AN ADAPTIVE MRF MODEL FOR BOUNDARY-PRESERVING SEGMENTATION OF MULTISPECTRAL IMAGES

Ciro D'Elia, Giovanni Poggi, Giuseppe Scarpa
Department of Electronic and Telecommunication Engineering
Università "Federico II" di Napoli, via Claudio 21, 80125 Napoli, Italy
Tel: +39 081 7683151; Fax: +39 081 7683149
email: {c.delia,poggi,scarpa}@unina.it

ABSTRACT

MRF models are widely used in remote-sensing image segmentation to take into account dependencies among neighboring pixels. Compared to non-contextual techniques, MRF-based techniques provide much smoother segmentation maps, as they are able to counter the effects of sensor noise. Because of finite resolution of sensors, however, many boundary pixels are mixed (comprise two different land covers) and are incorrectly classified as belonging to a third class. Here we propose an adaptive tree-structured MRF model, which largely reduces such classification errors and increases map smoothness without sacrificing classification fidelity.

1 INTRODUCTION

Image segmentation is an important processing step for a number of subsequent tasks, such as compression, classification, image understanding, and so on. This is all the more true for remote sensing images where the sheer amount of collected data calls for some form of compact representation of the information that makes it easier to handle and use it. Segmentation has two main, and partially contrasting, requirements:

1. to obtain a simple representation, with as few regions as possible, separated by smooth boundaries;
2. to obtain a faithful map of the actual land covers of the region.

Reaching these goals becomes quite difficult in the presence of a strong sensor noise (such as in SAR systems) which reduces the reliability of observables and leads to oversegmenting the image. In order to counter the effects of noise it is very helpful to model the *a priori* knowledge on the image and incorporate it in a MAP (maximum a posteriori probability) estimation procedure.

Building an accurate and manageable statistical model of an image is not an easy task, and still an open research field. In recent years, there has been a growing interest on MRF (Markov random field) models [1,2] because of their flexibility, sound mathematical framework, and fairly good numerical tractability. MRF-based segmentation techniques (e.g., [3,4]) have proven much superior than conventional clustering algorithms, provided that a good model of the image is available.

The main drawback of the MRF-based algorithms is their computational burden, which can become quite substantial depending on image size and model complexity, especially when a large number of classes is considered. For this reason, we have proposed [5,6] a tree-structured Markov random field (TS-MRF) model where the segments (homogeneous regions) of the image are organized in a binary tree. This model allows one to deal only with binary fields, and to segment smaller and smaller regions as the procedure goes on, with a huge reduction of complexity. Tree structured segmentation provides additional advantages [7]: since each region is associated with a different binary MRF, one can easily deal with nonstationary behaviors, and also gain insight about the region properties through the associated field parameters. Moreover, a tree-structured segmentation can be easily tracked by a supervisor and become a handy tool for interactive use.

In this paper we show that the TS-MRF model can be easily adapted to prevent the fragmentation of region boundaries, a phenomenon due essentially to the finite resolution of the sensors and observed both in flat and tree-structured MRFs. The adaptive model does not increase complexity and provides much smoother maps, especially useful for subsequent image compression [4,8], without compromising the segmentation fidelity. After a more detailed description of TS-MRF in Section 2, we present the adaptive model in Section 3 together with some experimental results.

2 TREE-STRUCTURED MRF

Image segmentation can be easily formulated as a MAP estimation problem. Suppose each pixel of the image $\mathcal{S}$ belongs to one of K different classes, and let $x_s \in \{1, \dots, K\}$ indicate the class of pixel s. Then $\mathbf{x} = \{x_s, s \in \mathcal{S}\}$ is the segmentation of the image $\mathcal{S}$ in K classes. Of course, $\mathbf{x}$ is unknown, and must be estimated from the observable data $\mathbf{y} = \{y_s, s \in \mathcal{S}\}$. We

