Multiscale Markov Random Field models for Parallel Image Classification

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

Abstract

In this paper, we are interested in multiscale Markov Random Field (MRF) models. It is well known that multigrid methods can improve significantly the convergence rate and the quality of the final results of iterative relaxation techniques. Herein, we propose a new hierarchical model, which consists of a label pyramid and a whole observation field. The parameters of the coarse grid can be derived by simple computation from the finest grid. In the label pyramid, we have introduced a new local interaction between two neighbor grids. This model gives a relaxation algorithm which can be run in parallel on the entire pyramid. On the other hand, the new model allows to propagate local interactions more efficiently giving estimates closer to the global optimum for deterministic as well as for stochastic relaxation schemes. It can also be seen as a way to incorporate cliques with far apart sites for a reasonable price.

1 Introduction

Herein, we are interested in the following general problem: we are given a set of sites $\mathcal{S} = \{s_1, s_2, \ldots, s_N\}$ and $\mathcal{F} = \{f_s\}_{s\in\mathcal{S}}$ a set of image data. Each of these sites may belong to any one 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 $\mathcal{S}$. $(\omega, \mathcal{F})$ is a MRF with respect to a chosen neighborhood-system $\mathcal{G} = \{\mathcal{G}_s\}_{s\in\mathcal{S}}$. Using Hammersley-Clifford theorem [1], $(\omega, \mathcal{F})$ follows a Gibbs distribution with an energy function $U(\omega, \mathcal{F}) = \sum_{C\in\mathcal{C}} V_C(\omega, \mathcal{F})$. C denotes a clique of $\mathcal{G}$, and $\mathcal{C}$ the set of all cliques (also $C_s = \{C : s \in C\}$). The restriction of ω to the sites of a given clique C is denoted by ω_C. The potential function $V_C(\omega_C)$ is defined for every C in $\mathcal{C}$ and every ω in Ω, where $\Omega = \Lambda^N$ is the set of all possible L^N discrete labelings. It expresses the local interaction between the labels of the clique. The Maximum A Posteriori (MAP) estimate of the label field is defined by:

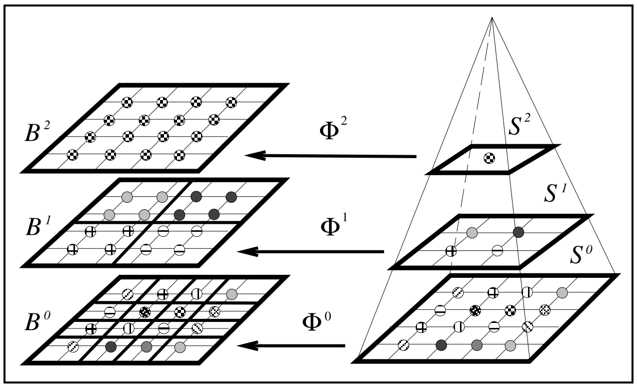

Figure 1: The isomorphism Φ^i between $\mathcal{B}^i$ and $\mathcal{S}^i$.

$$\hat{\omega} = \arg\max_{\omega\in\Omega} P(\omega \mid \mathcal{F}) = \arg\min_{\omega\in\Omega} U(\omega, \mathcal{F}) \qquad (1)$$

The initial problem is reduced to a combinatorial optimization problem, namely to the minimization of a non-convex energy function. Several approaches have been proposed to solve this task, such as Simulated Annealing (SA) [2, 3], Iterated Conditional Mode (ICM) [4], Modified Metropolis Dynamics (MMD) [5]... Multigrid schemes have also been proved to be very efficient for energy minimization [6]. In this paper, we propose a hierarchical multiscale model, called "3D multiscale" since we have interactions in three dimensions. The base of our approach is a classical multiscale model extensively studied by F. Heitz et al. in [7] and in [8]. Finally, we compare the results obtained with a monogrid, a multiscale and a 3D multiscale image classification model on noisy synthetic data and SPOT images.

