Constrained multiscale Markov random fields and the analysis of visual motion

Fabrice Heitz, Patrick Bouthemy, Patrick Pérez

UNITÉ DE RECHERCHE
INRIA-RENNES

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay-Cedex
France
Tél.:(1) 39 63 55 11

Rapports de Recherche

1992
25ème anniversaire

N° 1615

Programme 4
Robotique, Image et Vision

CONSTRAINED MULTISCALE MARKOV RANDOM FIELDS AND THE ANALYSIS OF VISUAL MOTION

Fabrice HEITZ
Patrick PEREZ
Patrick BOUTHEMY

Février 1992

RR_1615

IRISA

INSTITUT DE RECHERCHE EN INFORMATIQUE
ET SYSTEMES ALEATOIRES

Campus Universitaire de Beaulieu
35042 - RENNES CEDEX FRANCE
Tél. : 99 84 71 00 - Télex : UNIRISA 950 473 F
Télécopie : 99 38 38 32

Constrained Multiscale Markov Random Fields and the Analysis of Visual Motion

Champs markoviens multiéchelles, optimisation sous contraintes et analyse du mouvement

Publication Interne n°627-Janvier 1992-40 pages-Programme 4

Fabrice HEITZ, Patrick PEREZ† and Patrick BOUTHEMY

IRISA/INRIA, †IRISA/CNRS,

Campus Universitaire de Beaulieu

35042 Rennes Cedex, France

Abstract

The use of Markov Random Field (MRF) models within the framework of global bayesian decision has recently brought new powerful solutions to most of static and dynamic image analysis issues.

Use of MRF models with the Maximum A Posteriori criterion leads to the minimization of a global energy function which may exhibit local minima. This minimization is generally performed using deterministic, or stochastic relaxation algorithms which can be sped up significantly by using multigrid techniques.

In this paper we investigate a new approach to multigrid image analysis based on MRF models. The multigrid algorithm under concern relies on constrained optimization. The global optimization problem associated to MRF modeling is solved over a sequence of nested subsets of the original configuration space. Those subsets consist of allowed configurations constraining the desired solution at different scales. The constrained optimization can be implemented via a coarse-to-fine multigrid algorithm defined on a sequence of consistent multiscale MRF models.

The proposed multiscale paradigm yields fast convergence towards high quality estimates when compared to standard monoresolution or multigrid relaxation schemes. It reveals also far less sensitive to local minima than standard relaxation algorithms.

The efficiency of the approach is demonstrated on several relevant problems in image sequence analysis : motion detection, optical flow measurement and motion-based segmentation. Results are presented on real world sequences including several moving objects and camera motion.

Submitted to : "CVGIP : Image Understanding" in January 1992.

Champs markoviens multiéchelles, optimisation sous contraintes et analyse du mouvement

Résumé

L'introduction des champs markoviens en modélisation d'images a récemment permis d'apporter des solutions nouvelles et robustes à nombre de problèmes en analyse d'images statiques ou dynamiques.

L'estimation des primitives induite *in fine* par la modélisation markovienne met toutefois en jeu des algorithmes de relaxation qui peuvent donner lieu à des temps de calcul importants. Ceci motive l'introduction de techniques multigrilles, par ailleurs classiques en analyse numérique. Ces techniques permettent en effet une accélération substantielle des algorithmes de relaxation.

Mais pour l'heure il n'existe pas de support théorique permettant d'associer stratégie multigrille et modèles statistiques markoviens. Cet article propose donc une approche "multiéchelle" de la modélisation markovienne, qui est à la fois mathématiquement cohérente et aisément implémentable. Le modèle présenté s'appuie sur une technique d'optimisation sous contraintes sur le champ des primitives à estimer. Le problème d'optimisation globale associé à la modélisation markovienne est résolu sur une succession de sous-espaces de configurations emboîtés. Ces sous-espaces correspondent à des configurations autorisées contraignant la solution désirée à différents niveaux de résolution. Le schéma d'optimisation sous contraintes associé est implanté sur une structure multigrille définie sur une succession de champs markoviens multirésolutions dont les paramètres et la structure de voisinage s'expriment aisément à partir du modèle initial.

Le schéma d'optimisation multigrille ainsi défini présente des propriétes de convergence rapide vers des estimées de qualité proche de celles obtenues par de coûteux algorithmes de relaxation stochastique. Il s'avère également moins sensible aux minima locaux que les algorithmes de relaxation multigrille classiques.

L'efficacité de l'approche est illustrée sur plusieurs problèmes représentatifs en analyse de séquences d'images : détection du mouvement, mesure du mouvement apparent et segmentation au sens du mouvement. Nous présentons des résultats sur des séquences réelles comprenant à la fois des objets mobiles et des mouvements de caméra.

1 Introduction

Markov Random Fields (MRF) models have been successfully introduced in several important issues of still image processing such as image restoration, [3, 9, 17], edge detection, [16], image segmentation, [10, 16], multisource image analysis, [22], stereovision [2], computed tomography, [12], surface reconstruction, [10, 29] or scene interpretation, [31]. The mathematical framework is a statistical one : entities of interest in a given task are described by statistical models (Markov Random Fields) and bayesian estimation theory is used to extract the relevant information from the observed images. By defining comprehensive, global non-linear models, the MRF theory leads to significant improvement with respect to former local methods.

On the other hand, the analysis of image sequences has motivated a large number of investigations in the last decade, [1, 6, 33]. We have recently extended MRF models to the analysis of image sequences in problems such as motion detection, [27], visual motion estimation, [20, 21] and motion-based segmentation, [7, 14] (for other studies using similar statistical models, see [11, 23, 25, 32, 36]). Markov Random Fields appear as an efficient and powerful framework for specifying non-linear interactions between features of the same nature or of different one (motion vectors, motion discontinuities and occlusions for instance, [20]). They help to combine and organize spatial and temporal information by introducing strong generic knowledge on the features to be estimated. When they are associated to the Maximum A Posteriori criterion, they lead to the minimization of a global energy function which may exhibit local minima, [17]. This minimization is generally performed using deterministic, [3] or stochastic, [17] relaxation algorithms.

Besides, it is known that multigrid methods can improve significantly the convergence rate of iterative relaxation schemes, [37]. As will be seen here, they may also be useful when the energy to be minimized presents many local minima. It has indeed been conjectured that multiresolution analysis may, to a certain extent, *smooth* the energy landscape. Deterministic relaxation schemes can then be used at coarse scales to get a good initial guess, which may be refined over increasing resolution.

Multigrid relaxation algorithms have been considered early for visual motion analysis models based on partial differential equations, [19, 37] as well as on MRF, [24]. Several investigations have been pursued on multigrid methods for optical flow estimation, [13, 21, 24]. Yet, in multigrid implementation of statistical models such as MRF, the key problem remains the derivation of the model parameters at different scales. When global mathematical consistency is not guaranteed, the parameters and the neighborhood structure associated to the model can only be adjusted over scale in an ad-hoc way. Gidas, [18] has recently described a consistent mathematical framework for multiscale Markov modeling, based on the *renormalization group* approach. Unfortunately, this standard technique in statistical physics does not lead to tractable computational schemes (apart from simple cases such as the binary Ising model, [18].) Moreover, it is shown in [26] that simple resolution transformations (sampling or block averaging) on gaussian MRF result in the loss of Markovianity.

The main contribution of this paper is the derivation of a general multiscale algorithm suited to MRF models, which is both consistent and tractable. The algorithm is related to a multiscale "constrained" exploration of the set of solutions of the original optimization problem. The global optimization problem is solved within a sequence of particular subsets of the original configuration space. Those subsets contain constrained configurations describing the expected primitive maps at different scales. They generate a new class of consistent multiscale MRF statistical models.

This constrained optimization is shown to be equivalent to a multigrid algorithm defined on a sequence of multiscale models.

In [5], Bouman *et al.* have recently presented an approach to multiresolution MRF-based estimation using a pyramidal structure close to the one reported here. They describe a multiple resolution algorithm, defined on a quadtree, for maximizing the a posteriori distribution associated with a MRF. This leads to the same estimation structure as the one derived from "constrained multiscale optimization", but in [5], the adjustment of parameters over scale is adressed in a rather ad-hoc way. Expressing the multiresolution analysis as a constrained optimization problem enables us to generalize the scheme reported in [5] and to determine a consistent set of parameters for the MRF at different scales. It also allows to define other classes of multiscale MRF models associated to more general constraints on the configuration space. Such a generalization is presented in [34].

Finally, our approach also extends a recent work of B. Levy, [28] on multiscale linear Gauss-Markov models to general non-linear Markov Random Field models.

In this paper, the constrained multiscale approach will be examplified on different issues dealing with visual motion analysis. The remainder of the paper is organized as follows. The theory of Markov Random Field and its application to visual motion analysis are first recalled in Section 2. The multiscale MRF models are introduced in Section 3, along with a coarse-to-fine relaxation strategy for computing the solutions. In Section 4, the multiscale method is compared to monoresolution schemes and standard multigrid algorithms in the case of motion detection, optical flow computation and motion-based segmentation. The multiscale algorithm exhibits fast convergence properties, comparable to multigrid deterministic relaxation, and the final estimates are close to the near optimal solutions obtained by time consuming sequential stochastic methods. Results are presented on synthetic sequences as well as on real-world outdoor dynamic scenes.

2 A general framework for visual motion analysis

The general mathematical framework under concern in this paper is global bayesian estimation associated with a priori models corresponding to MRF models. To extract

labels describing motion from image sequences, the method proceeds as follows :
• One or more specialized modules extract from the image sequence, features (spatio-temporal gradients, edges...) that will be used as observations in the estimation process.
• Observations are combined using local photometric and structural relations with a priori generic knowledge on the expected result (regularization scheme), in order to derive estimates of the unknown labels.

2.1 MRF models

Let $O = (O_s,\ s \in S)$ designate an observation field defined on a rectangular lattice S (in image sequence analysis observations are usually related to the spatiotemporal variations of the intensity function). Let $E = (E_s,\ s \in S)$ denote the unobserved (hidden) motion label field, defined on the same lattice S [1]. Realizations of fields O (resp. E) will be denoted $o = \{o_s, s \in S\}$ (resp $e = \{e_s, s \in S\}$). Let Λ be the (discrete) state space of random variable e_s and Ω the (finite) set of all possible label configurations $e = \{e_s, s \in S\}$.

(E, O) is a Markov random Field with respect to a chosen neighborhood system $\mathcal{G} = \{\mathcal{G}_s, s \in S\}$ if :

$$p(e, o) > 0 \tag{1}$$

and :

$$\forall s \in S,\ \ p(e_s | e_r, r \neq s, o) = p(e_s | e_r, o_r, r \in \mathcal{G}_s) \tag{2}$$

(and a similar definition for o_s). The conditional probabilities (equ. (2)) are sufficient to specify uniquely the joint probability $p(e, o)$ that the system is in a given configuration (e, o). Owing to a theorem by Hammersley-Clifford, one can show that (E, O) follows a Gibbs distribution, [17] :

$$\begin{aligned} p(o, e) &= \frac{1}{Z}\ e^{-U(o,e)} \\ \text{where } U(o, e) &\triangleq \sum_{c \in C} V_c(o, e) \\ \text{and } Z &\triangleq \sum_{(e,o)} e^{-U(o,e)} \end{aligned} \tag{3}$$

U is called the *energy function* (the lowest energies correspond to the most likely configurations). $\mathcal{C}$ denotes the set of cliques associated to neighborhood system $\mathcal{G}$. Cliques c are subsets of sites which are mutual neighbors. The potential function V_c is locally defined on clique c and expresses the local interactions between the different variables of the clique (see [17] for an extensive theory of MRF). The form

[1] Different lattices for E and O can also be adopted, see [17]

of the potential functions V_c is of course problem dependent and defines the local and global properties of the model.

The Maximum A Posteriori (MAP) estimate of the label field is defined by :

$$\begin{aligned} \hat{e} &\stackrel{\Delta}{=} \arg\max_{e\in\Omega} p(e|o) \\ &= \arg\max_{e\in\Omega} \frac{p(e,o)}{p(o)} \\ &= \arg\min_{e\in\Omega} U(o,e) \end{aligned} \tag{4}$$

Hence, when associated to MAP estimation, MRF-based image analysis reduces to the minimization of energy function $U(o,e)$ with respect to the motion label field e, [17]. Minimizing the global energy function U is often an intricate problem : the number of possible label configurations is generally *very large* and, moreover, the global energy function U may exhibit local minima. Computationally demanding stochastic relaxation algorithms are therefore generally necessary to compute MAP solutions. Less cpu-time consuming deterministic relaxation algorithms such as ICM, [3], HCF (Highest Confidence First, [10]) can often be used instead, when a good initial guess is available (other deterministic approaches include GNC : Graduated Non Convexity, [4] and MFA : Mean Field Annealing, [38]). Deterministic approaches converge to configurations corresponding to local minima of the global energy function.

