A NOVEL MARKOV RANDOM FIELD MODEL FOR MULTISOURCE DATA FUSION IN URBAN AREAS

Giovanna Trianni[(1)], **Paolo Gamba**[(2)]

[(1)] ***Department of Electronics, University of Pavia, via Ferrata 1, 27100 Pavia (Italy), giovanna.trianni@unipv.it***

[(2)] ***As (1) above, but Email: paolo.gamba@unipv.it***

ABSTRACT

In this work we propose a new statistical method based on a Markov Random Field for a multisource classification of remotely sensed data, with particular attention to urban classification. During the classification process the MRF model allows to handle each pixel taking into account its neighbourhood, exploiting in this way the spatial context of each pixel in the image. Markov Random Fields, which include prior information in a statistically rigorous manner, provide a methodological framework which allows the image from different sensors to be merged in a consistent way. Taking into account some simple local interactions at the scale of a single pixel and its neighbours, MRF models show a complex global behaviour, which is the principal reason of their popularity among the scientific community. The performance of our method is evaluated by considering Landsat TM, ERS-1 and ASAR images. We found that in many cases the overall accuracy of the classification performed with the MRF model is higher (from 1% to 5%) than the one obtained by a neural classifier with a Fuzzy Artmap structure [1]; in particular, the accuracy of the urban class increases even of the 10%; moreover the joint classification of Landsat and SAR data resulted in a more precise discrimination among the different land cover classes present in the images.

INTRODUCTION

Nowadays, the increasing number of image sources providing different kinds of information about the earth for oceanic, meteorological, terrestrial applications has been causing the need of new methods to perform the fusion of multisensorial data. Since land cover, especially in some areas, is various and complicated, data coming from a single sensor are often not enough to obtain information of greater quality, or rather to generate a satisfactory and reliable interpretation of the scene. We propose here a new statistical method based on Markov Random Fields model and tested it by performing a joint classification of SAR and Landsat data. Radar data has very different characteristics from visible/infrared remotely sense data such as Landast TM, SPOT and ASTER. For example, radar can penetrate vegetation but it provides only single band image in most situation and is difficult to process and analyze, while LANDSAT is an advanced multispectral imager which covers a spectral region with 7 bands from the visible to the thermal infrared; with spatial resolution of 30 m, except the thermal infrared one, which has spatial resolution of 60. Moreover, Landsat has a panchromatic band with resolution of 15 m.
In the classification of an image, the spatial context play an important role in the interpretation of the scene. In fact, isolated pixels provide less information than the same pixels considered together with contextual information which can be of spectral, spatial or temporal kind. Spectral information refers to the use of different bands coming from a single multispectral sensor or from different sensors and often provide a good improvement in the classification of different land cover classes. Spatial context is obtained by considering adjacent pixels in a spatial neighbourhood, while the temporal context is obtained by considering acquisitions of the same area at different dates .

MRF CLASSIFICATION MODELING

Here we present the general classification model based on the Markov Random Field approach. Let consider a set of features or images coming from n sensors; then, let consider the $M \times N$ image acquired by sensor r as made up of MN pixels or feature vectors $X_r(1, 1), \ldots, X_r(M, N)$, $r = 1, 2, \ldots n$, where $X_r(i, j) = (x_r(i, j, 1), \ldots, x_r(i, j, B_r))$, and B_r is the number of spectral bands or features for sensor r. We assume to know that K classes $c_1, c_2, .., c_K$ are present in the images, and that they have prior probabilities $P(c_1), P(c_2), \ldots, P(c_K)$. Let denote with $C(i, j)$ the class for pixel (i, j); we indicate with X_r the set of the pixels of the whole image $X_r = \{X_r(i, j); 1 \le i \le M, 1 \le j \le N\}$ and with $C = \{C(i, j), 1 \le i \le M, 1 \le j \le N\}$ the set of labels for the same scene; in practice for a given pixel (i, j), $C(i,j) \in \{c_1, c_2, \ldots, c_K\}$.
If we indicate with $P(X_1, \ldots, Xn \mid C)$ the conditional probability density of feature vectors $X_1, X_2, \ldots, X_n$ given the scene labels set C, and with $P(C \mid X_1, \ldots, X_n)$ the posterior probabilities, the classification task consists in assigning each pixel to that class that maximizes the posterior probabilities. Of course, there is a relation between the data (measurements or features) and the prior information, which can be represented in a Bayesian formulation as:

$$P(C \mid X_1,....., X_n) = \frac{P(X_1,...., X_n \mid C)P(C)}{P(X_1,....X_n)} \tag{1}$$

where *P(C)* represents the prior model for the class labels.
Thus, we want to maximize the likelihood function

$$L(X_1,...., X_n \mid C) = P(X_1 \mid C)^{\alpha_1}....P(X_n \mid C)^{\alpha_n} P(C) \tag{2}$$

