CVGIP: IMAGE UNDERSTANDING
Vol. 59, No. 1, January, pp. 125–134, 1994

NOTE

Multiscale Minimization of Global Energy Functions in Some Visual Recovery Problems

F. HEITZ, P. PEREZ* AND P. BOUTHEMY

*IRISA/INRIA, *IRISA/CNRS, Campus Universitaire de Beaulieu, 35042 Rennes Cedex, France*

Received January 23, 1992; revised November 12, 1992

Many image analysis and computer vision problems have been expressed as the minimization of global energy functions describing the interactions between the observed data and the image representations to be extracted in a given task. In this note, we investigate a new comprehensive approach to minimize global energy functions using a multiscale relaxation algorithm. The energy function is minimized over nested subspaces of the original space of possible solutions. These subspaces consist of solutions which are constrained at different scales. The constrained relaxation is implemented via a coarse-to-fine multiresolution algorithm that yields fast convergence towards high quality estimates when compared to standard monoresolution or multigrid relaxation schemes. It also appears to be far less sensitive to local minima than standard relaxation algorithms. The efficiency of the approach is demonstrated on a highly nonlinear combinatorial problem which consists of estimating long-range motion in an image sequence on a discrete label space. The method is compared to standard relaxation algorithms on real world and synthetic image sequences. © 1994 Academic Press, Inc.

1. INTRODUCTION

Many tasks in computer vision and image analysis have recently been expressed as global optimization problems [3, 5, 8]. The general issue is to find the global minimum of an objective (also called energy) function which describes the interaction between the different variables modeling the image features in a given problem [8]. Two kinds of variables are generally considered: *observation variables* which correspond to the representation of the observed data and *hidden variables* (or labels) which are the representations to be extracted from the original images. Energy functions involve generally two components, one of which expresses the interaction between the hidden labels and the observations and the other which encodes constraints on the desired solution [8]. To keep the problem tractable, the energy is often decomposed as a sum of local interaction functions defined on a neighborhood [8]. Standard regularization approaches as well as MRF-based image analysis lead to the minimization of such global energy functions [8] (Markov random field models define a global probability distribution that is maximized when an energy function is minimized [9]). The choice of these energy functions is either heuristic or may be guided by a statistical modeling of the interaction between the variables.

Defining global energy functions is a powerful tool for specifying nonlinear interactions between different image features (luminance, edges, region labels, etc.). They help to combine and organize spatial and temporal information by introducing strong generic knowledge about the features to be estimated. For instance, global energy functions have been successfully introduced in image restoration [2, 9], edge detection [8], image segmentation [8], stereovision [1], computed tomography, surface reconstruction [5], visual motion analysis [7, 10, 12, 13, 16], and scene interpretation [15].

However, minimizing a global energy function is often an intricate problem: the number of possible label configurations is generally *very large* and the global energy function may exhibit many local minima. Computationally demanding stochastic relaxation algorithms are generally necessary to compute optimal solutions. Less cpu intensive deterministic relaxation algorithms such as ICM [2], HCF (highest confidence first [5]), or GNC (graduated nonconvexity [3]) can often be used instead, when a good initial guess is available. Deterministic approaches converge to configurations corresponding to local minima of the global energy function. On the other hand, it is known that multigrid methods can significantly improve the convergence rate of iterative relaxation schemes [19]. As will be seen here, multigrid methods may also be useful when the energy to be minimized has many local minima, as is often the case with nonlinear models. It has, indeed, been conjectured that multigrid analysis may, to a certain extent, *smooth* the energy landscape. Deterministic relax-

1049-9660/94 $6.00

ation schemes can then be used at coarse scales to obtain a good initial guess that may be refined at the finer scales.

Multigrid relaxation techniques have been considered before for image analysis models based on partial differential equations [6, 19] as well as on MRF [1, 12]. Yet, in multigrid implementations of relaxation algorithms devoted to the minimization of global nonlinear energy functions, the choice of the energy functions (and the adjustment of their parameters) at different scales remains a key problem. A standard choice is to adopt the same function at all scales, even though the interactions between variables are often resolution dependent.

The main contribution of this paper is the development of a new multiresolution algorithm (called "multiscale relaxation") that is applicable to the minimization of global energy functions used in image analysis or computer vision tasks. This algorithm is not equivalent to the multigrid approaches proposed in [6, 12, 19]. It 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 subspaces of the original space of configurations. Those subspaces contain constrained configurations describing the expected solutions at different scales. Each subspace defines a new "coarse" energy function whose parameters are derived from the original (full resolution) objective function. This constrained optimization is implemented using a coarse-to-fine procedure on a pyramidal structure.

In [4], Bouman *et al.* have recently presented a multiresolution relaxation technique, based on MRF models; this technique introduces a pyramidal structure close to the one reported here. However, in [4], the parameters of the objective function are kept the same at all scales. Expressing the multiresolution analysis as a constrained optimization problem enables us to generalize the scheme reported in [4] and to determine a consistent set of parameters for the objective function at the different scales. It also permits the definition of more general constraints on the configuration space. Such generalizations have been presented in [17]. The remainder of this paper is organized as follows: the multiscale relaxation is described in Section 2; in Section 3, the multiscale relaxation is compared to standard (monoresolution, multigrid, deterministic, and stochastic) relaxation schemes applied to a highly nonlinear optimization problem. Specifically, the problem is to estimate long-range motion (in an image sequence) on a discrete space of label configurations. This problem is a good benchmark problem since it yields a very complex energy landscape.

2. THE MULTISCALE RELAXATION ALGORITHM

The optimization of a global energy function is a powerful approach to extract relevant representations (labels) from an image. The method proceeds as follows:

- One or more specialized modules extract features (spatio-temporal gradients, edges, etc.) from the images that will be used as observations in the optimization process.
- Observations are combined using local photometric and structural relations with generic a priori knowledge on the label configurations (regularization scheme) in order to derive estimates of the unknown labels [5, 8].

