Engineering Journal of the University of Qatar, Vol. 14, 2001, pp. 77 – 104

AN EVALUATION OF THE EXPLICIT FUZZY METHOD USING PARAMETRIC AND NON-PARAMETRIC APPROACHES FOR SUPERVISED CLASSIFICATION OF MULTISPECTRAL REMOTE SENSING DATA

Farid Melgani, Bakir A.R. Al Hashemy and Saleem M.R. Taha
Electrical Engineering Department
College of Engineering
University of Baghdad, Iraq

ABSTRACT

Fuzzy Classification is of great interest because of its capacity to provide more useful information for Geographic Information Systems. This paper describes an Explicit Fuzzy Supervised Classification method, which consists of three steps. The explicit fuzzyfication is the first step where the pixel is transformed into a matrix of membership degrees representing the fuzzy inputs of the process. Then, in the second step, a MIN fuzzy reasoning rule followed by a rescaling operation are applied to deduce the fuzzy outputs, or in other words, the fuzzy classification of the pixel. Finally, a defuzzyfication step is carried out to produce a hard classification. The classification results of Landsat TM data show the promising performance of the method and, particularly, the classification time. These results are compared with those produced by the Maximum Likelihood method and a non-parametric method based on the use of Artificial Neural Networks.

KEY WORDS: Remote Sensing, Fuzzy Classification, Fuzzyfication, Fuzzy Reasoning Rule, Artificial Neural Networks.

1. INTRODUCTION

A wide range of pattern recognition techniques has appeared during the last two decades for information extraction from remotely sensed data. The most popular technique is the Maximum Likelihood Classification method known for its good performance and robustness. Also, an approach to the problem of classification by the use of Artificial Neural Networks has been developed by the following research works [1,2,3 and 4]. But as Heermann and Khazenie [1] say about the A.N.N.: " This changes the solution process from finding ways to understand and represent the problem in a computer language to a task of providing examples of the problem for the network to learn. ". The main advantage of the A.N.N. Classification methods is that they are "Distribution Free", but they involve many problems such that no general criteria for defining a suitable network architecture, dependence on training conditions and the slow training time [5].

But all these methods are used to carry out a hard classification based on the principle of " One pixel - One class. ". In other words, it means that a pixel is either a full member or not of a class. This logic of processing does not represent the best way to deal with data that are naturally mixed, and so imprecise in nature. The advent of the Fuzzy Set theory [6] has invaded a lot of fields such as Fuzzy Control Systems, Fuzzy Image Processing [7] and, inevitably, Fuzzy Classification of remote sensed data. Fuzzy Classification, or pixel unmixing, estimates the contribution of each class in the pixel. It assumes that a pixel is not an indecomposable unit in the image analysis and, consequently, works on a new principle : " One pixel - Several classes " to provide more information about the pixel unlike the hard classification methods which are poor in information extraction. A fuzzy c-means clustering algorithm has been developed in [8] and in [9] for unsupervised classification of remote sensing data. Wang [10] modified the traditional Maximum Likelihood method by implementing a fuzzy mean and a fuzzy covariance matrix beforehand calculated. Then, the fuzzy membership degrees are computed by applying the Maximum Likelihood procedure defined on fuzzy class signatures. Another parametric classification method has been proposed in [11] to express the class proportions. In [12] and [13], models based on a linear mixing are introduced. Of course , Artificial Neural Networks have been exploited in the fuzzy classification to relate their outputs to the class contribution in a given pixel [14], [15].

In this paper, an Explicit Fuzzy Classification method is proposed for multispectral remote sensing images carried out in three steps. An Explicit Fuzzyfication step is obtaining an estimation of the class contributions in each band assuming a Gaussian distribution of the classes. The second step applies a fuzzy reasoning rule on these fuzzy inputs to obtain, after a rescaling operation, the fuzzy classification of the scene. Finally, a hard classification can be deduced in the defuzzyfication step in order to illustrate the whole of the cover classes in the same map. For comparison purpose, the Maximum Likelihood algorithm is chosen because it represents a widely used "standard" that provides minimum total classification error for Gaussian class distribution. The method using Artificial Neural Networks is a non-parametric approach known for its high ability to reproduce the boundaries between the classes. This comparison will put under light the reliability of the Gaussian distribution model we use for defining the class fuzzy membership functions, to deal with the classification problems of actual applications. The proposed method is applied on Landsat TM data, bands 2,3,4 and 7, representing a part of Mafraq Region in Jordan (Fig.4-a). Results of the fuzzy classification and the comparison with a Neural Network and the Maximum Likelihood classifiers of the hard classification results show the efficiency of the method.

The rest of this paper is organized as follows: Section II provides a description of our approach and the methodology used. Section III presents the Neural Network approach, whereas the application and results are revealed in Section IV. Finally, Section V ends the paper by adequate conclusions.

