partly the same), the considered energy incorporates so-called *robust* penalty functions. These functions allow in the one hand to deal with the different kinds of large deviations w.r.t. chosen models, and, in the other hand, they offer a simple way to eventually couple the different variables of the problem through additional "auxiliary variables".

As already mentioned, energetic formulations can be viewed either from a *continuous* angle or from a *discrete* one. Former kind of approaches imply continuous functionals, variational calculus, deterministic partial differential equations, and discretization schemes (finite differences, finite elements), whereas the latter type of formalism is often related to Markov random fields and Bayesian inference. The two viewpoints provide different insights into a given problem, as well as different mathematical tools to cope with the various issues at hand. Despite their differences, it is known that they often lead to very similar discrete implementations. Although our presentation relies more on a discrete philosophy, we shall provide some hints about the connection between the two points of view as concerns the models themselves, as well as the algorithmic treatment.

The paper is divided into three main parts. It first focuses in Section II on a robust energy-based model for the incremental dense estimation of the apparent motion field with preservation of its discontinuities. To cope with the associated minimization we introduce an efficient tailor-made hierarchical technique which combines different and varying parameterizations of the unknown field. The compromise between local dense methods and global parametric approaches is thus introduced via the minimization process. In the second part of the paper (Section III) we show how the former energy function can be enriched to simultaneously estimate a motion-based segmentation of the scene. The

resulting joint estimation-segmentation model introduces another mix between local smoothness and region-wise parameterization. The intricate global minimization one ends up with is performed with a natural extension of the hierarchical optimization technique developed in the previous part. The last part (Section IV) is devoted to experimental results. The two approaches are validated qualitatively and quantitatively on real world sequences involving fluid phenomena or rigid mobile objects.

II. Robust estimation of dense optic flow

Dense motion estimation aims at recovering the apparent 2D displacement field $\boldsymbol{w} = \{\boldsymbol{w}_s, s \in S\}$ over the rectangular pixel lattice S, based on luminance function $f(t) = \{f(s,t), s \in S\}$ at two consecutive instants t and $t+1$. Assuming the temporal constancy of the brightness for a physical point between the two images, standard optic flow estimation relies on the *optic flow constraint equation* (OFCE):

$$\boldsymbol{\nabla} f(s,t+1)^T \boldsymbol{w}_s + f_t(s) = 0, \tag{1}$$

where $\boldsymbol{\nabla} f$ stands for the spatial gradient of f and $f_t(s) \triangleq f(s,t+1) - f(s,t)$ denotes the luminance variation at location s. This equation issues from a linearization of the brightness constancy assumption $f(s+\boldsymbol{w}_s, t+1) - f(s,t) = 0$. Seeing f_t as an approximation of the temporal derivative and tacking spatial gradient at time t instead of $t+1$, left hand side of (1) can also be interpreted as the material derivative $\frac{df}{dt}$ of f (i.e., the rate of change of f as observed when moving with point s).

As a Taylor series expansion, the OFCE also assumes that the unknown displacement $\boldsymbol{w}_s$ remains in the "domain of linearity" of the luminance function at location s. This is particularly unlikely to hold around sharp edges (where large gradients imply reduced linearity domain), and for large displacements. These limitations are usually circumvented by conducting an *incremental* estimation through a multiresolution hierarchy of sequences [6, 15]. Such an incremental multiresolution technique can be related to non-linear least squares Gauss-Newton method [4, 24]. We, too, stick to that multiresolution setup involving a pyramidal decomposition of the images. Even if we do not make it explicit, we shall assume throughout to be working at a given resolution of the multiresolution structure. One has to keep in mind that all definitions and derivations are thus meant to be reproduced at each resolution level according to a coarse-to-fine strategy.

A. *Incremental energy-based model*

Incremental estimation assumes that a rough estimate $\boldsymbol{w} = \{\boldsymbol{w}_s, s \in S\}$ of the unknown vector field is available (e.g., from an estimation at lower resolution or from a previous estimation). A refinement is sought in terms of an *increment field* $\mathsf{d}\boldsymbol{w} \in \Omega \subset (\mathbb{R} \times \mathbb{R})^S$. Based on the linearization of constancy brightness assumption from time t to $t+1$ with respect to that increment, refinement estimation consists in:

$$\widehat{\mathsf{d}\boldsymbol{w}} = \arg\min_{\mathsf{d}\boldsymbol{w}} [H_1(\mathsf{d}\boldsymbol{w}; f, \boldsymbol{w}) + \alpha H_2(\mathsf{d}\boldsymbol{w}; \boldsymbol{w})], \tag{2}$$

with [6, 24]:

$$H_1(\mathsf{d}\boldsymbol{w}; f, \boldsymbol{w}) \triangleq \sum_{s \in S} \rho_1[\boldsymbol{\nabla} f(s + \boldsymbol{w}_s, t+1)^T \mathsf{d}\boldsymbol{w}_s + f_t(s, \boldsymbol{w}_s)], \tag{3}$$

$$H_2(\mathsf{d}\boldsymbol{w}; \boldsymbol{w}) \triangleq \sum_{\langle s,r \rangle \in \mathcal{C}} \rho_2 \left[\| (\boldsymbol{w}_s + \mathsf{d}\boldsymbol{w}_s) - (\boldsymbol{w}_r + \mathsf{d}\boldsymbol{w}_r) \| \right], \tag{4}$$

where $\alpha > 0$, $\mathcal{C}$ is the set of neighboring site pairs lying on grid S equipped with some neighborhood system ν, $f_t(s, \boldsymbol{w}_s) \triangleq f(s + \boldsymbol{w}_s, t+1) - f(s,t)$ is the displaced frame difference, and ρ_1 and ρ_2 are standard *robust M-estimators*. They respectively penalize the *deviations* both from the OFCE and from the first-order smoothing prior. Unlike the quadratic penalty, these robust functions exhibit a saturating property (their derivatives are negligible at infinity in front of $2x$) in presence of large residuals. From a practical point of view, these functions, which are often non-convex for improved robustness [13], can be replaced by a so-called semi-quadratic formulation: under certain simple conditions on ρ (mainly concavity of $\phi(v) \triangleq \rho(\sqrt{v})$, see [8, 11, 16] for a complete account), there exists an increasing function ψ such that $\rho(x) = \min_{z \in (0,1]} [\tau z x^2 + \psi(z)]$, where $\tau \triangleq \lim_{v \to 0+} \phi'(v)$, i.e, ρ is the inferior envelope of a family of parabolas continuously indexed by *auxiliary variable* (or weight) z lying in $(0, 1]$.[1] The minimizer is given by $\arg\min_{z \in (0,1]} [\tau z x^2 + \psi(z)] = \frac{\rho'(x)}{2\tau x} = \frac{1}{\tau} \phi'(x^2)$.

[1] This function is defined as $\psi(z) \triangleq \phi \circ \phi'^{-1}(\tau z) - \tau z \phi'^{-1}(\tau z)$ [11, 16, 24]. It is strictly decreasing since $\psi'(z) = -\tau \phi'^{-1}(\tau z) < 0$.

Using this reformulation result, the minimization of $H_1 + \alpha H_2$ can be replaced by the minimization in $(\mathsf{d}\boldsymbol{w}, \delta, \beta)$ of augmented cost function $\mathcal{H} \triangleq \mathcal{H}_1 + \alpha\mathcal{H}_2$ where $\mathcal{H}_1$ and $\mathcal{H}_2$ are respectively:

$$\mathcal{H}_1(\mathsf{d}\boldsymbol{w}, \delta; f, \boldsymbol{w}) = \sum_{s \in S} \left[\tau_1 \delta_s \left[\boldsymbol{\nabla} f(s+\boldsymbol{w}_s, t+1)^T \mathsf{d}\boldsymbol{w}_s + f_t(s, \boldsymbol{w}_s) \right]^2 + \psi_1(\delta_s) \right], \tag{5}$$

$$\mathcal{H}_2(\mathsf{d}\boldsymbol{w}, \beta; \boldsymbol{w}) = \sum_{\langle s,r \rangle \in \mathcal{C}} \left[\tau_2 \beta_{sr} \| (\boldsymbol{w}_s + \mathsf{d}\boldsymbol{w}_s) - (\boldsymbol{w}_r + \mathsf{d}\boldsymbol{w}_r) \|^2 + \psi_2(\beta_{sr}) \right], \tag{6}$$

and $\delta = \{\delta_s,\ s \in S\}$, $\beta = \{\beta_{sr}, \langle s,r \rangle \in \mathcal{C}\}$ are two sets of auxiliary variables lying within $(0,1]$. This new minimization can then be led *alternatively* with respect to $\mathsf{d}\boldsymbol{w}$ and to the weights: energy $\mathcal{H}$ is quadratic w.r.t. $\mathsf{d}\boldsymbol{w}$ and the corresponding minimization amounts to a standard *weighted least squares* problem; conversely, $\mathsf{d}\boldsymbol{w}$ being frozen, the best weights are obtained in closed form [24]. Convergence of this alternate scheme is guaranteed to a global minimum if ρ is convex [11], and to a local minimum otherwise [13].

This robust energy-based modeling provides a generic dense estimator which can be applied to image sequences of various natures as far as they contain sufficient photometric contrast or texture and they are not too much noisy. In case neither condition is met within large areas of the image, the robust smoothness term is not strong enough to propagate the sensible estimates obtained at the border of these regions toward inner locations where data cannot be exploited. As a by-product of the hierarchical piece-wise parametric minimization we now introduce, these problems of insufficient or unreliable information will hopefully be circumvented. Other gains will be obtained in terms of global quality of estimates and computational load.

B. *Piece-wise parametric constraint*

Let us assume that the pixel grid is subdivided into a collection of patches. Let $\mathcal{B} \triangleq \{\mathcal{B}_n,\ n = 1 \dots N\}$ be this partition and $G^* = [\{1 \dots N\}, \nu^*]$ the associated connectivity graph with N vertices (Fig. 1).[2] Piece-wise parametric increment fields for this partition are defined as:

$$\forall n = 1 \dots N,\ \forall s \in \mathcal{B}_n, \quad \mathsf{d}\boldsymbol{w}_s = \Phi_n(\boldsymbol{\theta}_n, s), \tag{7}$$

[2] In case $\mathcal{B}$ is a regular partition into square patches, G is the N-site rectangular lattice with same neighborhood system as the original lattice.

where $\boldsymbol{\theta}_n$ is a parameter vector and Φ_n's are interpolation functions which can be different from one patch to another. The whole increment can then be expressed $\mathsf{d}\boldsymbol{w} = \Phi(\boldsymbol{\theta})$ with $\boldsymbol{\theta}^T = (\boldsymbol{\theta}_1^T \dots \boldsymbol{\theta}_N^T)$ lying in parameter space Γ. The full-rank function Φ is the *interpolator* between the reduced subspace Γ and the original configuration space Ω. It is a one-to-one mapping from Γ into the constrained configuration subset $\mathrm{Im}\Phi \subset \Omega$.

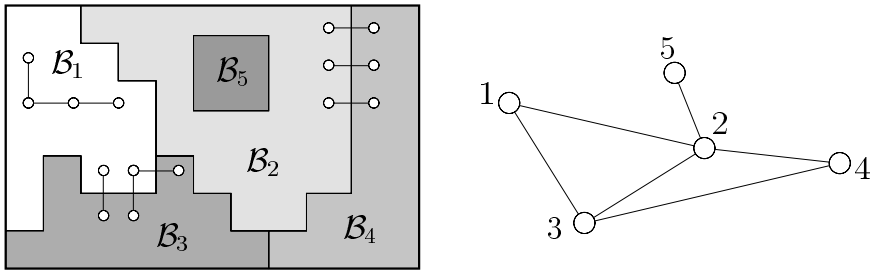

Fig. 1. Example of image partition $\mathcal{B} = \{\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{B}_4, \mathcal{B}_5\}$ (with some cliques of $\mathcal{C}_1$, $\mathcal{C}_{13}$, and $\mathcal{C}_{24}$), and associated adjacency graph $G^* = [\{1 \dots 5\}, \nu^*]$.

The constrained minimization of $\mathcal{H}$ in $\mathrm{Im}\Phi$ is equivalent to a new minimization defined on Γ:

$$\min_{\mathsf{d}\boldsymbol{w} \in \mathrm{Im}\Phi} \mathcal{H}(\mathsf{d}\boldsymbol{w}, \delta, \beta; f, \boldsymbol{w}) = \min_{\boldsymbol{\theta} \in \Gamma} \underbrace{\mathcal{H}(\Phi(\boldsymbol{\theta}), \delta, \beta; f, \boldsymbol{w})}_{\triangleq \mathcal{H}^*(\boldsymbol{\theta}, \delta, \beta; f, \boldsymbol{w})}.$$

New energy function $\mathcal{H}^*$ is readily derived from the original one (5-6). Denoting $\mathcal{C}_n \triangleq \{\langle s,r\rangle \in \mathcal{C} : \langle s,r\rangle \subset \mathcal{B}_n\}$ the set of neighboring site pairs included in patch $\mathcal{B}_n$ and $\mathcal{C}_{nm} \triangleq \{\langle s,r\rangle \in \mathcal{C} : s \in \mathcal{B}_n, r \in \mathcal{B}_m\}$ the set of neighboring site pairs straddling adjacent patches $\mathcal{B}_n$ and $\mathcal{B}_m$ (see Fig. 1), one can show that this new energy is similarly composed of two terms, $\mathcal{H}^* = \mathcal{H}_1^* + \alpha\mathcal{H}_2^*$, defined as:

$$\mathcal{H}_1^*(\boldsymbol{\theta}, \delta; f, \boldsymbol{w}) = \sum_n \sum_{s \in \mathcal{B}_n} \tau_1 \delta_s \left[\boldsymbol{\nabla} f(s + \boldsymbol{w}_s, t+1)^T \Phi_n(\boldsymbol{\theta}_n, s) + f_t(s, \boldsymbol{w}_s)\right]^2 + \psi_1(\delta_s), \tag{8}$$

$$\begin{aligned} \mathcal{H}_2^*(\boldsymbol{\theta}, \beta; \boldsymbol{w}) = &\sum_{\langle n,m\rangle} \sum_{\langle s,r\rangle \in \mathcal{C}_{nm}} \tau_2 \beta_{sr} \left\|(\boldsymbol{w}_s + \Phi_n(\boldsymbol{\theta}_n, s)) - (\boldsymbol{w}_r + \Phi_m(\boldsymbol{\theta}_m, r))\right\|^2 + \psi_2(\beta_{sr}) \\ &+ \sum_n \sum_{\langle s,r\rangle \in \mathcal{C}_n} \tau_2 \beta_{sr} \left\|(\boldsymbol{w}_s + \Phi_n(\boldsymbol{\theta}_n, s)) - (\boldsymbol{w}_r + \Phi_n(\boldsymbol{\theta}_n, r))\right\|^2 + \psi_2(\beta_{sr}), \end{aligned} \tag{9}$$

where $\langle n, m\rangle$ denotes neighboring pairs of region vortices in graph G^*. Note that the first term in the definition of $\mathcal{H}_2^*$ is reminiscent of the "skin and bones" model introduced by Ju *et al.* [20].