Besides, standard multigrid computational schemes have proved efficient for minimizing energy function $U(e,o)$ in visual motion analysis applications, [24]. In this paper we describe a new multiscale paradigm which has revealed particularly attractive as far as the minimization of $U(e,o)$ is concerned.

First we present different MRF models suited for motion detection, optical flow computation and motion-based segmentation. These models will be used for comparison purposes between monoresolution, standard multigrid relaxation schemes and the multiscale MRF aproach proposed in this paper (Section 4). These different algorithms will be compared according to two criteria : convergence speed and quality of the final estimated solutions. This quality can be assessed by comparing the final energies computed on the different solutions, once convergence is reached.

The presented models are of interest for real-world applications involving image sequence analysis (other models for visual motion analysis have been presented in [7, 11, 14, 20, 23, 25, 27, 32]). Besides, they are also representative of the different models which can be found in the litterature as far as the neighborhood structure and the energy function is concerned, [2, 3, 9, 10, 12, 17, 16, 22, 29]. They involve more or less large configuration spaces and model different kind of image features such binary variables (motion detection), region numbers (motion-based segmentation) or vectors (in motion measurement). As a consequence they are associated to more or less hard optimization problems. For instance, the optimization in the case of motion measurement is really intricate with the model introduced here, as noticed

in [25]. Hence, it appears as a good benchmark for comparing different optimization schemes.

2.2 MRF models for visual motion analysis

We consider the following basic problems in visual motion analysis :

- *Motion detection*, [27] : the camera is assumed to be static and we are interested by detecting the moving objects in the scene with respect to the background. Hence, motion detection is a binary problem, in which it has to be decided whether a point in the image belongs to the static background or to the mask of a moving object.

- *Visual motion measurement*, [20] : camera and objects may be mobile. The goal is to estimate a dense velocity field at every pixel location in the image plane. This problem is also known as the "optical flow" measurement problem. In this case, the information to be recovered are displacement vectors.

- *Motion segmentation*, (with moving camera and objects), [7, 14] : the image sequence is here subdivided into its different constituent (moving) parts. The aim is to get a partition of the original image related to motion-based criteria. The information to be recovered corresponds to a labeling of the image with discrete region numbers.

The solution of those "low-level" image processing issues is an important preliminary step in most applications involving a kinematical interpretation of the behavior of moving objects in a scene, [7, 14]. We describe now three MRF-based models which have been designed to solve these different issues. The proposed models describe the interactions between the spatio-temporal variations of the observed image sequence and the features of interest to be determined in the different problems under concern : binary variables in the detection case, displacement vectors in the measurement case and region labels for the motion-based segmentation.

2.2.1 A model for motion detection

We consider here a MRF model for motion detection in the case of a static camera. This model is a modified version of a model presented in [8]. It is suited to situations in which there is no overlapping between the masks of a given moving object at successive times (this assumption holds if an appropriate temporal undersampling of the image sequence is performed). A more sophisticated model taking into account overlapping is described in [27].

Let $f_t(s)$ denote the observed intensity function, where $s = (x, y)$, $s \in S$ designates the 2D spatial image coordinates, t the time axis and δt the temporal difference between two successive frames. Motion detection basically relies on the analysis of the variations in time of the intensity distribution. The observations

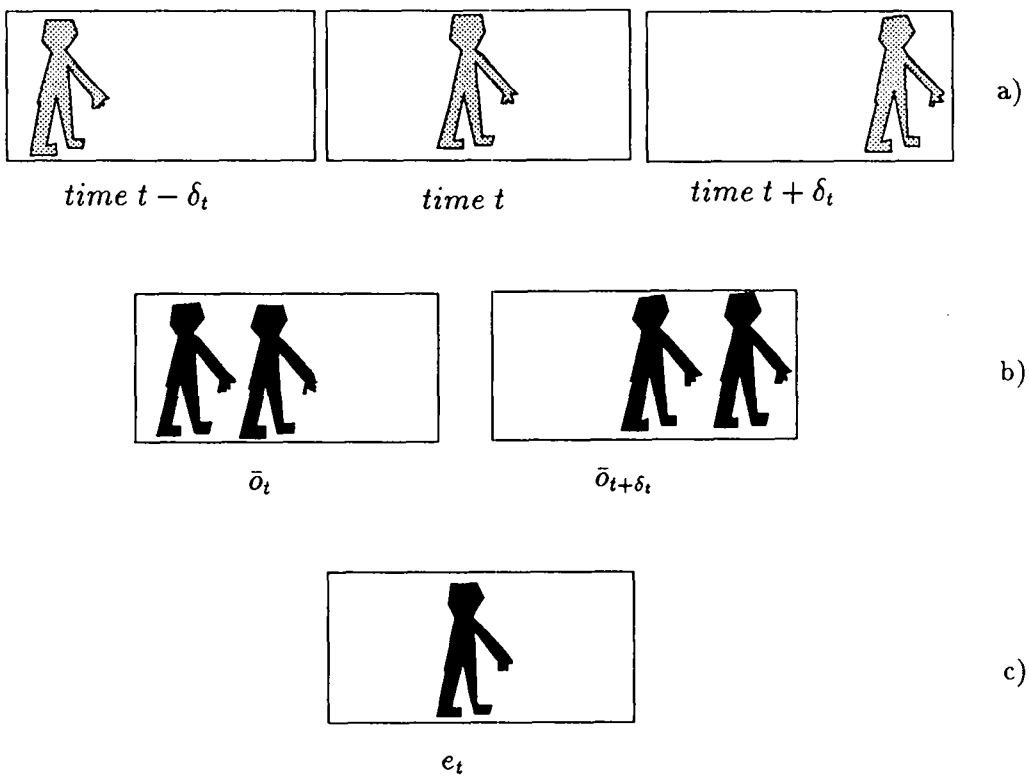

Figure 1: *Motion detection in the case of no overlapping between object masks*

consist in two observation arrays denoted o_t and $\bar{o}_t$. o_t is defined by :

$$o_t(s) \triangleq |f_t(s) - f_{t-\delta t}(s)| \tag{5}$$

and $\bar{o}_t$ is a logical map of temporal changes between time $t - \delta t$ and time t.

$\bar{o}_t(s) = 1$ if a temporal change of the intensity function is validated at point s, $\bar{o}_t(s) = 0$ otherwise. It is obtained using a likelihood test in which intensity is locally modeled by a linear function with an additive white gaussian noise of constant variance, (see [27] for details).

The label field $e_t(s)$ is binary-valued and describes the state of point $s = (x, y)$ at time t :

$e_t(s) = 0$ if s belongs to the static background,

$e_t(s) = 1$ if it is located on the mask of a mobile object at time t.

The MRF model introduced here is associated to a 8-neighborhood structure (Fig. 2) and describes spatial interactions between the temporal derivatives of the intensity function and the binary label field. For the motion detection model, we consider only two-element cliques (doubletons).

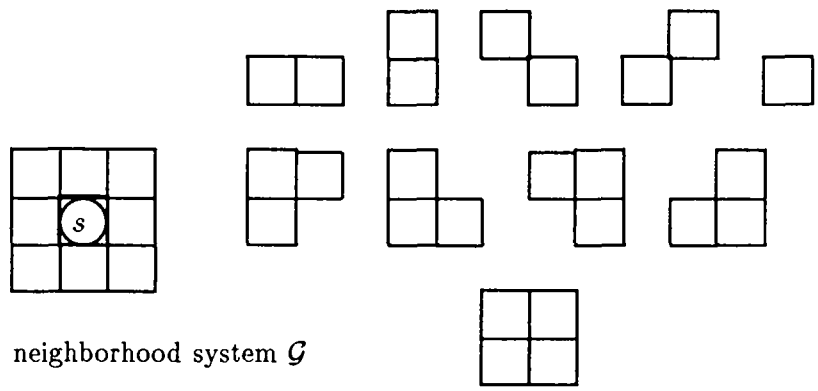

Figure 2: *8-neighborhood structure and associated cliques*

The global energy function is the sum of three terms :

$$U(o, e) = \underbrace{U_{11}(o_t, o_{t+\delta t}, e_t) + U_{12}(\bar{o}_t, \bar{o}_{t+\delta t}, e_t)}_{U_1(o_t, o_{t+\delta t}, \bar{o}_t, \bar{o}_{t+\delta t}, e_t)} + U_2(e_t) \tag{6}$$

with :

$$\begin{aligned} U_{11} &\triangleq \sum_{s \in S} \{\frac{1}{2\sigma^2}(o_{t+\delta t}(s) - \Psi(e_t(s)))^2 + \frac{1}{2\sigma^2}(o_t(s) - \Psi(e_t(s)))^2\} \\ U_{12} &\triangleq \sum_{s \in S} V_\tau(e_t(s), \bar{o}_t(s), \bar{o}_{t+\delta t}(s)) \\ U_2 &\triangleq \sum_{\{s_1, s_2\} \in C} V_{\{s_1, s_2\}}(e_t(s_1), e_t(s_2)) \end{aligned} \tag{7}$$

where :

$$V_{\{s_1,s_2\}}(e_t) \triangleq \begin{cases} +\beta_s & if \quad e_t(s_1) \neq e_t(s_2) \\ -\beta_s & if \quad e_t(s_1) = e_t(s_2) \end{cases} \tag{8}$$

$$\Psi(e_t(s)) \triangleq \begin{cases} 0 & if\ e_t(s) = 0 \\ \mu & if\ e_t(s) = 1 \end{cases} \tag{9}$$

and $V_\tau(e_t(s), \bar{o}_t(s), \bar{o}_{t+\delta t}(s))$ is defined in Table 1. The parameters of the model are : $\sigma^2, \beta_s > 0, \beta_\tau > 0$ and $\mu > 0$.

$\bar{o}_t$	e_t	$\bar{o}_{t+\delta t}$	V_τ
0	0	0	$-\beta_\tau/2$
0	0	1	$-\beta_\tau/2$
0	1	0	β_τ
0	1	1	β_τ
1	0	0	$-\beta_\tau/2$
1	0	1	β_τ
1	1	0	β_τ
1	1	1	$-\beta_\tau/2$

Table 1 : *definition of potential* V_τ

The energy includes a term U_1 describing pointwise interactions between labels and observations ($U_1 = U_{11} + U_{12}$) and a term U_2 describing the effect of neighboring labels on $e_t(s)$. In this expansion each term expresses a different interaction, which brings its partial contribution to the global estimation process. The different terms can be interpreted as follows :

- energy term U_2 stands for spatial interactions and favors homogeneous masks, (i.e. it tends to eliminate isolated points and to fill in missing points inside the masks).
- energy U_{11} expresses the consistency between the temporal variations (changes) in the sequence (o_t, $o_{t+\delta t}$) and the current estimates of the label field e_t. It helps detecting the temporal changes between successive frames. However detecting changes is not sufficient since it creates two kind of masks corresponding to the part of the background uncovered by the moving object in the previous frame and to the area covered in the current frame by the same object (see (Fig. 1b)).
- Energy U_{12} is used to reconstruct the mask of a mobile object at a given time (i.e. to filter out the areas which do not correspond to the object mask at time t : compare Fig. 1b and 1c).

Different results on real world outdoor scenes obtained using this model may be found in [8].

2.2.2 A model for optical flow measurement

We consider a standard model for optical flow measurement which has been used in [24] (for other models see [20, 21]). In this model the labels belong to a discrete state space Λ. We choose a discrete formulation of the optical flow computation problem because it is known to lead to a complex energy landscape exhibiting numerous *local minima*, [25]. Hence this model appears as a difficult example as far as optimization is concerned and will be useful for comparison purpose between standard and multiscale models.

Again $f_t(s)$ denotes the observed intensity function, where $s = (x, y)$, $s \in S$ designates the 2D spatial image coordinates, t the time axis and δt the temporal difference between two successive frames. The velocity vector at time t and site s is denoted $\vec{\omega}_t(s) = (u_t(s), v_t(s))$, with $u_t(s) = \frac{dx}{dt}(s), v_t(s) = \frac{dy}{dt}(s)$ and $\vec{\omega}_t = \{\vec{\omega}_t(s),\ s \in S\}$. In the model considered here velocities are defined on the same grid S as the pixels and the velocities are discretized according to a discrete state space $\mathcal{W} = (-u_{max} : u_{max}, -v_{max} : v_{max})$ with a step size of δ. The MRF model is associated to an 8-neighborhood system (Fig. 2) anf two element cliques (doubletons) are only considered. The global energy is specified by the following energy function :

$$U(f_t, f_{t+\delta t}, \vec{\omega}_t) = U_1(f_t, f_{t+\delta t}, \vec{\omega}_t) + U_2(\vec{\omega}_t) \tag{10}$$

where :

$$\begin{aligned} U_1(f_t, f_{t+\delta t}, \vec{\omega}_t) &= \sum_{s \in S} \{\ f_t(s) - f_{t+\delta t}(s + \vec{\omega}_t(s).\delta t)\ \}^2 \\ U_2(\vec{\omega}_t) &= \alpha^2 \sum_{\{s,r\} \in C} \| \vec{\omega}_t(s) - \vec{\omega}_t(r) \|^2 \end{aligned} \tag{11}$$

The first term in the energy (known as the "displaced frame difference") expresses the constant brightness assumption for a physical point over time along the motion trajectory. The second term balances the first one through the weighting parameter α; it can be interpreted as a regularization term which favors smooth solutions.