2. METHODOLOGY

Our approach, illustrated in Figure 1, analyzes the data in order to extract a maximal number of information, through an explicit fuzzyfication process, for making the decision rule more easy and reliable. Then, a MIN reasoning rule is carried out to pick out for each class its worst contribution from the set of membership degrees of the class among the different bands. The inferred raw fuzzy outputs are rescaled to provide a fuzzy classification of the pixel. The jump from a fuzzy classification to a hard classification is easily done by a MAX operation representing the defuzzyfication process.

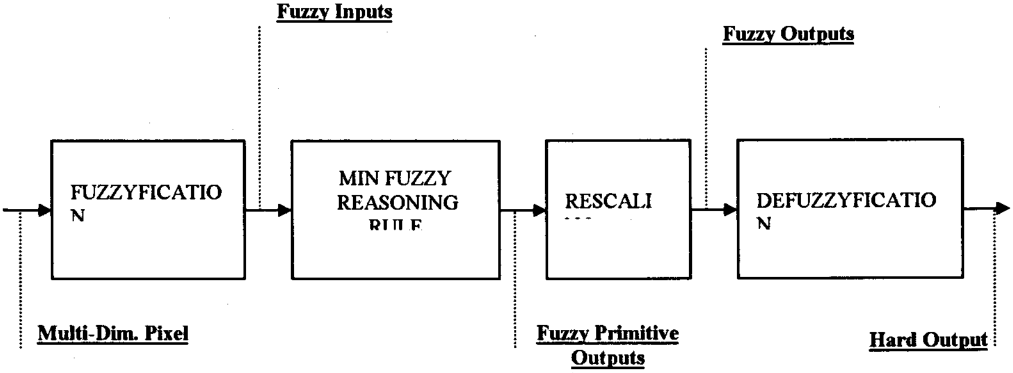

Fig. 1. Block diagram of the explicit fuzzy method.

Explicit Fuzzyfication Process

The fuzzy domain is consisting of several fuzzy sets representing the bands, and each fuzzy set (band) contains fuzzy subsets representing the cover classes. Each fuzzy subset (cover class c), in a given fuzzy set (band b) is fully defined by its membership function $f_{b,c}(x_b)$ where x_b is the brightness value of the spectral pixel $\underline{X}$ in the band b. The pixel vector $\underline{X}$ in the B-dimensional space is :

$$\underline{X} = [\, x_1, x_2, \ldots, x_b, \ldots, x_B \,]^T$$

We have oriented the choice of the membership function to the Gaussian distribution because it represents a powerful general distribution model and involves a minimal extraction computational cost from the statistical characteristics of the signatures. Furthermore, the linear models are not able to express correctly the natural non-linear distribution of the classes. The class distributions are totally independent of each others and, consequently, we do not impose any mixing assumption. The independence between the fuzzy subsets (classes) involves a non respect of the formal fuzzy requirement:

$$\forall\, x_b \in [0\,,255], \quad \sum_{i=1}^{N} f_{b,i}(x_b) = 1 \tag{1}$$

where N is the number of fuzzy subsets.

Of course, we can implement a normalization operation to satisfy (1) but we have not done it for three main reasons:

1. The MIN fuzzy reasoning rule is insensitive to this requirement.

Proof

The normalization applies the following transformation :

$$f_{i,c}(x_i) \qquad\qquad \tilde{f}_{i,c}(x_i) = \frac{f_{i,c}(x_i)}{\sum_{j=1}^{N} f_{i,j}(x_i)}$$

Of course, for $\forall\, x_i \in [0\,,255]$, $\sum_{j=1}^{N} f_{i,j}(x_i) \neq 1$

However, for $\forall\, x_i \in [0\,,255]$, $\sum_{j=1}^{N} \tilde{f}_{i,j}(x_i) = 1$

We deduce easily that the normalization operation is not necessary for the MIN reasoning rule because:

$\forall$ b $\in\{1,2, \ldots, B\}$and $\forall$ c $\in\{1,2, \ldots, N\}$,

$$\text{if } f_{b,c}(x_b) \leq f_{b,c}(y_b) \Rightarrow \frac{f_{b,c}(x_b)}{\sum_{i=1}^{N} f_{b,i}(x_b)} \leq \frac{f_{b,c}(y_b)}{\sum_{i=1}^{N} f_{b,i}(y_b)} \Rightarrow \tilde{f}_{b,c}(x_b) \leq \tilde{f}_{b,c}(y_b)$$

2. We alleviate the burden of unuseful computations due to the normalization operation.

3. The class distributions keep their natural aspect.

In the algorithm, we have introduced the notion of modulated Gaussian distribution if a prior knowledge of the class extents in the scene is available. The facultative prior knowledge of the class extents in the study area is reflected in the size of the class signatures, necessary to carry out a statistics extraction in order to compute two important parameters to define completely the membership function:

1. The mean (μ) of the class signature which represents the ideal pixel of the class, or in other words, the only point without any ambiguity about its membership.

