

# Unsupervised Hierarchical Image Segmentation based on the TS-MRF model and Fast Mean-Shift Clustering

R. Gaetano, G. Scarpa, G. Poggi, J. Zerubia

► **To cite this version:**

R. Gaetano, G. Scarpa, G. Poggi, J. Zerubia. Unsupervised Hierarchical Image Segmentation based on the TS-MRF model and Fast Mean-Shift Clustering. Proc. European Signal Processing Conference, EUSIPCO 2008, Aug 2008, Lausanne (CH), Switzerland. 2008. <inria-00503198>

**HAL Id: inria-00503198**

**<https://hal.inria.fr/inria-00503198>**

Submitted on 16 Jul 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# UNSUPERVISED HIERARCHICAL IMAGE SEGMENTATION BASED ON THE TS-MRF MODEL AND FAST MEAN-SHIFT CLUSTERING

Raffaele Gaetano<sup>†</sup>, Giuseppe Scarpa<sup>†</sup>, Giovanni Poggi<sup>†</sup>, and Josiane Zerubia<sup>‡</sup>

<sup>†</sup>Dip. Ing. Elettronica e Telecomunicazioni,  
Università “Federico II” of Naples, Italy  
phone/fax: + (39) 0817683151/49,  
email: *firstname.lastname@unina.it*

<sup>‡</sup>Ariana Research Group,  
INRIA - I3S, Sophia Antipolis, France  
phone/fax: + (33) 492387865/7643,  
email: *josiane.zerubia@sophia.inria.fr*

## ABSTRACT

Tree-Structured Markov Random Field (TS-MRF) models have been recently proposed to provide a hierarchical multiscale description of images. Based on such a model, the unsupervised image segmentation is carried out by means of a sequence of nested class splits, where each class is modeled as a local binary MRF.

We propose here a new TS-MRF unsupervised segmentation technique which improves upon the original algorithm w.r.t. several critical issues, from the selection of a suitable tree structure to the elimination of spurious classes. The major improvements come from resorting to the Mean-Shift procedure, used to estimate the number of pdf modes at each node, thus allowing for a generic (non-binary) tree, and to obtain a more reliable initial clustering for the subsequent MRF optimization. To this end, we devised a new reliable and fast clustering algorithm based on the Mean-Shift technique. Experimental results on a SPOT satellite image prove the potential of the proposed method.

## 1. INTRODUCTION

Along with the advances of research in the image analysis and processing fields, the problem of segmentation is assuming an ever growing importance in many applications, such as medical image analysis, remote-sensing image classification, content based image retrieval, etc. Given the large-spectrum goal of image segmentation, that is, providing a partition of image pixels into some regions according to certain homogeneity criteria, it is easily understood that such a problem can be addressed with a wide variety of approaches. This typically leads to application-specific solutions that can also make sense at different levels of abstraction. In this widely varying scenario, MRF-based image modeling, first introduced in the 80's [1, 2], still remains a very popular approach, mainly because of its effectiveness and flexibility in defining local dependencies among adjacent pixels, thus encompassing prior knowledge in the segmentation process with a reasonable complexity.

In order to improve the description capabilities of conventional MRF models and reduce the overall complexity of the derived segmentation algorithms, a new hierarchical MRF has been recently proposed, the Tree-Structured MRF (TS-MRF) model [3, 4], that proved to be quite effective and reliable, especially for the analysis and classification of various types of remote sensing images.

The main rationale for such a model is the observation that, especially in certain domains, images are often characterized by a distinctive hierarchical structure, with regions

that interact with one another in different ways and at different scales of observation.

The TS-MRF allows to model such a behavior by defining a suitable tree structure for the image of interest, and associating to each inner node of the tree a different image region and a different MRF, which is completely local to the corresponding region and has its dedicated parameters. This approach guarantees a much higher local adaptivity than classical MRFs. In addition, based on such a model, the segmentation problem can be formulated recursively, reducing a general  $K$ -ary segmentation procedure to a sequence of steps with just a few classes each, with a significant reduction of complexity.