2.2.3 A model for motion-based segmentation

The goal of motion-based segmentation is to separate regions in the image corresponding to different motion entities. Once this partition is obtained, a kinematical analysis can be performed in each region, [14]. A motion entity is supposed to correspond to a given description level of real physical motion.

A version of the MRF-based motion segmentation algorithm presented here has already been described in [15].

The label values to be estimated in the motion-based segmentation process will consist in *region numbers* r, $r = 1, R$. Let S be the set of sites s where these labels are defined.

For the model under concern here, [15], the observations are given by the spatio-temporal derivatives of the intensity function :

$$O = \{O_s,\ s \in S\} = \{\nabla f(s), \frac{\partial f}{\partial t}(s)\} \tag{12}$$

where f is the intensity function, ∇f its spatial gradient, $\frac{\partial f}{\partial t}$ its temporal derivative.

Motion-based segmentation only makes sense with respect to a motion model. For the while, it is assumed that the model is parametrized by a global parameter vector θ_r, within a given region r, $r = 1, R$ (2D linear motion models will be used in practice). Let $\theta = \{\theta_r, r = 1, R\}$. To derive the unknown label field e along with the number of regions R and the set of model parameters θ from the observed fields $(\nabla I, I_t)$, the following energy function U is minimized :

$$U(\nabla f, \frac{\partial f}{\partial t}, e, \theta, R) = U_1(\nabla f, \frac{\partial f}{\partial t}, e, \theta, R) + U_2(e). \tag{13}$$

U_2 expresses the expected spatial properties of the resulting label field, (homogeneous regions are favoured), while U_1 corresponds to the adequacy between observations and labels.

Prior model, energy term U_2

Again we consider a second-order neighborhood system, (Fig. 2) and cliques used are doubletons, denoted $c = \{s_1, s_2\}$. Let C be the set of all doubletons in the image. Energy U_2 is defined as follows:

$$U_2(e) = \sum_{\{s_1,s_2\} \in C} V(e_{s1}, e_{s2}) = \sum_{\{s_1,s_2\} \in C} \mu\ (1 - \delta_{e_{s1}=e_{s2}}) \tag{14}$$

where δ designates the Kronecker symbol and μ is a weighting parameter ($\mu > 0$).

This energy function is a standard one which favours homogeneity of the label field, [17].

Observation-label interaction model, energy term U_1

U_1 represents the interaction model between observations and labels. The starting idea is to compare the velocity vectors corresponding to the considered motion model to the true velocity vectors. A region can be stated as coherent with respect to motion cue if its motion content can be described through one model parametrization. Conversely a point can belong to region r if its true velocity vector $\vec{\omega}$ is close to the one induced at this point by the motion model parametrized by vector θ_r. Let us denote the model-related velocity vector field by $\vec{\omega}_{\theta_r}$. Formally, this leads to take into account the difference: $\vec{\omega} - \vec{\omega}_{\theta_r}$. If the velocity vector field $\vec{\omega}$ is not available, this quantity can not be directly evaluated. However the spatio-temporal derivatives of the intensity function are related to the apparent velocity vector by the well-known image flow constraint equation, [35]:

$$\nabla f(s).\vec{\omega}(s) + \frac{\partial f}{\partial t}(s) = 0$$

Therefore let us consider the following quantity at each point s:

$$\epsilon(s) = \nabla f(s).(\vec{\omega}_{\theta_r}(s) - \vec{\omega}(s))$$

Using the above mentioned image flow constraint, this can be written as follows:

$$\epsilon(s) = \nabla f(s).\vec{\omega}_{\theta_r}(s) + \frac{\partial f}{\partial t}(s)$$

If correctly parametrized, the variables ϵ are assumed to follow a zero-mean gaussian law of variance σ^2. Moreover they are supposed to be independent. In practice, 2D linear models are used:

$$\theta_r = (a, b, \alpha, \beta, \gamma, \delta)^T$$

and at point (x, y), the velocity vector generated by the linear motion model parametrized by θ_r is given by:

$$\vec{\omega}_{\theta_r}(x, y) = (a + \alpha.\Delta x + \gamma.\Delta y,\ b + \beta.\Delta x + \delta.\Delta y) \tag{15}$$

where Δx, and Δy designate the differential shift with respect to a reference point, which will be the image centre. The choice of such a linear model represents a good trade-off between model complexity and model efficiency.

The energy term U_1 can now be expressed as follows:

$$U_1(\nabla f, \frac{\partial f}{\partial t}, e, \theta, R) = \sum_{s \in S} \left(\frac{1}{2\sigma^2} (\nabla f(s).\vec{\omega}_{\theta_{e_s}}(s) + \frac{\partial f}{\partial t}(s))^2 \right) \tag{16}$$

The choice of expression (16) must be related to above mentioned assumptions tied to variables ϵ.

Finally, the global energy function U can be expressed as :

$$U = U_1 + U_2 = \sum_{s \in S} \left(\frac{1}{2\sigma^2} (\nabla f(s).\vec{\omega}_{\theta_{e_s}}(s) + \frac{\partial f}{\partial t}(s))^2 \right) + \mu \sum_{\{s_1, s_2\} \in C} (1 - \delta_{e_{s1} = e_{s2}}) \tag{17}$$

The minimization of energy function U is carried out according to a deterministic relaxation. The estimation of label field e and the identification of global parameters θ are performed alternately. Briefly, within each currently delimited region r the optimal parameter values $\hat{\theta}_r$ are estimated; then a number of minimization iterations concerning only label field e are performed; then the $\hat{\theta}_r$'s are updated, and so on.

The main reason for this is that parameters θ_r are global ones (i.e. each $\hat{\theta}_r$ is computed within each current region r), whereas label field e is updated site by site. The estimation of the parameter vector θ_r for each region r is achieved by minimizing energy term U_1 while freezing the other variables, i.e. the label field e. This leads to a least-square estimation of the parameters θ_r within each region r.

Besides, in the presented model, the right number of regions $\hat{R}$ in the final partition is not assumed to be known *a priori* (this number is usually predetermined in markovian approaches dealing with image segmentation issues). The number of regions is updated dynamically as explained hereafter :

- At the very beginning the label field is initialized by one single region when no *a priori* knowledge is available. The segmentation can also start from an initial partition derived for instance from the final partition computed on the previous frame, [15]. In multigrid techniques the initialization may also be provided by the initial guess obtained at coarser resolution. This will be the case here;

- Over time *new regions* can be created owing to the introduction in the relaxation process of a supplementary label ρ corresponding to an "outlier class". This label is assigned during the relaxation process to a given site s whenever the discrepancy between the model-related velocity vectors $\vec{\omega}_{\theta_r}(s)$ and the "true" velocity $\vec{\omega}(s)$ is too large, i.e. if none of the already existing labels accounts for the observed spatio-temporal variations.

  In practice, first the label value $\hat{r}_i$ among the ones existing in the *neighborhood* of s is determined which minimizes the local energy given by (instead of U, we consider $2\sigma^2 U$):

$$\Delta U_s(r) = (\nabla f(s).\vec{\omega}_{\theta_r}(s) + \frac{\partial f}{\partial t}(s))^2 + A \sum_{s_2 \in \mathcal{G}_s} (1 - \delta_{e_{s2}=r})$$

  where $\mathcal{G}_s$ designates the set of neighbors of s. The local energy associated to label ρ is defined by :

$$\Delta U_s(\rho) = \Phi \hat{\sigma}^2_{\hat{r}_i} + A \sum_{s_2 \in \mathcal{G}_s} (1 - \delta_{e_{s2}=\rho})$$

  where $\hat{\sigma}^2_{\hat{r}_i}$ is the variance of variable ϵ *a posteriori* estimated within region $\hat{r}_i$ and Φ is a positive parameter, which can be explicitly determined using tables associated to χ^2 statistical law. Indeed this is related to the probability of not selecting $\hat{r}_i$ whereas it is the right label. The label value finally selected is the one between $\hat{r}_i$ and ρ which minimizes the local energy ΔU_s.

  After a given number of iterations, new region labels are assigned to the connected areas of sufficient size labeled by ρ. The current number of regions is updated accordingly.

The different MRF models which have been presented in this section will be used to compare the proposed multiscale paradigm to the existing standard relaxation algorithms. Standard implementations of a multiresolution algorithm based on MRF have for instance been reported in [24]. We describe our own approach to the problem in the next section.

3 Constrained Multiscale MRF

3.1 Multiscale exploration of the configuration space

Our approach is based on a constrained optimization scheme related to a "multiscale exploration" of the space of solutions of the optimization problem at hand (equ. 4). We show that this exploration is equivalent to a multigrid algorithm defined on a set of multiscale MRF models. This multigrid algorithm has several attractive properties :

- it allows to propagate local MRF interactions more efficiently between distant points, yielding faster convergence;
- the energy functions at different scales do not present the same "landscape". As a matter of fact it has been conjectured, [24] that in low resolution images the energy landscape is "smoother" and may contain fewer local minima than the energy associated to the original full resolution image. Whereas the deterministic relaxation algorithm gets stuck in the first local minimum near the starting configuration, estimates closer to the global optimum may be expected from a multigrid method starting at a low resolution level where less local minima exist.

When considering a multiscale representation based on MRF statistical models, the main problem to be solved is the adjustment of the model parameters over scale. As shown in [18] the neighborhood structure may also vary from one scale to another. In [18] Gidas has adapted the renormalization group approach of statistical physics to derive consistent multiscale MRF models. The renormalization group approach essentially assumes that the energy of a realization of the MRF model does not change over scale (constant energy principle). Unfortunately the method described by Gidas is tractable only for very simple MRF models. Bouman *et al.*, [5] have recently presented a different model based on a single observation level and a label pyramid but the consistency of the representation is not guaranteed and the parameters of the model are adjusted in a rather ad-hoc manner. The approach we report here is both consistent and tractable. This paradigm relates to a recent work on coarse scale linear Gauss-Markov models by B. Levy, [28].

In the remainder we will focus on a particular class of MRF models with the following energy function :

$$U(e,o) \;=\; U_1(e,o) \;+\; U_2(e) \tag{18}$$

with :

$$\begin{aligned} U_1(e,o) &\triangleq \sum_{s\in S} V_1(e_s,o_s) \\ U_2(e) &\triangleq \sum_{c\in C} V_c(e) \end{aligned} \tag{19}$$

where a 8-neighborhood is adopted for the neighborhood system $\mathcal{G}$. The cliques $c \in \mathcal{C}$ associated to this neighborhood structure are shown in Fig. 2 and contain at most four sites. The energy functions presented in Sections 2.2 obviously belong to this class. Moreover, this class comprises most usual energy functions used in still and dynamic image analysis problems, [2, 17, 20] (the extension to larger neighborhoods is straightforward, although not only a notational matter).

The multiscale MRF model can be described as follows. First recall that the MRF model assigns an energy to all possible configurations of a field $e = \{e_s,\ s \in S\}$, $e \in \Omega$. Let us assume that the size of the full resolution lattice is $2^m \times 2^{m'}$. To generate a "coarse grid" MRF model, let us divide the initial grid S into cells (blocks) of 4 (2x2) pixels. Among all configurations in Ω we consider a subset Ω_1 where e is constant over each of the above defined 2x2 cells (Fig. 3). These configurations will describe the MRF at scale 1, at which the size of a "constant value cell" is $2^1 \times 2^1$. Scale i ($i \leq inf(m, m')$) is defined similarly by considering label configurations which are constant over cells of size $2^i \times 2^i$ (Fig. 3.)

