MULTITEMPORAL IMAGE CLASSIFICATION WITH AUTOMATIC BUILDING OF TREE-STRUCTURED MRF MODELS

Raffaele Gaetano, Giovanni Poggi, and Giuseppe Scarpa

Dipartimento di Ingegneria Elettronica e delle Telecomunicazioni
Università Federico II di Napoli – via Claudio 21, 80125 Napoli, Italy
{gaetano,poggi,giscarpa}@unina.it

ABSTRACT

In this work we deal with the classification of remote-sensing images following a statistical approach. To take into account prior information on the class of images of interest we model the image as a tree-structured Markov random field (TS-MRF), so as to fit the intrinsic structure of the data. TS-MRF models are defined recursively and, as such, lead to the formulation and solution of the segmentation task as a recursive problem, so that the original K-ary segmentation is decomposed into a sequence of reduced-dimensionality steps, and hence to a much simpler and more manageable segmentation. Here, we propose a method to build automatically the underlying tree structure of the model, based on a metric which compares class features in order to establish the hierarchical relationships among classes, and apply the technique to the segmentation of multitemporal remote-sensing images.

Key-words: Image classification, remote-sensing images, Markov random fields, hierarchical models.

1. INTRODUCTION

Segmentation aims at the partition of an image in disjoint regions, each one homogeneous with respect to some properties like intensity, texture, shape, *etc*. Such a task is needed in many high-level processing and applications in such diverse fields as remote-sensing, medical imaging, image restoration, or video coding.

In remote-sensing applications, the segmentation task is very difficult because of the presence of significant noise components and the intrinsic complexity of the images. Hence, data modeling becomes quite critical, and the statistical approach may result advantageous with respect to other non-stochastic approaches. In this work, we resort to the Markov Random Field (MRF) probability model [1], and in particular to a class of MRF models, named Tree-Structured MRFs (TS-MRF) [2, 3], which has proven to be a powerful and manageable tool to address such segmentation problems.

Indeed, TS-MRF models are defined recursively and, as such, lead to the formulation and solution of the segmentation task as a recursive problem. As a consequence, the original K-ary segmentation is decomposed into a sequence of reduced-dimensionality (typically binary) steps, and hence to a much simpler and more manageable segmentation. Each step subdivides a large region into two or more component regions leading to a tree-structured representation of the image. Each image region is associated with a node in a tree of classes, with the elementary regions, associated with the leaves of such a tree, corresponding to the final segmentation map.

In this work we propose a technique to single out the underlying tree structure of the model, based on the progressive merging of classes, guided by a statistical measure, the merging gain [2], which quantifies similarity among classes. We then apply the proposed technique to the problem of segmenting multitemporal remote-sensing images, so as to detect changed and unchanged regions.

2. TS-MRF MODELS AND SEGMENTATION

A random field X defined on a lattice S is said to be a MRF with respect to a given neighborhood system if the Markovian property holds for each site s. The distribution of a positive MRF can be proved to have a Gibbs form [1], that is

$$p(x|\theta) = \frac{1}{Z}\exp[-U(x,\theta)], \tag{1}$$

with $U(x,\theta) = \sum_{c\in C} V_c(x_c,\theta)$, where x is the realization of the field X, θ is the set of parameters of the model, the V_c functions are called potentials, U denotes the energy, Z is a normalizing constant that depends on θ, and c indicates a clique of the image. Note that each potential V_c depends only on the values taken on the clique sites $x_c = \{x_s, s \in c\}$ and, therefore, accounts only for local interactions. As a consequence, local dependencies in X can be easily modeled by defining suitable potentials $V_c(\cdot)$.

Let us now consider a K-class image segmentation problem and model the unknown label map with a random field X defined on the lattice S of the image y to be segmented. Furthermore, consider a generic tree $\overline{T} = T \cup \Lambda$, where Λ is the set of the K leaves of the tree, associated with the K classes present in the image, while T is the set of the internal nodes. Through such a tree we can represent the hierarchical relationships among classes, where each internal node represent a "virtual" class obtained by merging its descendant classes. An example is shown in Fig.1, where the whole 4-class image is associated with the root (node 1) of a binary tree, then the light+dark blue virtual class is associated with the left child (node 2), and finally the elementary classes with leaves 4 and 5 respectively. Likewise, the orange+brown virtual class is associated with node 3 and the elementary classes with leaves 6 and 7. For this simple synthetic image, it is clear that a hierarchical representation is certainly appropriate, but also natural images often exhibit this kind of structures, although not so evident, which is why a tree-structured representation can be of interest.

