NEW CLASSIFICATION TECHNIQUES FOR ANALYSIS OF REMOTE SENSING INTEGRATED DATA

Elena Console*, Marie Catherine Mouchot**
*** T.E.A. s.a.s.**
Via Bruno Chimirri 28, 88100 - Catanzaro (Italy)
phone: +39.961.74.75.28, fax: +39.961.72.56.36, e-mail: console_e@abramo.it
**** Département I.T.I., ENST de Bretagne**
Technopole de Brest-Iroise, BP 832, 29285 Brest Cedex (France)
phone: +33.298.00.13.58, fax: +33.298.00.10.98, e-mail: mc.mouchot@enst-bretagne.fr

Abstract -- The aim of this work is the application of techniques based on fuzzy logic for integrating and classifying spectral data and spatial information. The analysis was performed on a Landsat-TM image and on a set of map field data.

Firstly fuzzy membership values of Landsat image pixels were calculated and a buffer area was defined around map field data for each considered class.

The membership values were then modulated using spatial information derived from known targets.

Finally the accuracy of the classification was assessed on a set of test data. The obtained results shown a significant improvement compared to more traditional methods of classification.

INTRODUCTION

Adequation of techniques for processing remote sensing imagery improves the knowledge of Earth' surface and contributes remarkably to the development of policies for planning and monitoring of environmental resources. The possibility of using different type of data can increase our ability to visually discriminate between different ground targets.

The fusion techniques of optical imagery with other information require appropriate procedures for data analysis. Furthermore it becomes necessary to determine suitable methodologies of classification which take into account the heterogeneity of the acquired data and the indeterminateness of considered targets.

In this work we used fuzzy logic for integrating and classifying spectral and spatial information. More particularly the analysis was performed on a Landsat-TM image of Southern Italy and on cartographic data.

The map field data are represented by geographical coordinates of training pixels and by vector files of hydrography of the considered area.

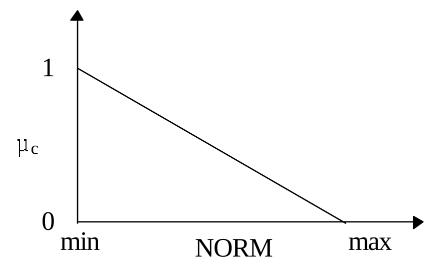
These known targets were used for modulating the fuzzy membership values of all pixels located in their neighbourhood. The size of the neighbourhood was defined differently for each class according to the experts's opinion.

The steps of the applied methodology were the followings:

a) choice of fuzzy membership function and membership values computing;
b) estimation of buffer around the set of known targets;
c) modulation of membership values in buffer areas;
d) hard classification from fuzzy classification;
e) accuracy assessment
f) mixed classification.

APPLIED METHODOLOGY

The chosen membership function for computing fuzzy membership values of each pixel to considered class has been the following (Figure 1):



if norm = min, then $\mu_c(x_i) = 1$
if norm = max, then $\mu_c(x_i) = 0$

Figure 1: Membership function

The adopted norm is defined as the absolute distance from the mean of the considered class normalized by the standard deviation:

$$norm(x_{i,c}) = \sum_{b=1}^{6} \frac{\left| x_{i,b} - \bar{x}_{c,b} \right|}{\sigma'_b}$$

where:
i = i*th* considered pixel
c = c*th* considered class
b = b*th* considered band

0-7803-3839-1/97 $10.00(c) IEEE

$\bar{x}_{c,b}$ = mean of c*th* class in the b*th* band

$$\sigma'_b = \frac{\sigma_b}{\sum_{b=1}^{6} \sigma}$$ normalized standard deviation

σ_b = standard deviation of the b*th* band.

Next, the physical distances of each pixel from map field data have been calculated (Figure 2a).

For each class a buffer area around the set of know targets (training pixels and hydrographic vector files) were defined on the ground based on expert's knowledge (Table 1). These windows delimit the pixels which the membership to the considered class is most probable (Figure 2b).