Minimizing $\mathcal{H}^*$ allows to recover a piece-wise parametric increment field where different parameterizations are combined. Note that, contrary to what is done in parametric *segmentation* approaches based on independent region-wise parametric models and in quadtree splines approach of Szeliski *et al.* [32]., here, the different parameter vectors $\boldsymbol{\theta}_n$'s do interact through the smoothness term that enforces continuity at patch frontiers.

In the next section we show how this constrained optimization can be easily embedded in a *hierarchical* optimization framework.

C. *Hierarchical constrained optimization*

We now consider a sequence of partitions. Previous constrained optimization can be successively applied with each of these partitions, thus providing a hierarchical optimization scheme: the original optimization problem of $\mathcal{H}$ is replaced by a succession of constrained minimizations. Let $\{\mathcal{B}^\ell, \ell = L \dots 0\}$ be the family of partitions with $\mathcal{B}^0 = S$, and $G^\ell = [S^\ell, \nu^\ell]$, $\ell = L \dots 0$, the associated connectivity graphs with N_ℓ-vertex sets S^ℓ. For each patch $\mathcal{B}^\ell$, an interpolator Φ^ℓ is chosen such that the size of $\Omega^\ell \triangleq \mathrm{Im}\Phi^\ell$ decreases as ℓ increases.[3] The constrained optimization in Ω^ℓ is equivalent to the minimization of the new energy function:

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

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

$$(\hat{\boldsymbol{\theta}}^\ell, \hat{\delta}, \hat{\beta}) = \underset{\boldsymbol{\theta}^\ell, \delta, \beta}{\mathrm{argmin}}\, \mathcal{H}^\ell(\boldsymbol{\theta}^\ell, \delta, \beta; f, \boldsymbol{w}^\ell)\ \ell = L \dots 0, \tag{10}$$

where the field to be refined at level ℓ, $\boldsymbol{w}^\ell \triangleq \boldsymbol{w}^{\ell+1} + \Phi^{\ell+1}(\hat{\boldsymbol{\theta}}^{\ell+1})$, is deduced from estimate at level $\ell+1$, and initial field $\boldsymbol{w}^L$ comes from an estimation at a coarser resolution or from a given initialization. Hence, the definition of energy at a given ℓ involves spatial and temporal luminance gradients computed

[3] A natural way of building this hierarchy of parametric representations is to consider nested partitions where $\mathcal{B}^\ell$ is made up from the subdivision of elements of $\mathcal{B}^{\ell+1}$. This nested structure is easily obtained with regular subdivision schemes (square or triangle subdivision). It is more difficult to design for irregular subdivision strategies. In section IV, we shall introduce an adaptative way to build square-based nested partitions.

with the second image being warped according to vector field $\boldsymbol{w}^\ell$, which changes from one level to the next one.

Each of these successive minimizations is processed in terms of iteratively reweighted least squares initialized by $\boldsymbol{\theta}^\ell \equiv \mathbf{0}$, yielding increment field $\widehat{\mathsf{d}\boldsymbol{w}}^\ell = \Phi^\ell(\widehat{\boldsymbol{\theta}}^\ell)$. The procedure is repeated until the finest level $\ell = 0$ is reached, and the motion field finally recovered is $\boldsymbol{w}^L + \sum_{\ell=L}^{0} \Phi^\ell(\widehat{\boldsymbol{\theta}}^\ell)$. This incremental minimization procedure can be viewed as a *hierarchical Gauss-Newton* minimization of $\sum_s \rho_1(f(s + \boldsymbol{w}_s, t+1) - f(s,t)) + \alpha \sum_{\langle s,r \rangle} \rho_2(\|\boldsymbol{w}_s - \boldsymbol{w}_r\|)$.[4]

In the same way, the incremental *multiresolution* setup for each single level of which previous hierarchical minimization was designed, can be interpreted as an *approximate hierarchical Gauss-Newton minimization* [4, 24] since successive linearizations concern different energy functions (based on different data).

D. *Linear parameterizations and energy minimization*

So far, we let the nature of parameterizations unspecified. In practice, interpolation functions are chosen linear:

$$\forall n \in S^\ell,\ \forall s \in \mathcal{B}_n^\ell,\ \mathsf{d}\boldsymbol{w}_s^\ell = P_n(s)\boldsymbol{\theta}_n^\ell, \tag{11}$$

where $P_n(s)$ is 2 by p_n matrix. The corresponding parameter spaces are $\Gamma^\ell = \times_{n=0}^{N_\ell} \mathbb{R}^{p_n}$. Standard parametric models used in motion analysis correspond to p_n=2, 4, 6 or 8 [1, 4]. In this work we will consider three possible parameterizations: the *constant model* (2 parameters of translation and $P_n(s) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$); the *simplified affine model* based on plane similarity (4 parameters of translation, scaling, and rotation, and $P_n(s) = \begin{bmatrix} 1 & 0 & x_s & y_s \\ 0 & 1 & y_s & -x_s \end{bmatrix}$); and the *affine model* (6 parameters and $P_n(s) = \begin{bmatrix} 1 & x_s & y_s & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_s & y_s \end{bmatrix}$), where, in previous expressions x_s and y_s stand for the coordinates of pixel s. As reported in section IV, we have investigated different combinations of these three parameterizations.

[4]When constrained subsets are nested, i.e., $\Omega^{\ell+1} \subset \Omega^\ell$, the succession of minimizations can be conducted in a slightly different way: the final estimate at a given level is not directly integrated in the main field to be refined at the next level, but simply used as an initialization for the iterative minimization process. More precisely, all $\boldsymbol{w}^\ell$ fields in (10) are the same, equal to some field $\boldsymbol{w}$, while $\widehat{\boldsymbol{\theta}}^{\ell+1}$ is now used to define the initial increment configuration at level ℓ, through $(\Phi^\ell)^{-1} \circ \Phi^{\ell+1}(\widehat{\boldsymbol{\theta}}^{\ell+1})$ (which makes sense since $\mathrm{Im}\Phi^{\ell+1} \subset \mathrm{Im}\Phi^\ell$). In this version, described in [24], the spatio-temporal luminance derivatives remain the same, i.e., computed with respect to $f(s,t)$ and $f(s+\boldsymbol{w}_s, t+1)$. In other terms, a single linearization of the brightness constancy assumption is considered (for a given resolution level), and the coarse-to-fine minimization turns out to be a standard *multigrid* scheme [17].

Introducing linear parameterizations within energies (8-9) yields

$$\mathcal{H}_1^\ell(\boldsymbol{\theta}^\ell, \delta; f, \boldsymbol{w}^\ell) = \sum_{n\in S^\ell} \sum_{s\in\mathcal{B}_n^\ell} \tau_1 \delta_s \left[\boldsymbol{\nabla} f(s+\boldsymbol{w}_s^\ell, t+1)^T P_n(s)\boldsymbol{\theta}_n^\ell + f_t(s, \boldsymbol{w}_s^\ell)\right]^2 + \psi_1(\delta_s), \tag{12}$$

$$\begin{aligned}\mathcal{H}_2^\ell(\boldsymbol{\theta}^\ell, \beta; \boldsymbol{w}^\ell) = &\sum_{\langle n,m\rangle} \sum_{\langle s,r\rangle\in\mathcal{C}_{nm}^\ell} \tau_2 \beta_{sr} \left\|(\boldsymbol{w}_s^\ell + P_n(s)\boldsymbol{\theta}_n^\ell) - (\boldsymbol{w}_r^\ell + P_m(r)\boldsymbol{\theta}_m^\ell)\right\|^2 + \psi_2(\beta_{sr}) \\ &+ \sum_{n\in S^\ell} \sum_{\langle s,r\rangle\in\mathcal{C}_n^\ell} \tau_2 \beta_{sr} \left\|(\boldsymbol{w}_s^\ell + P_n(s)\boldsymbol{\theta}_n^\ell) - (\boldsymbol{w}_r^\ell + P_n(r)\boldsymbol{\theta}_n^\ell)\right\|^2 + \psi_2(\beta_{sr}).\end{aligned} \tag{13}$$

The iteratively reweighted least squares minimization applied to this energy function amounts to alternate updates of weights and parameter vectors**.** The current parameter estimate $\boldsymbol{\theta}^\ell$ being fixed, we know that the optimal weight values are directly accessible. These values are:

$$\forall n,\ \forall s \in \mathcal{B}_n^\ell,\ \delta_s = \frac{1}{\tau_1}\phi_1'\left[\left(\boldsymbol{\nabla} f(s+\boldsymbol{w}_s^\ell, t+1)^T P_n(s)\boldsymbol{\theta}_n^\ell + f_t(s, \boldsymbol{w}_s^\ell)\right)^2\right], \tag{14}$$

$$\forall \langle n,m\rangle,\ \forall \langle s,r\rangle \in \mathcal{C}_{nm}^\ell,\ \beta_{sr} = \frac{1}{\tau_2}\phi_2'\left[\left\|(\boldsymbol{w}_s^\ell + P_n(s)\boldsymbol{\theta}_n^\ell) - (\boldsymbol{w}_r^\ell + P_m(r)\boldsymbol{\theta}_m^\ell)\right\|^2\right], \tag{15}$$

$$\forall n,\ \forall \langle s,r\rangle \in \mathcal{C}_n^\ell,\ \beta_{sr} = \frac{1}{\tau_2}\phi_2'\left[\left\|\boldsymbol{w}_s^\ell - \boldsymbol{w}_r^\ell + (P_n(s) - P_n(r))\boldsymbol{\theta}_n^\ell\right\|^2\right]. \tag{16}$$

It is worth noting that according to (16), the discontinuity variables β_{sr} located into patches of $\mathcal{B}^\ell$ (i.e., $\langle s,r\rangle \in \mathcal{C}_n^\ell$ for some $n \in S^\ell$) do not depend on the translational components of $\boldsymbol{\theta}^\ell$. In the *piece-wise constant* case they therefore depend only on $\boldsymbol{w}^\ell$, and can be computed right away within the first iteration at the current grid level. As soon as the values of all 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 (zeroing the gradient) 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 $\mathcal{B}_n^\ell$. This is detailed in the Appendix A for the three different possible parameterizations on $\mathcal{B}_n^\ell$.

Before explaining how the model can be enriched to deal with a joint segmentation process, it is interesting to emphasize the connection between this discrete formalization of the optic flow estimation problem and its possible continuous counterpart.

E. *Continuous formalization of optic flow estimation*

A continuous version of the discrete energy defined by (5-6) corresponds to the functional

$$\begin{aligned} J(\mathsf{d}\boldsymbol{w},\delta,\beta) = &\iint_{\mathcal{S}} \tau_1\delta(\boldsymbol{x})\left[\boldsymbol{\nabla}\tilde{f}(\boldsymbol{x})^T\mathsf{d}\boldsymbol{w}(\boldsymbol{x})+\tilde{f}_t(\boldsymbol{x})\right]^2+\psi_1(\delta(\boldsymbol{x}))d\boldsymbol{x}\\ &+\alpha\iint_{\mathcal{S}}\tau_2\beta(\boldsymbol{x})\|\boldsymbol{\nabla}(\boldsymbol{w}(\boldsymbol{x})+\mathsf{d}\boldsymbol{w}(\boldsymbol{x}))\|^2+\psi_2(\beta(\boldsymbol{x}))d\boldsymbol{x}, \end{aligned} \tag{17}$$

where $\boldsymbol{w}$ and $\mathsf{d}\boldsymbol{w}$ are, momentarily, two C^1-vector fields over continuous plane domain $\mathcal{S}\subset\mathbb{R}^2$, δ and β two scalar fields on the same domain, and $\boldsymbol{\nabla}\tilde{f}\triangleq\boldsymbol{\nabla}f(.+\boldsymbol{w},t+1)$, $\tilde{f}_t\triangleq f(.+\boldsymbol{w},t+1)-f(.,t)$.

The problem of minimizing this half-quadratic functional J can be addressed in terms of alternate minimization [14]. For fixed $\mathsf{d}\boldsymbol{w}$, Euler-Lagrange equations provide optimal expression of functions δ and β (using $\psi'(z)=\phi\circ\phi'^{-1}(\tau z)-\tau z\phi'^{-1}(\tau z)$):

$$\delta=\frac{1}{\tau_1}\phi_1'[(\boldsymbol{\nabla}\tilde{f}^T\mathsf{d}\boldsymbol{w}+\tilde{f}_t)^2],\ \text{and}\ \beta=\frac{1}{\tau_2}\phi_2'[\|\boldsymbol{\nabla}(\boldsymbol{w}+\mathsf{d}\boldsymbol{w})\|^2]. \tag{18}$$

The natural discretization of first equation that consists in taking values of δ, $\tilde{f}_t$, and $\boldsymbol{\nabla}\tilde{f}$ at pixel locations $\boldsymbol{x}=s\in S$, is readily obtained and yields exactly the same update rule as the one stemming from the minimization of discrete energy $\mathcal{H}$ w.r.t. $\{\delta_s\}_{s\in S}$. The same discretization scheme can be adopted for the second equation (as in [27]). However, if gradients are approximated by finite differences on the dual grid, it is simpler to have function β discretized on the same edge lattice. The corresponding discretized update is then the same as the one that minimizes $\mathcal{H}$ w.r.t. $\{\beta_{sr}\}_{\langle s,t\rangle}$.