The problem is stated as the minimization of a global energy function describing the interactions between the different (observed and hidden) variables.

2.1. *Global Energy Functions*

Let $o = \{o_s, s \in S\}$ designate the observed data defined on a rectangular lattice S. Let $e = \{e_s, s \in S\}$ denote the unobserved (hidden) label field, defined on the same lattice S.[1] Let Λ be the (discrete) state space of variable e_s and Ω the (finite) set of all possible label configurations e. Let $\mathcal{G} = \{\mathcal{G}_s, s \in S\}$ define a neighborhood system on S.

In the following we assume that the interactions between observed data o and hidden labels e can be defined in terms of an energy function of the form

$$U(e, o) = U_1(e, o) + U_2(e), \tag{1}$$

with

$$U_1(e, o) \triangleq \sum_{s \in S} V_1(e_s, o_s), \tag{2}$$

$$U_2(e) \triangleq \sum_{c \in C} V_c(e).$$

The optimal labeling we seek is, by definition, the minimum of this energy function. The first component $U_1(e, o)$ of the energy expresses (pointwise) interaction between the hidden labels e and the observation o, the second term $U_2(e)$ encodes constraints on the desired solution. The set of cliques associated to the neighborhood system $\mathcal{G}$, is denoted by $\mathcal{C}$. Cliques c are subsets of sites which are mutual neighbors. The local interaction function V_c depends only on the variables of clique c and expresses the local interactions between these different variables. The form of the local interaction functions V_c is problem dependent. Local interaction models have several relevant advantages: they are easy to specify and they yield tractable iterative computational schemes in which the label configurations can be updated locally [2, 5, 9].

In the following, we consider a standard eight-neighborhood system for $\mathcal{G}$. The cliques $c \in \mathcal{C}$ associated to the eight-neighborhood are shown in Fig. 1 and contain at

[1] Different lattices for E and O can also be adopted.

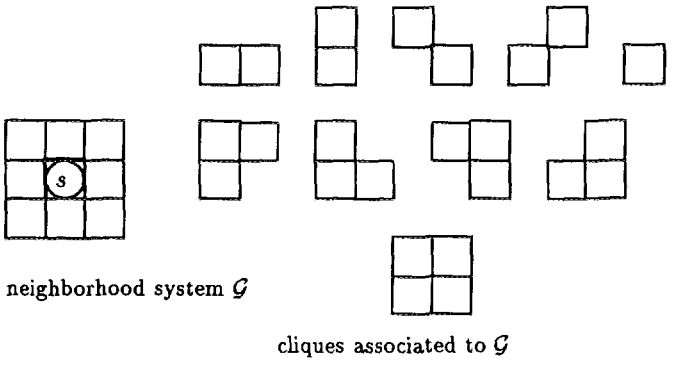

FIG. 1. Eight-neighborhood structure and associated cliques.

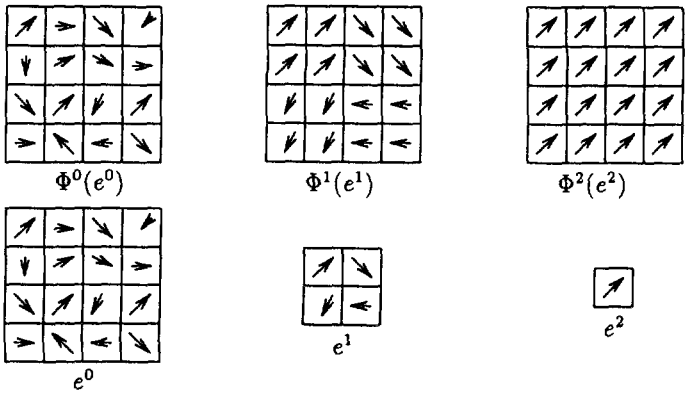

FIG. 3. A constrained labeling at scale i can be represented on a coarse grid S^i. Φ^i associates to a coarse configuration e^i the corresponding configuration in Ω_i (defined at full resolution).

most four sites. This energy class comprises most usual energy functions used in computer vision [1, 8, 10] (the extension of the method to larger neighborhoods is straightforward, although not only a notational matter).

2.2. *Multiscale Spaces of Configurations*

Let us assume that the size of the full resolution lattice is $2^m \times 2^m$. First recall that the energy function assigns an energy value to all possible configurations of the labeling $e = \{e_s, s \in S\}$, $e_s \in \Lambda$, $e \in \Omega$.

Instead of considering the minimization of energy function (2) directly on the full, generally huge, label configuration space Ω, let us consider the minimization of the energy function U on a hierarchy of nested subspaces Ω_i which are composed of label configurations constrained to be blockwise constant over cells of size $2^i \times 2^i$ (Fig. 2.)

These nested configuration spaces ($\Omega_i \subset \Omega_{i-1} \subset \cdots \subset \Omega_1 \subset \Omega_0 = \Omega$) are related to a description of the labeling at different scales. Scale 0 corresponds to the original configuration space Ω; scale i corresponds to configurations which are constant over cells of size $2^i \times 2^i$ (Fig. 2).

It is easy to see that at scale i, the number of independent labels is reduced by a factor $2^i \times 2^i$. Hence a configuration $e = \{e_s, s \in S\} \in \Omega_i$ can be represented on a coarse grid S^i of size $2^{m-i} \times 2^{m-i}$ (Fig. 3). The corresponding configuration on coarse grid S^i will be denoted $e^i = \{e^i_s, s \in S^i\}$.