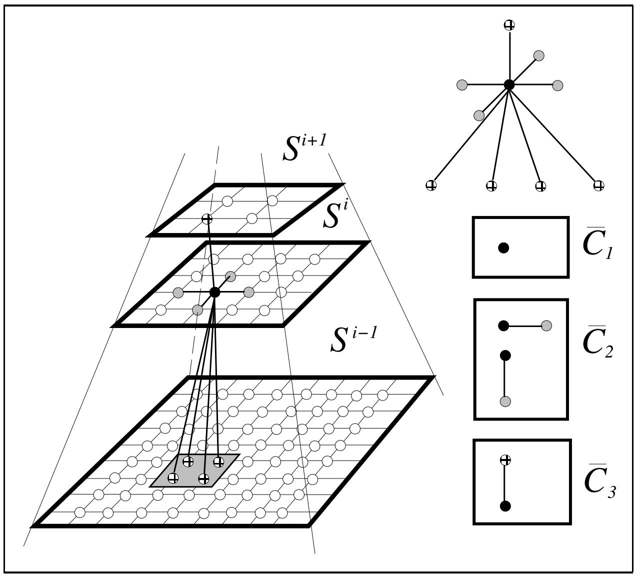

Figure 2: The neighborhood system $\bar{\mathcal{G}}$ and the cliques $\bar{\mathcal{C}}_1$, $\bar{\mathcal{C}}_2$ and $\bar{\mathcal{C}}_3$.

2 The 3D Multiscale Model

In the following, we will focus on a MRF with a first order neighborhood-system whose energy function is given by:

$$U(\omega, \mathcal{F}) = U_1(\omega, \mathcal{F}) + U_2(\omega) \tag{2}$$

where U_1 (resp. U_2) denotes the energy of the first order (resp. second order) cliques. To generate a multigrid MRF model, let us divide the initial grid into blocks of $n \times n$, typically 16 (4×4) neighbor pixels. We consider that the label assigned to the pixels of a block is constant over all the pixels of the block. 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 $\mathcal{B}^i = \{b^i_1, \ldots, b^i_{N_i}\}$ ($N_i = N/n^{2i}$) denotes 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 $\mathcal{S}$:

$$b^i_k \text{ and } b^i_l \text{ are neighbors} \iff \tag{3}$$
$$\begin{cases} b^i_k \equiv b^i_l \text{ or} \\ \exists C \in \mathcal{C} \mid C \cap b^i_k \neq \emptyset \text{ and } C \cap b^i_l \neq \emptyset \end{cases}$$

Let us partition the original set $\mathcal{C}$ into two disjoint subsets $\mathcal{C}^i_k$ (cliques which are included in b^i_k) and $\mathcal{C}^i_{k,l}$(cliques which sit astride two neighboring blocks $\{b^i_k, b^i_l\}$). It is obvious from this partition that our energy function (see Equation (2)) can be decomposed in the following way:

$$\begin{aligned}
U_1(\omega, \mathcal{F}) &= \sum_{s \in \mathcal{S}} V_1(\omega_s, f_s) \\
&= \sum_{b^i_k \in \mathcal{B}^i} \underbrace{\sum_{s \in b^i_k} V_1(\omega_s, f_s)}_{V_1^{\mathcal{B}^i}(\omega^i_k, \mathcal{F})} \\
&= \sum_{b^i_k \in \mathcal{B}^i} V_1^{\mathcal{B}^i}(\omega^i_k, \mathcal{F}) \qquad (4) \\
U_2(\omega) &= \sum_{C \in \mathcal{C}} V_2(\omega_c) \\
&= \sum_{b^i_k \in \mathcal{B}^i} \underbrace{\sum_{C \in \mathcal{C}^i_k} V_2(\omega_c)}_{V_k^{\mathcal{B}^i}(\omega^i_k)} \\
&\quad + \sum_{\{b_k, b_l\} neighbors} \underbrace{\sum_{C \in \mathcal{C}^i_{k,l}} V_2(\omega_c)}_{V_{k,l}^{\mathcal{B}^i}(\omega^i_k, \omega^i_l)} \\
&= \sum_{b^i_k \in \mathcal{B}^i} V_k^{\mathcal{B}^i}(\omega^i_k) \\
&\quad + \sum_{\{b_k, b_l\} neighbors} V_{k,l}^{\mathcal{B}^i}(\omega^i_k, \omega^i_l) \qquad (5)
\end{aligned}$$