The weight functions δ and β being fixed, one has to deal with the minimization of a quadratic functional of $\mathsf{d}\boldsymbol{w}$. This can be conducted first by writing down Euler-Lagrange equations as a necessary condition of optimality:

$$\tau_1\delta\boldsymbol{\nabla}\tilde{f}\boldsymbol{\nabla}\tilde{f}^T\mathsf{d}\boldsymbol{w}-\alpha\tau_2\mathrm{div}[\beta\boldsymbol{\nabla}(\boldsymbol{w}+\mathsf{d}\boldsymbol{w})]=-\tau_1\delta\tilde{f}_t\boldsymbol{\nabla}\tilde{f}. \tag{19}$$

If $\boldsymbol{w}$, $\mathsf{d}\boldsymbol{w}$, and δ are discretized on S, while β, $\boldsymbol{\nabla}\boldsymbol{w}$, and $\boldsymbol{\nabla}\mathsf{d}\boldsymbol{w}$ are discretized on the dual grid, and divergence operator is approximated by first-order central difference on S, this partial differential equation leads to a linear system that coincides with the one to be solved for minimizing $\mathcal{H}$ in $\mathsf{d}\boldsymbol{w}$.

We see that a standard discretization based on finite differences turns the minimization of continuous functional J into the same problem as the minimization of discrete energy $\mathcal{H}$. Continuous formalism, however, allows more flexibility in the choice of discretization scheme since the discretization step is

delayed compared to discrete energy case: whereas discrete modeling sticks right away to the pixel grid discretization, variational approaches offer other choices, especially when finite element method is used (within Euler-Lagrange formalism [12], or apart from it [29]). In many cases, however, discretization of the original continuous model is made as simple as possible w.r.t. pixel grid, thus yielding *in fine* the same discrete problems to be solved as those associated with the minimization of discrete energies.

As concerns the constrained minimization scheme introduced in section II.A when dealing with discrete energy $\mathcal{H}$, it can be viewed in continuous framework as a *Galerkin* technique for solving linear system arising from the discretization of (19). Denote $A\mathsf{d}\boldsymbol{w} = \boldsymbol{b}$ this system. Provided that interpolator Φ from Γ into Ω is linear, coarse-to-fine multigrid [17] relies on the resolution in Γ of so-called Galerkin system $\Phi^T A \Phi \boldsymbol{\theta} = \Phi^T \boldsymbol{b}$. The solution of this equation is obviously the minimizer of quadratic energy $||A\Phi\boldsymbol{\theta} - \boldsymbol{b}||^2$. In other terms it corresponds to the minimizer of $||A\mathsf{d}\boldsymbol{w} - \boldsymbol{b}||^2$ within subspace ImΦ. In case of simple discretization scheme mentioned earlier, $\mathcal{H}$ and $||A\mathsf{d}\boldsymbol{w} - \boldsymbol{b}||^2$ coincide up to an additive term independent from $\mathsf{d}\boldsymbol{w}$, and Galerkin system above provides the minimizer in $\boldsymbol{\theta}$ of reduced energy $\mathcal{H}^*$.

Finally, the whole approach to dense optic flow estimation we have introduced in this section could have been equivalently formulated, at the starting point, in a continuous fashion, as it is done in related works [12, 14, 29]. However, this does not hold for the augmented estimation-segmentation model to be presented.

III. Joint estimation-segmentation

In previous section we have described a general hierarchical method to estimate a dense optic flow field. It allows to mix different parameterizations of the unknown vector field. We shall see with experimental results that it provides a family of hierarchical motion estimators which give good results on sequences involving either fluid or rigid motions. Before reporting these experiments, we now introduce an extension of the model to couple the estimation process with a motion-based partition of the image.

Motion estimation and motion-based segmentation are two tightly interwoven problems: a good estimation of the motion field (or at least a sensible approximation of it) is required to obtain a good segmentation of the different apparent motions present in the scene; conversely, a good estimation of motion field cannot be obtained without an accurate estimation of the frontiers of the different moving objects. It is therefore natural to consider the resolution of these two problems as a whole.

This has been considered in a number of different ways and within a variety of methodological frameworks. Nevertheless, two main classes of estimation-segmentation approaches can be distinguished. The first one consists in an *unilateral* coupling between some motion estimates – such as sparse matchings, estimate of contour motions or dense optic flow estimate – and a segmentation process [1, 2, 9, 19, 26]. In that class of methods, motion cues are first extracted and then considered as data on which the segmentation is built. The second class of methods implies a real coupling between the estimation of motion and the extraction of a motion-based partition of the image within a joint procedure. This is usually achieved using a global energy function depending on both entities. In that context, different kinds of interactions have been recently proposed. In [31] the partition frontiers are estimated as a representation of the motion discontinuities. Except for the smoothness prior on the motion field inside each region, there is no stronger *a priori*. In [7, 10] the interaction consists in a cooperation between a dense motion field and a region-wise parametric polynomial flow. The motion is encouraged to have some similarity with the piece-wise parametric field associated with a segmentation.

The estimation-segmentation coupling we consider here belongs to the latter class. We aim at building, through a global discrete energy function, a cooperative method allowing to estimate simultaneously a dense motion field and a motion-based segmentation. The associated minimization is solved with an extension of the hierarchical optimization scheme described in section II.

A. *Compound energy*

Let $\mathcal{R}$ be a partition of S into an *unknown* number p of connected regions, $\mathcal{R} = \{\mathcal{R}_1 \dots \mathcal{R}_p\}$, and let $G(\mathcal{R}) = [\{1 \dots p\}, \mathcal{G}]$ be the associated p-site adjacency graph. We shall call "boundary" between regions $\mathcal{R}_i$ and $\mathcal{R}_j$ the set $\partial_{ij} \triangleq \{\langle s, r\rangle \in \mathcal{C} : s \in \mathcal{R}_i, r \in \mathcal{R}_j\}$. The set $\partial\mathcal{R} \triangleq \cup_{\langle i,j\rangle}\partial_{ij}$, where $\langle i, j\rangle$ denotes all neighboring site pairs of $G(\mathcal{R})$ (i.e., couples of neighboring regions from partition $\mathcal{R}$), stands therefore for the total frontier of the partition.

The extension of the energy-based optic flow estimation model is obtained by incorporating two terms to the global energy function $\mathcal{H}$.[5] The first one, E_{prior}, captures the *a priori* knowledge about the segmentation configuration. The second one, E_{interact}, specifies the mode of interaction between the segmentation and the rest of the estimation model (i.e., motion fields, weights and data). The mode of interaction we designed involves the motion field both at the frontiers and inside the regions: the

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

elements of the partition will interact with the estimation process through the discontinuity weights along the frontiers, and through a parametric goodness-of-fit criterion inside each region.

Energy component $E_{\text{interact}}(\mathsf{d}\boldsymbol{w}, \mathcal{R}, \beta; \boldsymbol{w})$ thus exhibits two terms: The first one is proportional to the sum of β_{sr} averages on sets ∂_{ij}. It then (a) favors low values (close to zero) of discontinuity weights along the borders and (b) guides the partition boundaries toward the most significant motion discontinuities. The second term enforces a polynomial parametric likeness of the dense motion field inside each region, *via* a robust penalization of the discrepancies. The global energy of the extended model is:

$$\mathcal{H}(\mathsf{d}\boldsymbol{w}, \delta, \beta; f, \boldsymbol{w}) + E_{\text{prior}}(\mathcal{R}) + E_{\text{interact}}(\mathsf{d}\boldsymbol{w}, \mathcal{R}, \beta; \boldsymbol{w}), \tag{20}$$

$$\text{where, } E_{\text{interact}}(\mathsf{d}\boldsymbol{w}, \mathcal{R}, \beta; \boldsymbol{w}) \triangleq \mu_1 \sum_{\langle i,j \rangle} \frac{1}{|\partial_{ij}|} \sum_{\langle s,r \rangle \in \partial_{ij}} \beta_{sr} + \mu_2 \sum_i \sum_{s \in \mathcal{R}_i} \rho_3(\|\boldsymbol{w}_s + \mathsf{d}\boldsymbol{w}_s - \boldsymbol{w}_s^i\|), \tag{21}$$

with some positive parameters μ_1 and μ_2 and the function ρ_3 being a robust M-estimator with hyper-parameter τ_3. The vector $\boldsymbol{w}_s^i$ is given by the affine model associated with the six-parameter vector $\boldsymbol{\theta}_i$ of region $\mathcal{R}_i$ (i.e., $\boldsymbol{w}_s^i \triangleq P(s)\boldsymbol{\theta}_i$, with 2×6 matrix $P(s)$ as defined in §II.D). This affine model is usually considered as a good trade-off between model complexity and versatility [9], and has thus been extensively used in motion analysis. The semi-quadratic re-writing yields:

$$\mathcal{E}_{\text{interact}}(\mathsf{d}\boldsymbol{w}, \mathcal{R}, \beta, \eta; \boldsymbol{w}) \triangleq \mu_1 \sum_{\langle i,j \rangle} \frac{1}{|\partial_{ij}|} \sum_{\langle s,r \rangle \in \partial_{ij}} \beta_{sr} + \mu_2 \sum_i \sum_{s \in \mathcal{R}_i} [\tau_3 \eta_s \|\boldsymbol{w}_s + \mathsf{d}\boldsymbol{w}_s - \boldsymbol{w}_s^i\|^2 + \psi_3(\eta_s)]. \tag{22}$$

The auxiliary variables η will be referred to as *parametric likeness weights* in the remainder.

The segmentation *a priori* term captures a loose geometric constraint according to standard Minimum Description Length (MDL) principle [22].

$$E_{\text{prior}}(\mathcal{R}) \triangleq \lambda |\partial \mathcal{R}| \text{ for some } \lambda > 0. \tag{23}$$

This energy term favors regions with short and smooth borders.

The whole energy function $\mathbb{H} \triangleq \mathcal{H} + E_{\text{prior}} + \mathcal{E}_{\text{interact}}$ has now to be minimized with respect to all the unknowns. A direct minimization of such function is obviously a very intricate problem. Nevertheless, it can be efficiently conducted through an extension of the hierarchical minimization strategy used for the motion estimation problem.

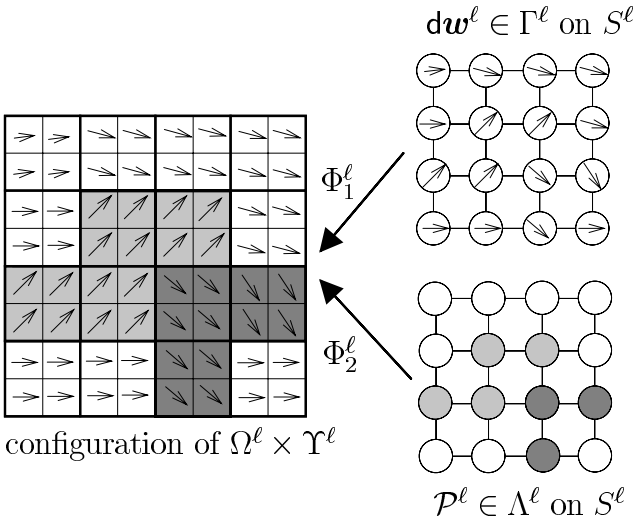

Fig. 2. A constrained increment field on a 2×2 block partition of S and a constrained segmentation into five regions (left); the associated increment field and partition on reduced grid S^ℓ (right).

B. *Hierarchical optimization*

As in §II.C, a family of increasing partitions $\mathcal{B}^\ell$, $\ell = L \ldots 0$, is specified and the optimization problem is decomposed as a recursive sequence of constrained minimizations based on these partitions. At level ℓ, the problem is the joint estimation of an increment field which is piece-wise parametric w.r.t. $\mathcal{B}^\ell$, and a segmentation made up from the elements of $\mathcal{B}^\ell$.

In this joint model, the use of partitions $\mathcal{B}^\ell$ with sophisticated variable shapes seems less relevant than for motion estimation only, since there is now an explicit segmentation beside to the field. Hence, we chose the simple nested family of $2^\ell \times 2^\ell$-block partitions along with the piece-wise constancy constraint on the increment field. At level ℓ, segmentation $\mathcal{R}$ is also constrained to fit the block partition $\mathcal{B}^\ell$. The corresponding joint configuration subset is denoted $\Omega^\ell \times \Upsilon^\ell$ (see instance of such constrained configurations on the left of figure 2).

Each constrained configuration of Ω^ℓ is equivalently described by a reduced increment field, $\mathsf{d}\boldsymbol{w}^\ell \in \Gamma^\ell$, lying on the grid S^ℓ. The one-to-one mapping from Γ^ℓ into Ω^ℓ will be denoted Φ_1^ℓ. In the same way a constrained partition of Υ^ℓ is associated with a partition into connected components of reduced grid $G^\ell = [S^\ell, \nu^\ell]$. If Λ^ℓ is the set of such partitions, the corresponding mapping from Λ^ℓ into Υ^ℓ is denoted Φ_2^ℓ, with $\Upsilon^\ell = \mathrm{Im}\Phi_2^\ell$ (see Fig. 2).

Constrained optimization in $\Omega^\ell \times \Upsilon^\ell$ is then equivalent to the minimization of the new energy function:

$$\mathbb{H}^\ell(\mathsf{d}\boldsymbol{w}^\ell, \mathcal{R}^\ell, \delta, \beta, \eta; f, \boldsymbol{w}) \triangleq \mathbb{H}(\Phi_1^\ell(\mathsf{d}\boldsymbol{w}^\ell), \Phi_2^\ell(\mathcal{R}^\ell), \delta, \beta, \eta; f, \boldsymbol{w}), \tag{24}$$

defined over $\Gamma^\ell \times \Lambda^\ell$, whereas the auxiliary variables, the data, and the field to be refined are still defined on the original grid S. We now deal with the cascade of optimization problems of reduced complexity:

$$\min_{\mathsf{d}\boldsymbol{w}^\ell, \mathcal{R}^\ell, \delta, \beta, \eta} \mathbb{H}^\ell(\mathsf{d}\boldsymbol{w}^\ell, \mathcal{R}^\ell, \delta, \beta, \eta; f, \boldsymbol{w}^\ell), \ \ell = L \ldots 0, \tag{25}$$

with, as in section II.C, $\boldsymbol{w}^\ell$ being defined according to the minimizer at previous level $\ell + 1$: $\boldsymbol{w}^\ell = \boldsymbol{w}^{\ell+1} + \Phi_1^\ell(\widehat{\mathsf{d}\boldsymbol{w}}^\ell)$. [6]

As with motion estimation alone, the new multigrid function $\mathbb{H}^\ell$ turns out to be composed of four terms similar to those of $\mathbb{H}$: $\mathbb{H}^\ell = \mathcal{H}_1^\ell + \alpha\mathcal{H}_2^\ell + \mathcal{E}_{\text{interact}}^\ell + E_{\text{prior}}^\ell$. The detailed expression of these different terms is presented in Appendix B.