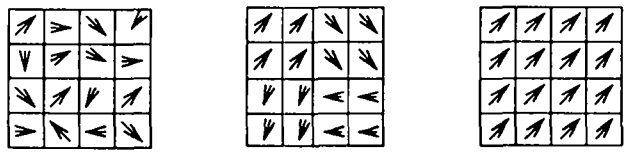

Figure 3: *Multiscale Markov Random Field : example of global configurations in $\Omega_0 = \Omega$ (scale 0), Ω_1 (scale 1) and Ω_2 (scale 2). Labels are represented by vectors for easier visual interpretation.*

The cells at scale i will be denoted B_k^i, $k = 1,.. N_i$ (where cells are ordered by row scanning for instance, and $N_i = 2^{m+m'-2i}$). Let B^i be the set of cells B_k^i and Ω_i designate the space of configurations with constant values on cells B_k^i ($\Omega_i \subset \Omega_{i-1} \subset \cdots \subset \Omega_1 \subset \Omega_0 = \Omega$) . The common value of e_s in cell B_k^i at scale i will be denoted e_k^i :

$$e \in \Omega_i \ \Leftrightarrow \quad \forall s \in B_k^i \quad e_s = e_k^i \qquad k = 1, \cdots, N_i \tag{20}$$

For a configuration $e \in \Omega_i$ the first term in the global energy function can then be decomposed on the cell structure according to equ. (20) :

$$\begin{aligned} U_1(e, o) &= \sum_{s \in S} V_1(e_s, o_s) \\ &= \sum_{B_k^i \in B^i} \sum_{s \in B_k^i} V_1(e_s, o_s) \\ &= \sum_{B_k^i \in B^i} V_1^i(e_k^i, o) \end{aligned} \tag{21}$$

with

$$V_1^i(e_k^i, o) \triangleq \sum_{s \in B_k^i} V_1(e_s, o_s) \tag{22}$$

For the second term $U_2(e)$, let us define on B^i the following neighborhood structure :

$$B_{k_1}^i \text{ and } B_{k_2}^i \text{ neighbors} \Leftrightarrow \begin{cases} B_{k_1}^i = B_{k_2}^i \\ \text{or} \\ \exists c \in \mathcal{C} : c \cap B_{k_1}^i \neq \emptyset \text{ and } c \cap B_{k_2}^i \neq \emptyset \end{cases} \tag{23}$$

Clearly, this definition reduces to the same neighborhood structure for cells B^i as the neighborhood system $\mathcal{G}$ defined on the fine grid model (in the case of a 8-neighborhood). The neighborhood on B^i will be denoted $\mathcal{G}^i$.

Now, consider a partition of the original clique set $\mathcal{C}$ into subsets corresponding to different locations with respect to cells B_k^i :

- *subset* $\mathcal{C}_k^i \subset \mathcal{C}$: c is included in B_k^i (Fig. 4a):

$$\mathcal{C}_k^i = \{c \in \mathcal{C} \mid c \subset B_k^i\}$$

- *subset* $\mathcal{C}_{k_1,k_2}^i \subset \mathcal{C}$: c sits astride two neighboring cells $\{B_{k_1}^i, B_{k_2}^i\}$ (Fig. 4b):

$$\mathcal{C}_{k_1,k_2}^i = \{c \in \mathcal{C} \mid c \subset \bigcup_{l=1}^{l=2} B_{k_l}^i \quad \text{and} \quad c \cap B_{k_l}^i \neq \emptyset \quad \text{for} \quad l = 1,2\}$$

- *subset* $\mathcal{C}_{k_1,k_2,k_3}^i \subset \mathcal{C}$: c sits astride three neighboring cells $\{B_{k_1}^i, B_{k_2}^i, B_{k_3}^i\}$ (Fig. 4c):

$$\mathcal{C}_{k_1,k_2,k_3}^i = \{c \in \mathcal{C} \mid c \subset \bigcup_{l=1}^{l=3} B_{k_l}^i \quad \text{and} \quad c \cap B_{k_l}^i \neq \emptyset \quad \text{for} \quad l = 1,2,3\}$$

- *subset* $\mathcal{C}_{k_1,k_2,k_3,k_4}^i \subset \mathcal{C}$: c sits astride four neighboring cells $\{B_{k_1}^i, B_{k_2}^i, B_{k_3}^i, B_{k_4}^i\}$ (Fig. 4d):

$$\mathcal{C}_{k_1,k_2,k_3,k_4}^i = \{c \in \mathcal{C} \mid c \subset \bigcup_{l=1}^{l=4} B_{k_l}^i \quad \text{and} \quad c \cap B_{k_l}^i \neq \emptyset \quad \text{for} \quad l = 1,2,3,4\}$$

This partition yields the following decomposition for energy function $U_2(e)$:

$$\begin{aligned} U_2(e) &= \sum_{c \in \mathcal{C}} V_c(e) \\ &= \sum_{B_k^i \in B^i} \sum_{c \in \mathcal{C}_k^i} V_c(e) \; + \end{aligned}$$

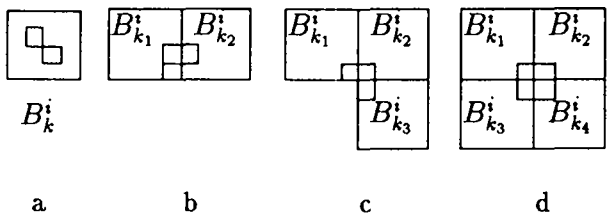

Figure 4: *Example of local clique configurations c with respect to cells* B^i_k

$$\sum_{\{B^i_{k_1},B^i_{k_2}\}\ neighbors} \sum_{c\in\mathcal{C}^i_{k_1,k_2}} V_c(e) + \sum_{\{B^i_{k_1},B^i_{k_2},B^i_{k_3}\}\ neighbors} \sum_{c\in\mathcal{C}^i_{k_1,k_2,k_3}} V_c(e) + \sum_{\{B^i_{k_1},B^i_{k_2},B^i_{k_3},B^i_{k_4}\}\ neighbors} \sum_{c\in\mathcal{C}^i_{k_1,k_2,k_3,k_4}} V_c(e) \tag{24}$$

which can be rewritten as :

$$\begin{aligned} U_2(e) \;=\; & \sum_{B^i_k\in B^i} V^i_k(e^i_k) \\ & + \sum_{\{B^i_{k_1},B^i_{k_2}\}\ neighbors} V^i_{k_1,k_2}(e^i_{k_1},e^i_{k_2}) \\ & + \sum_{\{B^i_{k_1},B^i_{k_2},B^i_{k_3}\}\ neighbors} V^i_{k_1,k_2,k_3}(e^i_{k_1},e^i_{k_2},e^i_{k_3}) \\ & + \sum_{\{B^i_{k_1},B^i_{k_2},B^i_{k_3},B^i_{k_4}\}\ neighbors} V^i_{k_1,k_2,k_3,k_4}(e^i_{k_1},e^i_{k_2},e^i_{k_3},e^i_{k_4}) \end{aligned} \tag{25}$$

with :

$$\begin{aligned} V^i_k(e^i_k) &\triangleq \sum_{c\in\mathcal{C}^i_k} V_c(e) \\ V^i_{k_1,k_2}(e^i_{k_1},e^i_{k_2}) &\triangleq \sum_{c\in\mathcal{C}^i_{k_1,k_2}} V_c(e) \\ V^i_{k_1,k_2,k_3}(e^i_{k_1},e^i_{k_2},e^i_{k_3}) &\triangleq \sum_{c\in\mathcal{C}^i_{k_1,k_2,k_3}} V_c(e) \\ V^i_{k_1,k_2,k_3,k_4}(e^i_{k_1},e^i_{k_2},e^i_{k_3},e^i_{k_4}) &\triangleq \sum_{c\in\mathcal{C}^i_{k_1,k_2,k_3,k_4}} V_c(e) \end{aligned} \tag{26}$$

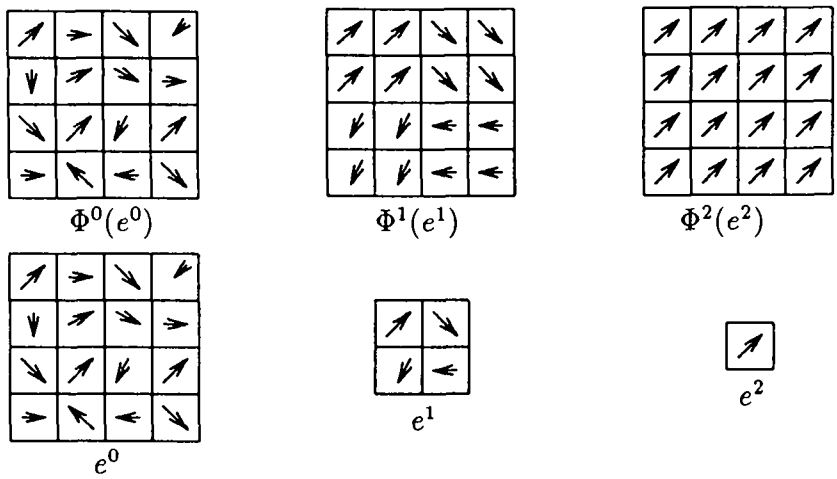

Figure 5: *Isomorphism Φ^i between a configuration e^i in Γ_i on which the coarse model is defined and the corresponding configuration $\Phi^i(e^i)$ in Ω_i*

Equivalent coarse resolution models

Considering equations (21) and (25), it turns out that the energy of a configuration $e \in \Omega_i$ can be expressed as a sum of potential functions, which are defined on mutually neighboring cells B^i_k and which only depend on the values e^i_k in those cells and on *the whole observation field o.* Thanks to that decomposition, the energy of a configuration $e \in \Omega_i$ can be computed very efficiently using an *equivalent coarse model* corresponding to a label field whose resolution has been reduced by a factor given by the cell size 2^ix2^i (Fig. 5). The resulting coarse grid, denoted S^i is isomorphic to B^i. The coarse label field $E^i = \{E^i_s, s \in S^i\}$ is associated to a reduced configuration space $\Gamma_i = \Lambda^{S^i}$ isomorphic to Ω_i, with $\Gamma_0 = \Omega_0 = \Omega$. Let Φ^i denote the bijection from Γ_i in Ω_i :

$$\begin{array}{rccl} \Phi^i & : & \Gamma_i & \longrightarrow & \Omega_i \\ & & e^i & \longmapsto & e = \Phi^i(e^i) \end{array} \tag{27}$$

defined as an interpolation from the coarse grid S^i to the fine grid $S^0 = S$ using simple repetition of the label estimates on the fine grid S (Fig. 5). The sequence of models on grids S^i, $i = n \cdots, 0$ defines a set of consistent multiscale Markov Random Field models, whose energy functions $U^i(e^i, o)$ are given by :

$$U^i(e^i, o) \;=\; U_1(\Phi^i(e^i), o) \;+ U_2(\Phi^i(e^i)) \tag{28}$$

Let us notice that, contrary to the approach proposed in [18], the coarse models have the same neighborhood structure as the original fine grid Markov Random Field (see equ(25)) and the potential functions at coarse resolution are derived by simple computations. The potential functions of the coarse grid models are defined from the characteristics of the original model, by equations (22) and (26).

This approach to multiscale MRF can be easily interpreted as a constrained exploration of subsets of the original (generally very large) configuration space Ω. In these subsets Ω_i configurations are constrained to remain constant in cells of size $2^i \times 2^i$.

A new multigrid relaxation algorithm

To take benefit from the sequence of multiscale models previously defined, instead of handling directly the original optimization problem :

$$\hat{e} = \arg\min_{e \in \Omega} U(e, o) \tag{29}$$

we consider the following sequence of problems defined on the constrained subsets Ω_i :

$$\hat{e}(i) = \arg\min_{e \in \Omega_i} U(e, o), \quad i = n, \cdots 0 \tag{30}$$

The constrained optimization over the blockwise constant configurations set Ω_i is equivalent to a minimization on the coarse MRF model U^i , i.e. :

$$\hat{e}^i = \arg\min_{e^i \in \Gamma_i} U^i(e^i, o), \quad i = n, \cdots 0 \tag{31}$$

with $\hat{e}(i) = \Phi^i(\hat{e}^i)$.

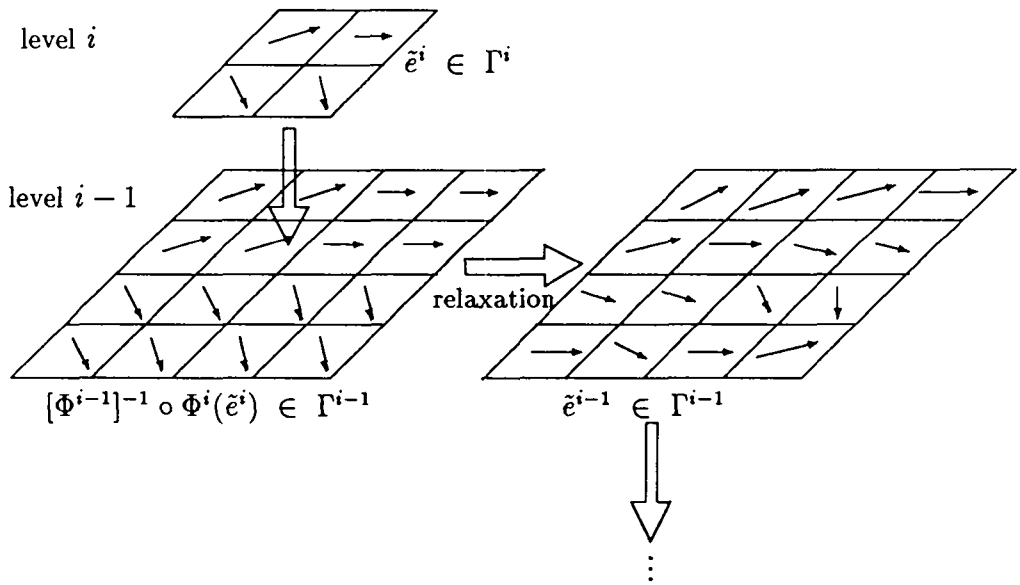

Figure 6: *A coarse-to-fine strategy for optimization in subsets* $\Gamma_i, i = n, \cdots 0$

We define a multigrid estimation strategy associated to the multiscale model which is a standard coarse-to-fine one, (Fig. 6). Starting from a coarse level n, the optimization problem (equ. (29)) is first approximately solved in subset Ω_n (equ.