With each internal node t of the tree, starting from the root, we can associate a K_t-class local field X^t (where $K_t \geq 2$ is the number of children of node t) which divides the corresponding region in K_t subregions. Hence, the whole field

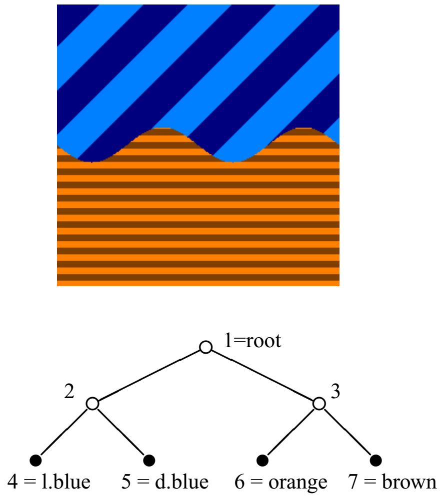

Figure 1: Example of synthetic image and associated tree.

x is recursively specified by the set of local fields $\{x^t\}_{t \in T}$, and is therefore a tree-structured field. Notice that the fields X^t are defined recursively, meaning that each of them can be defined only when all ancestor fields $X^{\omega(t)} = \{X^k\}_{k \in \omega(t)}$ are given[1] and, as a consequence, every probability law that concerns a local field X^t must be conditioned on the ancestor fields $X^{\omega(t)}$.

We say that the whole field $\{X^t\}_{t \in T}$ is a *tree-structured* MRF (TS-MRF) w.r.t. T if

$$p(x) = \frac{1}{Z}\exp[-U(x,\theta)] = \frac{1}{Z}\exp[-\sum_{t \in T} U_t(x^t)], \quad (2)$$

that is if each X^t is a MRF conditionally on $X^{\omega(t)}$, with Gibbs energy $U_t(x^t) = U_t(x^t; x^{\omega(t)}, \theta^t)$, and parameters θ^t [4]. As a consequence it is easy to show that T is an independency graph for the set of local fields $\{X^t\}_{t \in T}$, that is

$$p(x) = p(\{x^t\}_{t \in T}) = \prod_{t \in T} p(x^t | x^{\omega(t)}). \quad (3)$$

Let us now consider the problem of finding the MAP segmentation, $\hat{x} = \{\hat{x}^t\}_{t \in T}$, which maximizes the *a posteriori* distribution $p(y|x)p(x)$. Each component $\hat{x}^t$ can be written as

$$\hat{x}^t = \arg\max_{x^t} \left[\max_{\{x^k\}_{k \neq t}} p(y|x)p(x) \right] \quad (4)$$

which makes clear that the maximization at node t depends in general on all other nodes $k \neq t$. These in turn can be classified w.r.t. to t (see Fig.1) as ancestors, $\omega(t)$, descendants, $d(t)$, and non-related nodes or "others" $o(t)$. It is easily recognized that, given $\omega(t)$, node t is independent[2] of all nodes in $o(t)$. If t were also independent of $d(t)$, then 4 would reduce to

$$\hat{x}^t = \arg\max_{x^t} \left[\max_{\{x^k\}_{k \in \omega(t)}} p(y|x)p(x) \right] \quad (5)$$

[1] $\omega(t)$ indicates the set of ancestors of node t.

[2] by this we mean that all quantities pertaining to node t are independent of all quantities pertaining to node $k \in o(t)$.

and the maximization could be easily carried out one node at a time, from the root down to all the leaves of the tree. However, node t depends on its descendants both through the likelihood term (because the observed data for node t coincide with the observed data for its descendants), and through the field X^t (which defines the support for the descendant MRF's). The first dependence can be accounted for by suitably modifying the likelihood term as shown in [3]. As for the second dependence, in order to obtain a simple and fast maximization algorithm, we will simply neglect it. This choice, although arbitrary, is supported by experimental evidence which shows that the neglected terms do not improve significantly the segmentation quality, while much increasing the computational burden.