Segmentation based on the TS-MRF model has proven to be very successful in the supervised case [4], when the number of classes of interest and their synthetic parameters are known *a priori*. In the unsupervised case [3] results are also good, especially if compared with those of unstructured techniques, but some critical issues remain to be addressed. In fact, lacking any prior information, one is forced to estimate, by recursive optimization at each node, the very same tree structure underlying the data. If the optimization is inaccurate at some node, the whole tree structure might deviate from the most suitable one, with various types of undesirable effects, such as the fusion of different classes or the oversplitting of others.

In this work we propose an improved version of the TS-MRF unsupervised segmentation algorithm that effectively addresses the major problems briefly outlined above. The main improvements come from the use of a Mean-Shift based clustering. As a matter of fact, the Mean-Shift procedure [5] was already used in [6] to detect the number of modes, and hence the number of children for each node of the tree. In this work, however, its use is carried further, and besides finding the dominant modes for each class, it replaces the *Generalized Lloyd Algorithm* (GLA) [7] as the initial clustering technique, providing a much more reliable starting point for the subsequent MRF-based segmentation, and a much easier and stable detection of the correct tree-structure for the data. This is obtained through some significant modification of the Mean-Shift clustering itself, which now makes use of a variable-bandwidth strategy based on the *k-Nearest Neighbour* ( $k$ -NN) technique, and is implemented with a speed-up strategy that cuts significantly the computational complexity, otherwise intolerable for such applications.

In section 2, we first recall the basics of Mean-Shift analysis, and then describe the new Mean-Shift clustering algorithm, focusing in turn on the variable-bandwidth strategy, and then on the speed-up solutions introduced. In Section

3, after describing in more details the TS-MRF model, and the related segmentation algorithm, we show how the new clustering tool can be used to improve the performance for unsupervised segmentation tasks. Section 4 provides experimental evidence of the improved performances for remote sensing images. Finally, Section 5 draws conclusions.

## 2. FAST VARIABLE-BANDWIDTH MEAN-SHIFT CLUSTERING

### 2.1 Background

The Mean-Shift procedure for mode detection [5] is a robust and effective tool to compute local maxima of a probability distribution over a given feature space, based on the well known *Parzen Window* framework [8] for non-parametric density estimation.

The rationale behind this algorithm is that samples in a certain feature space can be easily associated with an empirical probability density function: briefly, if we consider a  $d$ -dimensional feature spaces and a set of  $n$  data points  $\{s_i\}_{i=1}^n$ , the following expression can be a reasonable estimation for the pdf:

$$\hat{p}(s) = \frac{c_{K_p}}{nh^d} \sum_{i=1}^n K_p \left( \left\| \frac{s-s_i}{h} \right\|^2 \right), \quad (1)$$

where  $K_p(\cdot)$  is a univariate strictly positive kernel profile function, such that a radially symmetric kernel can be generated from it through a rotation in  $\mathcal{R}^d$ ,  $c_{K_p}$  is a normalizing constant and  $h$  is the kernel size, often indicated in the literature as the “bandwidth” parameter, that controls the resolution at which modes are detected. It is demonstrated that the gradient of this expression can be written in the following form:

$$\nabla \hat{p}(s) = q(s) \mathbf{m}_{h,g}(s), \quad (2)$$

where  $q(\cdot)$  is a scalar function and

$$\mathbf{m}_{h,g}(s) = \frac{\sum_{i=1}^n s_i g \left( \left\| \frac{s-s_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left( \left\| \frac{s-s_i}{h} \right\|^2 \right)} - s \quad (3)$$

is the so called *mean shift* vector, where  $g(\cdot) = -K'_p(\cdot)$ . Therefore, based on the fact that mean shifts always point towards the maximum increase in the density, a gradient ascent procedure can be run, starting from any data point of the sample set, that will eventually converge to a stationary point in the distribution, that is, a mode of the pdf. Once a starting kernel center  $s$  is assigned, the procedure consists of two iterative steps:

1. compute the mean shift vector  $\mathbf{m}_{h,g}(s)$ ,
2. update the kernel center  $s = s + \mathbf{m}_{h,g}(s)$ ,

Clearly, to detect all significant modes, this procedure must be executed many times, each time with a different initialization, in order to cover most of the feature space.