(30)) by solving the equivalent problem (31). An estimate of $\hat{e}^n$ is obtained by a deterministic relaxation algorithm known as ICM, [3]. At level i if $\tilde{e}^i$ designates the estimate of $\hat{e}^i$ (obtained after convergence of the deterministic relaxation at that level), the algorithm at resolution level $i-1$ is initialized by $[\Phi^{i-1}]^{-1} \circ \Phi^i(\tilde{e}^i)$, which is just an interpolation of $\tilde{e}^i$ on the finer grid S^{i-1} (Fig. 6).

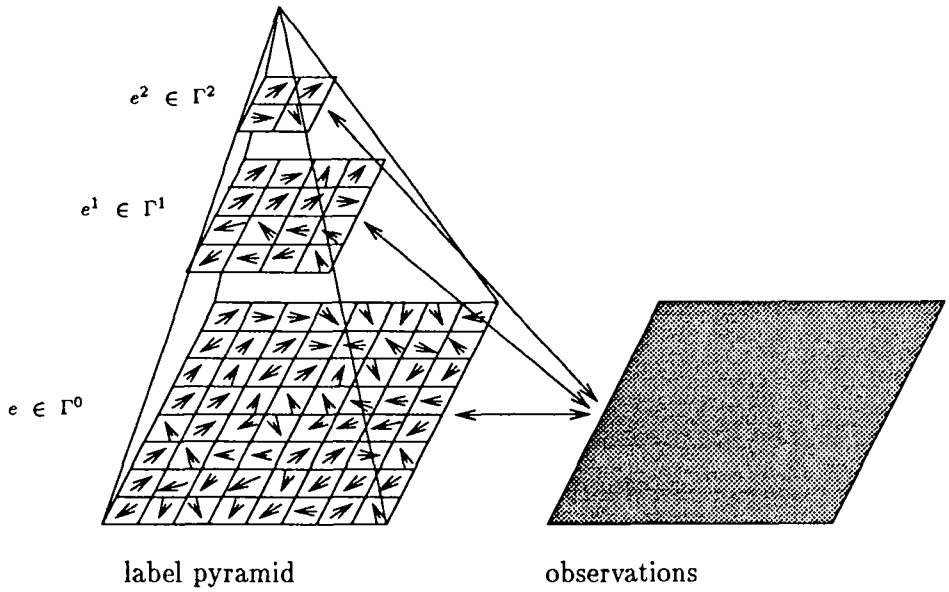

Figure 7: *The multiscale estimation structure*

Let us notice that whereas the algorithm uses a multigrid representation of labels, only one single (full resolution) level is used for the observation field o, as can be seen in equ. (31) (Fig. 7). We consider this to be an advantage of the method (which is also found in the scheme described in [5]), because *no multiresolution pyramid* has to be computed on observations. In standard multigrid schemes (Fig. 8), a pyramid of observations is usually obtained by low-pass filtering the original data, [13, 24]. The choice of the filters is generally quite arbitrary and the way they transform the original data is understood only qualitatively. Moreover, additional storage or computation is not necessary in our case for creating the data pyramid.

This approach to MRF-based multiscale image analysis has other advantages :
- it is globally consistent and is related to constrained optimization,
- it can be generalized to other constraints : in the present paper a piece-wise constant constraint on configurations has been introduced. Other examples (bilinear constraints for instance) can be found in [34].
- as already said, it involves only a label pyramid but does not require a pyramid of observations to be constructed : the observations at full resolution are used,

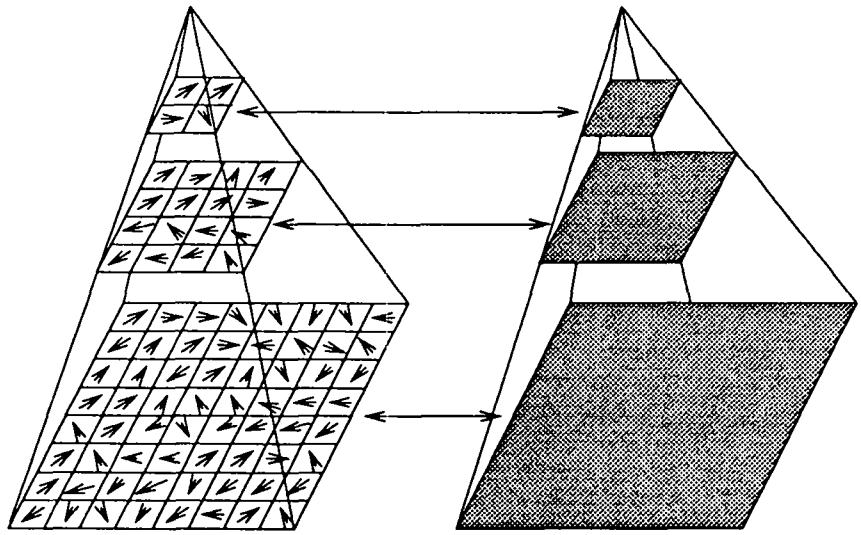

Figure 8: *The standard multigrid estimation structure*

• it is easy to implement and the multiscale models can be computed straightforwardly,
• it exhibits fast convergence properties when applied to non-linear optimization problems (see next section),
• as will be seen, it yields also, in many cases, better results (convergence to lower energy values) than standard multigrid methods.

4 Application to visual motion analysis

The multiscale algorithm has been implemented on the MRF models presented in section 2 for motion detection, optical flow measurement and motion-based segmentation. Four algorithms have been compared :

1. a standard monoresolution deterministic relaxation algorithm (**DR**), known as ICM, [3];

2. a monoresolution stochastic relaxation (**SR**) algorithm based on the Gibbs sampler, [17] ,

3. a standard coarse-to-fine multigrid relaxation (**MGR**) algorithm in which the same model is considered at each resolution (same parameters, same neighborhood system and same potential functions) and a pyramid of observations is constructed. A similar implementation of this scheme may be found in [24];

4. the proposed multiscale relaxation (**MSR**), in which the energy functions at each resolution level are computed in a consistent way as explained previously.

The temperature schedule used in the stochastic relaxation algorithm was : $T(j) = T_0.A^j$, with $A = 0.97$ and $T_0 = 300$, where j designates the number of sweeps on the image. The same parameters were chosen in every case for the finest resolution model. Up to four resolution levels were considered in the multigrid and multiscale methods.

For the multiresolution schemes (**MGR** and **MSR**), the same deterministic algorithm (ICM, [3]) was run at the different resolutions. The relaxation was performed until convergence was reached on the label fields (at each resolution). As a matter of fact, the process was stopped as soon as the number of sites changing their state between two complete image sweeps went below some specified threshold λ.

The equations giving the multiscale energies for motion detection, optical flow measurement and motion-based segmentation are derived from equations (21), (22), (25), (26) and (28). These multiscale models and the experimental results obtained with this new class of MRF are presented now.

4.1 Motion detection

4.1.1 Expression of the multiscale model

The multiscale model is derived from the single resolution model using equations (21), (22), (25), (26) and (28). The energy function on the coarse grid S^i is obtained from the equivalent energy function defined on the blockwise constant configuration space Ω_i, thanks to Equ. (28). The energy at coarse scale i can be developped as a sum of potential functions depending on cliques associated to an 8-neighborhood system defined on the coarse grid S^i. They depend on the full resolution observation field, as explained previously. The global energy at scale i becomes :

$$U^i(e_t^i, o_t, o_{t+\delta t}, \bar{o}_t, \bar{o}_{t+\delta t}) \triangleq U_{11}^i(e_t^i, o_t, o_{t+\delta t}) + U_{12}^i(e_t^i, \bar{o}_t, \bar{o}_{t+\delta t}) + U_2^i(e_t^i) \tag{32}$$

where :

$$U_{11}^i(e_t^i, o_t, o_{t+\delta t}) \triangleq \sum_{s \in S^i} \sum_{r \in B_s^i} \frac{[o_{t+\delta t}(r) - \Psi(e_t^i(s))]^2 + [o_t(r) - \Psi(e_t^i(s))]^2}{2\sigma^2} \tag{33}$$

$$U_{12}^i(e_t^i, \bar{o}_t, \bar{o}_{t+\delta t}) \triangleq \sum_{s \in S^i} \sum_{r \in B_s^i} V_\tau(e_t^i(s), \bar{o}_t(r), \bar{o}_{t+\delta t}(r)) \tag{34}$$

where the potential V_τ is defined in Table 1 and B_s^i denotes the block of the original grid on which $\Phi^i(e_t^i(s))$ is defined (see Fig. 5).

Energy U_2^i can be expressed as follows :

$$U_2^i(e_t^i) \triangleq \sum_{c \in C^i} V_c^i(e_t^i) \tag{35}$$

where C^i is the clique set associated to the 8-neighborhood system at scale i. The coarse scale potentials V_c^i are zero for cliques c of more than two sites. If $\mathcal{C}_h^i$ (resp. $\mathcal{C}_v^i$, and $\mathcal{C}_d^i$) denotes horizontal (resp. vertical and diagonal) doubleton (i.e. two elements) cliques, the potential functions of the coarse model are the following :

$$\begin{cases} for\ s \in S^i & V^i_{\{s\}}(e^i_t) \triangleq -p^i.\beta_s \\ for\ \{s_1,s_2\} \in \mathcal{C}_h^i \cup \mathcal{C}_v^i & V^i_{\{s_1,s_2\}}(e^i_t) \triangleq \begin{cases} +q^i_{hv}.\beta_s & if\ \ e^i_t(s_1) \neq e^i_t(s_2) \\ -q^i_{hv}.\beta_s & if\ \ e^i_t(s_1) = e^i_t(s_2) \end{cases} \\ for\ \{s_1,s_2\} \in \mathcal{C}_d^i & V^i_{\{s_1,s_2\}}(e^i_t) \triangleq \begin{cases} +q^i_{d}.\beta_s & if\ \ e^i_t(s_1) \neq e^i_t(s_2) \\ -q^i_{d}.\beta_s & if\ \ e^i_t(s_1) = e^i_t(s_2) \end{cases} \end{cases} \quad (36)$$

with $p^i = 2(2^i - 1)(2^{i+1} - 1)$, $q^i_{hv} = 2^i + 2(2^i - 1)$ and $q^i_d = 1$.

One can notice that for $i > 0$ the model becomes anisotropic. Indeed the diagonal clique parameter remains constant through scale whereas the one for horizontal and vertical cliques takes increasing values with scale.

4.1.2 Experimental results

The contribution of the multiscale algorithm is illustrated on a traffic scene (Fig. 9a) consisting of a highway observed from a bridge. The camera is static and the scene involves several moving cars with different apparent velocities and sizes in the image. In this problem, the final energies values obtained by the different algorithms are very close, which suggests that the energy landscape must be quite smooth (Table 2). Threshold λ (for testing convergence) was let to 1. The following set of parameters was used : $\sigma^2 = 1$, $\beta_s = 20, \beta_\tau = 100$ and $\mu = 10$. As expected, stochastic relaxation performs slightly better than the deterministic schemes, but to the detriment of computational efficiency. Besides, the standard **MGR** algorithm is significantly slower than the multiscale scheme **MSR**. As far as the final energies are concerned, the slight difference between **MGR** and **MSR** cannot be considered as significant for this example. However, one can notice a great difference in the quality of the final detection maps. For the **MGR** method, since the parameters of the model are not adjusted through scale, two moving vehicles at the top of the image have been lost in the final result (Fig. 10c). This is mainly due to a partial overlapping of the masks of these objects between two successive images. Since the objects are quite homogeneous, the detected temporal changes are weak[2]. Because the constant parameters are no longer adapted at low resolution, the model filters out those masks in the standard multigrid approach. This unpleasant behavior is a limitation of the standard multigrid algorithm because the parameters tuned for one scale are usually valid only for a restricted class of images characterized by specific spatial statistics which actually *are not scale invariant.* This is not the case

[2]More sophisticated models taking into account overlapping situations are described in [27]

for the multiscale approach which adapts parameters and model structure through scale. Results show almost the same accuracy as the monoresolution scheme for a reduced computational cost (Table 2) : 22 iterations were necessary for monoresolution deterministic relaxation whereas only 7.25 equivalent iterations are required in the multiscale approach with comparable results.

algorithm	SR	DR	MGR	MSR
nb-iter	294	22	17.8	7.25
final energy	-4032523	-4013916	-4021817	-4017871

Table 2 : Motion detection (*highway* sequence) : number of iterations and final energy considering the mono and multiresolution algorithms (SR : stochastic relaxation, DR : deterministic relaxation, MGR : standard multigrid relaxation, MSR : multiscale relaxation) *nb-iter* : computational equivalent to the number of complete sweeps through the image at the finest resolution