Now, we can define a 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 energy function on the grid $\mathcal{S}^i$ ($i = 0, \ldots, M$) is derived from Equations (4) and (5):

$$\begin{aligned}
U^i(\omega^i, \mathcal{F}) &= U^i_1(\omega^i, \mathcal{F}) + U^i_2(\omega^i) \\
&= U_1(\Phi^i(\omega_i), \mathcal{F}) + U_2(\Phi^i(\omega_i)) \\
\text{where } U^i_1(\omega^i, \mathcal{F}) &= \sum_{k \in \mathcal{S}^i} (V_1^{\mathcal{B}^i}(\omega^i_k, \mathcal{F}) + V_k^{\mathcal{B}^i}(\omega^i_k)) \\
&= \sum_{k \in \mathcal{S}^i} V^i_1(\omega^i_k, \mathcal{F}) \qquad (6) \\
\text{and } U^i_2(\omega^i) &= \sum_{\{k,l\} neighbors} V_{k,l}^{\mathcal{B}^i}(\omega^i_k, \omega^i_l) \\
&= \sum_{C^i \in \mathcal{C}^i} V^i_2(\omega^i_C) \qquad (7)
\end{aligned}$$

where C^i is a second order clique corresponding to the definition in Equation (4) and $\mathcal{C}^i$ is the set of cliques on the grid i.

Let $\bar{\mathcal{S}} = \{\bar{s}_1, \dots, \bar{s}_{\bar{N}}\} = \bigcup_{i=0}^{M} \mathcal{S}^i$ ($\bar{N} = \sum_{i=0}^{M} N^i$) denote the sites of the pyramid. We define the following function Ψ between two neighbor levels, which assigns to a site its descendants (that is the sites of the block corresponding to this site at the lower level):

$$\begin{array}{rcl} \Psi: & & \mathcal{S}^i \longrightarrow \mathcal{S}^{i-1} \\ \Psi(\bar{s}) & = & \{\bar{r} \mid \bar{s} \in \mathcal{S}^i \Rightarrow \bar{r} \in \mathcal{S}^{i-1} \text{ and } b_{\bar{r}}^{i-1} \subset b_{\bar{s}}^i\} \end{array} \quad (8)$$

It is clear that Ψ^{-1} will assign to a site its ancestor (that is the site at the upper level corresponding to the block of this site). Now we can define on these sites the following neighborhood-system (see Figure 2):

$$\bar{\mathcal{G}} = (\bigcup_{i=0}^{M} \mathcal{G}^i) \cup \{\Psi^{-1}(\bar{s}) \cup \Psi(\bar{s}) \mid \bar{s} \in \bar{\mathcal{S}}\} \quad (9)$$

where $\mathcal{G}^i$ is the neighborhood structure of the i^{th} level. 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). Let $\bar{\Omega}$ denotes the configuration-space of the pyramid:

$$\bar{\Omega} = \Gamma^0 \times \Gamma^1 \times \dots \times \Gamma^M = \{\bar{\omega} \mid \bar{\omega} = (\omega^0, \omega^1, \dots, \omega^M)\} \quad (10)$$

The model on the pyramid defines a MRF, whose energy function is given by:

$$\bar{U}(\bar{\omega}, \mathcal{F}) = \bar{U}_1(\bar{\omega}, \mathcal{F}) + \bar{U}_2(\bar{\omega}) \quad (11)$$

$$\begin{aligned} \bar{U}_1(\bar{\omega}, \mathcal{F}) &= \sum_{\bar{s} \in \bar{\mathcal{S}}} \bar{V}_1(\bar{\omega}_{\bar{s}}, \mathcal{F}) = \sum_{i=0}^{M} \sum_{s^i \in \mathcal{S}^i} V_1^i(\omega_{s^i}^i, \mathcal{F}) \\ &= \sum_{i=0}^{M} U_1^i(\omega^i, \mathcal{F}) \end{aligned} \quad (12)$$

$$\begin{aligned} \bar{U}_2(\bar{\omega}) &= \sum_{C \in \bar{\mathcal{C}}_2} \bar{V}_2(\bar{\omega}_C) + \sum_{C \in \bar{\mathcal{C}}_3} \bar{V}_2(\bar{\omega}_C) \\ &= \sum_{i=0}^{M} U_2^i(\omega^i) + \sum_{C \in \bar{\mathcal{C}}_3} \bar{V}_2(\bar{\omega}_C) \\ &= \sum_{i=0}^{M} \sum_{C \in \mathcal{C}^i} V_2^i(\omega_c^i) + \sum_{C \in \bar{\mathcal{C}}_3} \bar{V}_2(\bar{\omega}_C) \end{aligned} \quad (13)$$

The relaxation scheme on the pyramid consists of the minimization of the energy function $\bar{U}(\bar{\omega}, \mathcal{F})$.