3. BUILDING THE TS-MRF

Looking at the definition of a TS-MRF, several topics remain open, like the MRF models to associate with the nodes of the tree or the definition of the best tree structure to use. In this work we address the last question, under the conditions that only *binary* trees are permitted and only Potts models can be used as local fields. We also assume to know in advance the number of classes and some related statistics (class-wise mean vector and covariance matrix).

The proposed procedure builds the tree from the bottom, by associating the elementary classes with the leaves of the tree, and then iteratively merging couples of nodes. In fact the TS-MRF approach succeeds in separating highly correlated classes, difficult to split, by isolating them from the others and associating them to sibling nodes, so that a dedicated MRF can take charge of the task. Our conjecture, therefore, is that classes which are similar in terms of spectral response, or have a high degree of spatial adjacency, should be kept as "close" as possible in the tree. In the bottom-up perspective, this means that such classes should be merged first. The key point of the procedure is therefore a measure that quantifies class similarity and hence the goodness of a merge.

To this end we resort to a probabilistic measure named *merging gain*. This measure was originally introduced in [2], in an unsupervised context, to decide whether two classes had to remain isolated or be merged. It is defined as

$$M^t = \frac{p(y^t|\theta^t)}{p(y^t|\hat{x}^t)p(\hat{x}^t)} \quad (6)$$

where t is the candidate father of the two nodes under test (say t' and t''), and $\hat{x}^t$ is the realization of the MRF at node t which segments the region associated with t returning those associated with t' and t''. The merging gain is large when the classes are spectrally close because the single-node likelihood $p(y^t|\theta^t)$ is very close to likelihood $p(y^t|\hat{x}^t)$ computed after segmentation for the two separated classes. Moreover it is large when the two regions under test have an active boundary because the realization of the father MRF $\hat{x}^t$ is not very likely. Of course, to compute the merging gain, a preclassification of the image is needed, and here we used a simple maximum likelihood classifier to this end. For further details on the merging gain, the reader is referred to [2].

The tree structure is then singled out by a sequence of binary class merging starting from the pair of classes associated with the higher merging gain. By doing so, merging classes will progressively replace the original ones until a

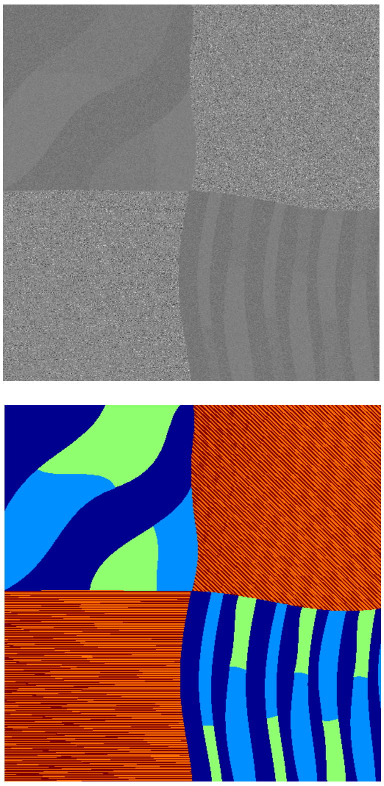

Figure 2: Band 1 of the synthetic image and ground truth.

single class that covers the whole image is reached and associated with the tree root.

4. EXPERIMENTAL RESULTS

We begin the experimental analysis by considering a simple synthetic image with 5 classes. Fig.2 shows one band of the synthetic image, clearly characterized by a very low signal-to-noise ratio, and its ground truth.