Table 1

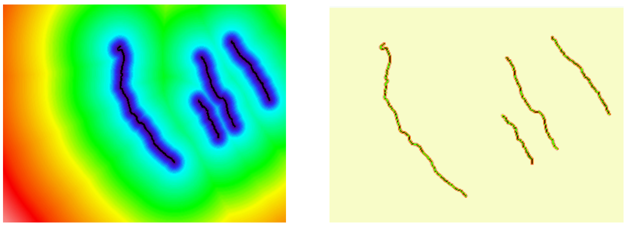
DIMENSION OF BUFFER AREA FOR EACH CLASS			
Class	**meters**	**Class**	**meters**
Grass covered area	200	Cultivated area	250
Orchards	250	Forest	100
Non vegetated area	100	Road network	60
Shady forest	500	Urban area	150
Hydrography	100		

We assigned these pixels to values between 0 (assigned to the central pixel of the buffer area) and 1 (assigned to the pixels of external edge) defined as modulation values, *m.v.* (Figure 3).

In order to integrate the spectral and the spatial information we assumed that the pixels located around the known targets belong most likely to the same class of the target. Hence the membership values of these neighbour pixels were modified using the modulation values drawn from spatial data:

$$\textit{final membership value} = \mu_c(x_i)^{m.v.}$$

The resulting images have been filtered applying a median filter, which size was identical to the class buffer size.



a) physical distances from streams b) buffer area

Figure 2: Hydrographic network

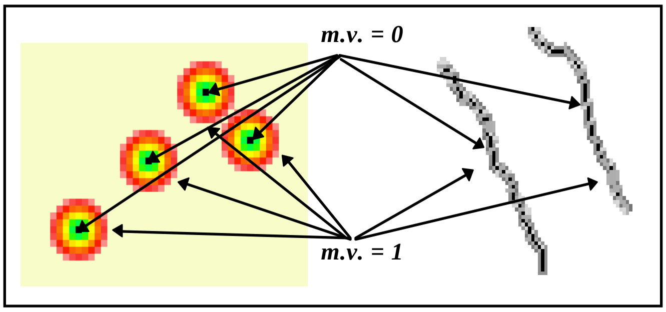


Figure 3 - Modulation values

RESULTS

In order to verify the accuracy of results, the error matrix and the K indices have been computed on hard classification derived from fuzzy classification against a sample of ground truth pixels.

The hard classification was produced by assigning each pixel to the class having the highest membership value. We considered a fixed threshold value, under which the pixel was labelled non classified:

$$\text{if } \mu_c(x_i) = \operatorname*{Max}_{c=1}^{8} \mu(x_i), \qquad \text{then } x_i \in \text{class c}$$

The results were compared with conventional classifications (minimum distance and maximum likelihood) and hard classification not integrating spatial information.

In the Table 2 we have reported the values of global index K and the conditional index K for classes "Urban area" and "Road Network". These two classes were chosen more particularly because they prove to exhibit a fairly high rate of false classifications in prior studies.

One can see that the integration of the spatial information significantly improved the accuracy of the classification both in term of global K and conditional K.

Table 2

COMPARISON OF RESULTS			
Algorithm	**Global K**	**Conditional K**	
		Urban a.	***Road netw.***
Minimum distance	0.37	0.26	0.27
Maximum likelihood	0.46	0.33	0.25
Fuzzy cl. non integrated with spatial information	0.30	0.44	0.19
Fuzzy cl. integrated with spatial information	**0.42**	**0.60**	**0.28**

Finally the mixed classification was considered; the mixed classes are represented by the combination of the two classes having the highest fuzzy membership values.

Superimposing the mixed classification image on the official cartography of the considered area it becomes obvious that correspondence is better than with maximum likelihood classification.

It is then confirmed that the “Urban area” and “Road Network” classes are better identified (Figure 4).