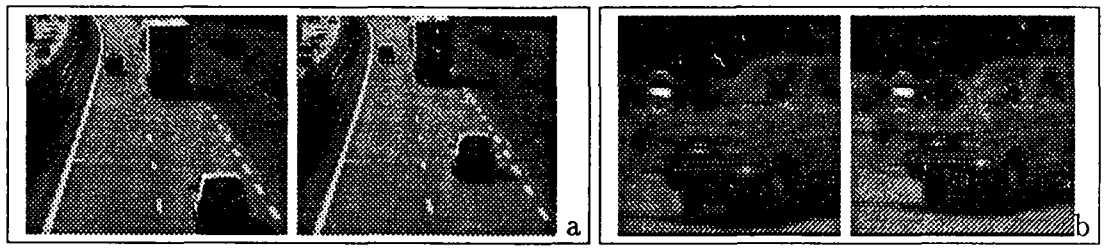

Figure 9: *Two frames from the original sequences : (a) "highway" sequence (256x256), (b) "camera zoom" sequence (100x100)*

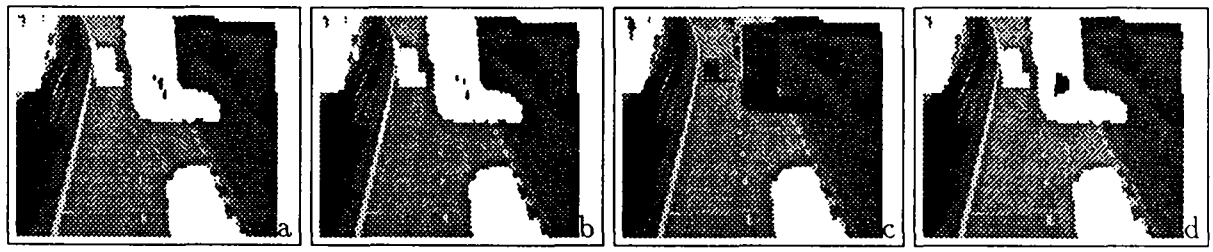

Figure 10: *Results of the detection algorithm ("highway" sequence) : (a) SR (b) DR (c) MGR (d) MSR*

4.2 Visual motion measurement

4.2.1 Expression of the multiscale model

In the case of the optical flow measurement model (section 2.2.2), the resulting multiscale energy function is given by :

$$U^i(\vec{\omega}_t^i, f_t, f_{t+\delta t}) \triangleq U_1^i(\vec{\omega}_t^i, f_t, f_{t+\delta t}) + U_2^i(\vec{\omega}_t^i) . \tag{37}$$

where :

$$U_1^i(\vec{\omega}_t^i, f_t, f_{t+\delta t}) \triangleq \sum_{s \in S^i} \sum_{r \in B_s^i} \{ f_t(r) - f_{t+\delta t}(r + \vec{\omega}_t^i(s).\delta t) \}^2 \tag{38}$$

and :

$$U_2^i(\vec{\omega}_t^i) \triangleq \sum_{c \in \mathcal{C}^i} V_c^i(\vec{\omega}_t^i) \tag{39}$$

where a 8-neighborhood structure is defined on S^i. The coarse scale potentials V_c^i are zero for cliques c of more than two sites. Let again $\mathcal{C}_h^i$ (resp. $\mathcal{C}_v^i$, and $\mathcal{C}_d^i$) denote horizontal (resp. vertical and diagonal) doubleton cliques. The potential functions V_c^i are defined as following :

$$\begin{cases} for\ c \in \mathcal{C} - (\mathcal{C}_h^i \cup \mathcal{C}_v^i \cup \mathcal{C}_d^i) & V_c^i(\vec{\omega}_t^i) \triangleq 0 \\ for\ \{s_1, s_2\} \in \mathcal{C}_h^i \cup \mathcal{C}_v^i & V_{\{s_1,s_2\}}^i(\vec{\omega}_t^i) \triangleq q_{hv}^i.\alpha \, \|\vec{\omega}_t^i(s_1) - \vec{\omega}_t^i(s_2)\|^2 \\ for\ \{s_1, s_2\} \in \mathcal{C}_d^i & V_{\{s_1,s_2\}}^i(\vec{\omega}_t^i) \triangleq q_d^i.\alpha \, \|\vec{\omega}_t^i(s_1) - \vec{\omega}_t^i(s_2)\|^2 \end{cases} \tag{40}$$

with $q_{hv}^i = 2^i + 2(2^i - 1)$ and $q_d^i = 1$.

As in the case of the motion detection model, an anisotropy on the coarse scale potential functions is observed.

4.2.2 Experimental results

In this case our experiments clearly show that the discrete formulation of motion measurement (equ. 11) yields an energy function which contains many local minima. The results reported here have been obtained with a random initialization for the label field. Two sequences are first presented : the first one shows a simulated camera zoom on a real image (Fig. 9b) which corresponds to a perfect divergent motion; the second one is the already described highway scene (Fig. 9a).

In Fig. 11 and 13 are drawn the velocity vector fields estimated by each of the four algorithms. The multiscale approach supplies fields close in quality to the one obtained by time consuming stochastic relaxation. Table 3 gives the number of iterations required in each case and the final energies. The global behavior of the different schemes can be appreciated by considering the energy plots reported in Fig. 12.

algorithm	SR	DR	MGR	MSR
nb-iter	468	42	8.75	5.38
final energy	87292	139042	116502	89355

camera zoom sequence

algorithm	SR	DR	MGR	MSR
nb-iter	259	32	16.88	7.46
final energy	766427	1117945	811468	749471

highway sequence

Table 3 : Motion measurement : number of iterations in the mono and multiresolution algorithms (SR : stochastic relaxation, DR : deterministic relaxation, MGR : standard multigrid relaxation, MSR : multiscale approach)
nb-iter : computational equivalent to the number of complete sweeps through the image at the finest resolution

The stochastic relaxation algorithm leads to the lowest energy value in the case of the "zoom" but requires 468 iterations to converge (Table 3). The proposed multiscale approach gives results close to the stochastic method with a gain of almost two orders of magnitude on the number of iterations. The standard methods (**DR, MGR**) remain stuck in rather high energy local minima. In the case of the "highway sequence" the stochastic relaxation performed even worse than the multiscale approach, when adopting the same temperature schedule as for the "zoom" sequence (Table 3).

These results have been confirmed by processing a benchmark of 41 sequences of two 64x64 frames obtained by applying the same synthetic motion (Fig. 14) to different real images (Fig. 15). The synthetic motion included translation, rotation and dilation on the same image (Fig. 14). Fig. 16 presents, for these 41 short sequences, the ratio of the final energy value reached by stochastic relaxation (**SR**) to the one obtained by the different other methods. The corresponding number of iterations required in each case is plotted in Fig. 17. In the average, the **MSR** method finds configurations close to the best estimates obtained by **SR**, with an improvement of up to two orders of magnitude on the number of iterations.

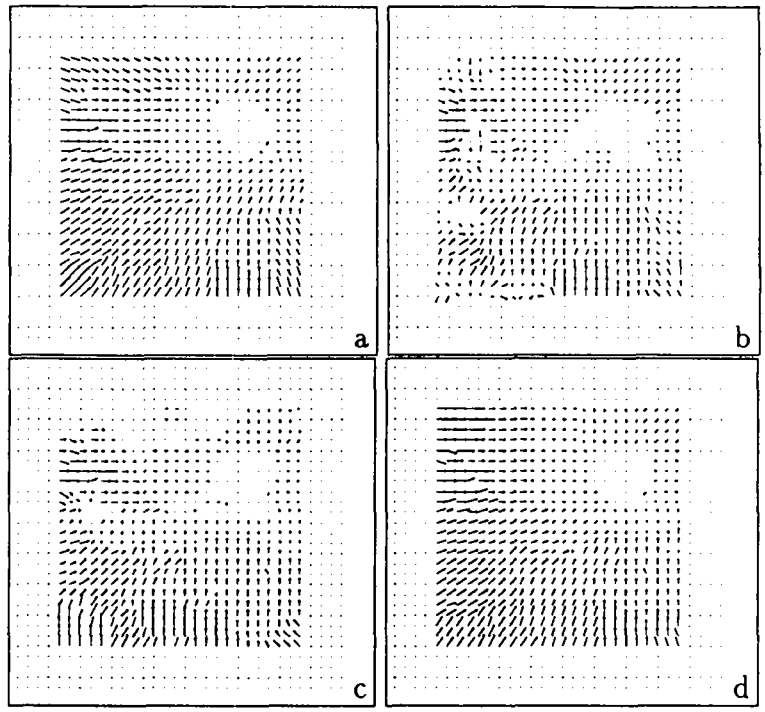

Figure 11: *Optical flow fields ("camera zoom" sequence) : (a) SR (b) DR (c) MGR (d) MSR*

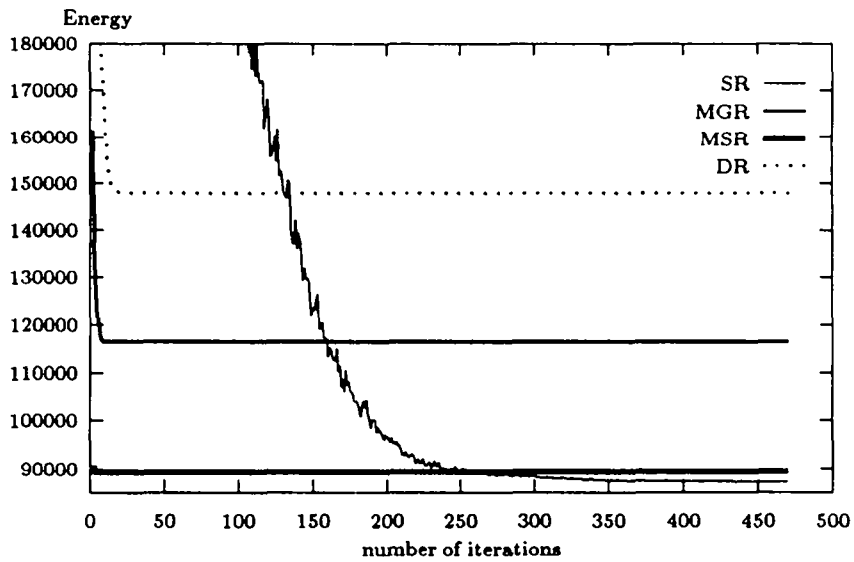

Figure 12: *Energy U versus iteration number : "camera zoom" sequence*

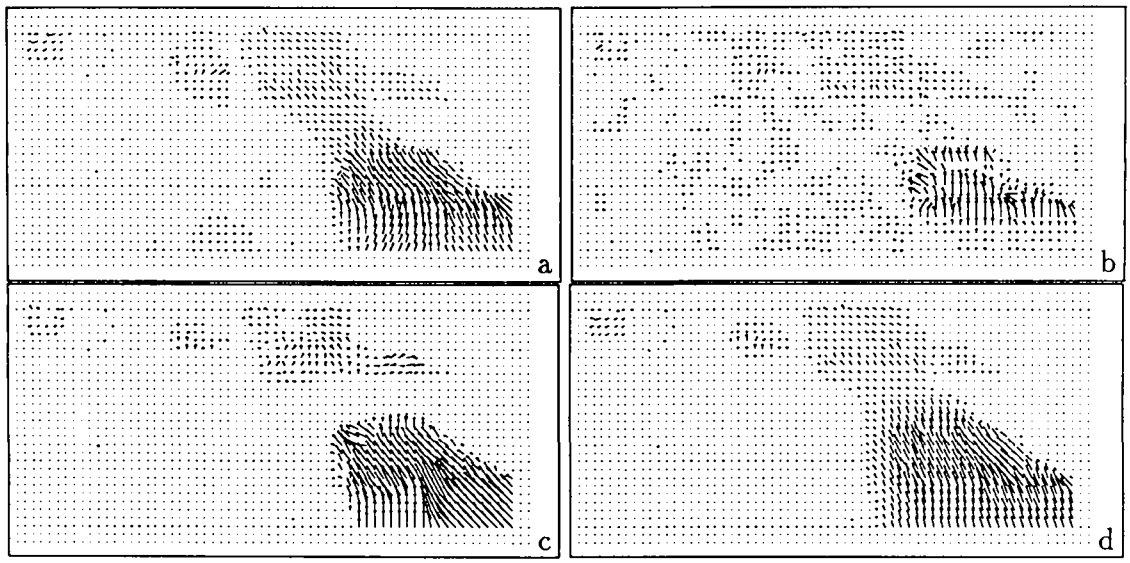

Figure 13: *Optical flow fields ("highway sequence") : (a) SR (b) DR (c) MGR (d) MSR*