### 2.2 Clustering by the mean-shift

The detection of modes through the Mean-Shift procedure determines an implicit clustering strategy over the feature space, since all the starting points of trajectories that converge towards the same mode (that is, belong to its “basin of attraction”) form a well defined cluster. However, this would require running the Mean-Shift procedure for each point of

the feature space, so as to identify the basin of attraction of all modes as clusters. Of course, this is unfeasible in practice, since for sample sets larger than several hundreds of data points computational time becomes extremely large for most of the possible applications. Hence, an efficient implementation is usually required, especially for data-intensive cases.

Another critical implementation issue is the choice of the kernel size, or *bandwidth* parameter, which plays a central role for density estimation since it determines the smoothness of the pdf and, consequently, the number of modes that the algorithm singles out<sup>1</sup>. Using too large a bandwidth leads to underestimating the number of modes, and the opposite for too small a value.

We propose here an implementation of Mean-Shift clustering which addresses the two problems briefly outlined above. In particular, the new algorithm is based on

- a data-dependent adaptive kernel size  $h$  that overcomes the instability observed for example in [6];
- a fast clustering technique, with a significant speed-up factor w.r.t. the basic procedure, that enable its use for real-world applications.

#### 2.2.1 $K$ -NN based adaptive bandwidth selection

The original Mean-Shift procedure proposed by Comaniciu [5] considers a fixed bandwidth parameter  $h$ , but this choice is clearly inappropriate whenever the density of points in the feature space varies wildly. In such cases, in fact, it is impossible to choose a value well suited for both high- and low-density areas.

To face this problem, we choose here to adapt the bandwidth parameter locally in the feature space by taking into account only the first  $k$ -Nearest Neighbors in the computation of the Mean Shift vector. This amounts to truncating the kernel at some distance from the center, but if  $k$  is not too small, this truncation will take place when the kernel has already a negligible value, independent of the local density. The bandwidth, instead, will clearly depend on the local density, being larger in low-density areas and smaller in high density ones.

More in detail, once a suitable value of  $k$  is selected, at each step of the procedure the set  $NN(s)$  of  $k$  points closest to  $s$  is singled out, and the kernel size is calculated as

$$h(s) = \sqrt{\frac{1}{k} \sum_{i \in NN(s)} \|s - s_i\|^2}, \quad (4)$$

This value is then used in (3) for the computation of the mean-shift vector where the summation is again restricted to the points in  $NN(s)$ .

It could be observed that this solution moves the problem from the estimation of parameter  $h$ , to that of parameter  $k$ , but it is well-known [9] that  $k$ -NN estimation is quite robust w.r.t. its parameter, and works quite well also in spaces of high dimensionality, which are instead quite challenging for the Mean-Shift. In next section, we propose a data dependent procedure for obtaining a stable estimate of the  $k$  parameter.

#### 2.2.2 Fast clustering strategy

Our speed-up strategy is based on the obvious consideration that all points that lie on the trajectory that goes from the

<sup>1</sup>Less critical is the kernel function, for which we select the quite usual Gaussian shape.

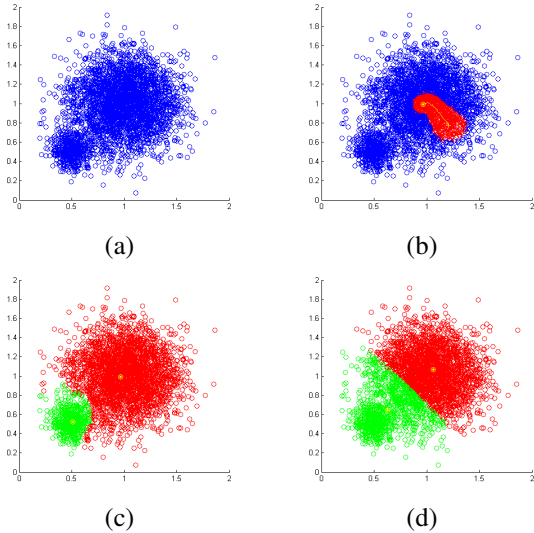


Figure 1: (a) bi-modal sample set, (b) Mean Shift trajectory with the corresponding “voting” points, (c) final clustering, (d) GLA-based clustering for comparison.

starting point to the corresponding mode belong necessarily to the same basin of attraction. Therefore, they could all be attributed, without error, to the same cluster.