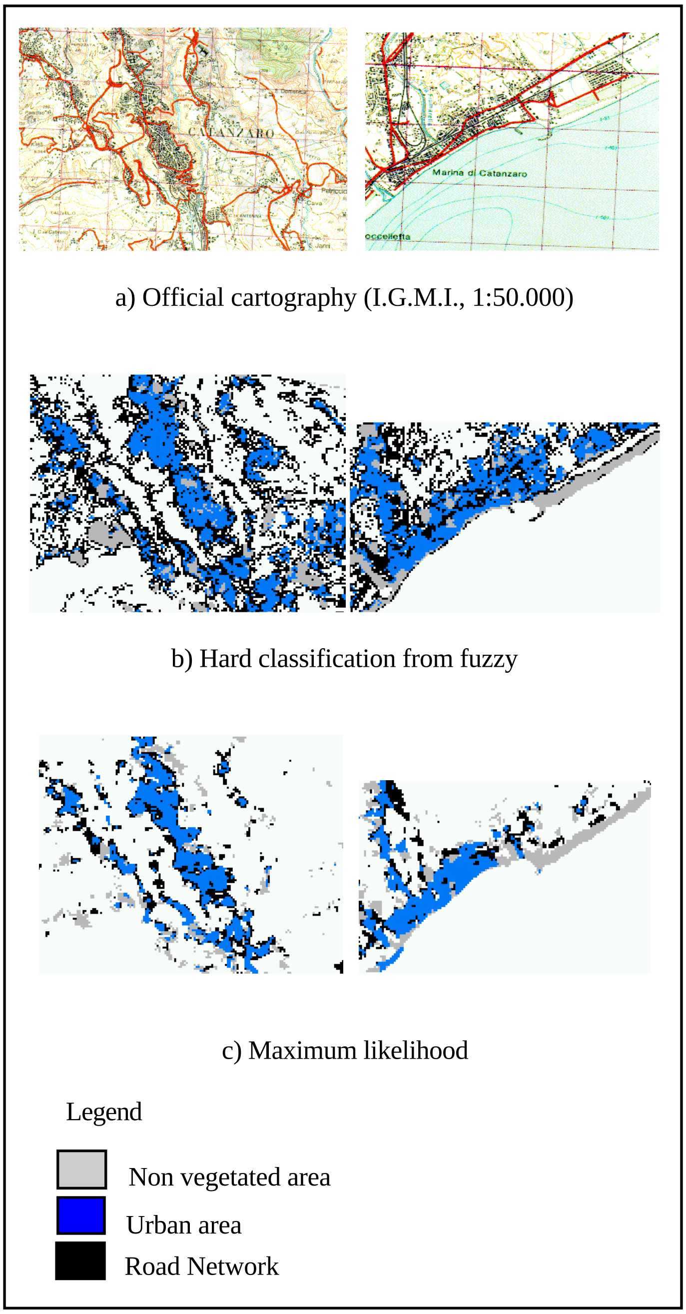


Figure 4: Comparison of results

CONCLUSIONS

The use of different data sources and their integration significantly improved the performance of the adopted fuzzy classification technique. In particular, the applied methodologie, by allowing the modulation of membership values of each class, enhanced its ability to discriminate among all considered soil coverage, especially in the difficult case of the urban agglomerates.

REFERENCES

[1] D. Altman, “Fuzzy Set Theoretic Approaches for Handling Imprecision in Spatial Analysis”, Int.Journal of Geographical Information Systems, vol. 8, n. 3, pp. 271-289, 1994

[2] S. Aronoff, “Classification Accuracy: a User Approach”, Photogrammetric Engineering and Remote Sensing, vol. 48, n. 8, pp. 1299-1307, 1982

[3] E. Console and M.C. Mouchot, “Fuzzy Classification Techniques in the Urban Area Recognition””, in Proceeding of IGARSS’96 - Remote Sensing for a Sustainable Future, IEEE Publications, vol.II, pp. 1373-1375, 1996

[4] L.A. Zadeh, “Fuzzy Sets”, Inf. Control., vol. 8, pp. 338-353, 1965