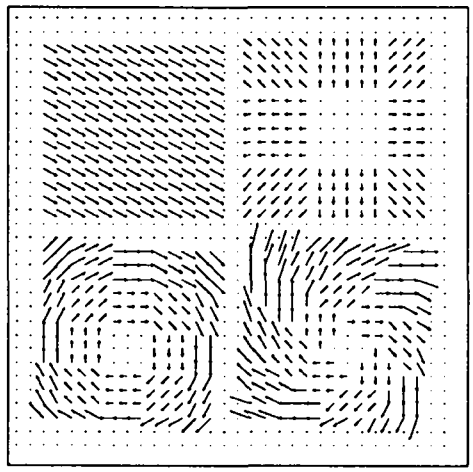

Figure 14: *Synthetic optical flow field applied on the 41 sequences*

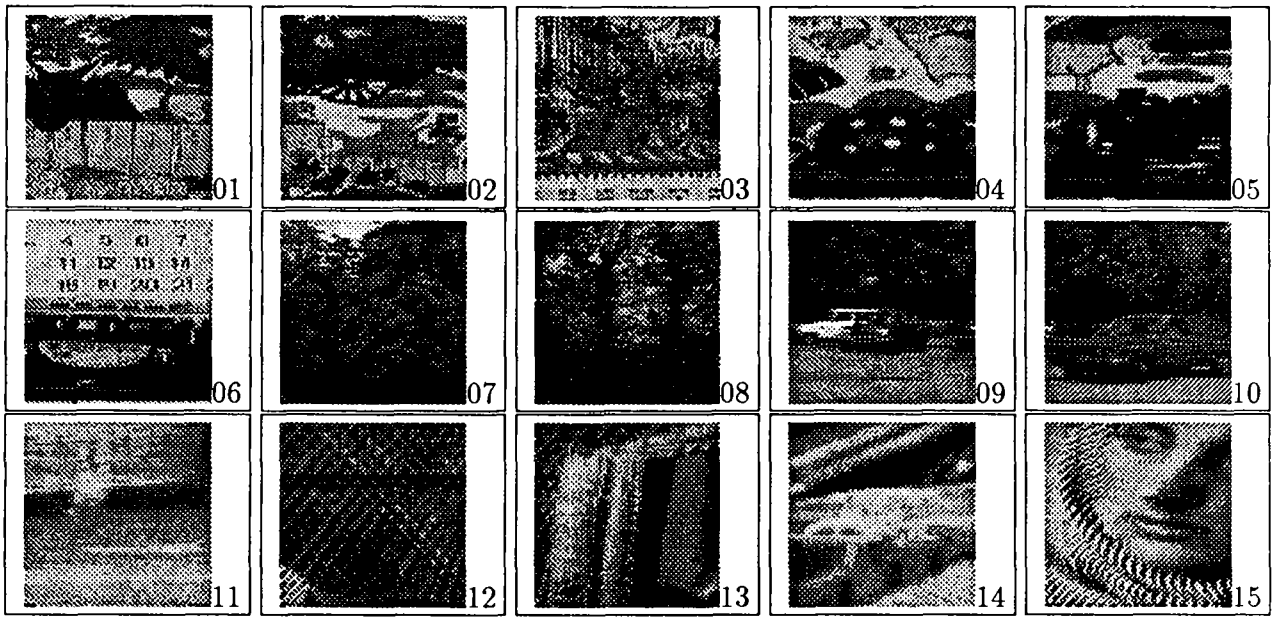

Figure 15: *Example of first frames in the 41 short sequences benchmark*

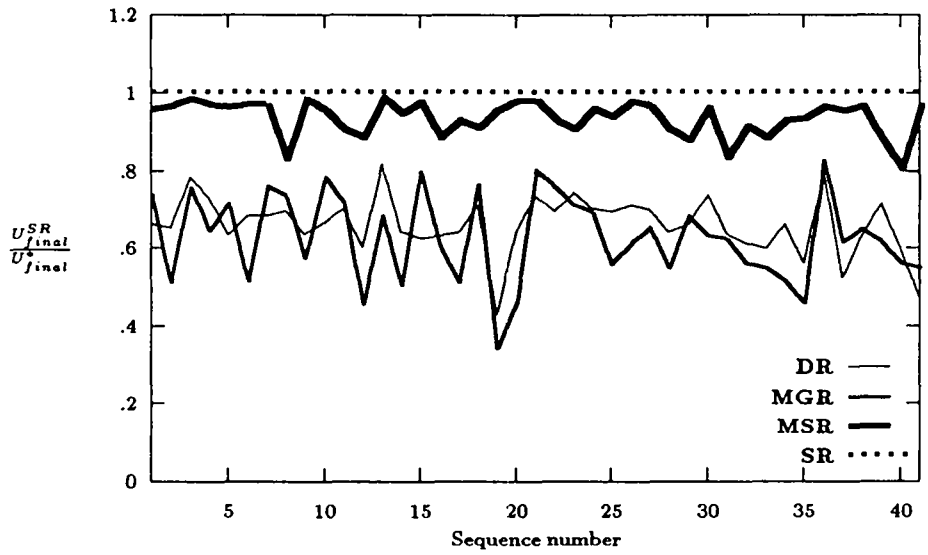

Figure 16: *Ratio of the final energy value reached with SR (U^{SR}_{final}) to the one obtained with the other methods (U^{*}_{final})*

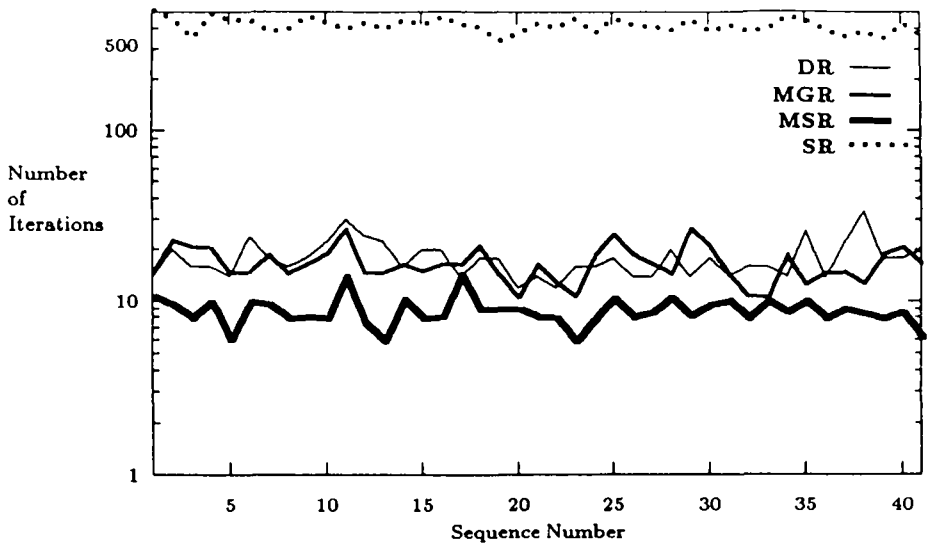

Figure 17: *Number of iterations at convergence*

4.3 Motion-based segmentation

4.3.1 Expression of the multiscale model

The equations giving the multiscale energies for motion-based segmentation are the following :

$$U^i(\nabla f, \frac{\partial f}{\partial t}, e^i, \theta, R) = U_1^i(\nabla f, \frac{\partial f}{\partial t}, e^i, \theta, R) + U_2^i(e^i) \tag{41}$$

with

$$U_1^i(\nabla f, \frac{\partial f}{\partial t}, e^i, \theta, R) = \sum_{s \in S^i} \sum_{r \in B_s^i} \left(\frac{1}{2\sigma^2} \left(\nabla f(r).\vec{\omega}_{\theta_{e_s^i}}(r) + \frac{\partial f}{\partial t}(r) \right)^2 \right) \tag{42}$$

$$U_2^i(e^i) = \sum_{\{s_1,s_2\} \in C^i} V_2^i(e_{s1}^i, e_{s2}^i) = \sum_{\{s_1,s_2\} \in C^i} p_{\{s_1,s_2\}}^i . \mu(1 - \delta_{e_{s1}^i = e_{s2}^i}) \tag{43}$$

where $p_{\{s_1,s_2\}}^i = 1$ when the clique is diagonal and $p_{\{s_1,s_2\}}^i = 2^i + 2(2^i - 1)$ when the clique is horizontal or vertical. They are derived from equations (21), (22), (25), (26) and (28). The parameter vectors θ_r associated to each region is computed on the full resolution configuration $\Phi^i(e^i)$.

4.3.2 Experimental results

For the motion segmentation model, only three algorithms have been implemented and compared :

1. the monoresolution deterministic relaxation algorithm (**DR**) described in subsection 2.2;
2. a standard coarse-to-fine multigrid relaxation (**MGR**) algorithm as explained previously;
3. our multiscale relaxation (**MSR**) algorithm.

It was not possible, in this case, to implement the stochastic relaxation algorithm since the number of regions, hence the global configuration space Ω, was not known *a priori* (this limitation does not exist for the multiscale method).

The same parameters were chosen in each case for the finest resolution model. Three resolution levels were considered in the multigrid and multiscale methods. For these schemes the same deterministic algorithm (described in section 2.2.3) was run at the different resolutions. The relaxation was performed until convergence was reached on the label fields (at each resolution). As a matter of fact, the process was stopped as soon as the number of sites changing their state between two complete image sweeps went below some specified threshold λ.

Figure 18: *three frames from the* parking lot *sequence (200x200) at time* t_1, t_5 *and* t_9.

Figure 19: *three frames from the* traffic *sequence (200x256) at time* t_1, t_3 *and* t_6.

We report results on two real world outdoor sequences (the *parking lot* sequence, Fig. 18 and the *traffic sequence*, Fig. 19). In both sequences the camera is moving.

The *parking lot* sequence, (Fig. 18), contains two manoeuvring cars. The first car in the foreground is rotating and moving toward the camera. The second one in the background is translating toward the camera. It partially hides a third car

which is static. The camera pans the scene from right to left. The image background consists of trees, some of them being stirred by the wind. The parameters chosen for this sequence are $\Phi = 15$ and $A = 150$.

In the *traffic* sequence, (Fig. 19), the camera is mounted on a car approaching a street crossing. A second car is moving across in the field of view from left to right. Another apparently moving object in the image is the traffic light on the right side which undergoes a different apparent motion than the background because of the significant difference in depth. The parameters chosen for this sequence are $\Phi = 40$ and $A = 200$.

The resulting segmentation for the two first frames of each sequence are shown in Fig. 20 and 21, for the different described algorithms. The corresponding number of iterations required in each case and the final energies are given in Table 4.

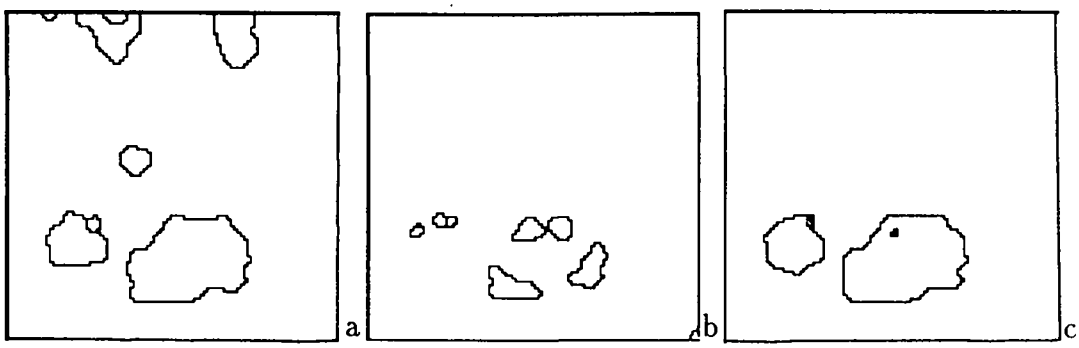

Figure 20: *segmentation map obtained for the* parking lot *sequence (200x200) a) deterministic monoresolution relaxation (DR) b) standard multigrid relaxation (MGR) c) multiscale relaxation (MSR)*

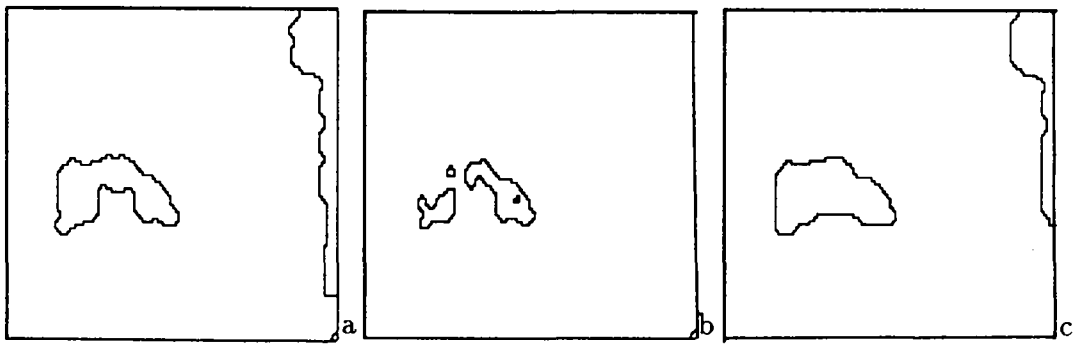

Figure 21: *segmentation map obtained for the* traffic *sequence (200x256) a) deterministic monoresolution relaxation (DR) b) standard multigrid relaxation (MGR) c) multiscale relaxation (MSR)*

In both cases the **MSR** algorithm leads to final segmentation maps with lower energy than the configurations obtained by the standard algorithms (**DR** or **MGR**). In the case of the *traffic* sequence, **MGR** remains stuck in a local minimum of high

algorithm	DR	MGR	MSR	
nb-iter	13	5.95	5.62	*parking lot sequence*
final energy	-573840	-569856	-583748	

algorithm	DR	MGR	MSR	
nb-iter	12	11.31	4.94	*traffic sequence*
final energy	-179998	+1030509	-217742	

Table 4 : Motion segmentation : number of iterations required to reach convergence and final energy values (nb-iter : computational equivalent to the number of complete sweep through the image at the finest resolution)

positive energy (see Table 4 and Fig. 21). In our experiments the standard coarse-to-fine mutigrid relaxation **MGR** gave generally results of lower quality than the multiscale method. This can be explained by the highly non-linear energy function which has to be minimized here : at low resolutions the **MGR** algorithm converges to rough configurations and the deterministic relaxation scheme is unable to get out of these configurations while proceeding to finer resolution levels. This is not the case of the **MSR** method which makes use of the *full resolution* observations, even at the coarsest scale, and thus provides segmentations of better quality (compare Fig. 20b and 20c, Fig. 21b and 21c). In the case of the *traffic* sequence the **MSR** is the only one which recovers the complete shape of the moving car. In the case of the *parking lot* sequence the two moving cars are neatly extracted.

