A Hierarchical Markov Random Field Model for Image Classification

Zoltan KATO, Marc BERTHOD and Josiane ZERUBIA. *INRIA - Sophia Antipolis - FRANCE*

Abstract. In this paper, we propose a hierarchical Markov Random Field (MRF) model. This model is based on a classical multiscale model, which consits 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 with a new annealing scheme: The Multi-Temperature Annealing (MTA) scheme, which consists to associate higher temperatures to higher levels beeing less sensitive to local minima at coarser grids. The model was tested on different synthetic and real images. The algorithms was implemented on a Connection Machine CM200.

1 Introduction

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 [1]:

Let $\mathcal{S} = \{s_1, s_2, \ldots, s_N\}$ be a set of sites or pixels and $\mathcal{F} = \{f_s\}_{s \in \mathcal{S}}$ the set of grey levels. 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 $\omega_s \in \Lambda$ to each site s in $\mathcal{S}$. Choosing a first order neighborhood-system, we can define a MAP estimator of the label field: $\hat{\omega} = \arg\max_{\omega \in \Omega} P(\omega \mid \mathcal{F}) = \arg\min_{\omega \in \Omega} U(\omega, \mathcal{F})$ where $U(\omega, \mathcal{F}) = U_1(\omega, \mathcal{F}) + U_2(\omega)$ 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:

$$U_1(\omega, \mathcal{F}) = \sum_{s \in \mathcal{S}} \left(\log(\sqrt{2\pi}\sigma_{\omega_s}) + \frac{(f_s - \mu_{\omega_s})^2}{2\sigma^2_{\omega_s}} \right) \quad (1)$$

$$U_2(\omega) = \sum_{C \in \mathcal{C}} V_2(\omega_C) \text{ where} \quad (2)$$

$$V_2(\omega_C) = V_{\{s,r\}}(\omega_s, \omega_r) = \begin{cases} -\beta & \text{if } \omega_s = \omega_r \\ +\beta & \text{if } \omega_s \neq \omega_r \end{cases} \quad (3)$$

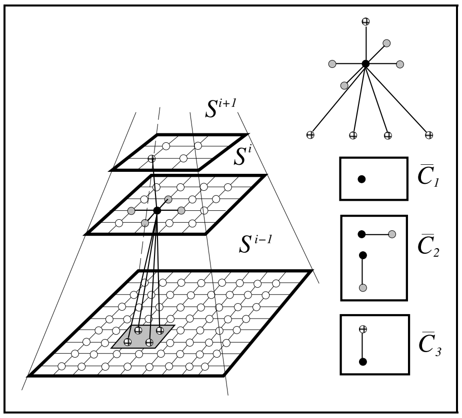

Figure 1: The label pyramid with the neighborhood system.

2 The Hierarchical Model

To generate a multigrid MRF model [2, 3], 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 $\mathcal{S}$ and the clique potentials are derived from the initial potentials: The first order potentials are the sum of the potentials on the finest grid which are included in the corresponding block. Similarly, the second order potentials are the sum of the potentials on the finest grid which are located between the two corresponding block.

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 $\mathcal{S}^0 = \mathcal{S}$. We can define the same neighborhood system on these grids as on the scales and the potentials are inherited from the blocks. We have then the following energy function at the level i:

$$\begin{aligned}
U_1^i(\omega^i, \mathcal{F}) &= \sum_{s^i \in \mathcal{S}^i} V_1^i(\omega^i_{s^i}, \mathcal{F}) \text{ where} \\
V_1^i(\omega^i_{s^i}, \mathcal{F}) &= \sum_{s \in b^i_{s^i}} V_1(\omega_s, f_s) + \sum_{C \in \mathcal{C}^i_{s^i}} V_2(\omega_C) \\
&= \sum_{s \in b^i_{s^i}} \left(\log(\sqrt{2\pi}\sigma_{\omega_s}) + \frac{(f_s - \mu_{\omega_s})^2}{2\sigma^2_{\omega_s}} \right) \\
&\quad -p^i \beta \qquad (4) \\
U_2^i(\omega) &= \sum_{C^i = \{r^i, s^i\} \in \mathcal{C}^i} V_2^i(\omega_{C^i}) \text{ where} \\
V_2^i(\omega_{C^i}) &= \sum_{\{r,s\} \in \mathcal{C}^i_{r^i, s^i}} V_2(\omega_r, \omega_s) \\
&= \begin{cases} -q^i \beta & \text{if } \omega_r = \omega_s \\ +q^i \beta & \text{if } \omega_r \neq \omega_s \end{cases} \qquad (5)
\end{aligned}$$

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 $q^i = n^i$.

Let $\bar{\mathcal{S}} = \{\bar{s}_1, \ldots, \bar{s}_{\bar{N}}\} = \bigcup_{i=0}^M \mathcal{S}^i$ ($\bar{N} = \sum_{i=0}^M N^i$) denote the sites of the pyramid. On these sites, we define the following neighborhood-structure (see Figure 1): 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 1). The configuration-space of the pyramid is the set of the configuration-spaces of 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 (5) and (4)). 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

$$\bar{V}_2(\bar{\omega}_C) = \bar{V}_{\{\bar{s}, \bar{r}\}}(\bar{\omega}_{\bar{s}}, \bar{\omega}_{\bar{r}}) = \begin{cases} -\gamma & \text{if } \bar{\omega}_{\bar{s}} = \bar{\omega}_{\bar{r}} \\ +\gamma & \text{if } \bar{\omega}_{\bar{s}} \neq \bar{\omega}_{\bar{r}} \end{cases} \qquad (6)$$

3 The Multi-Temperature Annealing

We have implemented a parallel Simulated Annealing [4] algorithm using the multi-temperature annealing scheme. First, we partition the pyramid 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). The hierarchical model, 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 complete theoretical study of the convergence has been conducted in [5].

4 Conclusion

New interactions have been introduced between neighboring grids in the label pyramid to generate a hierarchical model. These connections allow to propagate local interactions more efficiently however they make the model more complex, which demands computationally more expensive algorithms. We have also proposed a new annealing scheme, the Multi-Temperature Annealing, which is more efficient in our case than the classical ones.

References

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

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

[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] Z. Kato, M. Berthod, and J. Zerubia. A hierarchical markov random field model and multi-temperature annealing for parallel image classification. Research report, INRIA, 1993.