Where α_r , $0 \le \alpha_r \le 1$ is the reliability factor for sensor *r*. To obtain this expression for the likelihood function we assume conditional independence among the measurements coming from different sensors. As this assumption is difficult to verify in practice, we can say that it seems to be realistic for SAR and Landsat sensors, which use different wavelengths. In the single source classification we choose as α_r the overall accuracy of the classification of the images coming from a single source obtained with a neural Fuzzy Artmap classifier, while in the multisource classification we give α_r the overall accuracy of the joint classification of the two sensors, obtained with the same neural classifier.
We use a Markov Random Field [3],[4] to take into account the spatial context of each pixel of the image. If we indicate with G_{ij} the local neighbourhood of pixel *(i, j)* we can write:

$$\begin{aligned} &P(C(i,j) \mid C(k,l); \{k,l\} \neq \{i,j\} \\ &\quad = P(C(i,j) \mid C(k,l); \{k,l\} \in G_{ij}) \\ &\quad = \frac{1}{Z} e^{-U(C)/T} \end{aligned} \tag{3}$$

where *U* is the so called energy function, Z is a normalizing constant factor and T is a temperature term often used in statistical physics. In our model we always considered a second order neighbourhood, that is, the eight pixels closer to each single pixel of the image. If we want to maximize $P(C(i,j) \mid C(k,l); \{k,l\} \in G_{ij})$, we find that we need to minimize *U(C),* where

$$U(C(i,j)) = \sum_{\{k,l\} \in G_{ij}} \beta I(C(i,j), C(k,l)) \tag{4}$$

$$\begin{aligned} \text{and } I(C(i,j), C(k,l)) &= -1 \text{ if } C(i,j) = C(k,l) \\ &= \ \ 0 \text{ if } C(i,j) \neq C(k,l) \end{aligned} \tag{5}$$

with β parameter. This potential takes into account the neighbourhood of each pixel, by counting the pixels assigned to the class of pixel *(i, j)*.
To perform the classification we need to minimize

$$U(X_1, \cdots, X_n, C) = \alpha_r U_{spect}(X_r) + U_{sp}(C) \tag{6}$$

where U_{sp} is given by previous equation and U_{spectr} is defined as below.

$$U_{spectr}(X_r(i,j), C(i,j)) = \frac{B_r}{2} \ln \mid 2\pi \sum_k \mid + \frac{1}{2}(X_r(i,j) - \mu_k)^T \sum_k^{-1} (X_r(i,j) - \mu_k) \tag{7}$$

$\sum_k$ and μ_k are, respectively, the class-conditional covariance matrix and mean vector for class k and B_r is the number of spectral bands or features for source *r*.
When we classify an image we try to assign a class label to each pixel of the image, taking into account the values of the data as well as the prior information. To do this, there are many algorithms available, such as Simulated Annealing, MAP (Maximum A Posteriori Probability), MPM (Maximizing of the Posterior Marginals) and so on; these algorithms are computationally demanding; so, we choose the ICM, Iterated Conditional Mode, which allows to reach rapidly a local minimum of the energy function. We implemented the use of ICM algorithm by defining the following steps:

1. Initialize $\hat{C}$ for each pixel using the prior information, that is, a classification map of the same scene obtained with Fuzzy Artmap;

2. For all pixels *(i, j)*, update $\hat{C}(i, j)$ by the class that minimizes (7);
3. Repeat step (2) until the percentage of changed pixel between two subsequent iterations is smaller than a value chosen by the user.