As a consequence, the energy of a labeling $e \in \Omega_i$ may be rewritten as a function of labels defined on S^i. It turns out that at scale i one can derive from the energy function $U(e, o)$ an equivalent coarse energy function $U^i(e^i, o)$ which is a function of the coarse labeling e^i defined on S^i,

$$U^i(e^i, o) = U_1(\Phi^i(e^i), o) + U_2(\Phi^i(e^i)), \qquad (3)$$

where Φ^i is the isomorphism that associates to a configuration e^i defined on the coarse grid S^i, the corresponding full resolution configuration in Ω_i (Fig. 3). The multiscale energy functions U^i are defined for scales $i = 0, \ldots, n$.

Now, it is easy to verify that, if $\mathcal{G}$ is an eight-neighborhood system, the energy function U^i at scale i can be expressed as a sum of local interaction functions associated to *the same eight-neighborhood system on the coarse grid* S^i. Indeed, let us consider $c \in \mathcal{C}$, an arbitrary clique associated to the fine grid energy (c contains at most four mutual neighboring sites, see Fig. 1). The sites in c may be either included in a $2^i \times 2^i$ constant label cell, or they may sit astride two, three, or four different cells (see Fig. 4). As a consequence, the local interaction functions $V_1(e, o)$ and $V_c(e)$ associated to the original fine grid energy function can be rewritten on the coarse grid for cliques consisting of up to four neighboring sites on grid S^i. Hence U^i can be expressed as a sum of interaction functions depending on one, two, three, or four site cliques associated to an eight-neighborhood system.

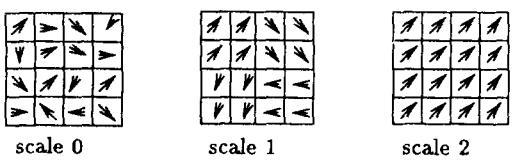

FIG. 2. Multiscale spaces of configurations: example of constrained configurations in $\Omega_0 = \Omega$ (scale 0), Ω_1 (scale 1), and Ω_2 (scale 2). Labels are represented by vectors for easier visual interpretation.

2.3. *A Coarse-to-Fine Relaxation Algorithm*

To exploit the sequence of multiscale energy previously defined (Eq. (3)), instead of minimizing the original energy function (Eq. (2)) over the full configuration space Ω, we consider the following sequence of optimization problems:

$$\hat{e}(i) = \arg\min_{e \in \Omega_i} U(e, o), \quad i = n, \cdots, 0. \qquad (4)$$

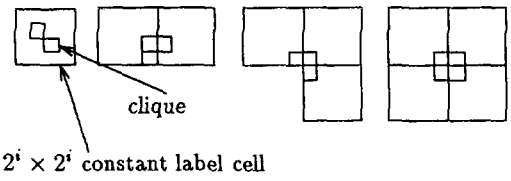

FIG. 4. Example of clique locations with respect to cells with constant labels.

By Eq. (3), this is equivalent to the minimization of the coarse energy functions:

$$\hat{e}^i = \arg\min_{e^i} U^i(e^i, o), \quad i = n, \cdots, 0. \tag{5}$$

These optimization problems are solved using a standard coarse-to-fine multigrid strategy (Fig. 5). Starting from a coarse scale n, the optimization problem is first solved in subset Ω_n (Eq. (4)) by solving the equivalent problem (Eq. (5)). An estimate of $\hat{e}^n$ is obtained by a deterministic nonlinear Gauss–Seidel relaxation algorithm known as ICM [2]. This defines a first (crude) solution to the original problem and will be refined on the subsequent finer levels.

At level i, let $\tilde{e}^i$ designate the estimate of $\hat{e}^i$ (obtained after the convergence of the deterministic relaxation at that level). The algorithm at resolution level $i - 1$ is initialized by the configuration $[\Phi^{i-1}]^{-1} \circ \Phi^i(\tilde{e}^i)$ which corresponds to an interpolation of $\tilde{e}^i$ on the finer grid S^{i-1} using a simple repetition of the estimated labels (see Fig. 5). Note that this interpolation defines a consistent prolongation (or interpolation operator) from coarse space Ω_i to space Ω_{i-1} (indeed, $\tilde{e}^i$ and its interpolation at level $i - 1$ are associated to the same configuration $\Phi^i(\tilde{e}^i)$ at full resolution and hence have identical energies).

After interpolation, a new relaxation step is performed

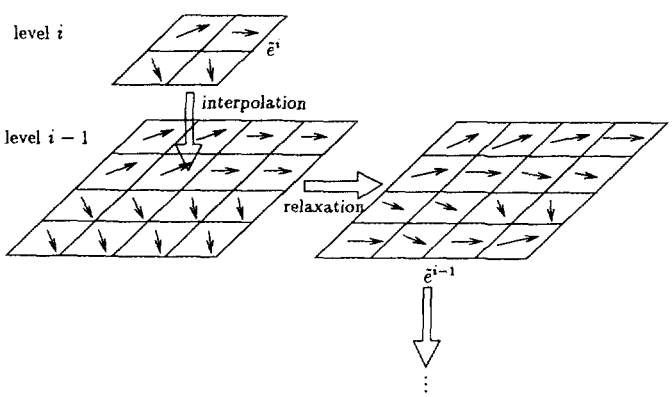

FIG. 5. A coarse-to-fine strategy for optimization at scales $i = n, \cdots, 0$.

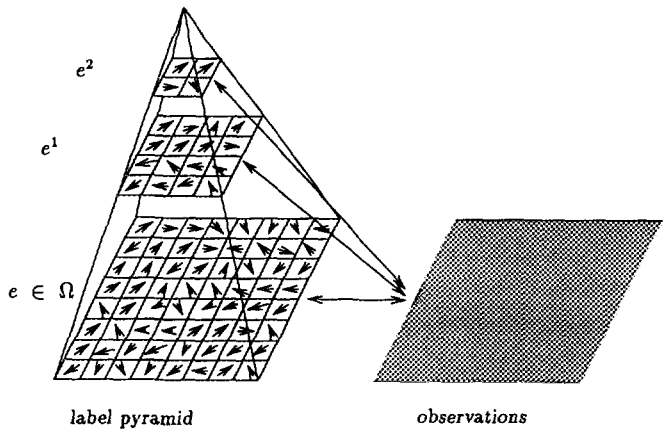

FIG. 6. The multiscale relaxation structure.

until convergence is obtained at scale $i - 1$; this procedure is repeated at the finer scales.