2. The standard deviation (σ) of the class signature, which will determine in a way the width of the fuzzy subset.

The membership function of a class c in a band b is:

$$f_{b,c}(x_b) = \exp\left(-\frac{(x_b - \mu_{b,c})^2}{2\sigma^{*2}_{b,c}} \right) \tag{2}$$

where, $\mu_{b,c}$ is mean of the class c in the band b and $\sigma^*_{b,c}$ is modulated standard deviation of the class c in the band b.

$$\sigma^*_{b,c}=\alpha_{b,c} \times \sigma_{b,c} \tag{3}$$

where, $\sigma_{b,c}$ is standard deviation of the class c in the band b $\alpha_{b,c}$ is modulation factor.

If no prior knowledge of the class extents in the study area is available, the modulation factor must be neutralized to work with a pure Gaussian distribution. It means that:

$$\forall\ b,c: \quad \alpha_{b,c} = 1$$

Otherwise, the modulation factor is calculated by the formula:

$$\alpha_{b,c} = \log(P_{b,c} + \beta_0) \qquad (4)$$

where, $P_{b,c}$ is expected extent of class c in band b β_0 is a constant.

In order to reduce the problem that the prior knowledge of the class extents in the scene is generally vague, we will not use it directly in the modulation factor. Instead, we propose to reflect it in the size of the class signatures and to introduce the notion of expected extent of the classes in each band by deducing the number of pixels expected to correspond to the mean of each class signature. For a given band b, $T_{b,c}$ represents the expected number of pixels, in the signature histogram, corresponding to the mean $\mu_{b,c}$ of the class c.

$$P_{b,c} = \frac{T_{b,c}}{\sum_{i=1}^{N} T_{b,i}} \qquad (5)$$

where, N is the number of classes.

The modulation principle is useful to favor the large extent classes over the lower extent ones and, particularly, in the spectral regions where overlaps make difficult the decision. We have deduced by empirical observation that a constant β_0=1.25 is the optimal one to satisfy the need to " not strengthen too much the strong classes and not weaken too much the weak classes ". Because this constant value leads to a quasi linearity in the definition interval [0 , 1] of the variable $P_{b,c}$, the modulation function can be approximated by a linear function.

$$\alpha_{b,c} = 0.6 \times P_{b,c} + 0.22 \qquad (6)$$

The fuzzyfication process, as illustrated in Figure 2, computes for a given pixel the membership degrees for each class c and band b from the membership functions fully defined using the Gaussian distribution and the results of the statistics