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).

A	iter.	total time	time/iter.	error	β	γ
Gibbs	89	10.39 sec.	0.117 sec.	2576	1.0	—
B	—	—	—	—	—	—
Gibbs	146	14.7 sec.	0.1 sec.	2118	1.0	—
C	—	—	—	—	—	—
Gibbs	42	460.9 sec.	10.97 sec.	1231	1.0	0.2

Table 2: The results on the noisy synthetic image with 16 classes (A-monogrid, B-multiscale, C-3D multiscale).

3 The Parallel Relaxation Scheme

Using the 3D multiscale model, we have implemented a parallel Simulated Annealing [2, 3, 9] 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. The theoretical study of the convergence in the case of an homogeneous or an inhomogeneous annealing has been conducted in [10].

4 Experimental Results

We have applied this model for the supervised image classification and we have compared the Gibbs

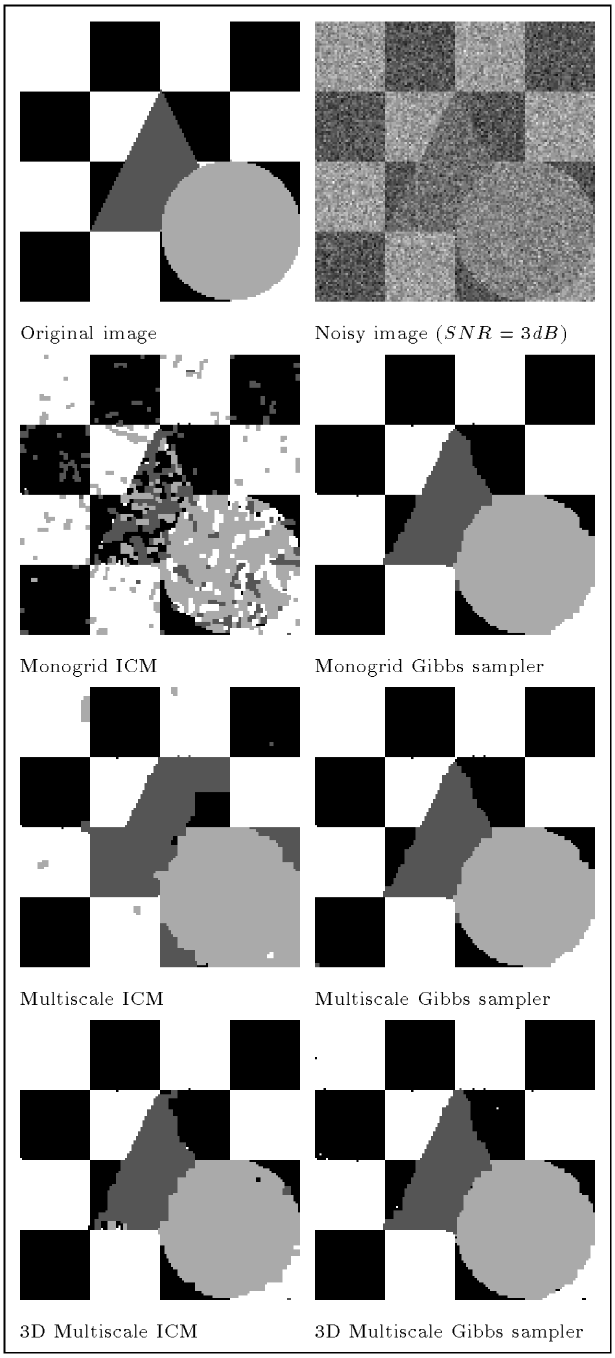

Figure 3: Results obtained on a noisy synthetic image (128×128) with 4 classes

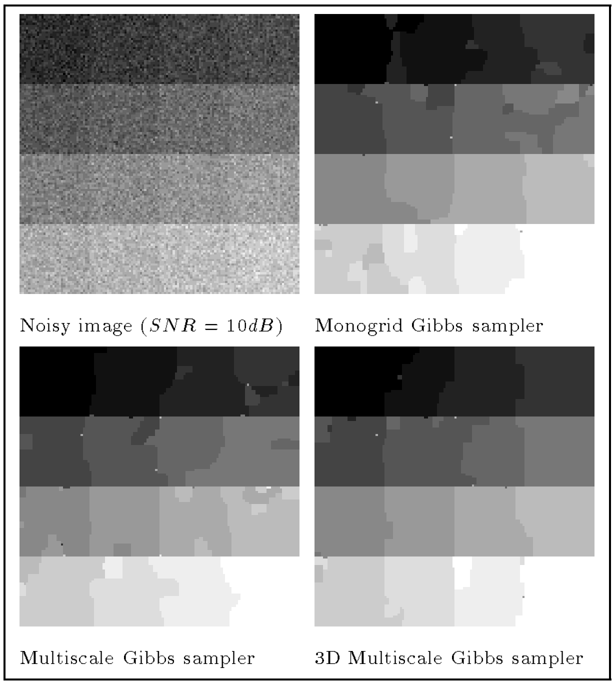

Figure 4: Results obtained on a noisy synthetic image (128×128, $SNR = 10dB$) with 16 classes

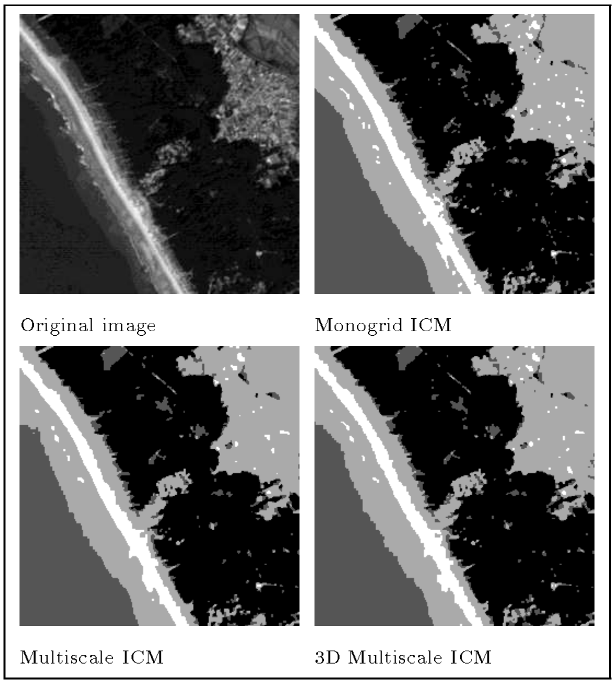

Figure 5: Results obtained on a SPOT image (256×256) with 4 classes

sampler [2] to Iterated Conditional Mode [4] using three models for each algorithm (monogrid, multiscale and 3D multiscale). The detailed equations corresponding to the classification problem can be found in [11].