We note that, whereas the algorithm uses a multigrid representation of labels, only one level (full resolution) is used for the observation field o, as can be seen in Eq. (3) (Fig. 6). We consider this to be an advantage of the method (also found in the scheme described in [4]), because *no multiresolution pyramid* has to be computed from the observations. In standard multigrid schemes, a pyramid of observations is usually obtained by low-pass filtering the data [6, 12]. The choice of the filters is generally quite arbitrary and the way the data is transformed is understood only qualitatively. Moreover, additional storage or computation is not necessary to create the data pyramid in the method presented here.

The multiscale relaxation algorithm has several other advantages:

- it is easily interpreted as a constrained optimization algorithm;
- it can be generalized to other constraints: in the present paper a piecewise constant constraint on configuration has been introduced. Other examples (bilinear constraints, for instance) can be found in [17];
- it is easy to implement, and the multiscale energy functions can be derived easily;
- it exhibits fast convergence properties when applied to nonlinear optimization problems (see next section) and it yields better solutions (convergence to lower energy values) than standard multigrid methods.

3. EXPERIMENTS

The multiscale relaxation is compared to standard (monoresolution, multigrid, deterministic, and stochastic) relaxation schemes on a highly nonlinear optimization problem. The issue is to estimate long-range motion in an image sequence. The problem is solved on a discrete space

of label configurations by minimizing a global energy function which involves two components: the first one corresponds to a similarity function that matches brightness patterns; the second term is a standard regularization term. This problem is known to generate a very complex energy landscape [12], and hence seems to be a good benchmark problem for comparing the various methods. This optimization problem is also associated with a very large configuration space and is thus cpu intensive. Other applications in the field of visual motion analysis (motion detection and motion-based segmentation) may be found in an extended version of this paper [11].

The various algorithms will be compared according to two criteria: convergence speed (number of iterations and cpu time) and the quality of the final estimated solutions. This quality can be assessed by comparing the final energy computed using the different solutions, once convergence is reached.

3.1. *Long-Range Motion Estimation Using Similarity Functions*

We assume that the labels, representing the velocity vectors, belong to a discrete state space Λ. 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. The velocity vector at time t and at site s is denoted $\boldsymbol{\omega}_t(s) = (u_t(s), v_t(s))$, with $u_t(s) = (dx/dt)(s)$, $v_t(s) = (dy/dt)(s)$ and $\boldsymbol{\omega}_t = \{\boldsymbol{\omega}_t(s), s \in S\}$. 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 energy function is associated to an eight-neighborhood system (Fig. 1) and we only consider two element cliques,

$$U(\boldsymbol{\omega}_t, f_t, f_{t+\delta t}) = U_1(\boldsymbol{\omega}_t, f_t, f_{t+\delta t}) + U_2(\boldsymbol{\omega}_t), \tag{6}$$

where

$$\begin{aligned} U_1(\boldsymbol{\omega}_t, f_t, f_{t+\delta t}) &= \sum_{s \in S} \{f_t(s) - f_{t+\delta t}(s + \boldsymbol{\omega}_t(s) \cdot \delta t)\}^2 \\ U_2(\boldsymbol{\omega}_t) &= \alpha^2 \sum_{\{s,r\} \in \mathcal{C}} \|\boldsymbol{\omega}_t(s) - \boldsymbol{\omega}_t(r)\|^2. \end{aligned} \tag{7}$$

The first term in the energy is a similarity function (also known as the "displaced frame difference") that expresses the constant brightness assumption for the image of a physical point over time, along the motion trajectory. The second term can be interpreted as a regularization term which favors smooth solutions. The parameter α controls the relative weight of the two terms. This approach to long-range motion estimation relates to standard methods based on the matching of iconic structures. The multiscale relaxation defines a hierarchical scheme to compute the velocity field for this model.

The multiscale energy functions are derived from the original full resolution energy (Eq. (6)) using Eq. (3). These energy functions are decomposed on coarse-scale cliques,

$$U^i(\boldsymbol{\omega}_t^i, f_t, f_{t+\delta t}) \triangleq U_1^i(\boldsymbol{\omega}_t^i, f_t, f_{t+\delta t}) + U_2^i(\boldsymbol{\omega}_t^i), \tag{8}$$

where

$$U_1^i(\boldsymbol{\omega}_t^i, f_t, f_{t+\delta t}) \triangleq \sum_{s \in S^i} \sum_{r \in B_s^i} \{f_t(r) - f_{t+\delta t}(r + \boldsymbol{\omega}_t^i(s) \cdot \delta t)\}^2, \tag{9}$$

$$U_2^i(\boldsymbol{\omega}_t^i) \triangleq \sum_{c \in \mathcal{C}^i} V_c^i(\boldsymbol{\omega}_t^i), \tag{10}$$

where an eight-neighborhood structure is defined on the coarse grid S^i and B_s^i denotes the constant label cell at full resolution corresponding to site $s \in S^i$ at coarse resolution (Fig. 3). The coarse scale interaction functions V_c^i are zero for cliques c of more than two sites. Let $\mathcal{C}_h^i$, $\mathcal{C}_v^i$, and $\mathcal{C}_d^i$ denote the set of horizontal, vertical, and diagonal two-element cliques, respectively. The interaction functions V_c^i are given by

$$\begin{aligned} &\text{for } c \in \mathcal{C} - (\mathcal{C}_h^i \cup \mathcal{C}_v^i \cup \mathcal{C}_d^i) \\ &\qquad V_c^i(\boldsymbol{\omega}_t^i) = 0, \\ &\text{for } \{s_1, s_2\} \in \mathcal{C}_h^i \cup \mathcal{C}_v^i \\ &\qquad V_{\{s_1,s_2\}}^i(\boldsymbol{\omega}_t^i) = q_{hv}^i \cdot \alpha \|\boldsymbol{\omega}_t^i(s_1) - \boldsymbol{\omega}_t^i(s_2)\|^2, \\ &\text{for } \{s_1, s_2\} \in \mathcal{C}_d^i \\ &\qquad V_{\{s_1,s_2\}}^i(\boldsymbol{\omega}_t^i) = q_d^i \cdot \alpha \|\boldsymbol{\omega}_t^i(s_1) - \boldsymbol{\omega}_t^i(s_2)\|^2, \end{aligned} \tag{11}$$