Although it is extremely difficult that any sample point will coincide *exactly* with a point of this path, one can reasonably assume that sample points that are *close* to the trajectory belong very likely to the same basin. By clustering all such points in one shot we can reduce drastically the complexity of clustering, but, due to the approximation, we also risk to cause some errors, especially for data points that are close to the watershed between two basins of attraction. Hence, in order to preserve the accuracy of clustering, we do not assign sample points on the fly, but rather implement a voting mechanism and decide only a posteriori, with a majority rule, when all sample points have been touched by at least one trajectory.

A more precise description of the modified procedure is given below:

1. *Initialization*: set all sample points as *non visited*.
2. *Mean-Shift*: run the procedure starting from a randomly selected non visited point: at each step along the trajectory, mark as *visited* all points  $s_i$  such that  $\|s - s_i\| < h(s)$ , and for each of them add a vote to the “final” mode.
3. *Mode validation*: once convergence is reached, compute the distance  $d_{\min}$  between the new tentative mode and the closest mode already detected:
  - if  $d_{\min} < h/2$  reject the new mode, and mark the closest mode as final;
  - otherwise accept the new mode, and mark it as final.
4. *Test*: if there are still non visited points, go to step 2.
5. *Clustering*: assign each visited point to the mode (and cluster) with the most votes.

An example of clustering provided by the described procedure is presented in Fig.1: the bivariate sample set of part (a), obtained as a mixture of two normally distributed data sets, is given as input to the clustering algorithm. In part (b) the effect of a single modified Mean-Shift procedure is

represented, where all the points in red are “giving a vote” to the final mode. Part (c) shows the final clustering, which appears to follow quite faithfully the underlying distribution and is certainly much better than the GLA-based clustering shown in part (d) where, in addition, the correct number of clusters had to be provided as a further input.

### 3. TS-MRF UNSUPERVISED SEGMENTATION BASED ON MEAN-SHIFT CLUSTERING

As already mentioned in Section 1 and discussed in details in [3, 4], TS-MRF modeling is based on the hypothesis that the data possess an inherent hierarchical structure in terms of spatial and spectral properties. Given *a priori* the number of classes  $K$  to be retrieved with their parameters, and a suitable tree  $T$  that describes the hierarchical structure of data, a “simple” MRF is associated with each inner node  $t$  and the segmentation can be carried out by top-down induction over the tree with a recursive optimization<sup>2</sup> of the different MRFs.

In the unsupervised case, however, no prior information is available on the image, and all parameters, including the tree structure, must be retrieved during the process. the segmentation is decomposed into a sequence of nested splits, starting from the whole image, and going on until all elementary regions are identified according to a given stopping criterion. The entire image is therefore associated with the root of a tree, and each split creates some new nodes, generating gradually the desired tree structure whose growth is governed by a suitable metric called *split-gain* [3]. Terminal nodes of the structure correspond to the final classes of the map. This approach provides therefore a hierarchical multi-scale segmentation of the image, with a *set* of finer and finer segmentation maps among which the user is free to select the most appropriate for the specific needs, and a synthetic high-level description given by the tree itself together with the class parameters.

Of course, the general segmentation strategy must be translated into a real-world functioning algorithm, where a number of implementation choices, sometimes driven by complexity concerns, might have a critical impact on the overall performance. One such choice, made in [3] to simplify the local optimization task, is to consider only binary tree structures, reducing the segmentation process to a sequence of nested binary splits controlled by a suitable stopping criterion. Such a constraint, however, might cause the detection of false contours as can happen when three or more balanced classes are present in the same region. In [6] we removed this constraint and resorted to the Mean-Shift procedure to detect the number of pdf modes in a class, and hence the number children at a given node. Another critical choice is the use of the GLA to carry out the initial clustering needed to perform the MRF optimization at each node. In fact, image pixels are often described by a complex and generally unbalanced probability distribution in the spectral domain, in which case the GLA can easily provide inaccurate results, as in the example of Fig.1(d).