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 number of misclassified pixels) and the clique-potentials for the second order cliques on the same level (β) and between levels (γ). The second image is a synthetic image with 16 classes and a $SNR = 10dB$ (see Figure 4). In Table 2, we give the values of the chosen parameters for the second image. It can be noticed that the Gibbs sampler with the 3D model gives a better result on this image, whereas with the previous image (Figure 3) all the models give nearly the same results. Finally, we deal with a SPOT image (see Figure 5) with 4 classes. For each images, the label pyramid has been generated with 4 levels.

5 Conclusion

In this paper, we have proposed a hierarchical multiscale MRF model. New interactions have been introduced between neighboring grids in the label pyramid. These 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] R. Azencott. Markov fields and image analysis. *Proc. AFCET, Antibes*, 1987.

[2] 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.

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

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

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

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

[7] F. Heitz, P. Perez, E. Memin, and P. Bouthemy. Parallel Visual Motion Analysis Using Multiscale Markov Random Fields. In *Workshop on Motion*, Princeton, October 1991.

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

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

[10] Z. Kato, M. Berthod, and J. Zerubia. Multiscale markov random field models for parallel image classification. Technical report, INRIA, 1993.

[11] Z. Kato, M. Berthod, and J. Zerubia. Parallel Image Classification using Multiscale Markov Random Fields. In *Proc. ICASSP*, Minneapolis, April 1993.