PARALLEL IMAGE CLASSIFICATION USING MULTISCALE MARKOV RANDOM FIELDS

Zoltan KATO, Marc BERTHOD and Josiane ZERUBIA
INRIA - 2004 Route des Lucioles - BP 93
06902 Sophia Antipolis Cedex - FRANCE
email: name@sophia.inria.fr - Fax (33) 93 65 76 43

ABSTRACT

In this paper, we are interested in massively parallel multiscale relaxation algorithms applied to image classification. First, we present a classical multiscale model applied to supervised image classification. The model consists of a label pyramid and a whole observation field. The potential functions of the coarse grid are derived by simple computations. Then, we propose another scheme introducing a local interaction between two neighbor grids in the label pyramid. This is a way to incorporate cliques with far apart sites for a reasonable price. Finally we present the results on noisy synthetic data and on a SPOT image obtained by different relaxation methods using these models.

INTRODUCTION

Markov Random Fields (MRF) have become more and more popular during the last few years in image processing [1, 2, 3, 4, 5]. A good reason for that is that such a modelization is the one which requires the less a priori information on the world model. On the other hand, the local behavior of MRF permits to develop highly parallel algorithms for the resolution of combinatorial optimization problems associated with such a model.

We consider herein a simple MRF-model for supervised image classification which operates only on grey levels without texture. The class parameters are known a priori, they are computed on a training set. We briefly describe this monogrid model, a more detailed explanation can be found in [3]:

Let $\mathcal{S} = \{s_1, s_2, \ldots, s_N\}$ be a set of sites or pixels and $\mathcal{F} = \{f_s\}_{s \in \mathcal{S}}$ a set of image data (grey levels here). Each of these sites may belong to any of L classes (or equivalently take any label from $\Lambda = \{1, 2, \ldots, L\}$). A global discrete labeling ω assigns one label ω_s ($\omega_s \in \Lambda$) to each site s in $\mathcal{S}$. Choosing a first order neighborhood-system, we can define the following MAP estimator of the label field:

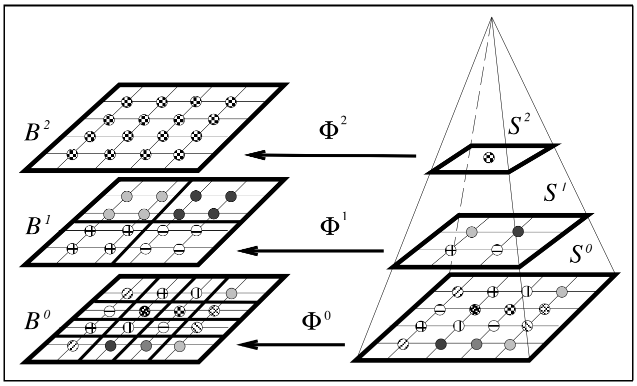

Figure 1: The isomorphism Φ^i between $\mathcal{B}^i$ and $\mathcal{S}^i$.

$$
\begin{aligned}
\hat{\omega} &= \arg\max_{\omega \in \Omega} P(\omega \mid \mathcal{F}) \\
&= \arg\min_{\omega \in \Omega} U(\omega, \mathcal{F}) \qquad (1) \\
U(\omega, \mathcal{F}) &= U_1(\omega, \mathcal{F}) + U_2(\omega) \qquad (2)
\end{aligned}
$$

where $U(\omega, \mathcal{F})$ is the energy function associated with the model, $U_1(\omega, \mathcal{F})$ (resp. $U_2(\omega)$) is the energy with respect to (w.r.t.) the first (resp. second) order cliques. Assuming $P(f_s \mid \omega_s)$ to be Gaussian, the class $\lambda \in \Lambda$ is represented by its mean value μ_λ and its deviation σ_λ:

$$
\begin{aligned}
U_1(\omega, \mathcal{F}) &= \sum_{s \in \mathcal{S}} \left(\log(\sqrt{2\pi}\sigma_{\omega_s}) + \frac{(f_s - \mu_{\omega_s})^2}{2\sigma^2_{\omega_s}} \right) \qquad (3) \\
U_2(\omega) &= \sum_{C \in \mathcal{C}} V_2(\omega_C) \text{ where} \qquad (4) \\
V_2(\omega_C) &= V_{\{s,r\}}(\omega_s, \omega_r) = \left\{ \begin{array}{ll} -\beta & \text{if } \omega_s = \omega_r \\ +\beta & \text{if } \omega_s \neq \omega_r \end{array} \right. \qquad (5)
\end{aligned}
$$

THE MULTISCALE MODEL

To generate the multiscale version of our model, let us divide the initial grid into blocks of size $n \times n$ (4×4 here) neighbor pixels. We consider that the label assigned to the pixels of a block is constant over all the pixels of the block [2]. These configurations will describe the MRF at scale 1. Scale i is defined similarly by considering labels which are constant over blocks of size $n^i \times n^i$.