METHODOLOGY

To achieve our classification aims we chose - as urban test site - the city of Pavia, which has already been widely analyzed for other purposes and of which we had a consistent number of SAR data (ERS-1 and ASAR) and a good knowledge thanks to a detailed ground truth, manually extracted from very high resolution images of the area, which delineates the urban area (red) against agricultural surround (green) and the Ticino river (blue) flowing South of the town. So, we collected some Landsat images of the city to perform our classification based on the MRF model. The resulting data set considered is made up by two Landsat images, acquired on 7th April 1994 and 8th October 2000, by two single polarization Precision Mode ASAR images, recorded on 25th November 2002 and on 8th December 2002, and by one Ers data acquired on 13th August 1992. After having collected these data of the test city and having registered the one to the others and to the related ground truth we first performed the single source and the joint classification of the two types of data trough the use of a neuro-fuzzy classifier, in order to provide a first analysis of the classification accuracy available using Landsat, SAR and the two sensors together and then we used these classification maps as priori information for the classifier based on the MRF model.

RESULTS AND CONCLUSIONS

In the first table shown below the first classification for each SAR acquisition, which is indicated by (A) is the one obtained by ARTMAP which has been used ad priori information for the MRF classifier. The second classification, indicated by (M), is the one obtained with this classifier.

Images	Water	Vegetation	Urban Areas	OA
08/13/93 (A)	75,25	88,13	52,6	83,2
08/13/93 (M)	94,31	85,77	65,57	83,35
11/25/02 (A)	56,48	71,37	56,33	69,06
11/25/02 (M)	45,29	72,32	61,83	70,31
12/08/02 (A)	29,73	72,96	40,59	67,72
12/08/02 (M)	58,79	77,18	46,6	72,76
11/25+12/08 (A)	61,58	80,05	56,44	76,54
11/25+12/08 (M)	45,07	81,41	60,86	77,87
11/25+12/08+08/13(A)	62,9	91,48	61,41	86,89
11/25+12/08+08/13(M)	46,75	95,98	63,07	90,53
11/25+12/08+08/13(SA)	66,43	93,43	61,85	88,69
11/25+12/08+08/13(M)	42,75	95,95	62,37	90,31

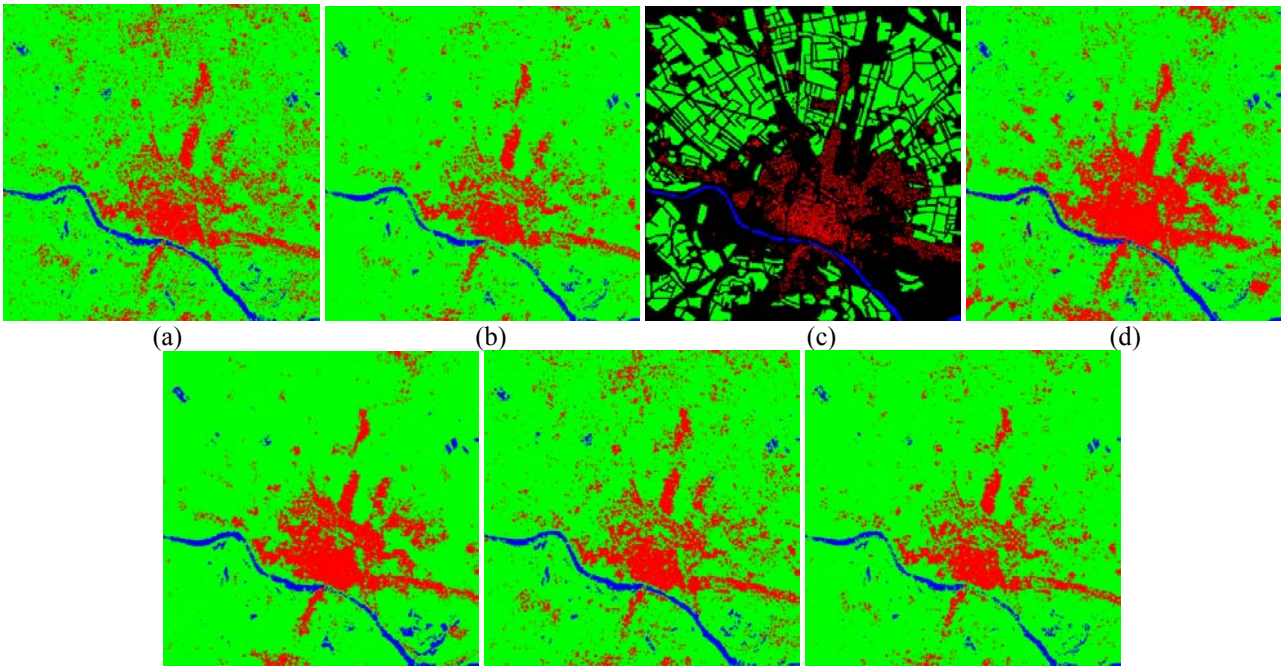
In the following table we show the classification results for Landsat data with the two classifiers (ARTMAP and MRF), from which we can say that the overall accuracies improve with the second approach.