The overall optimization process consists in an alternate minimization of the different weights (β, δ and η) and of the original variables of interest $\mathsf{d}\boldsymbol{w}^\ell$ and $\mathcal{R}^\ell$.

The reduced partition $\mathcal{R}^\ell$ and the parametric likeness weights η being fixed, one has to solve:

$$(\widehat{\mathsf{d}\boldsymbol{w}}^\ell, \hat{\beta}, \hat{\delta}) = \arg\min_{(\mathsf{d}\boldsymbol{w}^\ell, \beta, \delta)} [\mathcal{H}_1^\ell + \alpha\mathcal{H}_2^\ell + \mathcal{E}_{\text{interact}}^\ell]. \tag{26}$$

Apart from the interaction term, this is the same problem as in section II.C, and one can again resort to iteratively reweighted least squares, as follows.

Increment field $\mathsf{d}\boldsymbol{w}^\ell$ being fixed, let $\langle s, r \rangle$ be a clique in $\mathcal{C}$ and denote by m and n the block numbers (possibly identical) such that: $s \in \mathcal{B}_n^\ell$ and $r \in \mathcal{B}_n^\ell$. From β_{sr}'s point of view, the only change with respect to motion estimation case (15-16) occurs if $\langle s, r \rangle$ straddles two neighboring regions of $\Phi_2^\ell(\mathcal{R}^\ell)$.

[6] As mentioned in section II.C, this *hierarchical Gauss-Newton* minimization can be replaced by a *multigrid* minimization when constrained configuration subsets are nested, which is the case here ($\Omega^{\ell+1} \times \Upsilon^{\ell+1} \subset \Omega^\ell \times \Upsilon^\ell$): the final estimate at level $\ell + 1$ is projected at level ℓ through $[\Phi_1^\ell]^{-1} \circ \Phi_1^{\ell+1}$ (resp. $[\Phi_2^\ell]^{-1} \circ \Phi_2^{\ell+1}$) and used as an initial configuration at that level ℓ.

The optimal value of the discontinuity weights is given by:

$$\forall i,\ \forall \langle s,r\rangle \subset \mathcal{R}_i,\ \beta_{sr} = \frac{1}{\tau_2}\phi_2'\left[\left\|(\boldsymbol{w}_s^\ell + \mathsf{d}\boldsymbol{w}_n^\ell) - (\boldsymbol{w}_r^\ell + \mathsf{d}\boldsymbol{w}_m^\ell)\right\|^2\right], \tag{27}$$

$$\forall \langle i,j\rangle,\ \forall \langle s,r\rangle \in \partial_{ij},\ \beta_{sr} = \frac{1}{\tau_2}\phi_2'\left[\left\|(\boldsymbol{w}_s^\ell + \mathsf{d}\boldsymbol{w}_n^\ell)\ - (\boldsymbol{w}_r^\ell + \mathsf{d}\boldsymbol{w}_m^\ell)\right\|^2 + \frac{\mu_1}{\tau_2|\partial_{ij}|}\right]. \tag{28}$$

For a site pair in between two neighboring regions, (28) implies that the optimal value is decreased (since ϕ_2' is decreasing) compared to the "segmentation-free" case encountered in section II. The compound energy favors low discontinuity weights along the border of current segmentation.

The data weights δ are only involved in $\mathcal{H}_1$, therefore the update rule directly stems from (14) with piece-wise constant parameterization:

$$\delta_s = \frac{1}{\tau_1}\phi_1'\left[\left(\boldsymbol{\nabla} f(s+\boldsymbol{w}_s^\ell, t+1)^T \mathsf{d}\boldsymbol{w}_n^\ell + f_t(s, \boldsymbol{w}_s^\ell)\right)^2\right]. \tag{29}$$

Considering now that the weights β and δ are frozen, the energy function $\mathbb{H}^\ell$ is quadratic with respect to $\mathsf{d}\boldsymbol{w}^\ell$. Its minimization is equivalent to the resolution of a linear system which is very similar to the one obtained with the hierarchical (with constant model) estimation of motion alone. In case of Gauss-Seidel iteration, the only change comes from the influence of parametric field to which dense increment field is related within $\mathcal{E}_{\text{interact}}$ (see Appendix B).

Minimization of $\mathbb{H}^\ell$ w.r.t. the unknown partition $\mathcal{R}^\ell$ and the parametric likeness weights η, is conducted in the same alternate minimization spirit. First the partition is fixed, and the weights η_s's and the motion parameters $\boldsymbol{\theta}_i$'s are estimated using iterated reweighted least squares. For a given region $\mathcal{R}_i^\ell \in \mathcal{R}^\ell$, the update of the motion parameter vector results from least squares regression

$$\boldsymbol{\theta}_i = \left[\sum_{n\in\mathcal{R}_i^\ell}\sum_{s\in\mathcal{B}_n^\ell} \eta_s P(s)^T P(s)\right]^{-1} \sum_{n\in\mathcal{R}_i^\ell}\sum_{s\in\mathcal{B}_n^\ell} \eta_s P_s^T(\boldsymbol{w}_s^\ell + \mathsf{d}\boldsymbol{w}_n^\ell), \tag{30}$$

while the parametric likeness weights are updated according to:

$$\forall n \in \mathcal{R}_i^\ell,\ \forall s \in \mathcal{B}_n^\ell,\ \eta_s = \frac{1}{\tau_3}\phi_3'(\|\boldsymbol{w}_s^\ell + \mathsf{d}\boldsymbol{w}_n^\ell - \boldsymbol{w}_s^i\|^2). \tag{31}$$

Afterwards, the partition $\mathcal{R}^\ell$ is updated using both local and global deformations. Local updates consist in moving each point of the border $\partial\mathcal{R}^\ell$ within a small neighborhood in order to lower the

associated energy terms. Let $\mathcal{R}'^{\ell}$ be the candidate new segmentation (see Fig. 3). Assuming that the region-wise motion parameters $\boldsymbol{\theta}_i$'s and the adjacency graph of the partition remain the same in this local deformation, the associated energy variation is:

$$\mathbb{H}^{\ell}(\mathcal{R}'^{\ell}) - \mathbb{H}^{\ell}(\mathcal{R}^{\ell}) = \lambda 2^{\ell}(|\partial\mathcal{R}'^{\ell}| - |\partial\mathcal{R}^{\ell}|) + \sum_{\langle i,j\rangle} \left[\frac{\mu_1}{|\partial_{ij}^{\ell}|} \sum_{\langle s,r\rangle \in \partial_{ij}^{\ell}} \beta_{sr} - \frac{\mu_1}{|\partial'^{\ell}_{ij}|} \sum_{\langle s,r\rangle \in \partial'^{\ell}_{ij}} \beta_{sr} \right] + \sum_i \left[\sum_{n \in \mathcal{R}'_i \setminus \mathcal{R}_i} \sum_{s \in \mathcal{B}_n^{\ell}} \rho_3(\|\boldsymbol{w}_s^{\ell} + \mathsf{d}\boldsymbol{w}_n^{\ell} - \boldsymbol{w}_s^i\|) - \sum_{n \in \mathcal{R}_i \setminus \mathcal{R}'_i} \sum_{s \in \mathcal{B}_n^{\ell}} \rho_3(\|\boldsymbol{w}_s^{\ell} + \mathsf{d}\boldsymbol{w}_n^{\ell} - \boldsymbol{w}_s^i\|) \right]. \tag{32}$$

This local energy variation is easily computed. In practice, a new position is considered for each border element of the current partition $\mathcal{R}^{\ell}$. If this position corresponds to an energy decrease it is accepted and the partition is updated. In our experiments, a border element is allowed to move one site forward or backward in the direction perpendicular to the border. Let us note that these displacements may be quite large since they actually correspond to 2^{ℓ} pixels. Despite this, motion parameters are not re-estimated when evaluating energy variations in order to keep computational load reasonable.

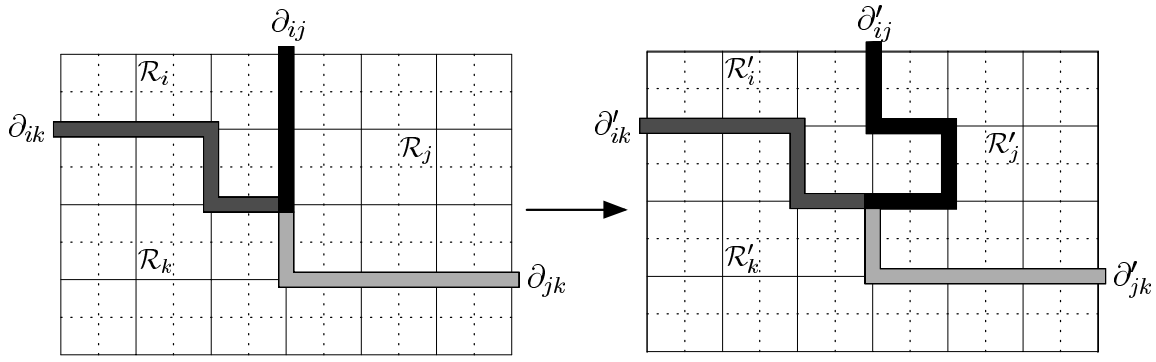

Fig. 3. Example of local deformation of a partition $\mathcal{R} = \Phi_2^{\ell}(\mathcal{R}^{\ell})$ into $\mathcal{R}' = \Phi_2^{\ell}(\mathcal{R}'^{\ell})$, with $\ell = 2$.

Beside local deformations, global updates allow to change at once a whole region as well as the topology of the partition (number and connectivity of regions). In this work we only consider global transformations based on creation of new regions and merging of adjacent regions.

The merging of two adjacent regions consists in removing their common boundary, when this yields a global energy decrease. $\mathcal{R}^{\ell}$ being the current block partition, the energy variation associated with the merging of two regions i and j is derived as for local deformation (except that one has to compute the new parametric model associated with $\mathcal{R}_i^{\ell} \cup \mathcal{R}_j^{\ell}$ to determine the actual energy of potential new

partition). This is done for each pair of adjacent regions. The boundary leading to the greatest energy decrease is removed and the corresponding regions are merged. This process is repeated until a complete stability is reached.

The inclusion of a new region is based on a classification of the parametric likeness weights $\eta = \{\eta_s, s \in S\}$ within two classes. The first class gathers the *outliers* to the parametric likeness model whereas the second one groups points where the affine model fits well the non-parametric vector field. A classic model for supervised image classification is specified through the following energy minimization problem:

$$\hat{x}^{\,\ell} \in \{0,1\}^{|S^\ell|} = \arg\min_{x^\ell}\{\sum_{n\in S^\ell}\sum_{s\in\mathcal{B}_n^\ell}\frac{1}{2\sigma^2(x_s^\ell)}\Big[m(x_s^\ell)-\eta_s\Big]^2+\log\sigma(x_s^\ell)+\sum_{\langle n,m\rangle}\mathbf{1}(x_n^\ell,x_m^\ell)\}, \qquad (33)$$

where $\mathbf{1}(x,y)=1$ if $x=y$ and $\mathbf{1}(x,y)=0$ otherwise. The parameters $\sigma(0)$ and $m(0)$ (respectively $\sigma(1)$ and $m(1)$) stand for the standard deviation and the mean of the outlier class (respectively the inlier class). In all experiments these parameters were set to $m(0)=0.2$, $m(1)=0.98$, and $\sigma(0)=\sigma(1)=0.1$. The minimizer is searched with deterministic ICM algorithm [5]. If a sufficiently large region of connected outliers is eventually recovered, it is incorporated as a new region in the partition.

Global deformations obviously involve far more expensive computations than local deformations. In practice, we only use them at the beginning of each level ℓ. Recall that the whole procedure is multiresolution (based on a pyramid of images), and hierarchical within each resolution. As concerns the initialization at the coarser resolution, the initial motion field is set to zero and associated with a partition composed of a unique region (the whole image).

C. *Continuous point of view*

Contrary to what we did for motion estimation, it is not easy to express the estimation-segmentation model within continuous formalism. The reason is twofold.

First, it is a complicated issue just to specify what a partition of the image plane is, in this framework. *General* definition implies the use of an unknown number of pieces of Jordan curves, connected at junction points [25]. This seems very delicate to be manipulated in practice. Hence, there is unfortunately no continuous counterpart to the simple definition of a discrete partition by a labeling of the discrete pixel grid into different connected components.

However, continuous formalization of the segmentation has been extensively used in the *particular* cases where only independent closed regions have to be extracted from the "background": the border of each region can then be defined either explicitly as a closed curve (as in "snakes" and "active contours" approaches [21]), or implicitly (as in level set techniques [28]). In these specific situations, continuous formalism constitutes an alternative to discrete models, and the interaction mechanism we introduced between segmentation border and discontinuity weights can be translated. Let $\gamma : [0,1] \rightarrow \mathcal{S}$ be a closed curve indexed by u, and $\gamma^\bullet$ the region of $\mathcal{S}$ with border γ. An interaction term analog to (22) could be:

$$\mu_1 \int_0^1 \beta(\gamma(u))du + \mu_2 \iint_{\gamma^\bullet} [\tau_3 \eta(\boldsymbol{x}) \| \boldsymbol{w}(\boldsymbol{x}) + \mathsf{d}\boldsymbol{w}(\boldsymbol{x}) - \boldsymbol{w}^i(\boldsymbol{x}) \|^2 + \psi_3(\eta(\boldsymbol{x}))] d\boldsymbol{x}.$$

Euler-Lagrange equation of the compound functional provides conditions on functions β and η which are the continuous counterpart of discrete update rules (27-28) and (31). Besides, with standard smoothing priors, evolution of γ is the one of a "snake" driven by external energy $\mu_1 \int_0^1 \beta(\gamma(u))du$. The interaction mechanism based on discontinuity auxiliary variables we propose could as well be considered in case region borders are implicitly defined in terms of level sets. This mechanism thus offers a new way to deal with generic issue of joint anisotropic diffusion and extraction of regions, in both discrete and continuous frameworks.

The second reason for which it seemed more natural to let our estimation-segmentation approach within the discrete realm is related to the choice of the minimization technique. Using partitions of the discrete pixel grid enables us to extend in a very simple way the efficient hierarchical minimization scheme previously introduced for the estimation alone.