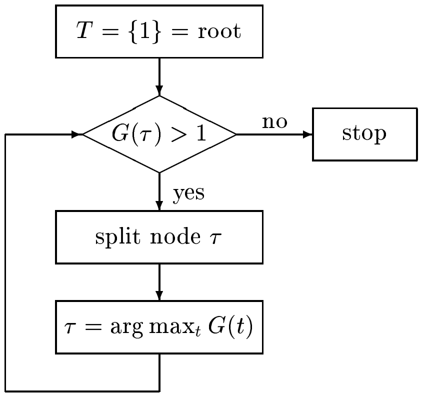

Figure 1: Flow chart of the TS-MRF algorithm

are interested in multispectral images, composed of several (up to some hundreds) spectral bands, and hence y_s, named the spectral signature of pixel s, is itself a vector. Modeling all quantities as random variable/fields (capital letters), we accept as our segmentation $\widehat{\mathbf{x}}$ the most likely realization of $\mathbf{X}$ given the field of observables, namely,

$$\widehat{\mathbf{x}} = \arg\max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}) = \arg\max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$$

The observables are modeled as conditionally independent Gaussian, given the class, namely $p(\mathbf{y}|\mathbf{x}) = \prod_s p(y_s|x_s)$, with $p(y_s|k) \sim N(\mu_k, \Sigma_k)$. As for the field of classes, it is convenient to model it as a Markov random field [1,2]. Indeed, this is a reasonably simple, yet general, model which keeps into account the spatial dependencies in the image through the conditional probability that a pixel belong to a given class given the classes of its neighbors. As a result, $\mathbf{X}$ has a Gibbs prior distribution

$$p(\mathbf{x}) = \frac{1}{Z} \exp[\sum_{c \in \mathcal{C}} V_c(\mathbf{x}, \theta)] \qquad (1)$$

where Z is a normalizing constant, and the $V_c(\cdot, \theta)$'s are potential functions, defined on suitable cliques c of the image, and depending on some hyperparameters θ.

Given this model, the segmentation problem amounts to maximizing the function $p(\mathbf{y}|\mathbf{x})p(\mathbf{x})$ over $\mathbf{x}$, where all the quantities K, μ_k, Σ_k and θ are in general unknown and must be estimated themselves from the data. Due to the inherent complexity of this problem, in practical applications one must resort to heuristics that reduce the search complexity, and accept suboptimal solutions.

To drastically reduce the search complexity, in [5] we have introduced a tree-structured MRF model, where the full segmentation is obtained through a sequence of binary segmentations. More precisely, the whole image is associated to the root node $t = 1$ of a tree T, and is segmented in two regions using a binary MRF model. The two new regions, associated with the children of the root, $t = 2$ and $t = 3$, can be likewise segmented by means of newly defined local binary MRFs, and the growth of the tree continues until a suitable stopping condition is met. Therefore, each node t of the tree is associated with a region of the image $\mathcal{S}_t$, a field of observables $\mathbf{Y}_t$ with realization $\mathbf{y}_t$, a binary MRF $\mathbf{X}_t$ with realization $\mathbf{x}_t$, and a set of parameters $\{\mu_t, \Sigma_t, \theta_t\}$. The leaves of the tree partition the image in K disjoint regions, namely provide the desired segmentation.

The growth of the tree can be based exclusively on local decisions. In fact, a split gain $G(t)$ is associated with each node t,

$$G(t) = \frac{p(\mathbf{y}_t|\mathbf{x}_t)p(\mathbf{x}_t)}{p(\mathbf{y}_t|\widetilde{\mathbf{x}}_t)p(\widetilde{\mathbf{x}}_t)}$$

defined as the likelihood ratio between the two hypotheses of splitting the region in two (according to the realization $\mathbf{x}_t$ of the local binary MRF) or leaving it unaltered, (namely, accepting a uniform field $\widetilde{\mathbf{x}}_t$ with all pixels in the same class). $G(t) > 1$ indicates that region $\mathcal{S}_t$ is better described by a two-class field rather than by a uniform field. When all gains are less than 1 the tree stops growing (see Fig.1).