Therefore, we now replace the GLA based clustering with the more accurate variable-bandwidth Mean-Shift based clustering described in the preceding section. Even though our fast implementation helps limiting the processing weight, plain Mean-Shift clustering would have an exceedingly high

<sup>2</sup>In this work we refer to the Potts MRF framework, and optimization is obtained using a *Maximum a Posteriori* criterion, see [10, 11].

computational complexity for the very large images we usually deal with, and hence we will eventually resort to a hybrid *Mean-Shift/Maximum Likelihood* (MS-ML) classifier. In more details, for each region to split, we extract a reasonably large random subset of pixels (from 1% of the region area to the entire region, depending on its size); the Mean-Shift clustering described in Section 2 is then applied to this sample set, to retrieve the number of classes and their statistics; this information is eventually used by a Maximum Likelihood classifier that runs over the whole region.

In Section 2, we did not address the problem of selecting a suitable value of  $k$  for the  $k$ -NN based bandwidth estimation of (4). A typical choice is to set  $k$  to a fraction, *e.g.*, 10%, of the sample set cardinality, which, given the robustness of  $k$ -NN, provides usually good results. For some nodes, however, this simple choice turned out to be unsatisfactory, causing a proliferation of modes in the Mean-Shift clustering and a certain instability in the segmentation. This is not surprising, after all, given that the same algorithms are used at all nodes, from the root, corresponding to the whole image, to terminal leaves corresponding sometimes to much smaller and much more fragmented regions.

To avoid this misleading behavior, we defined a simple procedure that increments the value of  $k$  (increasing stability but also computational burden) until the Mean-Shift provides reliable results. In detail, the initial value of  $k$  is set to  $\text{round}(\alpha_1 |S|)$ , with  $S$  the selected sample set. Then the Mean-Shift procedure, *without indirect clustering*, is run for a maximum of  $N_1$  times, keeping track of the number of detected modes: if the number of modes converges (remains the same for  $N_2$  times) the current value of  $k$  is accepted, otherwise it is increased,  $k = k(1 + \Delta)$ , and the procedure is repeated. In any case,  $k$  is not allowed to increase beyond  $\alpha_2 * |S|$ , when it is frozen anyway. Notice that this procedure provides a further criterion to decide whether to split a node or not, since the detection of a single mode qualifies the corresponding region as elementary.

Using a more reliable technique to carry out the initial clustering does certainly improve the subsequent MRF optimization, but there is a more subtle and important consequence in the context of hierarchical segmentation. In fact, the MS-ML clustering provides quite a reliable segmentation in the spectral domain, while the MRF model allows to take into account contextual information to regularize the final map. The points that change label during MRF optimization turn out to be “outliers” in the spectral domain for the final class  $\omega$ , that is, their statistics will be far apart from those of points originally attributed to  $\omega$  by the MS-ML technique. If class  $\omega$  is segmented again, such outliers can give origin to one or more separate clusters, leading to dramatic over-segmentation errors. We are now in the position to solve this dangerous phenomenon, by simply erasing such points from the new sample set. Notice that this was not possible with a GLA initialization, since the initial clusters were so far from the final segmentation (compare again Fig.1) that such erasure would amount to eliminate large valid chunks of data.

#### 4. EXPERIMENTAL RESULTS

We assess the performance of the improved unsupervised segmentation algorithm through a set of experiments on a SPOT satellite image of the Lannion Bay, in France, August 1997 (©SPOTImage/CNES), composed of three  $1480 \times$

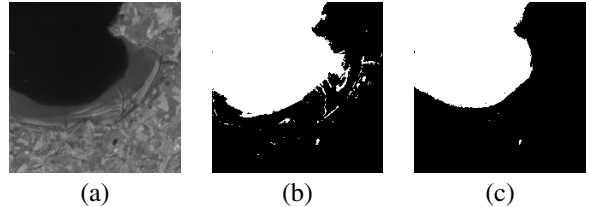


Figure 2: Detail of the XS3 channel (©SPOTImage/CNES) (a), initial *sea class* split using GLA (b), and MS-ML (c).

1024 bands and a spatial resolution of 20m.