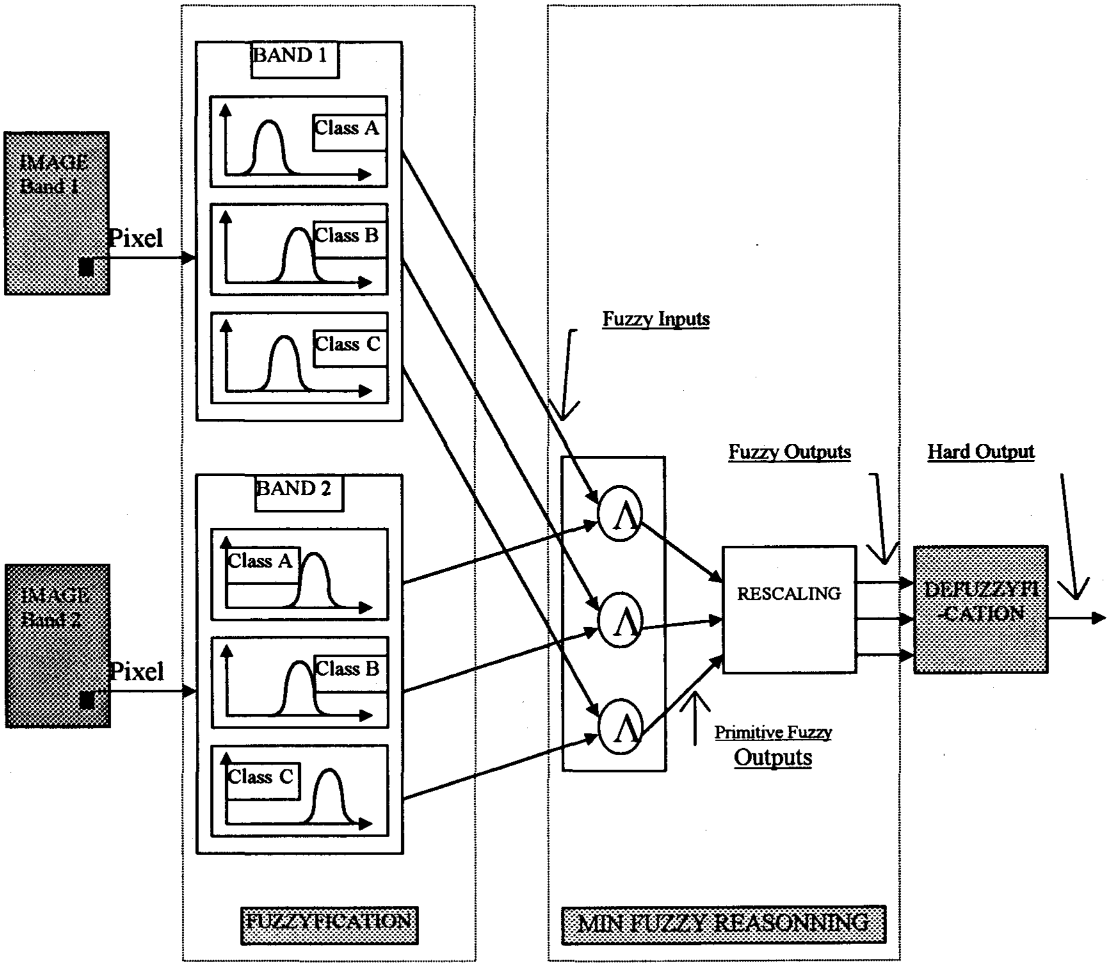

Fig. 2. Architecture of the explicit fuzzy classification method illustrated for two bands and three classes.

extraction. We obtain a matrix of fuzzy inputs $\underline{F}$ip of order B×N where N is the number of classes and B the number of bands.

For a multispectral pixel $\underline{X}$, the matrix of fuzzy inputs $\underline{F}$ip can be written :

$$\underline{F}ip = \begin{bmatrix} f_{1,1}(x_1) & f_{1,2}(x_1) & \dots & f_{1,N}(x_1) \\ f_{2,1}(x_2) & f_{2,2}(x_2) & \dots & f_{2,N}(x_2) \\ \dots & \dots & \dots & \dots \\ f_{B,1}(x_B) & f_{B,2}(x_B) & \dots & f_{B,N}(x_B) \end{bmatrix} \quad (7)$$

This matrix is then analyzed by the MIN reasoning rule which is the second step of the method to produce the fuzzy classification.

MIN Fuzzy Reasoning Rule

If we observe the behavior of the human being before its daily problems, we will certainly deduce that one, always, use a simple but very efficient principle which can be resumed in these few words "Work on the worst assumptions to find the best results". The fuzzy reasoning rule that satisfies entirely this principle is inevitably the MIN rule. In another side, the absence of covariance information makes the fuzzy partitions not necessarily optimal because classes with a high degree of covariance will not be limited to the high fuzzy membership areas. Therefore, high fuzzy membership values represent more a maximum possible value than a high prior probability of class membership. Consequently, the lowest fuzzy value for all the bands is more likely to be closest to the "true" fuzzy membership. The MIN reasoning rule, applied on the matrix of fuzzy inputs produced by the previous step, will consider, for each class, the membership degrees provided by the different fuzzy sets (bands) and pick out the minimal membership degree to represent the class extent in the pixel.

Applying the MIN operation, for which an illustration is given in the Fig. 2, on (7), we obtain a primitive fuzzy output vector.

$$\underline{F}'op = [\ F'_1(\underline{X}),\ F'_2(\underline{X}),\ \dots\dots,\ F'_N(\underline{X})\ ]^T \quad (8)$$

where

$$F'_i(\underline{X}) = \mathrm{Min}\ (\ f_{b,i}(x_B)\ ) \quad \text{with } b=1,2,\ \dots,\ B \quad (9)$$

The fast, simple and reliable step of the MIN fuzzy reasoning rule is immediately followed by a rescaling operation in order to normalize the class extents deduced from different fuzzy sets and sharing the same pixel.

The final vector of fuzzy outputs is,

$$\underline{F}op = [\, F_1(\underline{X}),\ F_2(\underline{X}),\ \ldots\ldots,\ F_N(\underline{X})\,]^T \tag{10}$$

where

$$F_i(\underline{X}) = \frac{F'_i(\underline{X})}{\sum_{j=1}^{N} F'_j(\underline{X})} \tag{11}$$

This vector represents the fuzzy classification showing the inferred class extents of the pixel, and from which a hard classification can be deduced by a defuzzyfication step.

We note that the extraction from the data of first order statistical characteristics for deducing the fuzzy membership functions and the application of the MIN fuzzy reasoning rule involve a partition of the spectral space into fuzzy parallelepiped regions which relate conceptually the method to a fuzzy parallelepiped classifier. Furthermore, the explicit fuzzyfication associated to the MIN fuzzy reasoning rule reveal the modular aspect of the method, as shown in Fig. 2, which allows to add and remove any band from the classification scheme of the study area because of the relative independence between the bands. The adding or removal of a cover class however, requires a minor computational cost due to the adjustment of the modulation factors if a prior knowledge of the class extents is available.