The use of binary fields only, together with the locality of the splitting (the segmentation of a region does not depend on other regions) leads to a significant reduction of the computational complexity with respect to the case where a flat K-class MRF is used. Moreover, the tree structure allows for a simpler interpretation and handling of the segmentation map [7].

On the other hand, the tree-structure introduces also some additional constraint, which tend to impair the segmentation performance. For example, in a three-class image, it can easily happen that in the very first step of the algorithm a single region be split in two because of noise, leading ultimately to an oversegmentation. This problem, however, can be easily solved in various ways. In [6] we proposed a split-and-merge procedure, based on a merge gain $M(t', t'')$ dual to the split gain $G(t)$, that effectively restores the segmentation quality.

Another annoying phenomenon that occurs both with tree-structured and flat models is the fragmentation of boundaries between region. In next Section we describe this problem and its origins and propose an effective solution based on a simple adaptive MRF model.

3 BOUNDARY FRAGMENTATION AND PROPOSED ADAPTIVE MODEL

In Fig.2 we show two bands of the multispectral image used for the experiments, together with the K-class segmentation map provided by the TS-MRF and an enlarged detail of the same map.

It clearly appears that several region boundaries are "fragmented", namely, there are many boundary pixels which are associated with none of the adjacent regions

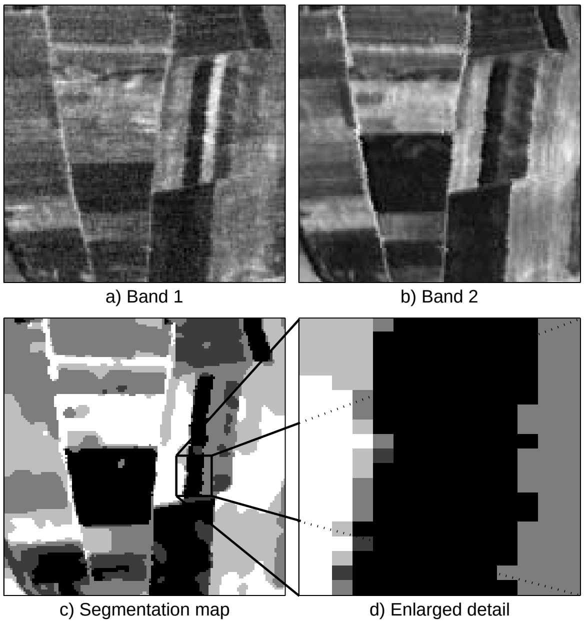

Figure 2: Boundary fragmentation

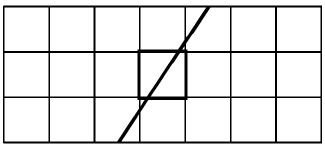

Figure 3: Cell integration. Center pixel is mixed

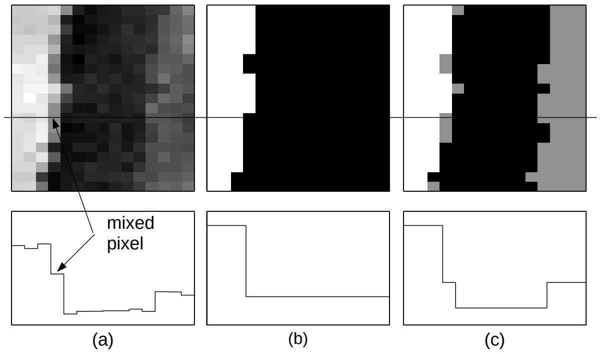

Figure 4: Segmentation error. A mixed-class pixel (a) is first associated with the "dark" region (b), and when this is split, with the wrong subregion (c).