IV. Experimental results

In this section we present results obtained for the dense optic flow estimator and for the joint estimation-segmentation approach. The first approach has been applied to sequences that seem not to admit any simple region-wise parametric description. The compromise between local dense representation and more global parametric one offered by the hierarchical minimization allows to get interesting results on fluid sequences for instance. At the opposite, the parametric description used in the joint estimation-segmentation approach is more appropriate to scenes involving rigid objects.

Note that in both approaches we selected Leclerc estimator for the different robust functions ($\rho(x) \triangleq 1 - \exp(-\tau x^2)$) [22].

A. *Results on optic flow estimation*

The experiments have been carried out both on synthetic image – for which a ground truth is known and comparative quantitative comparisons can be reported [3] – and on real world sequences. The synthetic sequence is the well known Yosemite sequence used in comparative benchmark of Barron *et al.* [3]. The two other sequences for which results are presented here involve fluid phenomena. The first one, named Smoke, shows smoke diffusing in front of a camera under tough lighting conditions (Fig. 6a). The second one, named Depression (Fig. 4a), is a meteorological video sequence involving large displacements. It includes a through of low pressure and different moving clouds. Parameter values used for the three sequences are reported in Table I.

As for hierarchical constrained minimization, both *regular* and *adaptative* partitions into square patches have been considered. In regular case, partition $\mathcal{B}^\ell$ is constituted of $2^\ell \times 2^\ell$ square blocks and associated adjacency graph $G^\ell = [S^\ell, \nu^\ell]$ is a regular lattice with same neighborhood system as the original grid $G = [S, \nu]$. In adaptative case, partition $\mathcal{B}^{\ell-1}$ is determined on-line, based on the previous partition $\mathcal{B}^\ell$ and on the associated final estimate $(\widehat{\mathsf{d}\boldsymbol{w}}^\ell, \widehat{\beta}^\ell, \widehat{\delta}^\ell)$. It is obtained by dividing some of the elements of $\mathcal{B}^\ell$ according to a certain criterion. It seems natural to base this division criterion upon the agreement of the current motion estimate with the luminance conservation assumption, measured on block $\mathcal{B}_n^\ell$ by $\sum_{s\in\mathcal{B}_n^\ell}[f(s+\boldsymbol{w}_s^\ell+\widehat{\mathsf{d}\boldsymbol{w}}_s^\ell, t+1) - f(s,t)]^2$ (sum of squared registration errors), or, in linearized form: $\sum_{s\in\mathcal{B}_n^\ell}[\boldsymbol{\nabla} f(s+\boldsymbol{w}_s^\ell, t+1)^T \widehat{\mathsf{d}\boldsymbol{w}}_s^\ell + f_t(s, \boldsymbol{w}_s^\ell)]^2$, where $\widehat{\mathsf{d}\boldsymbol{w}}_s^\ell = P_n(s)\widehat{\boldsymbol{\theta}}_n^\ell$. Instead of using this quantity which has to be computed, we use final data weights $\hat{\delta}_s$'s (which are function of the squares in previous sum, according to (14).). Experimental evidences indicated that it is more appropriate to consider how uniform (instead of how good) is the quality of the agreement within considered patch. A block is thus divided into four sub-blocks if the standard deviation of $\{\hat{\delta}_s^\ell, s \in \mathcal{B}_n^\ell\}$ exceeds a given threshold. In [12] another subdivision strategy is proposed (within a finite elements method), based on the so-called normal flow $-\frac{f_t}{\|\boldsymbol{\nabla} f\|}$. It thus does not depend on the current estimate and can be computed *a priori* from the data. In between latter work and what we propose, an on-line adaptative strategy based on *residual* normal flow $\frac{f(s+\boldsymbol{w}_s^\ell+\widehat{\mathsf{d}\boldsymbol{w}}_s^\ell, t+1)-f(s,t)}{\|\boldsymbol{\nabla} f(s+\boldsymbol{w}_s^\ell, t+1)\|}$ is proposed in [32].

As mentioned in §II.D, three different parameterizations (corresponding respectively to 2, 4, and 6 parameters) are considered. We use them within six different combinations denoted M_6, M_4, M_2, M_{64}, M_{62}, and M_{642}, where subscripts indicate allowed parameterizations. Models M_6, M_4 and M_2 deal with a single type of parameterization. In these three cases the hierarchical procedure is stopped when a certain minimal size (8×8, 4×4, and 1×1 resp.) is reached by the smallest patches of the current

TABLE I
PARAMETER VALUES FOR OPTIC FLOW ESTIMATION EXPERIMENTS.

	Yosemite	Smoke	Depression
number of resolution levels	2	2	3
patch size of the coarser partition	16	32	16
smoothing parameter α	300	300	200
parameters τ_1 tuning ρ_1	0.03	0.03	0.02
parameters τ_2 tuning ρ_2	2	0.25	4

TABLE II
RESULTS ON YOSEMITE.

	Regular partition			Adaptative partition		
Model	μ	σ	cpu	μ	σ	cpu
M_6	5.07^o	7.20^o	34s	5.09^o	7.20^o	28s
M_4	5.56^o	9.20^o	72s	6.57^o	9.20^o	58s
M_2	4.97^o	7.64^o	64s	6.53^o	7.64^o	42s
M_{64}	5.01^o	7.23^o	58s	5.17^o	7.62^o	44s
M_{62}	4.69^o	6.89^o	78s	5.25^o	7.87^o	72s
M_{642}	4.75^o	6.89^o	92s	5.31^o	7.86^o	73s

partition. In contrast, models M_{64}, M_{62}, and M_{642} mix different parameterizations in the following way: in the three cases affine model is used for each patch at least as large as 8×8. In M_{64} and M_{642}, patches of size 4×4 are equipped with the simplified affine model (and no smaller patches are considered in M_{64}). In M_{62} (resp. M_{642}) constant model is used for blocks of size 4×4 (resp. 2×2) and less.

Following [3], 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; it is 100% in our case). Table II lists, for the different versions of our model, the mean angular error (μ) and the associated standard deviation (σ). The CPU times, measured on a SUN ULTRA SPARC (200 Mhz), are also reported in there.

In comparison, Table III recalls some of the results presented by Barron *et al.* (see corresponding references therein). They concern an adaptation of Horn and Schunck's algorithm, 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).[7] The results obtained by the *multigrid* method we introduced in [24] are also included.

[7]Results on a sequence where the sky is removed are also reported by other authors [6, 20, 24]. We believe that the resulting motion is probably too simple to allow significant comparisons between the different state-of-the-art motion estimators.

TABLE III
COMPARATIVE RESULTS ON YOSEMITE

Technique	μ	σ	*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%
Mémin and Pérez [24]	5.37^o	8.19^o	100%

On Yosemite sequence hierarchical methods associated with regular grids provide a *dense* estimate almost as good as those obtained with the best (non-dense) mentioned methods, and slightly better that the ones obtained with multigrid minimization. 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 (M_{642}) or when the six-parameter affine model is coupled with the constant model (M_{62}). The M_6 model on adaptative grids gives the lowest computation times. Compared to the other models on irregular grids, it also yields the best results.

Based on these comparisons, multi-parametric models do not seem of much interest compared to mono-parametric ones, at least regarding the average angular discrepancy criterion. We shall see that it is actually of interest in real cases. Note also that the *global* criterion of the final estimate quality should be considered with caution. It is for instance not able to assess the ability of a method to preserve discontinuities of the motion field.

Note that in the case of M_2 associated with a regular subdivision, results are slightly improved compared to the pure top-down multigrid method [24]. Hierarchical Gauss-Newton with its successive inter-level warping seems to perform better on this particular example.

To complete our comparisons, we show results obtained on real world sequences involving fluid phenomena. Figure 4 presents for Depression the final motion fields respectively estimated by M_2 with regular division and M_6 with the adaptative division. The two vector fields are displayed the same way, namely subsampled by 6 and magnified by 4. We can notice that with the constant constraint, the flow is drastically under-estimated and over-smoothed compared to the one produced with the affine constraint. As a consequence on Depression, local features of interest such as the depression center in the 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.

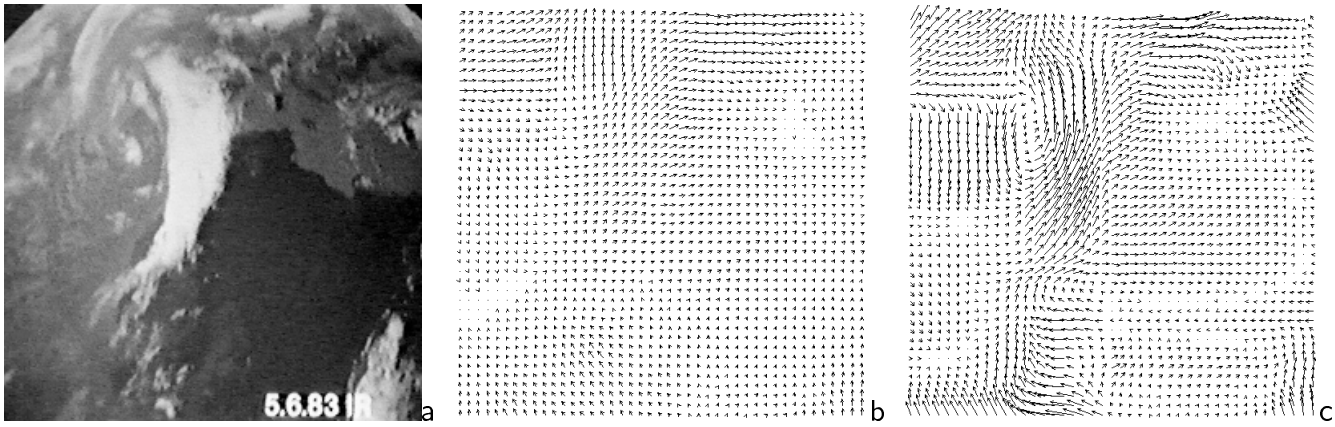

Fig. 4. Results on Depression (256×256): (a) one frame, (b) motion field estimated with M_2 and regular partitioning (cpu time: 105s), (c) motion field estimated with M_6 and adaptative partitioning (cpu time: 24s).

Smoke sequence constitutes a difficult example with large motion amplitude (up to 20 pixels) and low photometric gradients. The various versions of our model have provided quite different results (see M_2, M_6 and M_{62} estimates in Fig. 5). Visually, estimates obtained with M_{62} seemed the more compliant with the apparent motion in the scene. This was confirmed by reconstruction errors $|f(s+\widehat{\boldsymbol{w}}_s, t+1) - f(s,t)|$ (see samples in Fig. 5). Means of this squared error are respectively 75, 98, and 50 for M_2, M_6, and M_{62}, for error images showed in Fig. 5: in this example M_{62} out-performs M_6 and M_2 within as much cpu time as M_2 (around 700s for these 512×512 images). As expected, M_6, which does not use blocks smaller than 8×8, is much faster (around 200s), but provides poorest results. The final estimation partition obtained with M_{62} for one image of the sequence is given in Fig. 6b along with four consecutive motion fields estimated from the sequence.

Finally, note the whole multiresolution/hierarchical algorithm converges quickly, since only ten or so low cost iterations are required at each grid level. Furthermore, as demonstrated in [24] for the constant model associated with a multigrid minimization, the resulting motion estimator exhibits a low sensitiveness to parameter values.

B. *Results on joint estimation-segmentation*

In the case of the joint estimation-segmentation approach, we report experiments on Yosemite synthetic sequence (for comparison purpose) and on two real-world sequences. These two sequences involve rigid moving objects. The first sequence is a Parking lot sequence which involves two cars moving in the foreground while the camera pans the scene. (Fig. 8). The second one, named Calendar includes several

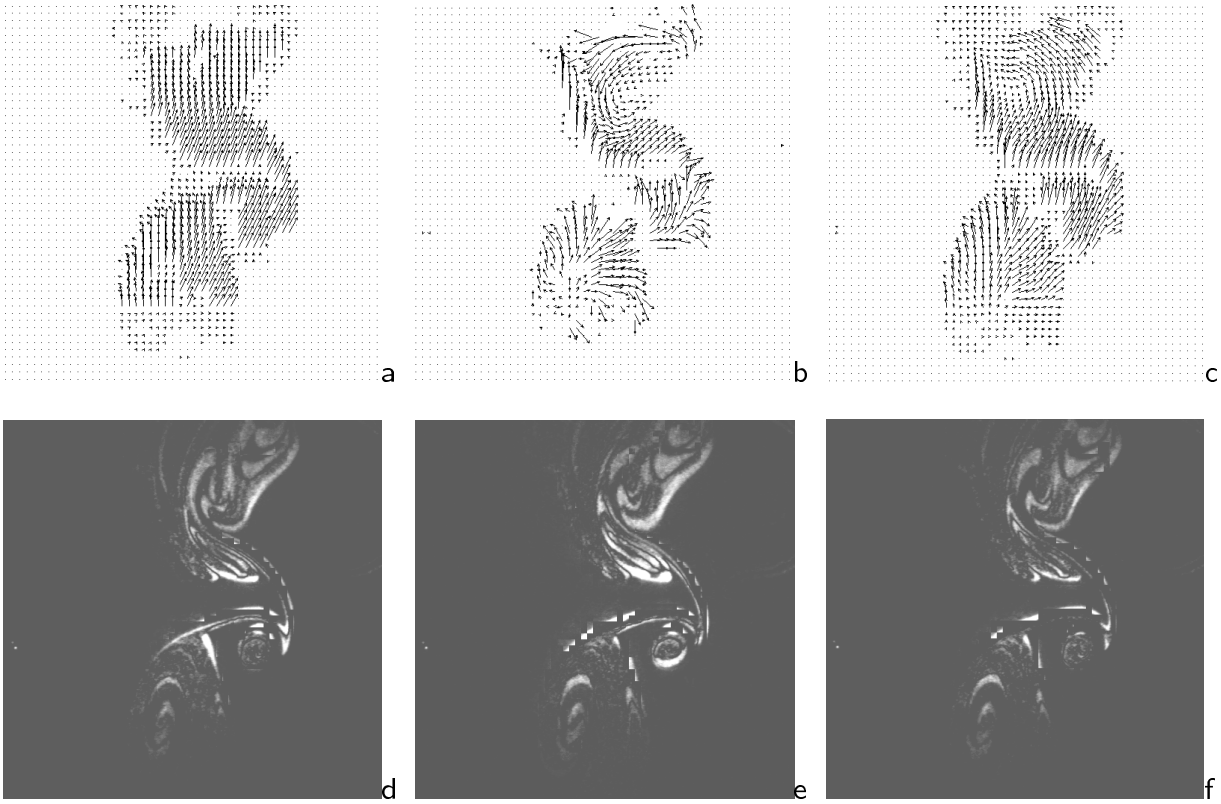

Fig. 5. Comparison of M_2, M_6 and M_{62} on Smoke (512×512): (a-b-c) final flow recovered with the adaptative M_2, M_6 and M_{62}, (d-e-f) corresponding reconstruction error images.

moving objects (a calendar moving vertically and a toy train pushing a ball) and a horizontal panning of the camera (Fig. 9).