Defuzzyfication

Next to the fuzzy classification, a hard classification is provided by performing a MAX operation to defuzzyfy the fuzzy output into a hard output. We select among the classes mixed in the pixel the class C with the highest extent such that

$\forall i \in \{1,2, \ldots, N\}$ and $i \neq c$,

$$F_C(\underline{X}) \geq F_i(\underline{X}) \tag{12}$$

3. THE NEURAL NETWORK APPROACH

The last decades have shown an interest in the mechanisms of the human brain in a hope to discover a new computing machine philosophy. The biological inspiration of the brain led to Artificial Neural Networks which are an architecture of elementary processing elements called neurons (nodes) carrying an activation function, working in parallel and highly interconnected by weighted links representing the network spirit.

Neural Network Architecture

As Hornik [16] proved in 1989, a three layer network with sigmoid activation function is capable to reproduce any decision function with any degree of accuracy provided that enough hidden units are available. Consequently, we direct our choice to a three layer architecture as illustrated in Figure 3.

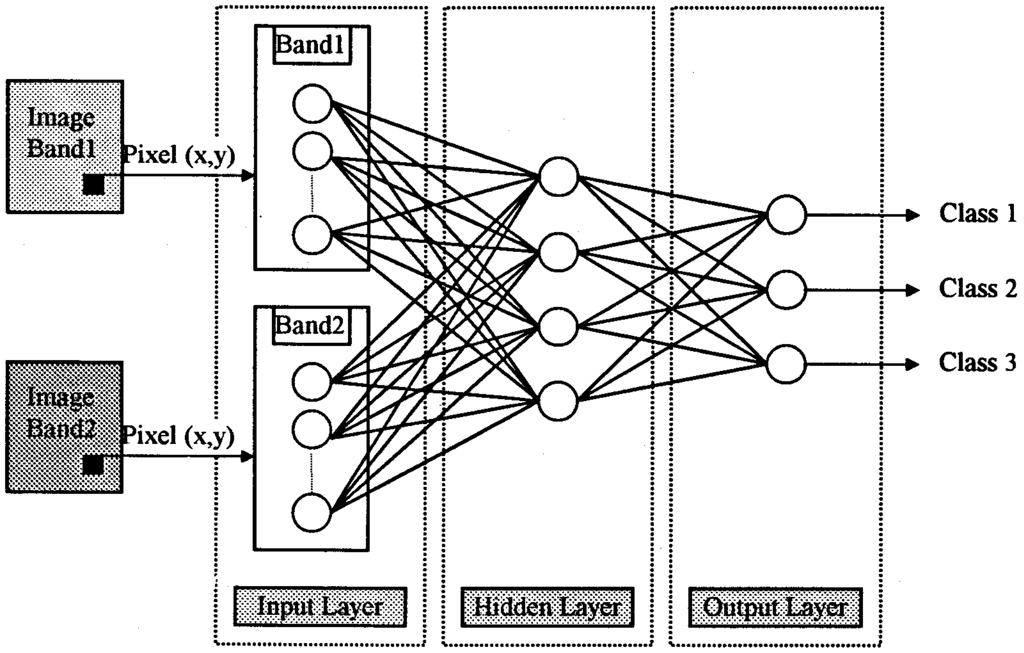

Fig. 3. Neural net architecture (two bands and three classes)

The input layer is the interface between the study area image (taken in different bands) and the network. The brightness value representing the pixel in a given band is beforehand coded into a set of ten values according to the fuzzy input coding. Concerning the hidden layer, despite some attempts to find a general formula that determines the hidden node number [17], nobody has succeeded to prove that its method produces the optimal number of hidden nodes. Unfortunately, a trial-and-error method based on its own experience of the A.N.N. behavior remains the only way to compromise between the necessity for obtaining a minimal number of hidden nodes and a maximal classification accuracy. The output layer does not pose any problem because the number of nodes is equal to the class number, each output node representing a class.

Fuzzy Input Coding

The goal of the input coding is to express the initial information (brightness value) in a more detailed form in order to increase the network sensitivity to a smallest variation of the input. The best input coding is that satisfying the following requirements:

1. Neighboring values must have a quasi-similar code to involve a smaller number of weight changes.

2. Neighboring values must also be sufficiently distinguishable.

In the fuzzy input coding, the input is not expressed in a digital form but is defined in the interval [0,1]. The Universe of Discourse (brightness value axis) is partitioned into ten fuzzy subsets with Gaussian membership functions. The empirical formula satisfying the above requirements used to deduce the fuzzy input for the K^{th} fuzzy subset and a pixel of brightness value x is:

$$F_k = \exp\left[-\frac{\left(x - 28(k - 1)\right)^2}{625} \right] \tag{13}$$

This information representation reveals a faster convergence of the neural network training compared with the bipolar or gray input coding techniques.