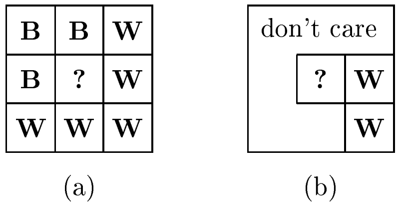

Figure 5: Use of prior information: (a) a two-class dilemma; (b) context gives a clear indication.

but are attached a different label. By visual inspection, it is clear that quite often this phenomenon does not correspond to a ground truth, namely, there is no actual "third" region in between the two neighboring regions. Instead, it is likely due to the finite resolution of the sensor which, in boundary cells, happen to integrate contributions from several land covers (see Fig.3), originating a spectral response that is quite different from those of the adjacent regions.

Therefore, in boundary cells, the observables are not very reliable. In a binary split, the pixel is attributed to either one of the two adjacent regions, but when this is further split chances are that this "uncertain" pixel is erroneously classified. Fig.4 further clarifies this point by showing how fragmentation arises in the detail of Fig.2(d). The first split separates the bright region on the left from the dark right region: some boundary pixels are classified as dark although they have intermediate characteristics. The second split operates on the dark region and further separates it: because of their mixed nature, some of the former boundary pixels are attributed to the wrong subregion. Obviously, this makes much less sense than choosing either one of the neighboring regions and should be avoided unless the observables give a clear indication.

It is clear that the data by their very nature tend to cause such wrong segmentations near region boundaries, and in fact this phenomenon arises just as well when flat (non-structured) MRF models are used.

The problem is not only that the observables are unreliable, but also that prior information does not help correctly resolving such ties. However, in the case of tree-structured MRF a simple fix is available to make better use of contextual information. Consider the example of Fig.5(a): here the context does not give much help, because the neighborhood of the target pixel is almost evenly divided (5 "white", 3 "black") and the observable *should* play a relevant role. In the case of Fig.5(b), however, the target pixel is surrounded by either white or "don't care" pixels, namely pixels outside the region of interest. Here, the relevant context is all "white" and it seems reasonable to favor strongly this hypothesis reducing the relative importance of the observables.

Based on these observation, we decided to change slightly the binary MRF model in order to ripristinate the relative importance of the context even when the number of relevant cliques is reduced.

We use a very simple model, with neighborhood η^2 and only binary cliques, defined on couples of 8-connected pixels. The potentials are defined as

$$V_c(\mathbf{x}) = \begin{cases} \beta & \text{if } x_s = x_{s'}, \quad s, s' \in c \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

with the edge penalty β as the only hyperparameter of the model. Trivial manipulations show that, by this model, the ratio of the two *a priori* probabilities given the neighborhood depends only on the number of