As for the parameter values, the number of resolution levels was respectively 2 for Calendar and Yosemite and 1 for Parking lot. The number of grid levels was fixed to 6 for Calendar and to 5 for the two others. The other parameters were tuned according to the amplitude of motion present in each sequence: $\alpha = 100$, $\tau_1 = 0.02$, $\mu_1 = 50$, and $\lambda = 5$ for the three sequences; whereas $\tau_2 = 3$, $\tau_3 = 3$ and $\mu_2 = 25$ for Calendar and Yosemite, they are $\tau_2 = 10$, $\tau_3 = 5$, and $\mu_2 = 5$ for Parking lot.

In Table IV results of the dense estimation alone (for hierarchical Gauss-Newton version, and multigrid version, resp.) are compared with those obtained on Yosemite within joint estimation-segmentation. Both parametric and dense motion fields provided by the latter method slightly improve mean angular error, while providing compact information on the scene within the final motion-based partition of the image plane (Fig. 7) and the associated parameter vectors.

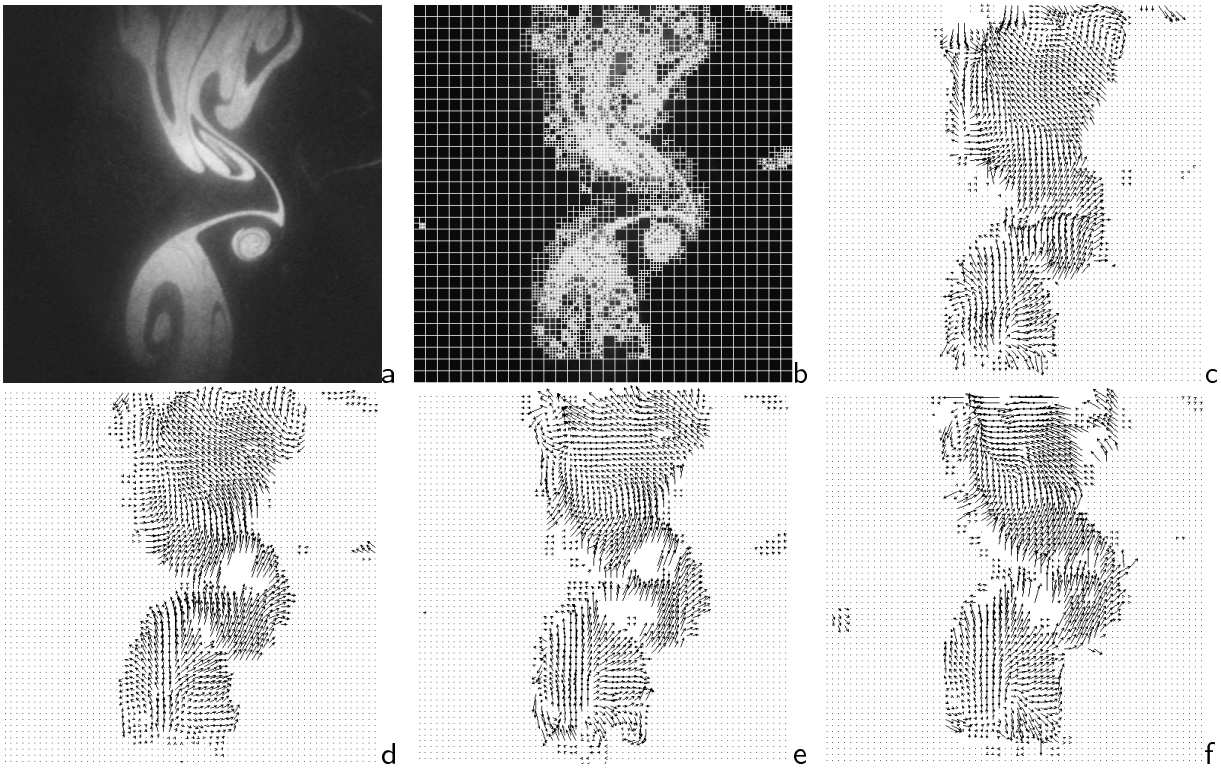

Fig. 6. Results on Smoke (512×512): (a) one frame, (b) final constraint partition, (c-d-e-f) four consecutive motion fields estimated with M_{62} along with adaptative partitioning.

TABLE IV
RESULTS ON YOSEMITE

Model	μ	σ
Estimation-segmentation		
parametric estimate	4.54^o	8.37^o
dense estimate	4.91^o	8.46^o
Dense estimation only		
hierarchical Gauss-Newton	4.97^o	7.64^o
multigrid [24]	5.37^o	8.19^o

The parametric field is actually a good approximation of the dense estimated flow field. This nice behavior of the method is confirmed by the results obtained on Calendar and Parking lot. For instance, in Calendar some difficult regions (such as the roof of the wagon or the space between the locomotive and the wagon) are fairly well recovered (Fig. 9). These remarks hold also in the Parking lot case where the method retrieves an interesting "structural approximation" of the front car (Fig. 8). Otherwise, as may be seen on the three sequences, the method has the ability to fit correctly the motion discontinuities.

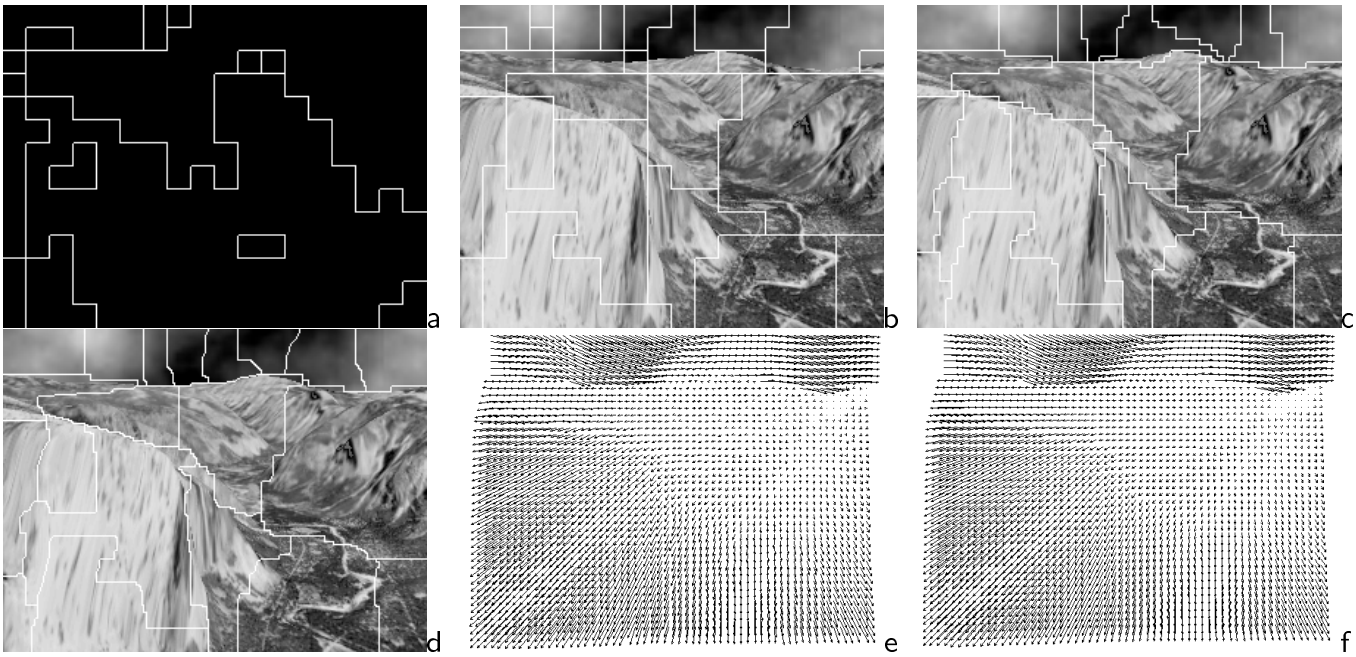

Fig. 7. Results on Yosemite (224×288): (a) segmentation initialization at the coarsest level $\ell = 4$, (b-c-d) final partitions ($\ell = 4, 2, 0$), (e) dense optic flow, (f) parametric flow (cpu time $\sim$ 8mn).

See for example the mountain crest in Yosemite, the front car in Parking lot, and the engine and the rolling ball in Calendar.

We should also outline that the method is not very sensitive to the initialization. More precisely, the method is able to recover meaningful partitions of the flow field from quite "distant" initializations (see for instance Fig. 9 and Fig. 7).

Rough estimates of the computation times (code not hand-optimized) obtained on a Sun Ultra Sparc are also given in the captions of figures 7, 8 and 9.

V. Conclusion

In this paper, we have presented a comprehensive energy-based framework for the incremental estimation, and the segmentation of the optic flow. Using robust cost functions, a dense discontinuity-preserving motion estimator has first been introduced, and a special care has been dedicated to its algorithmic implementation: a hierarchical constrained minimization framework is proposed which allows to mix different parameterizations with respect to a regular or an adaptative partitioning of the image. The ability of resulting estimator to recover intricate non rigid motions has been especially demonstrated on sequences involving moving fluids. For situations where motion-based segmentation of the sequence

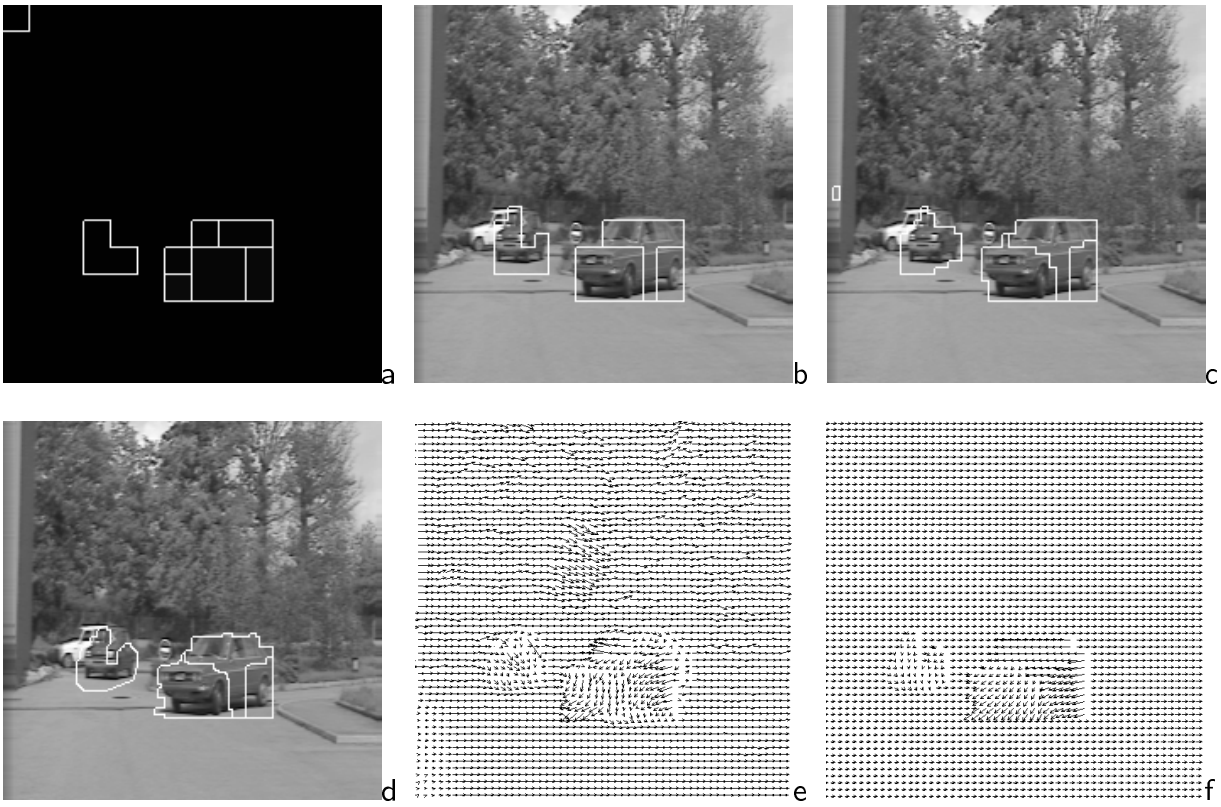

Fig. 8. Results on Parking lot (224×224): (a) segmentation initialization at level $\ell = 4$, (b-c-d) final segmentations ($\ell = 3, 2, 0$), (e) dense optic flow, (f) parametric flow (cpu time $\sim$ 4mn).

makes sense and is of interest, the previous model has been extended to simultaneously handle both issues. Dense estimation and region-wise parametric representation of same flow can thus be jointly recovered in an alternate cooperative way within hierarchical minimization. Of particular interest here, we propose a simple mechanism of interaction between a dense discontinuity-preserving estimator and a segmentation process, through auxiliary variables appearing in the half-quadratic formulation of robust cost functions. This mechanism could be used elsewhere (e.g., simultaneous restoration-segmentation of still images), and could probably be considered (and theoretically studied) from the anisotropic diffusion continuous point of view.

As for the dense estimation alone, we now investigate the introduction in our generic multi-parametric framework of representations which would be explicitly devoted to fluid dynamics. Concerning the joint estimation-segmentation issue, more sophisticated or complete interaction mechanisms should be investigated (e.g., to take into account photometric discontinuities as a useful information, or to handle explicitly the problem of occlusions at borders of motion regions).