Neural Network Training

During the training of the neural network, the weights are adjusted using the Back-Propagation Algorithm which is an extended version of the delta rule and has an objective to reduce the error between the actual and desired outputs in a gradient descent manner. Another important factor is the learning rate behavior during the network training process. The Exponential Learning technique is based on the idea to decrease the learning rate in an inverse exponential manner in order to profit from significant learning rate values in the first phase of learning to speed up the convergence process and, then, to drive the process carefully toward the optimal weight set in the last phase of training. Considering the following notations.

$$\begin{cases} k: & \text{Sweep number} \\ \alpha(0): & \text{Starting learning rate} \\ \alpha(k): & \text{Learning rate at } k^{th} \text{ sweep} \end{cases}$$

The formula is:

$$\alpha(k) = \alpha(0).\exp\left(\frac{-k}{100}\right) \qquad (14)$$

4. APPLICATION AND RESULTS

Data

Figure 4-a shows the study area (Mafrak -Jordan) taken in April 1994 through the band 7 of the Landsat TM with a spectral spatial resolution of 30×30 m^2. The 412×321 pixels to be classified are expressed in four spectral bands (Green band (0.52 ~ 0.60 μm), Red band (0.63 ~ 0.69 μm), Near-Infrared band (0.76 ~ 0.90 μm) and Infrared band (2.08 ~ 2.35 μm)) and contain five major cover classes, namely: Water, Village, Agricultural land, Forest 1 (Fir) and Forest 2 (Bush) with no available prior knowledge of their extent. A C language program is used to implement the three methods in a 166 MHz Pentium PC.

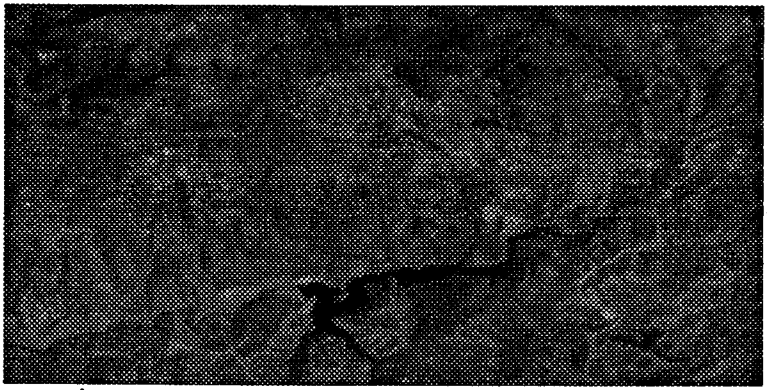

Fig. 4-a. Raw Landsat TM image of Mafrak (Jordan)

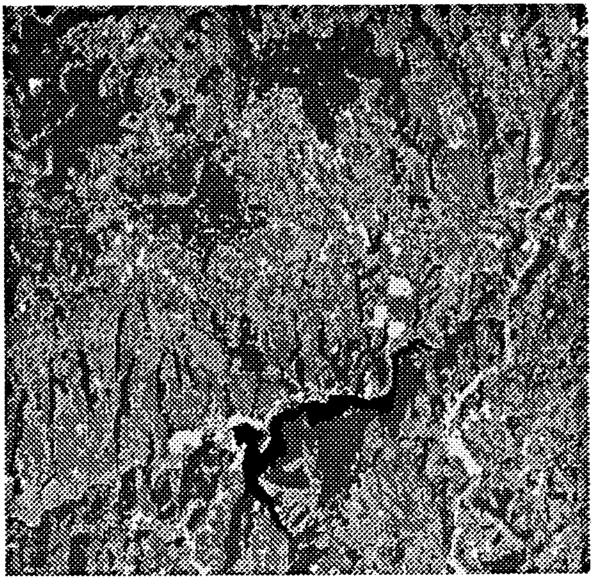

Fig. 4-b. By Maximum Likelihood Method taken in the Infrared Band.

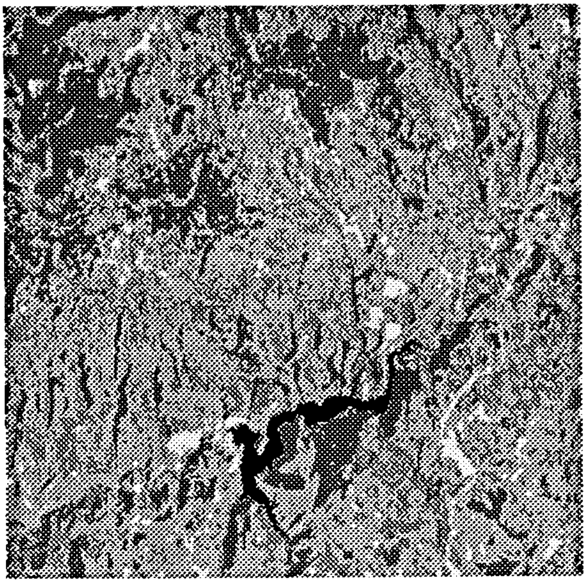

Fig. 4-c. By Explicit Fuzzy Method.