Let $\mathcal{B}^i = \{b^i_1, \ldots, b^i_{N_i}\}$ $(N_i = N/n^{2i})$ denote the set of blocks and Ω_i the configuration-space at scale i $(\Omega_i \subset \Omega_{i-1} \subset \cdots \subset \Omega_0 = \Omega)$. The label associated with block b^i_k is denoted by ω^i_k. We can define the same neighborhood structure on $\mathcal{B}^i$ as on $\mathcal{S}$: b^i_k and b^i_l are neighbors iff exists any clique C on the finest level which sits astride the two blocks $\{b^i_k, b^i_l\}$. Partitioning the original clique-set $\mathcal{C}$ into two disjoint subsets $\mathcal{C}^i_k$ (cliques included in b^i_k) and $\mathcal{C}^i_{k,l}$(cliques siting astride two neighboring blocks $\{b^i_k, b^i_l\}$), we get the following energy function:

$$\begin{aligned} U_1(\omega, \mathcal{F}) &= \sum_{s\in\mathcal{S}} V_1(\omega_s, f_s) = \sum_{b^i_k \in \mathcal{B}^i} \sum_{s \in b^i_k} V_1(\omega_s, f_s) \\ &= \sum_{b^i_k \in \mathcal{B}^i} V_1^{\mathcal{B}^i}(\omega^i_k, \mathcal{F}) \qquad (6) \\ U_2(\omega) &= \sum_{C\in\mathcal{C}} V_2(\omega_c) \\ &= \sum_{b^i_k \in \mathcal{B}^i} \sum_{C \in \mathcal{C}^i_k} V_2(\omega_c) + \sum_{\{b_k, b_l\}} \sum_{C\in\mathcal{C}^i_{k,l}} V_2(\omega_c) \\ &= \sum_{b^i_k \in \mathcal{B}^i} V_k^{\mathcal{B}^i}(\omega^i_k) + \sum_{\{b_k, b_l\}} V_{k,l}^{\mathcal{B}^i}(\omega^i_k, \omega^i_l) \qquad (7) \end{aligned}$$

Now, we can define our pyramid (see Figure 1) where level i contains the coarse grid $\mathcal{S}^i$ which is isomorphic to the scale $\mathcal{B}^i$. The coarse grid has a reduced configuration space $\Gamma_i = \Lambda^{N_i}$. The isomorphism Φ^i from $\mathcal{S}^i$ in $\mathcal{B}^i$ is just a projection of the coarse label field to the fine grid $\mathcal{S}^0 = \mathcal{S}$. The model on the grids $\mathcal{S}^i$ $(i = 0, \ldots, M)$ defines a set of consistent multiscale MRF models, whose energy functions derived from Equations (6) and (7):

$$\begin{aligned} U_1^i(\omega^i, \mathcal{F}) &= \sum_{s^i \in \mathcal{S}^i} V_1^i(\omega^i_{s^i}, \mathcal{F}) \text{ where} \\ V_1^i(\omega^i_{s^i}, \mathcal{F}) &= \sum_{s\in b^i_{s^i}} V_1(\omega_s, f_s) + \sum_{C \in \mathcal{C}^i_{s^i}} V_2(\omega_C) \\ &= \sum_{s\in b^i_{s^i}} \left(\log(\sqrt{2\pi}\sigma_{\omega_s}) + \frac{(f_s - \mu_{\omega_s})^2}{2\sigma^2_{\omega_s}} \right) \\ &\quad -p^i\beta \qquad (8) \\ U_2^i(\omega) &= \sum_{C^i = \{r^i, s^i\} \in \mathcal{C}^i} V_2^i(\omega_{C^i}) \text{ where} \\ V_2^i(\omega_{C^i}) &= \sum_{\{r,s\}\in \mathcal{C}^i_{r^i,s^i}} V_2(\omega_r, \omega_s) \\ &= \begin{cases} -q^i\beta & \text{if } \omega_r = \omega_s \\ +q^i\beta & \text{if } \omega_r \neq \omega_s \end{cases} \qquad (9) \end{aligned}$$