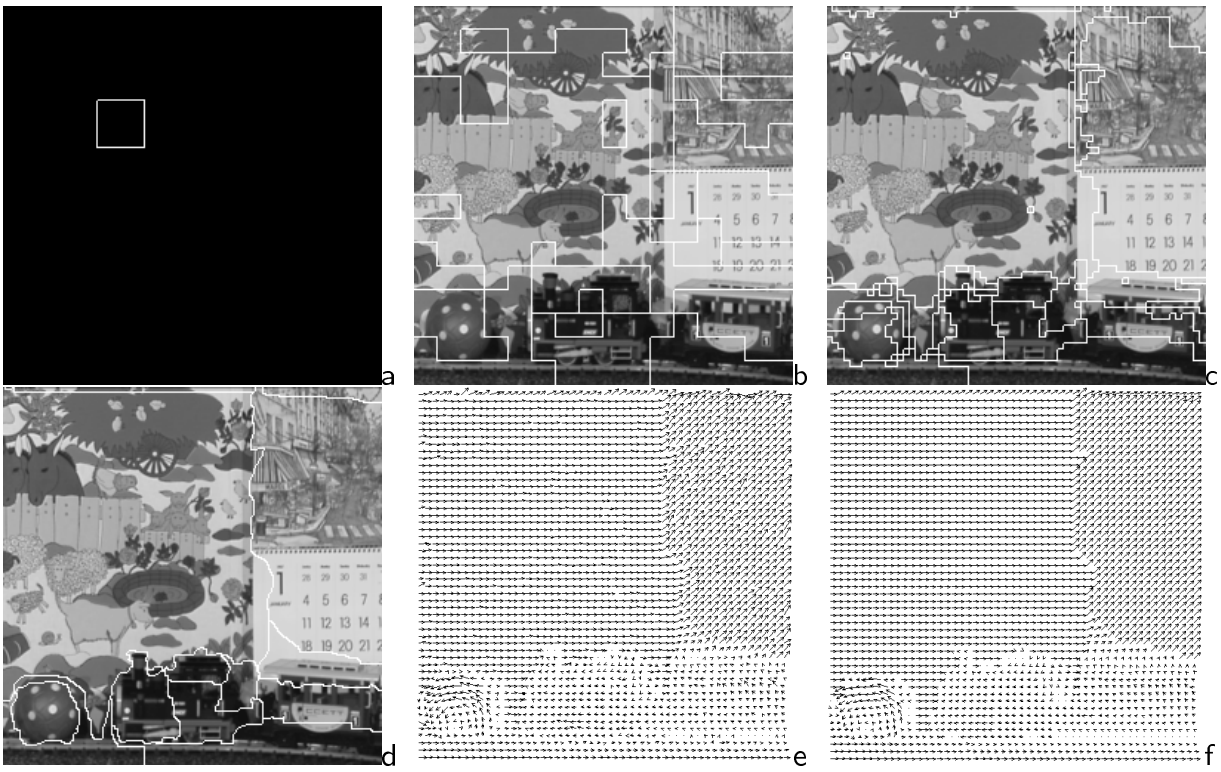

Fig. 9. Results on Calendar (256×256): (a) partition initialization at level $\ell = 4$, (b-c-d) final segmentations ($\ell = 4, 2, 0$), (e) dense optic flow, (f) parametric flow. (cpu time $\sim$ 8mn).

References

[1] G. Adiv. Determining three-dimensional motion and structure from optical flow generated by several moving objects. *IEEE Trans. Pattern Anal. Machine Intell.*, 7:384–401, 1985.

[2] S. Ayer and H.S. Sawhney. Layered representation of motion video using robus t maximum-likelihood estimation of mixture models and Mdl encoding. In *Proc. Int. Conf. Computer Vision*, pages 777–784, Cambridge, June 1995.

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

[4] J. Bergen, P. Anandan, K. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In G. Sandini, editor, *Proc. Europ. Conf. Computer Vision*, volume 558 of *LNCS*, pages 237–252. Springer-Verlag, 1992.

[5] J. Besag. On the statistical analysis of dirty pictures. *J. Royal Statist. Soc.*, B 48(3):259–302, 1986.

[6] M. Black and P. Anadan. The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):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.*, 18(10):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. Computer Vision*, 19(1):75–104, 1996.

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

[10] M. M. Chang, A. M. Tekalp, and M. I. Sezan. Simultaneous motion estimation and segmentation. *IEEE Trans. Image Processing*, 6(9):1326–1333, 1997.

[11] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Processing*, 6(2):298–311, 1997.

[12] I. Cohen and I. Herlin. Optical flow and phase portrait methods for environmental satellite image sequences. In *Proc. Europ. Conf. Computer Vision*, pages II:141–150, Cambridge, UK, April 1996.

[13] A. Delaney and Y. Bresler. Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography. *IEEE Trans. Image Processing*, 7(2):204–221, 1998.

[14] R. Deriche, P. Kornprobst, and G. Aubert. Optical flow estimation while preserving its discontinuities: a variational approach. In *Proc. Asian Conf. Computer Vision*, volume 1, pages 290–295, Singapore, December 1995.

[15] W. Enkelmann. Investigation of multigrid algorithms for the estimation of optical flow fields in image sequences. *Comput. Vision, Graphics, Image Proc.*, 43:150–177, 1988.

[16] D. Geman and G. Reynolds. Constrained restoration and the recovery of discontinuities. *IEEE Trans. Pattern Anal. Machine Intell.*, 14(3):367–383, 1992.

[17] W. Hackbusch. *Multi-Grid Methods and Applications.* Springer-Verlag, Berlin, 1985.

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

[19] Y. Huang, K. Palaniappan, X. Zhuand, and J. Cavanaugh. Optic flow field segmentation and motion estimation using a robust genetic partitioning algorithm. *IEEE Trans. Pattern Anal. Machine Intell.*, 17(12):1177–1190, 1995.

[20] X. Ju, M.J. Black, and A.D. Jepson. Skin and bones: Multi-layer, locally affine, optical flow and regularization with transparency. In *Proc. Conf. Comp. Vision Pattern Rec.*, pages 307–314, 1996.

[21] M. Kass, A. Witkin, and D. Terzopoulos. Snakes : active contour models. *Int. J. Computer Vision*, 1(4):321–331, 1988.

[22] Y. Leclerc. Constructing simple stable descriptions for image partitioning. *Int. J. Computer Vision*, 3:73–102, 1989.

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

[24] E. Mémin and P. Pérez. Dense estimation and object-based segmentation of the optical flow with robust techniques. *IEEE Trans. Image Processing*, 7(5):703–719, 1998.

[25] D. Mumford and J. Shah. Optimal approximation by piecewise smooth functions and associated variational problems. *Comm. Pure and Appl. Math.*, 42:577–685, 1989.

[26] D.W. Murray and H. Buxton. Scene segmentation from visual motion using global optimization. *IEEE Trans. Pattern Anal. Machine Intell.*, 9(2):220–228, 1987.

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

[28] S. Osher and J. Sethian. Fronts propagating with curvature dependent speed: algorithms based on Hamilton-Jacobi formulations. *J. Computational Physics*, 79:12–49, 1988.

[29] C. Schnorr, R. Sprengel, and B. Neumann. A variational approach to the design of early vision algorithms. *Computing Suppl.*, pages 149–165, 1996.

[30] B.G. Schunck. The image flow constraint equation. *Comput. Vision, Graphics, Image Proc.*, 35:20–46, 1986.

[31] C. Stiller. Object-oriented estimation of dense motion fields. *IEEE Trans. Image Processing*, 6(2):234–250, 1997.

[32] R. Szeliski and H.-Y. Shum. Motion estimation with quadtree splines. *IEEE Trans. Pattern Anal. Machine Intell.*, 18(12):1199–1210, 1996.

[33] R. Wildes, M. Amabile, A.M. Lanzillotto, and T.S. Leu. Physically based fluid flow recovery from image sequences. In *Proc. Conf. Comp. Vision Pattern Rec.*, pages 969–975, Puerto Rico, June 1997.

A Parametric Gauss-Seidel iteration

For the sake of concision, we shall denote $\nabla \tilde{f}(s) \triangleq \nabla f(s + \boldsymbol{w}_s^\ell, t+1)$ the spatial gradient in the second image, displaced according to $\boldsymbol{w}^\ell$, and $\tilde{f}_t(s) = f_t(s, \boldsymbol{w}_s^\ell)$ the displaced frame difference. S is partitioned according to $\mathcal{B}^\ell = \{\mathcal{B}_1^\ell \dots \mathcal{B}_{N_\ell}^\ell\}$, the associated adjacency graph being $G^\ell = [S^\ell, \nu^\ell]$. Let $\mathcal{B}_n^\ell$ be the current block in the iterative visit process implied by Gauss-Seidel method. One has simply to minimize $\mathcal{H}^\ell$ with respect to $\boldsymbol{\theta}_n^\ell$, the total field outside $\mathcal{B}_n^\ell$ being frozen. The fraction of energy actually concerned is:

$$\begin{aligned}\mathcal{H}_n^\ell(\boldsymbol{\theta}_n^\ell, \delta, \beta; f, \boldsymbol{w}) \stackrel{\triangle}{=} & \tau_1 \sum_{s \in \mathcal{B}_n^\ell} \delta_s \left[\nabla \tilde{f}(s)^T P_n(s) \boldsymbol{\theta}_n^\ell + \tilde{f}_t(s)\right]^2 \\ & + \alpha\tau_2 \sum_{\langle s,r\rangle \in \mathcal{C}_{\partial n}^\ell} \beta_{sr} \left\| \boldsymbol{w}_s^\ell + P_n(s)\boldsymbol{\theta}_n^\ell - \boldsymbol{w}_r \right\|^2 \\ & + \alpha\tau_2 \sum_{\langle s,r\rangle \in \mathcal{C}_n^\ell} \beta_{sr} \left\| (\boldsymbol{w}_s^\ell + P_n(s)\boldsymbol{\theta}_n^\ell) - (\boldsymbol{w}_r^\ell + P_n(r)\boldsymbol{\theta}_n^\ell) \right\|^2,\end{aligned} \tag{34}$$

where $\mathcal{C}_{\partial n}^\ell \stackrel{\triangle}{=} \cup_{m \in \nu^\ell(n)} \mathcal{C}_{nm}^\ell$ is the set of cliques of $\mathcal{C}$ straddling the border of $\mathcal{B}_n^\ell$. The increment field in the neighborhood of $\mathcal{B}_n^\ell$ is a mix of various parameterizations relative to different parts of the (possibly) irregular grid S^ℓ: for $m \in \nu^\ell(n)$ and $\langle s, r\rangle \in \mathcal{C}_{nm}$, $\mathsf{d}\boldsymbol{w}_r^\ell = P_m(r)\boldsymbol{\theta}_m^\ell$. However, the only thing of actual interest when updating $\boldsymbol{\theta}_n^\ell$ is the total field $\boldsymbol{w}_r \triangleq \boldsymbol{w}_r^\ell + P_m(r)\boldsymbol{\theta}_m^\ell$. As a consequence, in the following computations, the neighboring parameterizations do not appear explicitly in the regularization part of the update. Their are simply hidden within the total field on the neighboring patches. Writing that

the partial derivative of this piece of energy vanishes, one gets:

$$\begin{aligned}\frac{\partial \mathcal{H}_n^\ell(\boldsymbol{\theta}_n^\ell, \beta, \delta; f, \boldsymbol{w})}{\partial \boldsymbol{\theta}_n^\ell} =& \tau_1 \sum_{s \in \mathcal{B}_n^\ell} \delta_s P_n(s)^T \boldsymbol{\nabla} \tilde{f}(s) \left[\boldsymbol{\nabla} \tilde{f}(s)^T P_n(s) \boldsymbol{\theta}_n^\ell + \tilde{f}_t(s) \right] \\ &+ \alpha\, \tau_2 \sum_{\langle s,r \rangle \in \mathcal{C}_{\partial n}^\ell} \beta_{sr} P_n(s)^T \left[\boldsymbol{w}_s^\ell + P_n(s) \boldsymbol{\theta}_n^\ell - \boldsymbol{w}_r \right] \\ &+ \alpha\, \tau_2 \sum_{\langle s,r \rangle \in \mathcal{C}_n^\ell} \beta_{sr} (P_n(s) - P_n(r))^T \left[\boldsymbol{w}_s^\ell - \boldsymbol{w}_r^\ell + (P_n(s) - P_n(r)) \boldsymbol{\theta}_n^\ell \right] = 0.\end{aligned} \tag{35}$$

A compact vectorial formulation of this equation can be achieved by introducing the following matrices and vectors indexed respectively by the sites of block $\mathcal{B}_n^\ell$, the cliques inside the block, and the cliques straddling the border of the block:

$$A_n \triangleq \begin{bmatrix} \vdots \\ \boldsymbol{\nabla} \tilde{f}(s)^T P_n(s) \\ \vdots \end{bmatrix}_{s \in \mathcal{B}_n^\ell}, \quad \boldsymbol{F}_n \triangleq \begin{bmatrix} \vdots \\ \tilde{f}_t(s) \\ \vdots \end{bmatrix}_{s \in \mathcal{B}_n^\ell}, \quad \Delta_n \triangleq \operatorname{diag}(\cdots,\ \delta_s,\ \cdots)_{s \in \mathcal{B}_n^\ell},$$

$$C_n \triangleq \begin{bmatrix} \vdots \\ P_n(s) - P_n(r) \\ \vdots \end{bmatrix}_{\langle s,r \rangle \in \mathcal{C}_n^\ell}, \quad B_n \triangleq \operatorname{diag}(\cdots,\ \beta_{sr} \mathbb{I}_2,\ \cdots)_{\langle s,r \rangle \in \mathcal{C}_n^\ell},$$

$$C_{\partial n} \triangleq \begin{bmatrix} \vdots \\ P_n(s) \\ \vdots \end{bmatrix}_{\langle s,r \rangle \in \mathcal{C}_{\partial n}^\ell}, \quad \text{and } B_{\partial n} \triangleq \operatorname{diag}(\cdots,\ \beta_{sr} \mathbb{I}_2,\ \cdots)_{\langle s,r \rangle \in \mathcal{C}_{\partial m}^\ell},$$

where $\mathbb{I}_2 \triangleq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, as well as the following block-wise and border-wise averages:

$$\overline{\boldsymbol{\theta}}_{\partial n}^\ell \triangleq \frac{1}{b_{\partial n}} \sum_{\langle s,r \rangle \in \mathcal{C}_{\partial n}^\ell} \beta_{sr} P_n(s)^T (\boldsymbol{w}_r - \boldsymbol{w}_s^\ell), \text{ with } b_{\partial n} \triangleq \sum_{\langle s,r \rangle \in \mathcal{C}_{\partial n}^\ell} \beta_{sr}$$

$$\overline{\boldsymbol{\theta}}_n^\ell \triangleq \frac{1}{b_n} \sum_{\langle s,r \rangle \in \mathcal{C}_n^\ell} \beta_{sr} (P_n(s) - P_n(r))^T (\boldsymbol{w}_r^\ell - \boldsymbol{w}_s^\ell), \text{ with } b_n \triangleq \sum_{\langle s,r \rangle \in \mathcal{C}_n^\ell} \beta_{sr}.$$

Linear equation (35) then reads:

$$\left[\tau_1 A_n^T \Delta_n A_n + \alpha \tau_2 C_{\partial n}^T B_{\partial n} C_{\partial n} + \alpha \tau_3 C_n^T B_n C_n \right] \boldsymbol{\theta}_n^\ell = -\tau_1 A_n^T \Delta_n \boldsymbol{F}_n + \alpha \tau_2 b_{\partial n} \overline{\boldsymbol{\theta}}_{\partial n}^\ell + \alpha \tau_3 b_n \overline{\boldsymbol{\theta}}_n^\ell. \tag{36}$$

The direct resolution of this linear system provides the updated value of parameter vector $\boldsymbol{\theta}_n^\ell$. In this equation, matrices A_n, C_n and $C_{\partial n}$, and vectors $\overline{\boldsymbol{\theta}}_n^\ell$ and $\overline{\boldsymbol{\theta}}_{\partial n}^\ell$ depend on the type of parameterization associated with block $\mathcal{B}_n^\ell$. Let us give their expressions (when simplified forms are available) for the three different parameterization.