For this simple image, the proposed algorithm finds easily the correct tree-structure, shown in Fig.3 together with the corresponding TS-MRF segmentation map. In Fig.4, for comparison, we show the segmentation map obtained with a "wrong" tree structure, also shown in Fig.4. The global performance parameters show only minor, although consistent, changes: the overall accuracy τ, that is the percentage of sample pixels that are correctly classified, is 87.9% with the correct structure and 85.8% with the wrong one. Likewise, the normalized accuracy τ^{norm}, which weighs the accuracy against the relative abundance of the various classes, goes from 91.5% to 88.4%, and also the Kappa parameter reduces from 84.5% to 81.8%. This 2-3% difference is certainly significant, but the visual analysis of the segmentation maps is even more interesting as it shows that the correct tree structure helps preserving the fine structure of the image, especially in the most critical areas characterized by high spatial activity. In the brown-orange areas, for example, where most errors are concentrated because of the rapid alternation of classes with similar statistics, the map of Fig.3 exhibits a statistical behavior that, despite the errors, closely resembles that of the original map. The same does not happen with the wrong structure, where the orange areas become predominant, the quasi-periodic behavior is almost lost, and some pixels are even attributed to other (blue) classes. Similar results are obtained when no structure at all is used, namely, an ordinary 5-class MRF model is used to carry out the segmentation.

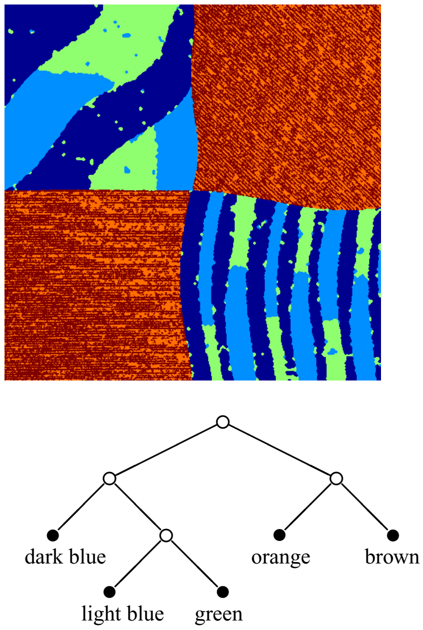

Figure 3: Segmentation with the correct tree structure.

We now turn to experimenting with a real-world multi-temporal image. The data set used for the experiments refers to an agricultural area in the Basin of the Po River (northern Italy), and is composed by two registered images acquired by the Landsat TM sensor (bands 1-6) in April and May 1994 (a 512 $\times$ 512 square section, April, band 3, is shown in Fig. 5). For this image a ground-truth was available to provide land cover composition and class-parameters estimation. The land classes present in the April data set are wet rice fields (WR), bare soil (BS), cereals and wood, while in May we find corn, wet rice fields, dry rice fields (DR), cereals and wood.

From ground-truth inspection, it results that changes involve only some specific classes: wet rice fields which partially became dry rice fields, and bare soil that disappears thoroughly to become rice fields (wet or dry) or corn (this last class being absent in April), while, on the other hand, other two classes (cereals and wood) have no change. Based on these observations, change-classes have been defined and added to no-change classes, as listed in Fig. 6.

The "binary" TS-MRF model used in experiments was built with simple Potts [1] local fields at each node of the tree, while single splits were performed by combining an ICM procedure with the MPL method for the estimation of prior parameters [5].

The tree structure of the model singled out by the proposed method is drawn in Fig. 6. Subsequently, a joint segmentation of the two images has been performed on the whole data set by considering such a global class hierarchy and the resulting classification map is shown Fig. 7.

The results are very close (both visually and in terms of classification accuracy) to those provided by a reference TS-MRF segmenter in which the tree structure was however built

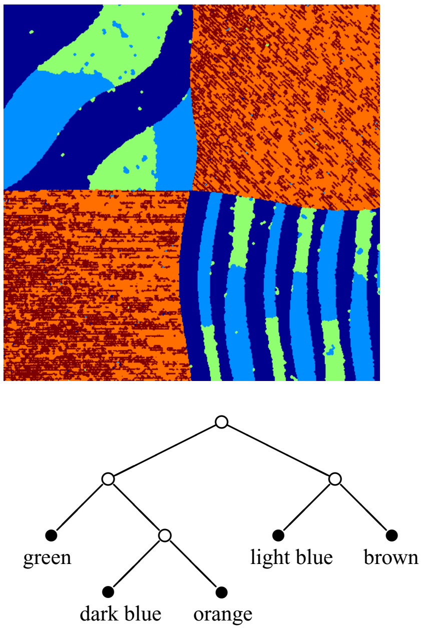

Figure 4: Segmentation with the wrong tree structure.