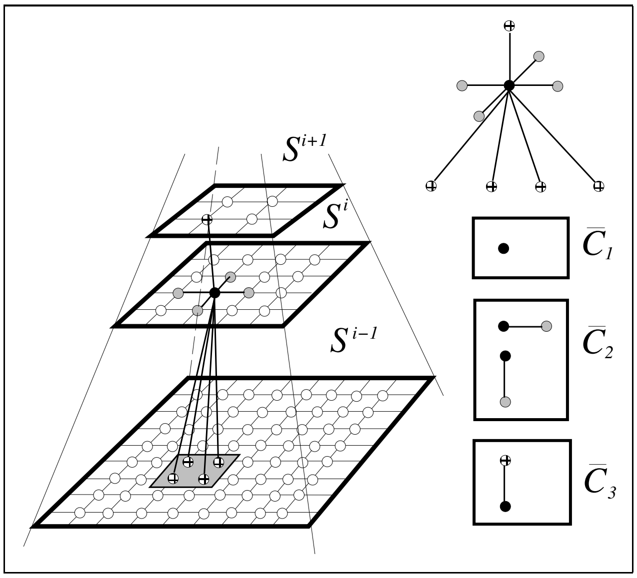

Figure 2: The neighborhood system $\bar{\mathcal{G}}$ and the cliques $\bar{\mathcal{C}}_1$, $\bar{\mathcal{C}}_2$ and $\bar{\mathcal{C}}_3$.

where p^i is the number of cliques included in the same block at scale i $(\mathcal{B}^i)$ and q^i is the number of cliques between two neighboring blocks at scale i $(\mathcal{B}^i)$. Considering blocks of $n \times n$ and a first order neighborhood system, we get $p^i = 2n^i(n^i - 1)$ and $q^i = n^i$. Using a top-down strategy in the pyramid, we solve the problem first at a higher level i, then the level $i-1$ is initialized by $(\Phi^{i-1})^{-1} \circ \Phi^i(\hat{\omega}^i)$, where $\hat{\omega}^i$ is obtained at the convergence of a relaxation algorithm at the level i (see Figure 1).

THE 3D MULTISCALE MODEL

Keeping the pyramid described in the previous section, we introduce local interactions between two neighbor grids. In order to define these interactions, we first modify the neighborhood structure. Let $\bar{\mathcal{S}} = \{\bar{s}_1, \ldots, \bar{s}_{\bar{N}}\} = \bigcup_{i=0}^{M} \mathcal{S}^i$ denotes the sites of the pyramid. On these sites, we define the following neighborhood-structure (see Figure 2): Any site $\bar{s}$ interacts with its ancestor (that is the site at the upper level corresponding to the block of $\bar{s}$) and its descendants (that is the sites of the block at the lower level corresponding to $\bar{s}$). We will consider herein only the first and second order cliques. Clique-potentials for the other cliques are supposed to be 0. Let $\bar{\mathcal{C}}$ denotes the set of these cliques which can be partitioned into three disjoint subsets $\bar{\mathcal{C}}_1, \bar{\mathcal{C}}_2, \bar{\mathcal{C}}_3$ corresponding to first order cliques, second order cliques which are on the same level and second order cliques which sit astride two neighboring levels (see Figure 2). The configuration-space of the pyramid is the set of the configuration-spaces of

A	iter.	total time	time/iter.	error	β	γ
Gibbs	68	3.01 sec.	0.044 sec.	183	1.0	—
ICM	9	0.15 sec.	0.016 sec.	2948	1.0	—
B	—	—	—	—	—	—
Gibbs	101	3.86 sec.	0.038 sec.	176	1.0	—
ICM	17	0.22 sec.	0.013 sec.	1657	0.9	—
C	—	—	—	—	—	—
Gibbs	41	141.97 sec.	3.463 sec.	191	0.7	0.1
ICM	11	30.17 sec.	2.743 sec.	293	0.8	0.5

Table 1: The results on the noisy synthetic image with 4 classes (A-monogrid, B-multiscale, C-3D multiscale).