For *constant model* $P_n \equiv \mathbb{I}_2$, yielding $A_n^T = [\cdots \boldsymbol{\nabla}\tilde{f}(s) \cdots]_{s\in\mathcal{B}_n^\ell}$, $C_{\partial n}^T = [\cdots \mathbb{I}_2 \cdots]$, $C_n = 0$, $\overline{\boldsymbol{\theta}}_n^\ell = 0$, and $\overline{\boldsymbol{\theta}}_{\partial n}^\ell = \frac{1}{b_{\partial n}} \sum_{\langle s,r\rangle\in\mathcal{C}_{\partial n}^\ell} \beta_{sr}(\boldsymbol{w}_r - \boldsymbol{w}_s^\ell)$. Equation (36) simplifies as follows:

$$
\begin{aligned}
(36) \quad &\Leftrightarrow \quad \left(\tau_1 A_n^T \Delta_n A_n + \alpha\tau_2 b_{\partial n}\mathbb{I}_2\right)\boldsymbol{\theta}_n^\ell = -\tau_1 A_n^T \Delta_n \boldsymbol{F}_n + \alpha\tau_2 b_{\partial n}\overline{\boldsymbol{\theta}}_{\partial n}^\ell \\
&\Leftrightarrow \quad \left(\frac{1}{\gamma}A_n^T\Delta_n A_n + \mathbb{I}_2\right)\boldsymbol{\theta}_n^\ell = \overline{\boldsymbol{\theta}}_{\partial n}^\ell - \frac{1}{\gamma}A_n^T\Delta_n\boldsymbol{F}_n, \text{ with } \gamma \triangleq \frac{\alpha\tau_2 b_{\partial n}}{\tau_1} \\
&\Leftrightarrow \quad \boldsymbol{\theta}_n^\ell = \overline{\boldsymbol{\theta}}_{\partial n}^\ell - \frac{\gamma A_n^T\Delta_n(A_n\overline{\boldsymbol{\theta}}_{\partial n}^\ell + \boldsymbol{F}_n) + \det\mathbb{A}_n\overline{\boldsymbol{\theta}}_{\partial n}^\ell + \mathrm{com}\mathbb{A}_n A_n^T\Delta_n\boldsymbol{F}_n}{\gamma(\gamma + \mathrm{trace}\mathbb{A}_n) + \det\mathbb{A}_n}, \qquad (37)
\end{aligned}
$$

with $\mathbb{A}_n \triangleq A_n^T\Delta_n A_n$.

For *simplified affine model* $P_n(s) = [\mathbb{I}_2\ p(s)]$ with $p(s) \triangleq \begin{bmatrix} x_s & y_s \\ y_s & -x_s \end{bmatrix}$, yielding:

$$
\begin{aligned}
C_{\partial n}^T B_{\partial n} C_{\partial n} &= \sum_{\langle s,r\rangle\in\mathcal{C}_{\partial n}^\ell} \beta_{sr}\begin{bmatrix} \mathbb{I}_2 & p(s) \\ p(s) & (x_s^2+y_s^2)\mathbb{I}_2 \end{bmatrix} \\
C_n^T B_n C_n &= b_n\,\mathrm{diag}(0,0,1,1) \\
b_n\overline{\boldsymbol{\theta}}_n^\ell &= \sum_{\langle s,r\rangle\in\mathcal{C}_n^\ell(\vdots)} \beta_{sr}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}(\boldsymbol{w}_r^\ell - \boldsymbol{w}_s^\ell) + \sum_{\langle s,r\rangle\in\mathcal{C}_n^\ell(\text{-}\bullet\text{-})} \beta_{sr}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}(\boldsymbol{w}_r^\ell - \boldsymbol{w}_s^\ell),
\end{aligned}
$$

where $\mathcal{C}_n^\ell(\vdots)$ (resp. $\mathcal{C}_n^\ell(\text{-}\bullet\text{-})$) contains cliques of $\mathcal{C}_n^\ell$ lying along the x-direction (resp. y-direction).

For *affine model* $P_n(s) = \mathbb{I}_2 \otimes \boldsymbol{e}(s)^T$, with $\boldsymbol{e}(s)^T \triangleq [1\ x_s\ y_s]$, yielding the following expressions for the matrices and vectors involved in equation (36):

$$
\begin{aligned}
A_n^T\Delta_n A_n &= \sum_{s\in\mathcal{B}_n^\ell} \delta_s\left(\boldsymbol{\nabla}\tilde{f}(s)\boldsymbol{\nabla}\tilde{f}(s)^T\right)\otimes\left(\boldsymbol{e}(s)\boldsymbol{e}(s)^T\right) \\
C_{\partial n}^T B_{\partial n} C_{\partial n} &= \mathbb{I}_2 \otimes \sum_{\langle s,r\rangle\in\mathcal{C}_{\partial n}^\ell} \beta_{sr}\boldsymbol{e}(s)\boldsymbol{e}(s)^T \\
C_n^T B_n C_n &= \mathbb{I}_2 \otimes \sum_{\langle s,r\rangle\in\mathcal{C}_n^\ell} \beta_{sr}(\boldsymbol{e}(s)-\boldsymbol{e}(r))(\boldsymbol{e}(s)-\boldsymbol{e}(r))^T \\
&= \mathbb{I}_2 \otimes \mathrm{diag}\left(0, \sum_{\langle s,r\rangle\in\mathcal{C}_n^\ell(\vdots)} \beta_{sr}, \sum_{\langle s,r\rangle\in\mathcal{C}_n^\ell(\text{-}\bullet\text{-})} \beta_{sr}\right)
\end{aligned}
$$

$$
\begin{aligned}
A_n^T \Delta_n \boldsymbol{F}_n &= \sum_{s \in \mathcal{B}_n^\ell} \delta_s f_t(s) \boldsymbol{\nabla} \tilde{f}(s) \otimes \boldsymbol{e}(s) \\
b_{\partial n} \overline{\boldsymbol{\theta}}_{\partial n}^\ell &= \sum_{\langle s,r \rangle \in \mathcal{C}_{\partial n}^\ell} \beta_{sr} (\boldsymbol{w}_r - \boldsymbol{w}_s^\ell) \otimes \boldsymbol{e}(s) \\
b_n \overline{\boldsymbol{\theta}}_n^\ell &= \sum_{\langle s,r \rangle \in \mathcal{C}_n^\ell} \beta_{sr} (\boldsymbol{w}_r^\ell - \boldsymbol{w}_s^\ell) \otimes (\boldsymbol{e}(s) - \boldsymbol{e}(r)) \\
&= \sum_{\langle s,r \rangle \in \mathcal{C}_n^\ell (\vdots)} \beta_{sr} (\boldsymbol{w}_r^\ell - \boldsymbol{w}_s^\ell) \otimes [0\ 1\ 0]^T + \sum_{\langle s,r \rangle \in \mathcal{C}_n^\ell (\bullet\!\!-\!\!\bullet)} \beta_{sr} (\boldsymbol{w}_r^\ell - \boldsymbol{w}_s^\ell) \otimes [0\ 0\ 1]^T.
\end{aligned}
$$

B Constrained estimation-segmentation at level ℓ

Data term: Using same block-wise notations as in Appendix A, it is easy to get the following compact expression:

$$
\mathcal{H}_1^\ell(\mathsf{d}\boldsymbol{w}^\ell, \delta; f, \boldsymbol{w}^\ell) = \sum_{n \in S^\ell} \left[\tau_1 (A_n \mathsf{d}\boldsymbol{w}_n^\ell + \boldsymbol{F}_n)^T \Delta_n (A_n \mathsf{d}\boldsymbol{w}_n^\ell + \boldsymbol{F}_n) + \sum_{s \in \mathcal{B}_n^\ell} \psi_1(\delta_s) \right]. \tag{38}
$$

For each site of S^ℓ, one gets a sort of block-wise optical flow constraint expression involving aggregated observations.

Smoothing term: Considering the piece-wise constant constraint on the increment field, the prior energy may be written as [24]:

$$
\mathcal{H}_2^\ell(\mathsf{d}\boldsymbol{w}^\ell, \beta; \boldsymbol{w}^\ell) = \mathcal{H}_2(\mathbf{0}, \beta; \boldsymbol{w}^\ell) + \tau_2 \sum_{\langle n,m \rangle \in \mathcal{C}^\ell} \left[\beta_{nm} \|\mathsf{d}\boldsymbol{w}_n^\ell - \mathsf{d}\boldsymbol{w}_m^\ell\|^2 + 2(\mathsf{d}\boldsymbol{w}_n^\ell - \mathsf{d}\boldsymbol{w}_m^\ell)^T \overline{\Delta \boldsymbol{w}}_{nm}^\ell \right], \tag{39}
$$

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

Parametric likeness term: $\mathcal{R}^\ell \in \Lambda^\ell$ being the current partition of S^ℓ, we denote $\mathcal{R} = \Phi_2(\mathcal{R}^\ell)$ the associated constrained partition of S and ∂_{ij} the pieces of frontiers between adjacent regions $\mathcal{R}_i$ and $\mathcal{R}_j$ of $\mathcal{R}$. We have:

$$
\mathcal{E}_{\text{interact}}^\ell(\mathsf{d}\boldsymbol{w}^\ell, \mathcal{R}^\ell, \beta, \eta; \boldsymbol{w}^\ell) = \mu_1 \sum_{\langle i,j \rangle} \frac{1}{|\partial_{ij}|} \sum_{\langle s,r \rangle \in \partial_{ij}} \beta_{sr} + \mu_2 \sum_i \sum_{n \in \mathcal{R}_i^\ell} \sum_{s \in \mathcal{B}_n^\ell} [\eta_s \|\boldsymbol{w}_s^\ell + \mathsf{d}\boldsymbol{w}_n^\ell - \boldsymbol{w}_s^i\|^2 + \psi_3(\eta_s)], \tag{40}
$$

which, like previously, reduces to:

$$\mathcal{E}^{\ell}_{\text{interact}}(\mathsf{d}\boldsymbol{w}^{\ell},\mathcal{R}^{\ell},\beta,\eta;\boldsymbol{w}^{\ell}) = \mathcal{E}^{\ell}_{\text{interact}}(\mathbf{0},\mathcal{R}^{\ell},\beta,\eta;\boldsymbol{w}^{\ell}) + \mu_2 \sum_i \sum_{n\in\mathcal{R}_i} [z_n ||\mathsf{d}\boldsymbol{w}^{\ell}_n||^2 + 2(\mathsf{d}\boldsymbol{w}^{\ell}_n)^T \overline{\Delta \boldsymbol{w}}^{\ell}_n(\boldsymbol{\theta}_i)], \quad (41)$$

where $z_n \triangleq \sum_{s\in\mathcal{B}^{\ell}_n} \eta_s$ and $\overline{\Delta \boldsymbol{w}}^{\ell}_n(\boldsymbol{\theta}_i) \triangleq \sum_{s\in\mathcal{B}^{\ell}_n} \eta_s(\boldsymbol{w}^{\ell}_s - \boldsymbol{w}^{i}_s)$.

Partition a priori term: The reduced expression of this term is:

$$E^{\ell}_{\text{prior}}(\mathcal{R}^{\ell}) = E_{\text{prior}}(\Phi^{\ell}_2(\mathcal{R}^{\ell})) = \lambda 2^{\ell} |\partial \mathcal{R}^{\ell}|. \quad (42)$$

Gauss-Seidel iteration w.r.t. $\mathsf{d}\boldsymbol{w}^{\ell}_n$: Setting to zero the derivative of $\mathbb{H}^{\ell}$ w.r.t. to $\mathsf{d}\boldsymbol{w}^{\ell}_n$, one gets the same update equation as in (37) (with $\mathsf{d}\boldsymbol{w}^{\ell}_n = \boldsymbol{\theta}^{\ell}_n$, for it is the constant model), but with slightly modified definitions:

$$\overline{\boldsymbol{\theta}}^{\ell}_{\partial n} \triangleq \frac{\sum_{m\in\nu(n)} \left(\beta_{nm}\mathsf{d}\boldsymbol{w}^{\ell}_m - \overline{\Delta \boldsymbol{w}}^{\ell}_{nm}\right) - \frac{\mu_2\tau_3}{\alpha\tau_1}\overline{\Delta \boldsymbol{w}}^{\ell}_n(\boldsymbol{\theta}_i)}{b_{\partial n} + \frac{\mu_2\tau_3}{\alpha\tau_1} z_n}, \quad \gamma \triangleq \frac{\alpha\tau_2 b_{\partial n} + \mu_2\tau_3 z_n}{\tau_1}, \quad (43)$$

where now appears the parametric field of the region $\mathcal{R}_i$ to whom block n belongs. For $\eta \equiv 0$ (which implies $z_n = 0$ and $\overline{\Delta \boldsymbol{w}}^{\ell}_n(\boldsymbol{\theta}_i) \equiv 0$), the expression coincides with (37), as expected, since this amounts to removing parametric fitness energy term.