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

3.2. *Experimental Results*

In our experiments, four different algorithms have been compared:

1. a standard monoresolution deterministic relaxation algorithm (DR), known as ICM [2];
2. a monoresolution stochastic relaxation (SR) algorithm based on the Gibbs sampler [9],
3. a standard coarse-to-fine multigrid relaxation (MGR) algorithm in which *the same energy* is considered at each resolution (same parameters, same neighborhood system, and same local interaction functions) and a pyramid of observations is constructed. A similar implementation of this scheme may be found in [1], and

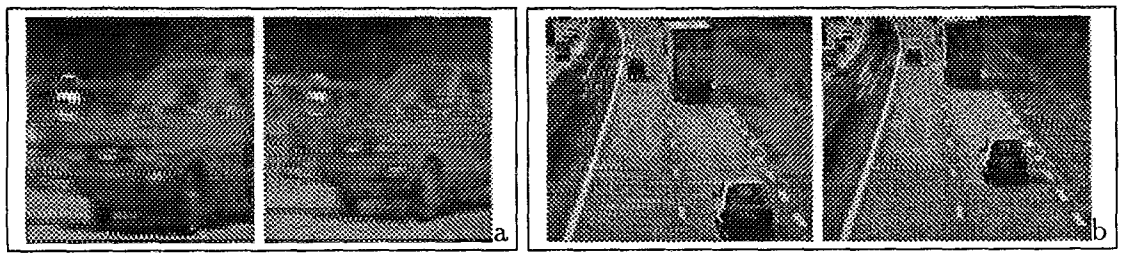

FIG. 7. Two frames from the original sequences: (a) "camera zoom" sequence (100 × 100); (b) "highway" sequence (256 × 256).

4. the proposed multiscale relaxation (MSR), in which the energy function and its parameters are derived at each resolution level from the original (full resolution) energy function.

The temperature schedule used in the stochastic relaxation algorithm was: $T(j) = T_0 \cdot A^j$, 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. Four resolution levels were considered in the multigrid and multiscale methods. A discrete label space $\mathcal{W} = (-4:+4, -4:+4)$ with a step size of $\delta = 1$ has been adopted at the full resolution.

For the multiresolution schemes (MGR and MSR), the same deterministic algorithm (ICM [2]) was run at the different resolutions. The relaxation was performed until convergence was reached on the label fields (at each resolution). In 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. Besides, the listed number of iterations corresponds to the "equivalent number of iterations at full resolution" nb_{eq}, defined as

$$nb_{eq} = \sum_{i=0}^{n} \frac{1}{4^i} nb_i, \tag{12}$$

where nb_i is the number of iterations (i.e., the number of full sweeps through the image) at scale i.

For a more fair comparison between MSR and MGR, the parameters of the standard multigrid method MGR could also have been adjusted at each scale. Let us, however, note that this adjustment is necessarily supervised. The multiscale relaxation does not have this drawback since the parameters of the energy functions at each scale are derived from the original energy function. Besides, in our experiments we were not able to find, for the MGR method, a set of parameters which would have led to results of the same quality as MSR.

Our experiments clearly show that the minimization of the similarity function (Eq. (7)) yields an energy function that contains many local minima. The results reported here have been obtained with a random initialization of the label field. Two image sequences were used: The first one is obtained by a simulated camera zoom on a real image (Fig. 7a); this corresponds to a perfectly divergent motion. The second one is a (real-world) highway scene showing moving cars (Fig. 7b).

In Figs. 8 and 10 the velocity vector fields estimated by each of the four algorithms are shown. The multiscale approach supplies fields qualitatively similar to the one obtained using the time-consuming stochastic relaxation procedure. Table 1 gives the number of iterations and the cpu time required in each case, as well as the final energy. The cpu times are given for a C implementation on a workstation (Sparc 1). The global behavior of the different schemes can also be studied by considering the energy plots shown in Fig. 9.

The stochastic relaxation algorithm leads to the lowest energy value in the case of the "zoom" but requires 468 iterations to converge (Table 1 and Fig. 9). The proposed multiscale approach gives results close to the stochastic method with a gain of almost two orders of magnitude

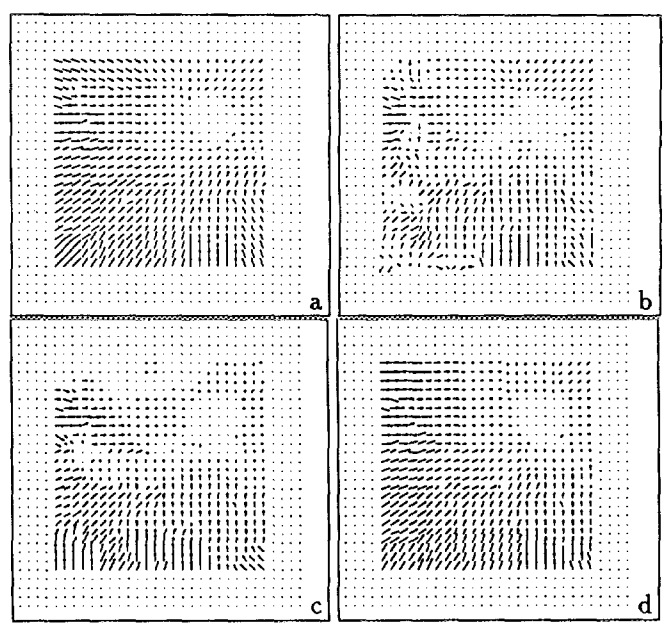

FIG. 8. Optical flow fields ("camera zoom" sequence): (a) SR, (b) DR, (c) MGR, (d) MSR.