each level: $\bar{\Omega} = \Gamma^0 \times \Gamma^1 \times \cdots \times \Gamma^M$. It remains to define the energy function of the model. This can be done easily using the above partition. It is clear from this partition, that the energy w.r.t. the cliques in $\bar{\mathcal{C}}_1$ and $\bar{\mathcal{C}}_2$ is simply the sum of the energy of each level: $\sum U^i(\omega, \mathcal{F})$ (see Equations (9) and (8)). The energy w.r.t. $\bar{\mathcal{C}}_3$ can be expressed as the sum of its clique-potentials: $\sum_{C \in \bar{\mathcal{C}}_3} \bar{V}_2(\bar{\omega}_C)$ where

$$\bar{V}_2(\bar{\omega}_C) = \bar{V}_{\{\bar{s},\bar{r}\}}(\bar{\omega}_{\bar{s}}, \bar{\omega}_{\bar{r}}) = \begin{cases} -\gamma & \text{if } \bar{\omega}_{\bar{s}} = \bar{\omega}_{\bar{r}} \\ +\gamma & \text{if } \bar{\omega}_{\bar{s}} \neq \bar{\omega}_{\bar{r}} \end{cases} \quad (10)$$

THE PARALLEL RELAXATION SCHEME

Using the 3D multiscale model, we have implemented a parallel Simulated Annealing [4, 6, 7] algorithm. First, we partition the pyramid $\bar{\mathcal{S}}$ into disjoint updating sets so that pixels which belong to the same set are conditionally independent, given the data of all the other sets. This enable us to update different levels at the same time. Then, to the higher levels, we associate higher temperatures which enable to be less sensitive to local minima. However at a finer resolution, the relaxation is performed at a lower temperature (at the bottom level, it is closed to 0). Such an algorithm, compared to its monogrid version, demands 1.5 time more processors and (n^2+1) more communications (n is the block size) for each virtual processor.

EXPERIMENTAL RESULTS

We have compared the Gibbs sampler [4] and Iterated Conditional Mode [5] using three models for each algorithm (monogrid, multiscale and 3D multiscale). All tests have been conducted on a Connection Machine CM200 with $8K$ processors.

The first image is a synthetic image with 4 classes and a $SNR = 3dB$ (see Figure 3). In Table 1, we give for each model and for each method the number of iterations, the computer time, the error of the classification (= the nunber of misclassified pixels) and the parameter β (see Equations (5), (8), (9)) and γ (see Equation (10)).

The second image is a SPOT image (see Figure 4) with 4 classes.

For both images, the label pyramid was generated with 4 levels.

CONCLUSION

In this paper, we have proposed a new multiscale MRF model modifying a classical one. The proposed modification consists of a new interaction scheme between two neighbor grids in the label pyramid. The new connections allow to propagate local interactions more efficiently. Another advantage is that the relaxation algorithm can be run in parallel on the entire pyramid. On the other hand, these interactions make the model more complex, which demands computationally more expensive algorithms.

References

[1] C. Bouman. A Multiscale Image Model for Bayesian Image Segmentation. Technical Report TR-EE 91-53, Purdue University, 1991.

[2] P. Perez and F. Heitz. Multiscale Markov Random Fields and Constrained Relaxation in Low Level Image Analysis. In *Proc. ICASSP*, San-Francisco, March 1992.

[3] Z. Kato, J. Zerubia, and M. Berthod. Satellite image classification using a modified Metropolis dynamics. In *Proc. ICASSP*, San-Francisco, California, USA, March 1992.

[4] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.

[5] J. Besag. On the statistical analysis of dirty pictures. *Jl. Roy. Statis. Soc. B.*, 1986.

[6] P.J. Van Laarhoven and E.H. Aarts. Simulated annealing : Theory and applications. *Reidel Pub., Dordrecht, Holland*, 1987.

[7] R. Azencott. *Parallel Simulated Annealing : an overview of basic techniques*, pages 37–46. Wiley, 1992.