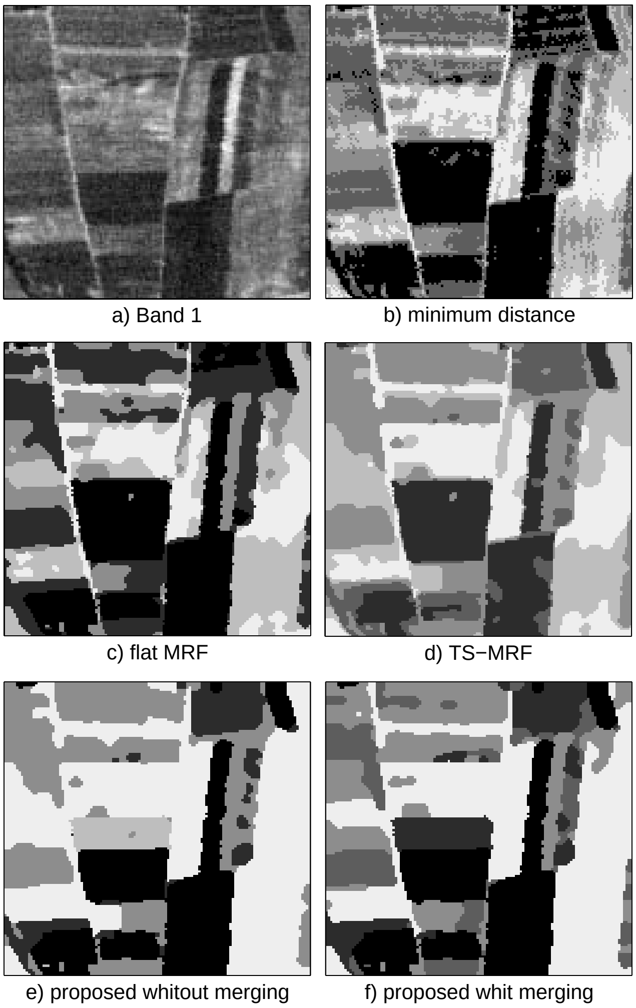

Figure 6: Some segmentation results.

surrounding "black" and "white" pixels, N_B and N_W, namely,

$$\frac{\Pr(X_s = W|\eta_s^2)}{\Pr(X_s = B|\eta_s^2)} = \exp[\beta(N_W - N_B)] \qquad (3)$$

With this definition, a 5-3 and a 2-0 context are equivalent.

To restore the value of prior information, we simply increase adaptively the edge penalty value when the context is not complete. In particular, we use

$$\beta' = \frac{8}{N_B + N_W}\beta$$

by which a 2-0 context becomes equivalent to an 8-0 one. Of course, other choices are also possible, and generally work reasonably well, such as doubling β for boundary pixels, or setting it to a rather large value (this requires some prior experiments).

This simple modification has provided very good results in the experiments. To assess the performance of the proposed technique we take into account both smoothness and fidelity of the segmentation map, that are often contrasting requirements, as pointed out in the introduction. In Fig.6 we show again a band of the test image (a) together with its 5-class segmentation based on: minimum-distance clustering (b), a flat-MRF model (c), the tree-structured MRF model of [5] (d), and the proposed adaptive TS-MRF model without (e) or with (f) merging [6]. By visual inspection, we note that both the proposed technique (with or without merging) and flat-MRF guarantee quite a faithful segmentation map, as far as one can judge without the help of a ground truth. The TS-MRF map (d) is also acceptable, while the minimum distance map (b) is clearly degraded by noise. As for the smoothness, both flat-MRF and TS-MRF exhibit widespread boundary fragmentation, whereas the proposed technique all but eliminates this problem, providing a simpler map, much more convenient for subsequent compression [8].

In conclusion, thanks to the tree-structured MRF model, the proposed segmentation algorithm has much lower computational complexity than those based on flat-MRF. In addition, the use of an adaptive model prevents the fragmentation of region boundaries, and leads to smoother and more faithful maps, well suited for further higher-level processing.

REFERENCES

[1] S.Geman, D.Geman: "Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images", *IEEE Trans. Patt. Anal. Machine Intell.*, pp.721-741, Nov.1984.

[2] S.Z.Li: *Markov random field modeling in image analysis*, Springer-Verlag, 2001.

[3] J.K.Fwu, P.M.Djuric: "Unsupervised vector image segmentation by a tree-structure ICM algorithm," *IEEE Trans. Medical Imaging*, pp.871-880, Dec.1996.

[4] G.Gelli, G.Poggi, A.R.P.Ragozini: "Compression of multispectral images based on tree-structured MRF segmentation and transform coding", *Proc. IEEE IGARSS 99*, pp.1167-1170, June 1999.

[5] G.Poggi, A.R.P.Ragozini: "Image segmentation by tree-structured Markov random field", *IEEE Signal Proc. Letters*, pp.155-157, July 1999.

[6] C.D'Elia, G.Poggi, G.Scarpa: "Advances in the segmentation and compression of multispectral images" *Proc. IEEE IGARSS 01*, pp., June 2001.

[7] L.Breiman, J.H.Friedman, R.A.Olshen, C.J.Stone, *Classification and regression trees*, Wadsworth, Belmont (CA) 1984.

[8] G.Gelli, G.Poggi: "Compression of multispectral images by spectral classification and transform coding", *IEEE Trans. Image Proc.*, pp.476-489, Apr.1999.