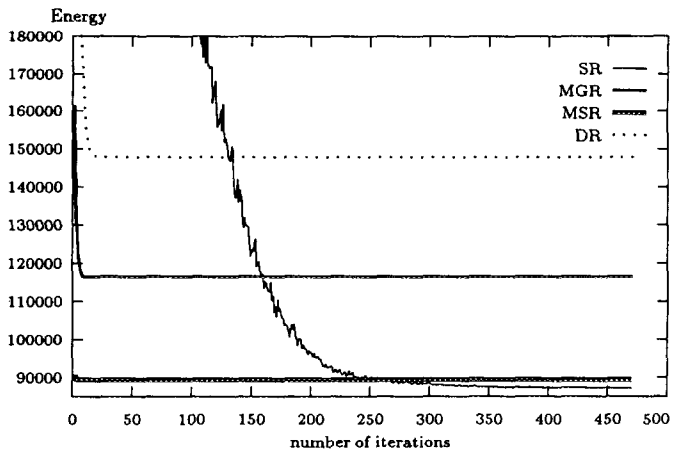

FIG. 9. Energy U versus iteration number: "camera-zoom" sequence.

TABLE I
Motion Measurement

Algorithm	SR	DR	MGR	MSR
	(a) Camera Zoom Sequence			
nb_{eq}	468	42	8.75	5.38
Final energy	87,292	139,042	116,502	89,355
Cpu time	9h7m33s	43m17s	10m12s	6m14s
	(b) Highway Sequence			
nb_{eq}	259	32	16.88	7.46
Final energy	766,427	1,117,945	811,468	749,471
Cpu time	104h48m7s	7h35m15s	4h13m57s	2h28m50s

Note. The number of iterations in the monoresolution and multiresolution algorithms (SR, stochastic relaxation; DR, deterministic relaxation; MGR, standard multigrid relaxation; MSR, multiscale relaxation); nb_{eq}, equivalent number of iterations at full resolution (see text).

in the number of iterations. The standard methods (DR, MGR) become stuck at rather high energy local minima. It is interesting to note that in the case of the "highway sequence," the stochastic relaxation performed worse than the multiscale approach, when adopting the same temperature schedule as for the "zoom" sequence (Table 1).

A very significant gain in the cpu time is also obtained for the MSR method, compared to all the other methods (it is faster than the standard multigrid method in both examples). Note that the method enables us to save cpu time even though the algorithm uses full-resolution observations at each scale. This is mainly due to a faster convergence at the finest resolution, where most of the cpu time is consumed.

These points are best illustrated by another set of experiments in which we process 41 sequences, each of two 64 × 64 frames, obtained by applying the same synthetic motion (Fig. 11) to different real images (Fig. 12). The synthetic motion includes translation, rotation, and dilation (Fig. 11). Figure 13 presents, for these 41 short sequences, the ratio of the final energy value reached by stochastic relaxation (SR) to the one obtained by the various other methods. The corresponding number of iterations and cpu times required by each method are plotted in Fig. 14 and Fig. 15. On the average, the MSR method

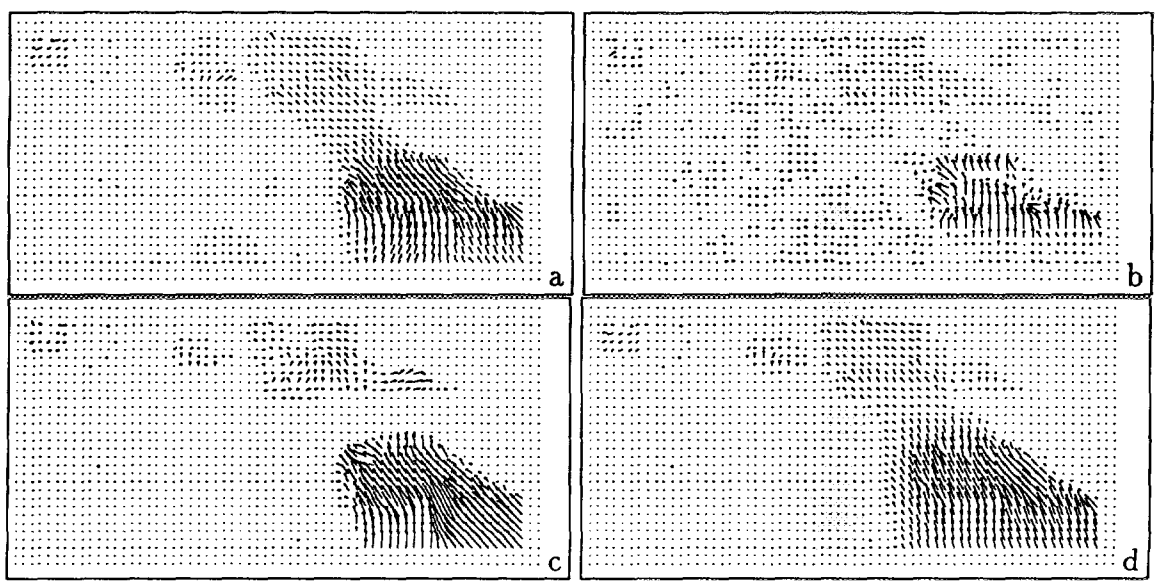

FIG. 10. Optical flow fields ("highway sequence"): (a) SR, (b) DR, (c) MGR, (d) MSR.