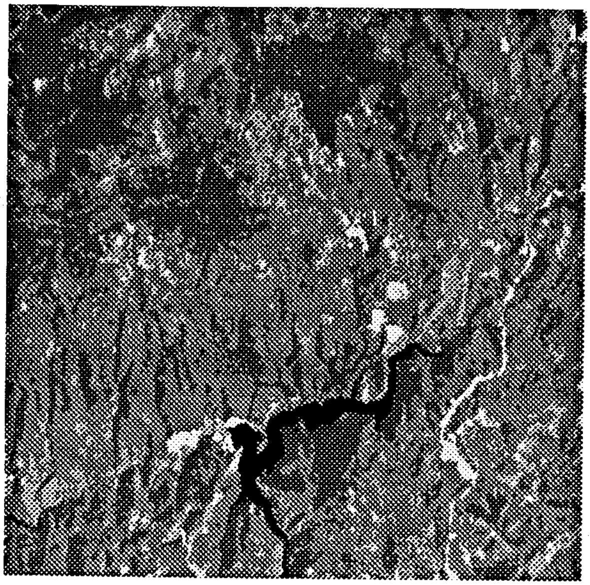

Fig. 4-d. By Neural Network Method.

Fig.4. Hard Classification of the Study Area.White (Village), Lightgray (Agric.), Gray (Forest 2),Darkgray (Forest 1), Dark (Water).

Discussion

For each class, we have selected a set of pixels from which two thirds are used as training site and the last third as test site (Table 1). After applying a statistics extraction from the data, we can observe, from the one dimensional fuzzy partition of the spectral space (Fig. 5), that the five classes are concentrated in the first third part of the brightness value axis involving a strong overlap between them. The effect of the MIN fuzzy reasoning rule is investigated by a two dimensional fuzzy partition of the spectral space to prove the capability of this rule to provide reliable results despite the overlapping degree between the classes. Figure 6 reveals the results of the rule applied on band 3 versus band 2 leading to a remaining conflict zone at the neighborhood of the point (B. V. band 2 = 50, B. V. band 3 = 50) between the classes : Village, agricultural land and Forest 2. In another side, the figure illustrates clearly the partition of the spectral space into fuzzy parallelepiped regions, which represent the multispectral references for the fuzzy and hard classifications of the scene.

Table 1. Pixel Number of Training and Test sites.

	TRAINING PIXELS	TEST PIXELS
WATER	75	34
VILLAGE	174	87
AGRIC.	230	115
FOREST 1	104	52
FOREST 2	132	72
TOTAL	**715**	**360**

Table 2. Neural Network Information.

Input Layer Nodes	4×10
Hidden Layer Nodes	10
Output Layer Nodes	5
Input Coding	FUZZY
Learning Method	EXPN
Rate Value	0.2
Momentum Value	0
Sweep Number	199
Learning Time [s]	567.82

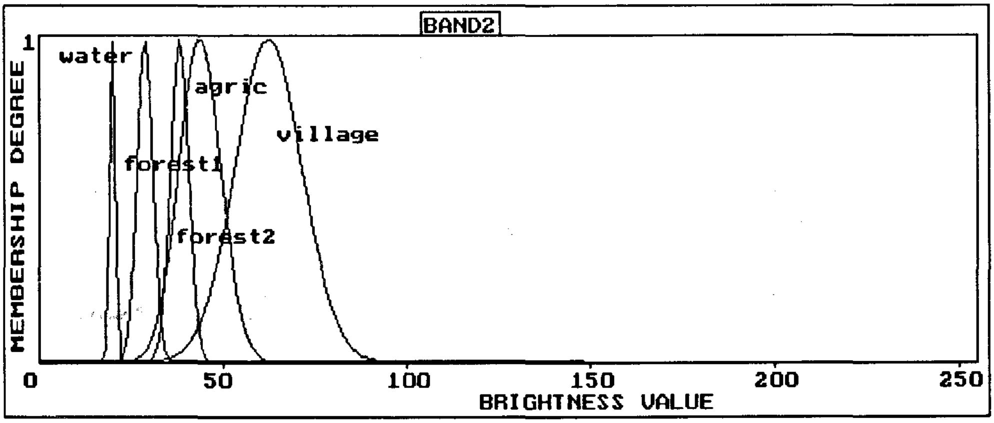

Fig. 5-a. One dimensional fuzzy partition of the band 2.