For both the original TS-MRF algorithm and the new version proposed here we use the same settings for the MRF optimization part, and stop the tree growth at 8 classes. The mode detection procedure uses  $\alpha_1 = 0.08$ ,  $\alpha_2 = 0.12$ ,  $\Delta = 0.05$ ,  $N_1 = 20$  and  $N_2 = 10$ .

The improvements due to the use of the MS-ML are quite clear since the first stages of segmentation. In Fig.2(a) we show a detail of the source image, along with two maps that, for both the original (b) and new version (c) of the algorithm, show the “sea” class (in white) as identified by the top-level clustering, before any MRF regularization. The errors introduced by the GLA are quite evident in Fig.2(b), as well as the very high accuracy of the MS-ML classification of Fig.2(c). Such a good initialization will likely improve, and certainly simplify the subsequent optimization process (making up for the increased complexity of the MS-ML clustering). Moreover, it will allow to single out easily the few label-switching points to eliminate in further spectral clustering steps.

Fig.3(a) shows the complete image (again XS3 channel), provided with the available ground truth (©COSTEL), not reported for sake of brevity, used to compute quality figures. The segmentation maps obtained with the original and improved TS-MRF algorithms are reported in Fig.3(b) and (d), respectively. Fig.3(c) and (e) instead, show the tree structures detected by both algorithms, where the leaves are associated *a posteriori* to the eight semantic classes so as to maximize the overall accuracy as computed on the ground truth.

At a visual analysis, results provided by the proposed version are much more accurate than those of the original algorithm: no major losses are noticeable, at least on top level classes, unlike in the map of Fig.3(c) where a serious oversplitting of the “forests” class sticks out. Numerical results confirm such empirical observations: the overall classification rate goes from around 60% to 74.4% mainly due to the more precise detection of some large classes, such as the “forests” and “urban areas” classes, as appears from the user’s and producer’s accuracies<sup>3</sup> reported in Tab.1. Such an improvement can be likely ascribed to the improved segmentation accuracy obtained in the first steps, also due to the more flexible tree structure. As can be seen in Fig.3(e), in fact, the new technique, by resorting directly to a 3-class top-level split, immediately detects and validates the “forests” class, preventing it from being oversplit in later stages.

Finally, we present an interesting result concerning the TS-MRF based *supervised* segmentation technique described in [4] and referred to therein as TS/U. The supervised procedure has been run here replacing the original binary tree-structure selected in [4] by visual inspection with the tree

<sup>3</sup>Complete confusion matrices could not be shown for lack of space.

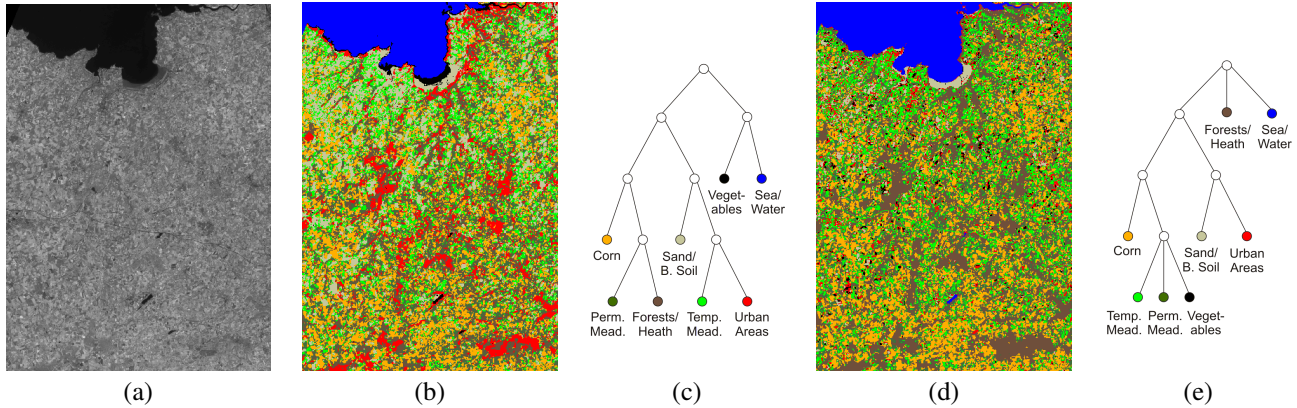


Figure 3: XS3 channel of an image of Lannion Bay, France (©SPOTImage/CNES) (a), unsupervised segmentation by the original TS-MRF algorithm (b) and the corresponding tree (c), the new version (d) and the corresponding tree (e).