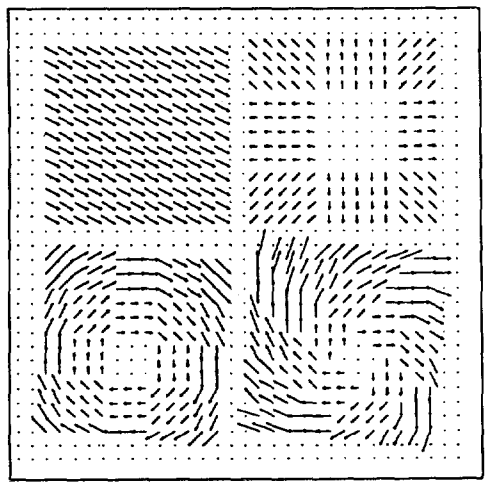

FIG. 11. Synthetic optical flow field applied on the 41 sequences.

finds configurations close to the best estimates obtained by SR, with an average gain of about 60, both on the number of iterations and cpu time (Table 2).

This benchmark illustrates clearly that cpu time is saved with the MSR method, even though it uses full resolution observations at each scale. When compared to monoresolution schemes, time is saved when computing the regularizing term of the energy function, since the labels are processed on a multigrid structure. Time is also saved with respect to MGR, even though the MSR scheme uses full resolution observations at each scale. This is due to a faster convergence of the multiscale relaxation at high resolutions: on the average, for the 41 sequences, MGR spends 14.6 iterations at full resolution, whereas MSR only requires 5.0 iterations to converge.

TABLE 2

Average Number of Iterations, CPU Time and the Ratio of the Final Energy Value Reached by Stochastic Relaxation (SR) to the One Obtained by the Other Methods

Algorithm	SR	DR	MGR	MSR
Average nb_{eq}	409.32	16.46	16.32	6.62
Average cpu time	3h4min11s	7min25s	6min48s	3min13s
Average U^{RS}_{final}/U_{final}	1.000	0.663	0.631	0.940

Note. The average is computed on the 41 sequences.

The multiscale algorithm has also been applied to other problems in visual motion analysis (motion detection and motion segmentation) with similar qualitative and quantitative improvements. These results can be found in [11].

4. CONCLUSION

In this paper we have described a multiscale approach for the multiresolution analysis of images based on global optimization. The approach relies on constrained optimization and has been demonstrated on a highly nonlinear global optimization problem. The multiscale relaxation algorithm presents several attractive features when compared to standard multigrid schemes: simple implementa-

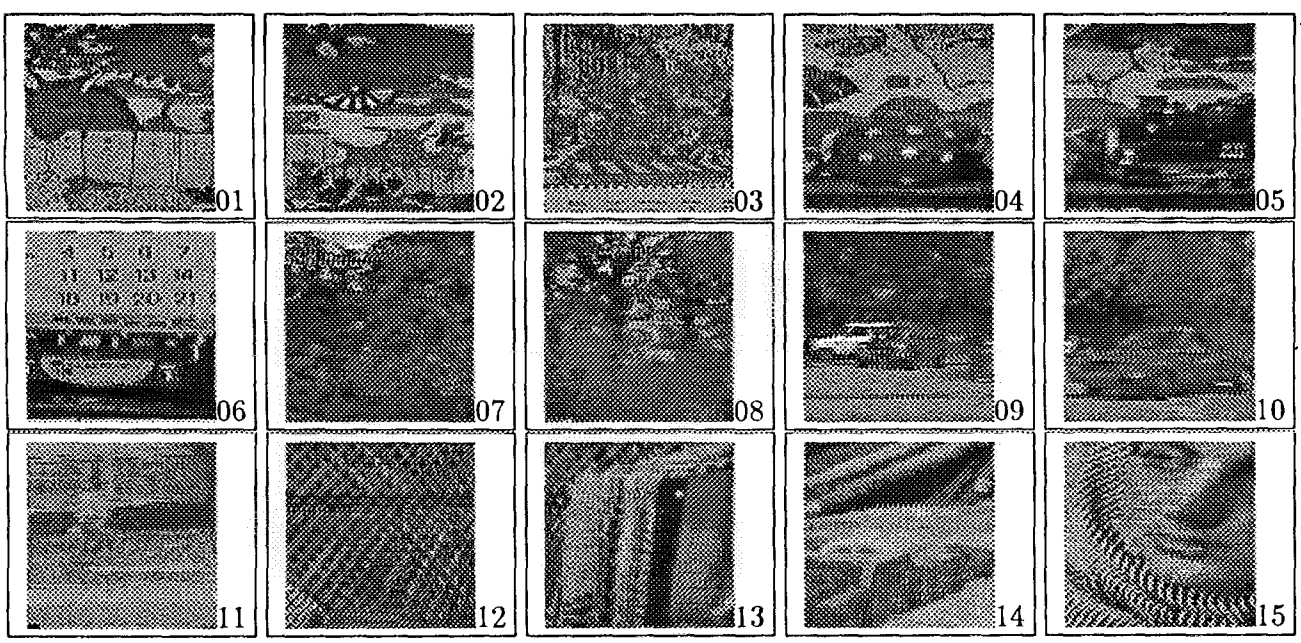

FIG. 12. Examples of first frames in the 41 short sequences.

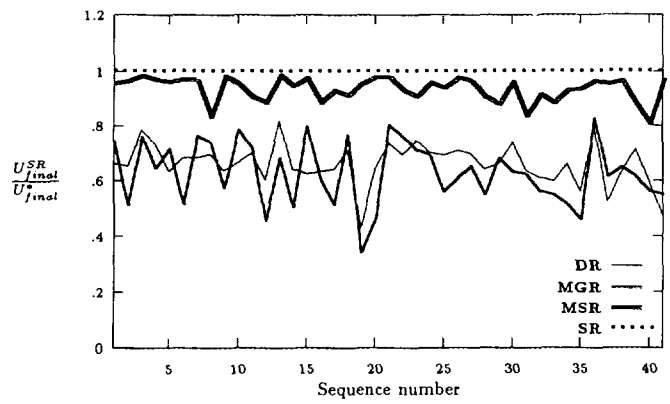

FIG. 13. Ratio of the final energy value reached with SR (U^{SR}_{final}) to the one obtained with the other methods (U^{*}_{final}).