Images	Water	Vegetation	Urban Areas	OA
04/07/94 (A)	82,88	60,29	91,26	64,85
04/07/94 (M)	90,81	60,78	90,73	65,39
10/08/00 (A)	91,13	65,41	89,12	69,11
10/08/00 (M)	91,25	74,06	87,71	76,24
04/07+10/08 (A)	94,31	73,61	88,46	76,03
04/07+10/08 (M)	76,77	79,6	89,24	80,79

Images	Water	Vegetation	Urban Areas	OA
ERS+ASAR+LAND 10/08 (A)	87,09	69,24	87,27	72,01
ERS+ASAR+LAND 10/08 (M)	48,99	82,53	84,14	81,94
ERS+ASAR+LAND 06/21(A)	60,06	92,9	67,67	90,14
ERS+ASAR+LAND 06/21 (M)	54,07	98,23	71,58	93,71

Finally, in the last table we show the classification results of the whole available data set considered together obtained with both the classifiers. In all these tables we compare the classification results obtained with Fuzzy Artmap and with

the MRF model. We considered - in all of these classifications - α as the overall accuracy of the Artmap classification, $\beta = 1.5$, a second order neighbourhood and a pixels' threshold percentage of 5%. We can say that with these parameters the ICM algorithm reach the convergence in two, three or at most in four iterations.



(a) (b) (c) (d)

(e) (f) (g)

Fig. 1 . Classification maps of the three SAR available images obtained with ARTMAP (a) and with the MRF approach (b); the ground truth used to compare classification results (c); classification maps of the joint classification of the LANDSAT image of 21st June and the three SAR images with ARTMAP (d) and MRF (e). Classification maps of the three SAR available images obtained with spatial ARTMAP (f) and with the MRF approach (g).

As we can see from these results and from the classification maps shown in Fig. 1, the MRF approach results in a sometimes slightly but in any case higher classification overall accuracy, and in particular the urban areas are often better recognized with this method. We want to remark here that we also considered as priori information for the MRF model the classification obtained with spatial Fuzzy Artmap obtained considering a spatial window of 3×3 pixels around each pixel of the image. The classification results of this approach for the classification of the three SAR data are shown in the last two lines of the first table, while the classification maps are shown in Fig.1 (f) and (g). This approach, tested on the whole data set, allow to reach the same classification results as the ones obtained using as priori information the fuzzy Artmap map, but it allows to stop the MRF based classification after one iteration. So, it results in a faster classification. From the tests done we can say that the joint classification of LANDSAT and SAR data provides a better accuracy in urban areas characterization in regard to the use of SAR data alone. These accuracies are comparable with the ones obtained from the classification of Landsat data alone, but in the joint classification of the two sensors we obtained a better discrimination of all the classes. In conclusion, we want to emphasize the use of more than one sensor in land cover mapping and the use of a MRF model based approach as a fast and promising method.

REFERENCES

[1] G. Amici, F. Dell'Acqua, P. Gamba, G. Pulina, "Fuzzy, neural and neuro-fuzzy classification of pre- and post event SAR images for flood monitoring and disaster mitigation", *Proc. Of the First International Workshop on Multi-temporal Analysis of Remote Sensing Images,* Trento, Italy, 13-14 Sept. 2001, pp. 100-107.

[2] O. Jakubowicz, "Autonomous reconfiguration of sensor systems using neural nets", in *Proc. SPIE: Sensor Fusion*, Orlando, FL. Apr. 1988, pp. 197-203.

[3] R. Chellappa and A. Jain, Eds., "Markov Random Fields: Theory and Application", New York: Academic, 1993.

[4] R. C. Dubes and A. K. Jain, "Random field models in image analysis", *J. Appl. Statist.*, vol. 16, pp. 131-163, 1989.