Classes	TS-MRF w/GLA		TS-MRF w/MS-ML	
	U.A.	P.A.	U.A.	P.A.
Water	100%	75.6%	100%	94.2%
Bare Soil	75.5%	95%	97.5%	69.1%
Urban	4%	6.3%	49.2%	40.8%
Forests	97.1%	46.2%	97.1%	95.6%
Temp. Mead.	38.5%	30.2%	25.3%	33.2%
Perm. Mead.	23.1%	22%	33.7%	28.8%
Vegetables	0%	0%	3.4%	4%
Corn	63.9%	95.5%	65.2%	94.5%
Overall Acc.	<b>59.8%</b>		<b>74.4%</b>	

Table 1: Per-class and overall accuracies for the classification of Fig.3(b) and Fig.3(c) respectively.

structure of Fig.3(e) detected by the unsupervised technique proposed in this paper. A quite accurate segmentation map has been obtained, and, most notably, the overall accuracy is now 85.3% as opposed to the 82.3% obtained with the hand-picked tree-structure. This seems to show that the tree-structure detected here does fit well the source data and could well be used as a preliminary tool in supervised TS-MRF segmentation, eliminating the need for a heavy user intervention.

## 5. CONCLUSION AND FUTURE WORK

Unsupervised image segmentation based on the TS-MRF model relies heavily on the detection of a tree structure that correctly describe the data structure and on the accurate optimization of MRFs at each node.

The segmentation algorithm proposed in [3] proves unsatisfactory under both respects because of some important simplifications that we remove here. In particular, we allow for the use of generic (rather than binary) trees, and improve the MRF initialization at each node, resorting in both cases to the Mean-Shift procedure. In the first case, Mean-Shift allows us to estimate the number of pdf modes at each node, and hence the number of children nodes, while in the latter it is used, together with a Maximum-Likelihood classifier, to replace the much less reliable GLA clustering.

To this end, a fast new Mean-Shift clustering algorithm is proposed, whose main features are the adaptive kernel size selection via a *k*-Nearest Neighbors approach and a speed-up strategy which reduces the computational burden with little harm for the clustering accuracy.

Land classification experiments prove the effectiveness of the proposed solution, both for unsupervised segmentation and as a tool for the automatic definition of a suitable tree structure in the context of supervised segmentation [4].

Further studies on the proposed Mean-Shift clustering algorithm and its application to TS-MRF are currently ongoing, along with the enrichment of the experimental evidence, to definitively assess performances of the proposed method.

## REFERENCES

- [1] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 1, pp. 721–741, Nov. 1984.
- [2] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society, series B*, vol. 48, pp. 259–302, 1986.
- [3] C. D'Elia *et al.*, "A tree-structured Markov random field model for Bayesian image segmentation," *IEEE Trans. on Im. Proc.*, vol. 12, no. 10, pp. 1259–1273, Oct. 2003.
- [4] G. Poggi *et al.*, "Supervised segmentation of remote sensing images based on a tree-structured MRF model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 8, pp. 1901–1911, Aug. 2005.
- [5] D. Comaniciu and P. Meer, "Mean Shift: a robust approach toward feature space analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, May 2002.
- [6] R. Gaetano *et al.*, "Identification of image structure by the Mean Shift procedure for hierarchical MRF-based image segmentation," in *Proc. EUSIPCO 2006*, Florence, Italy, Sept. 2006.
- [7] A. Gersho and R. Gray, *Vector quantization and signal compression*. Boston, MA: Kluwer, 1992.
- [8] R. O. Duda *et al.*, *Pattern Classification*. Wiley, 2000.
- [9] D. W. Scott, *Multivariate Density Estimation*. Wiley, 1992.
- [10] G. Winkler, *Image analysis, random fields and dynamic Monte Carlo methods*, 1st ed. Springer-Verlag, 1995.
- [11] S. Z. Li, *Markov random field modeling in image analysis*, 1st ed. Springer-Verlag, 1995.