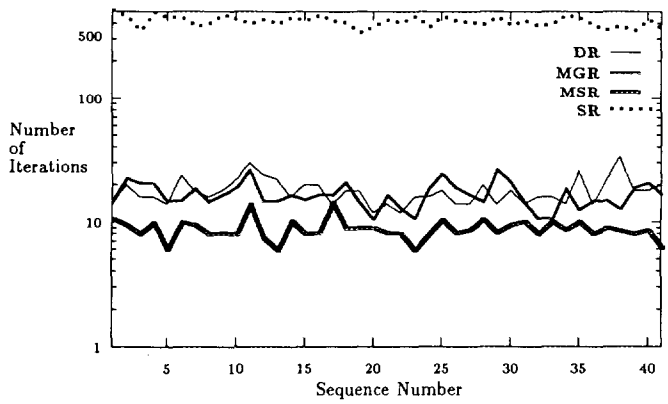

FIG. 14. Number of iterations at convergence.

tion, fast convergence towards high quality solutions. Gains of up to two orders of magnitude in the convergence speed with respect to stochastic relaxation have been observed. The estimates are close in quality to those obtained using time-consuming stochastic relaxation algorithms.

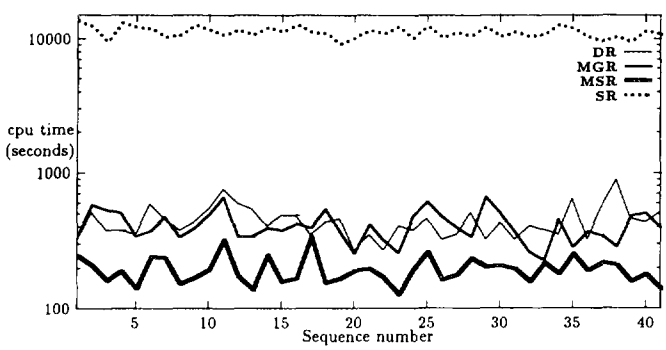

FIG. 15. Cpu times for the 41 sequences.

Let us emphasize that the proposed approach is general and can be applied to other image analysis problems whose solution is based on the minimization of global energy functions. It is adapted to sophisticated nonlinear models which have been developed recently in image analysis, for instance, in image segmentation [8] or in structure from motion problems [18]. Finally, this algorithm is a good candidate for parallel implementation on regular arrays consisting of locally interconnected processors [14].

ACKNOWLEDGMENTS

This work has been partly supported by MRT (French Ministry of Research and Technology) and CNRS in the context of the PRC Program "Man Machine Interface" under Contract PMFE 88F1 548, by the GDR TdSI 134, and by the Brittany County Council under a contribution to the student grant. We thank B. Levy for fruitful discussions.

REFERENCES

1. S. T. Barnard, Stochastic stereo matching over scale, *Int. J. Comput. Vision* **3**, 1989, 17–32.
2. J. Besag, On the statistical analysis of dirty pictures, *J. R. Statist. Soc. Ser. B* **48**(3), 1986, 259–302.
3. A. Blake and A. Zisserman, *Visual Reconstruction*, MIT Press, Cambridge, MA, 1987.
4. C. Bouman and B. Liu, Multiple resolution segmentation of textured images, *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(2), 1991, 99–113.
5. P. B. Chou and C. M. Brown, The theory and practice of bayesian image modeling, *Int. J. Comput. Vision* **4**, 1990, 185–210.
6. W. Enkelmann, Investigations of multigrid algorithms for the estimation of optical flow fields in image sequences, *Comput. Vision Graphics, Image Process.* **43**, 1988, 150–177.
7. 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, Hawaii, June 3–6 1991*, pp. 166–172.
8. D. Geman, S. Geman, C. Graffigne, and D. Pong, Boundary detection by constrained optimization, *IEEE Trans. Pattern Anal. Machine Intell.* **12**(7), 1990, 609–628.
9. S. Geman and D. Geman, Stochastic relaxation, Gibbs distributions and the bayesian restoration of images, *IEEE Trans. Pattern Anal. Machine Intell.* **6**(6) 1984, 721–741.
10. F. Heitz and P. Bouthemy, Multimodal motion estimation and segmentation using Markov random fields, in *Proc. 10th Int. Conf. Pattern Recognition, Atlantic City, June 1990*, Vol. 1, pp. 378–383.
11. F. Heitz, P. Perez, and P. Bouthemy, *Constrained Multiscale Markov Random Fields and the Analysis of Visual Motion*, Technical Report 1615, INRIA, Rennes, Feb. 1992.
12. J. Konrad and E. Dubois, Multigrid Bayesian estimation of image motion fields using stochastic relaxation, in *Proc. 2nd Int. Conf. Computer Vision, Tarpon Springs Florida, Dec. 1988*, pp. 354–362.
13. 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, Barcelona, 1990*, pp. 947–950.
14. 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, Paris, Dec. 1991*, pp. 309–320.

15. J. W. Modestino and J. Zhang, A Markov Random Field model-based approach to image interpretation, *IEEE Trans. Pattern Anal. Machine Intell.* **14**(6), 1992, 606–615.

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

17. 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, March 1992*, pp. III-61–III-64.

18. 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, Atlantic City, June 1990*, Vol. 1, pp. 390–397.

19. D. Terzopoulos, Image analysis using multigrid relaxation methods, *IEEE Trans. Pattern Anal. Machine Intell.* **8**(2), 1986, 129–139.