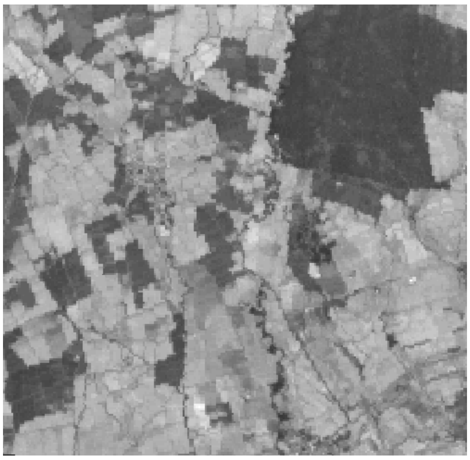

Figure 5: Band 3 (April) of the Landsat image.

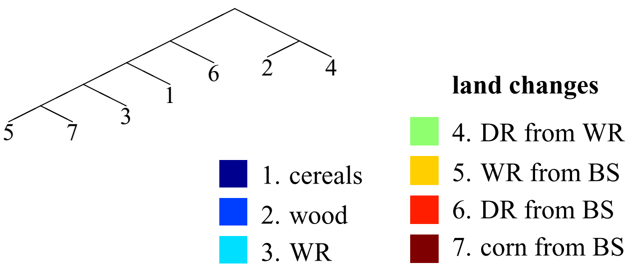

Figure 6: Tree structure and classes (WR, wet rice fields; DR, dry rice fields; BS, bare soil).

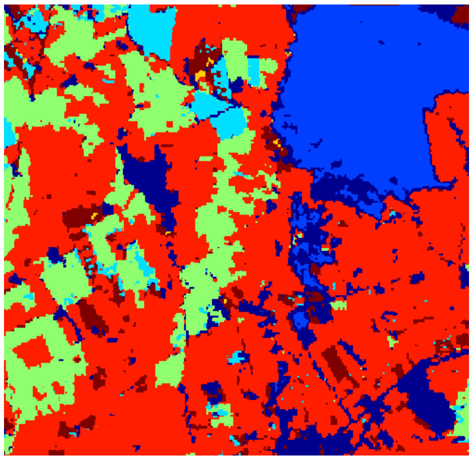

Figure 7: Segmentation map provided by TS-MRF algorithm.

ad hoc by a human operator based both on statistical data and on the semantics of the classes (see [6]). The tree building procedure is therefore fully satisfactory since it seems able to provide the same results that were obtained through a painstaking construction of the tree by the experimenter. It is worth underlining that both TS-MRF classifiers compare favourably with other MRF-based non-structured classifier (see again [6]) even though the image used is not so "structured", as in other experiments like those presented in [3].

5. CONCLUSIONS

In this work we have considered a class of MRF models, the TS-MRF, and shown some properties which allow for the fast recursive segmentation and classification of images. We have then proposed a simple procedure to build the tree structure of the model so as to follow the underlying statistical structure of the image. The procedure is based on a probabilistic measure, the merge gain, that accounts for the similarity of classes by weighting their spatial and spectral correlations. The proposed technique has been finally applied, with good results, to the segmentation of a multitemporal remote-sensing image.

REFERENCES

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

[2] C.D'Elia, G.Poggi, G.Scarpa, "A Tree-Structured Markov random field model for Bayesian image segmentation," *IEEE Trans. on Image Proc.*, vol.12, no.10, pp.1259-73, Oct. 2003.

[3] G.Poggi, G.Scarpa, J.Zerubia, "Supervised segmentation of remote-sensing images based on a tree-structured MRF model," to be published in *IEEE Trans. on Geosc. and Rem. Sen.*, 2005.

[4] G.Scarpa, "Tree-structured Markov random field models for segmentation of noisy images with application to remote sensing," PhD Thesis, University Federico II of Naples (I), 2004.

[5] J.Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Stat. Soc.*, Series B, vol.48, pp.259-302, 1986.

[6] L.Cicala, G.Poggi, G.Scarpa, "Supervised segmentation of remote-sensing multitemporal images based on the tree-structured Markov random field model," *Proc. IEEE IGARSS'04*, Anchorage (Alaska), September 20-24, 2004.