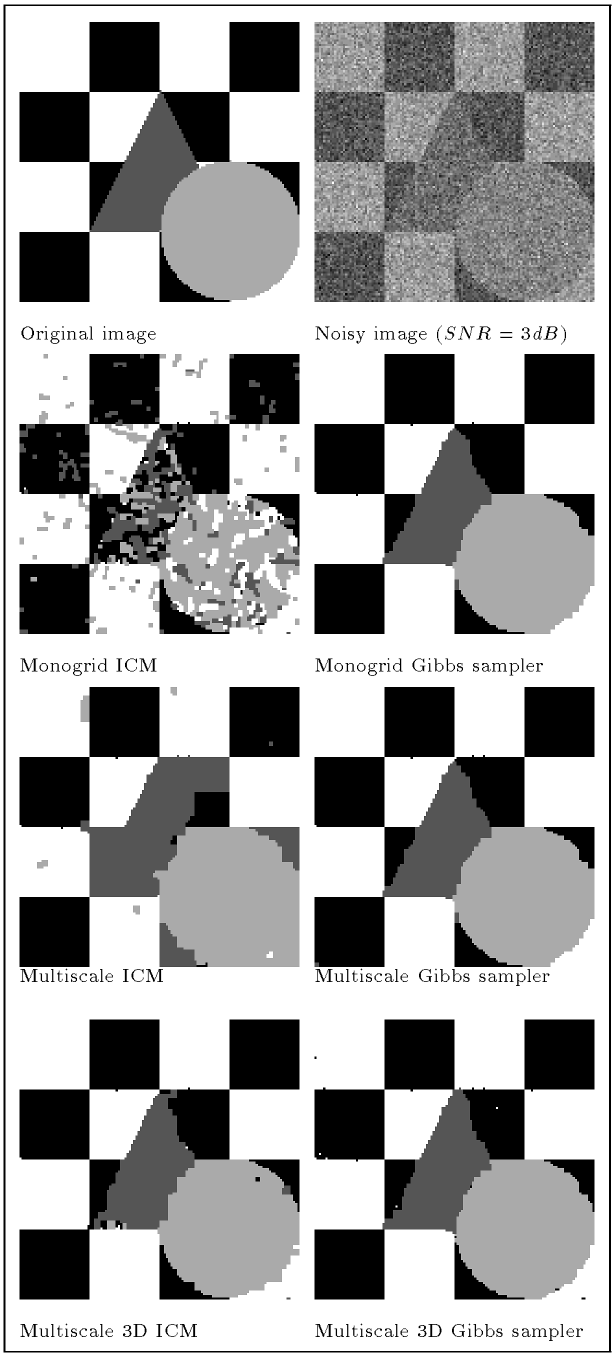

Figure 3: Results obtained on a noisy synthetic image (128×128) with 4 classes

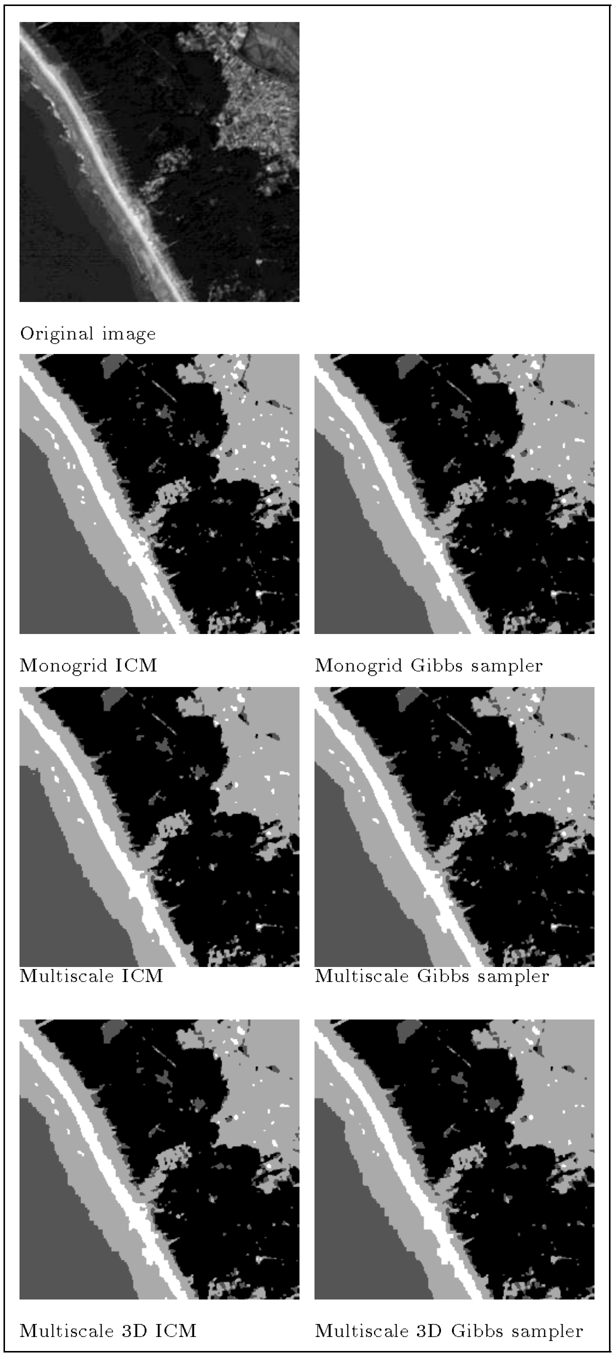

Figure 4: Results obtained on a SPOT image (256×256) with 4 classes