Besides, in the case of **MSR**, the number of iterations required to converge is reduced by half with respect to standard monoresolution relaxation **DR** (Table 4). In the case of the traffic sequence, convergence of **MGR** is as slow as the one of the monoresolution relaxation.

Those results, obtained on different MRF models demonstrate the efficiency of the proposed multiscale paradigm when an energy function, presenting a great number of local minima, has to be minimized. We are currently extending the method towards :

- more general constraints (bilinear constraints are reported in [34]),
- a coupled MRF model, generating a complex energy landscape, enabling both continuous motion measurement, [20] and motion-based segmentation, [14].
- a multiscale scheme leading to a temporaly-linked segmentation for a long sequence.

Besides, we also study general parallelization approaches associated to the multiscale algorithm. This algorithm is indeed a good candidate for parallel implementation on regular arrays consisting of locally interconnected processors, [30].

5 Conclusion

In this paper we have described a new multiscale approach for the multigrid analysis of images based on Markov Random Field modeling. The approach relies on constrained optimization and has been demonstrated on several relevant issues in image sequence analysis : motion detection, motion measurement and motion-based segmentation. The multiscale MRF approach presents several attractive features when compared to standard multigrid schemes : mathematical consistency, simple implementation. It leads to efficient global optimization algorithms. Gains of up to two orders of magnitudes in the convergence speed with respect to simulated annealing have been observed. The estimates are close in quality to those obtained with time consuming stochastic relaxation algorithms in the case of optical flow measurement. The motion-based segmentation maps obtained with the multiscale scheme appear also more robust than those extracted with standard methods.

Let us emphasize that the proposed approach is general and can be applied to other low-level image analysis problems based on discrete global statistical models. It is adapted to sophisticated non-linear models which have been developed recently in image analysis for instance in image segmentation, [16] or structure from motion, [36]. Finally, this algorithm is a good candidate for parallel implementation on regular arrays consisting of locally interconnected processors, [30].

Acknowledgments

This work has been partly supported by MRT (French Ministery of Research and Technology) and CNRS in the context of the PRC Program "Man Machine Interface" under contract PMFE 88F1 548, by MRT in the context of the EUREKA european project PROMETHEUS, under PSA-contract VY/85241753/14/Z10 and by Brittany County Council under contribution to student grant. We thank B. Levy and E. François for fruitful discussions.

References

[1] J.K. AGGARWAL and N. NANDHAKUMAR. – On the computation of motion from sequences of images - a review. – *Proc. IEEE*, Vol. 76, No 8: pages 917–935, 1988.

[2] S.T. BARNARD. – Stochastic stereo matching over scale. – *Int. J. Comp. Vis.*, Vol 3: pages 17–32, 1989.

[3] J. BESAG. – On the statistical analysis of dirty pictures. – *J. Royal Statist. Soc.*, Vol. 48, Serie B, No 3: pages 259–302, 1986.

[4] A. BLAKE and A. ZISSERMAN. – *Visual Reconstruction.* – MIT Press, 1987.

[5] C. BOUMAN and B. LIU. – Multiple resolution segmentation of textured images. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 13, No 2: pages 99–113, Feb. 1991.

[6] P. BOUTHEMY. – Extracting dense motion information from an image sequence : Optic flow estimation and related segmentation issues. – In G.E. Taylor, editor, *Kinematic and Dynamic Issues on Sensor Based Control*, volume NATO ASI Series, Vol F57, pages 223–250. Springer, 1990.

[7] P. BOUTHEMY and E. FRANCOIS. – Motion segmentation and qualitative dynamic scene analysis from an image sequence. – *submitted to Int. Jal. Computer Vision*, June 1991.

[8] P. BOUTHEMY and P. LALANDE. – Motion detection in an image sequence using Gibbs distributions. – In *Proc. 14th Int. Conf. Acoustics, Speech and Signal Processing*, pages 1651–1654, Glasgow, 1989.

[9] B. CHALMOND. – Image restoration using an estimated Markov model. – *Signal Processing*, Vol. 15, No 2: pages 115–129, September 1988.

[10] P.B. CHOU and C.M. BROWN. – The theory and practice of bayesian image modeling. – *Int. J. Comp. Vis.*, Vol 4: pages 185–210, 1990.

[11] P.A. DEVIJVER. – Real-time modeling of image sequences based on hidden Markov Mesh random field models. – Technical Report M-307, Philips Research Lab., June 1989.

[12] J.M. DINTEN. – Tomographic reconstruction of axially symmetric objects : Regularization by a markovian modelization. – In *Proc. 10th Int. Conf. Pattern Recognition*, volume 2, pages 153–158, Atlantic City, June 1990.

[13] W. ENKELMANN. – Investigations of multigrid algorithms for the estimation of optical flow fields in image sequences. – *Comput. Vision, Graphics, Image Processing*, Vol 43: pages 150–177, 1988.

[14] E. FRANCOIS and P. BOUTHEMY. – Multiframe-based identification of mobile components of a scene with a moving camera. – In *IEEE Int. Conf. Computer Vision Pattern Recognition*, pages 166–172, Hawaii, June 3-6 1991.

[15] E. FRANCOIS and P. BOUTHEMY. – Temporally-linked motion-based segmentation in a long image sequence. – In *7th Scandinavian Conf. Image Analysis*, pages 1134–1141, Aalborg, Denmark, August 1991.

[16] D. GEMAN, S. GEMAN, C. GRAFFIGNE, and D. PONG. – Boundary detection by constrained optimization. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 12, No 7: pages 609–628, July 1990.

[17] S. GEMAN and D. GEMAN. – Stochastic relaxation, Gibbs distributions and the bayesian restoration of images. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 6, No 6: pages 721–741, November 1984.

[18] B. GIDAS. – A renormalization group approach to image processing problems. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol 11 No 2: pages 164–180, February 1989.

[19] F. GLAZER. – Multilevel relaxation in low-level computer vision. – In A. Rosenfeld, editor, *Multiresolution image processing and analysis*, pages 312–330. Springer, 1984.

[20] F. HEITZ and P. BOUTHEMY. – Multimodal motion estimation and segmentation using Markov random fields. – In *Proc. 10th Int. Conf. Pattern Recognition*, volume 1, pages 378–383, Atlantic City, June 1990.

[21] F. HEITZ and P. BOUTHEMY. – Multimodal estimation of discontinuous optical flow using Markov Random Fields. – Technical Report 1367, INRIA-Rennes, submitted to IEEE Trans. PAMI, January 1991.

[22] F. HEITZ, H. MAITRE, and C. de COUESSIN. – Event detection in multisource imaging using contextual estimation. – In *Proc. Int. Conf. Acoust., Speech, Signal Processing*, pages 1647–1650, Glasgow, May 1989.

[23] J. HUTCHINSON, C. KOCH, J. LUO, and C. MEAD. – Computing motion using analog and binary resistive networks. – *Computer*, Vol. 21: pages 52–63, March 1988.

[24] J. KONRAD and E. DUBOIS. – Multigrid Bayesian estimation of image motion fields using stochastic relaxation. – In *Proc. 2nd Int. Conf. Computer Vision*, pages 354–362, Tarpon Springs, Florida, Dec. 1988.

[25] J. KONRAD and E. DUBOIS. – A comparison of stochastic and deterministic solution methods in bayesian estimation of 2D motion. – In *Proc. First European Conference on Computer Vision*, pages 149–160, Antibes, France, April 1990. Springer.

[26] S. LAKSHMANAN and H. DERIN. - Gaussian markov random fields at multiple resolutions. - In R. Chellappa and A.K. Jain, editors, *Markov Random Fields : Theory and Applications*. Academic Press, 1991.

[27] P. LALANDE and P. BOUTHEMY. - A statistical approach to the detection and tracking of moving objects in an image sequence. - In *Proc. Conf. EUSIPCO'90*, pages 947–950, Barcelona, 1990.

[28] B.C. LEVY. - Multiscale models and estimation of discrete Gauss-Markov random fields. - In *2nd SIAM Conf. Linear Algebra in Systems Control and Signal Processing*, San Francisco, California, Nov. 1990.

[29] J.L. MARROQUIN, S. MITTER, and T. POGGIO. - Probabilistic solution of ill-posed problems in computational vision. - *J. American Statis. Assoc.*, Vol 82: pages 76–89, 1987.

[30] E. MEMIN, F. CHAROT, and F. HEITZ. - Parallel algorithms and architectures for multiscale Markov Random Field-based image analysis. - In *Workshop on Computer Architecture for Machine Perception*, pages 309–320, Paris, Dec. 1991.

[31] J.W. MODESTINO and J. ZHANG. - A Markov Random Field model-based approach to image interpretation. - In *IEEE Conf. Comp. Vision Pattern Rec.*, pages 458–465, June 1989.

[32] D.W. MURRAY and H. BUXTON. - Scene segmentation from visual motion using global optimization. - *IEEE Trans. Pattern Anal. Machine Intell.*, Vol 9, No 2: pages 220–228, March 1987.

[33] NAGEL, H.H. - Image sequences - ten (octal) years - from phenomenology towards a theoretical foundation. - *Int. J. Pattern Rec. Artif. Intel.*, Vol 2, No 3: pages 459–483, 1988.

[34] P. PEREZ and F. HEITZ. - Multiscale markov random fields and constrained relaxation in low level image analysis. - In *Proc. Int. Conf. Acoust., Speech, Signal Processing*, San Francisco, to appear, March 1992.

[35] SCHUNCK, B.G. - The image flow constraint equation. - *Computer Vision, Graphics and Image Processing*, Vol. 35: pages 20–46, 1986.

[36] J. SUBRAHMONIA, Y.P. HUNG, and D.B. COOPER. - Model-based segmentation and estimation of 3D surfaces from two or more intensity images using Markov random fields. - In *Proc. 10th Int. Conf. Pattern Recognition*, volume 1, pages 390–397, Atlantic City, June 1990.

[37] D. TERZOPOULOS. – Image analysis using multigrid relaxation methods. – *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 8, No 2: pages 129–139, March 1986.

[38] YUILLE A.L. and BULTHOFF H. – Stereo integration, mean field theory and psychophysics. – In *Proc. First European Conference on Computer Vision*, pages 73–82, Antibes, France, April 1990. Springer.

LISTE DES PUBLICATIONS INTERNES IRISA 1992

PI 624 SIGNAL AS A MODEL FOR REAL-TIME AND HYBRID SYSTEMS
Albert BENVENISTE, Michel LE BORGNE, Paul LE GUERNIC
Janvier 1992, 22 pages.

PI 625 ON THE CENTRAL-LIMIT THEOREM FOR TRACKING ESTIMATORS WITH SMALL GAIN - INFINITE HORIZON CASE
Bernard DELYON, Anatoli JUDITSKY
Janvier 1992, 16 pages.

PI 626 A MONTE CARLO METHOD BASED ON ANTITHETIC VARIATES FOR NETWORK RELIABILITY COMPUTATIONS
Mohamed EL KHADIRI, Gerardo RUBINO
Janvier 1992, 28 pages.

PI 627 CONSTRAINED MULTISCALE MARKOV RANDOM FIELDS AND THE ANALYSIS OF VISUAL MOTION
Fabrice HEITZ, Patrick PEREZ, Patrick BOUTHEMY
Janvier 1992, 40 pages.

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique.

ISSN 0249-6399