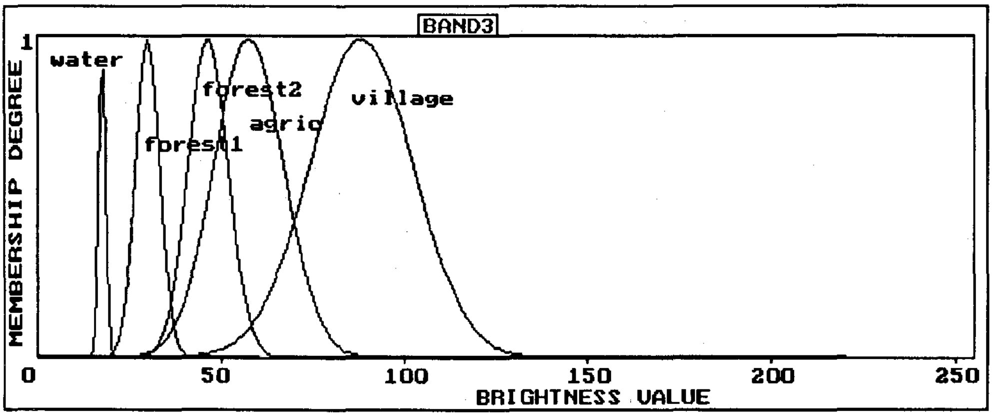

Fig. 5-b. One dimensional fuzzy partition of the band 3.

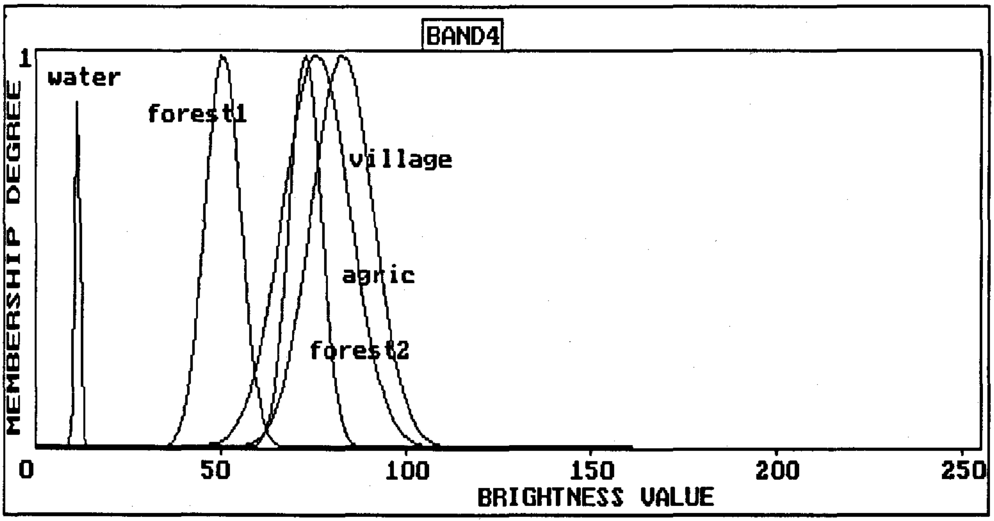

Fig. 5-c. One dimensional fuzzy partition of the band 4.

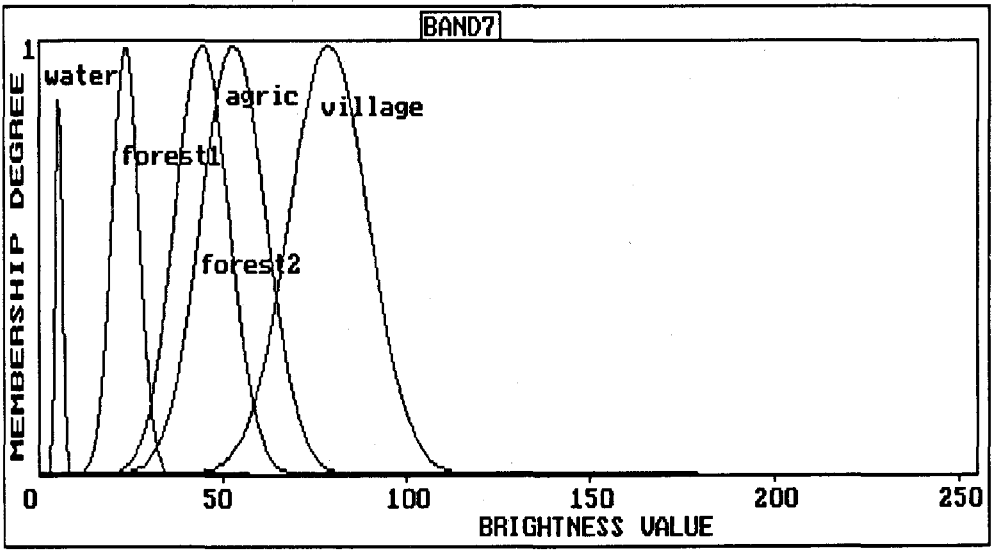

Fig. 5-d. One dimensional fuzzy partition of the band 7.

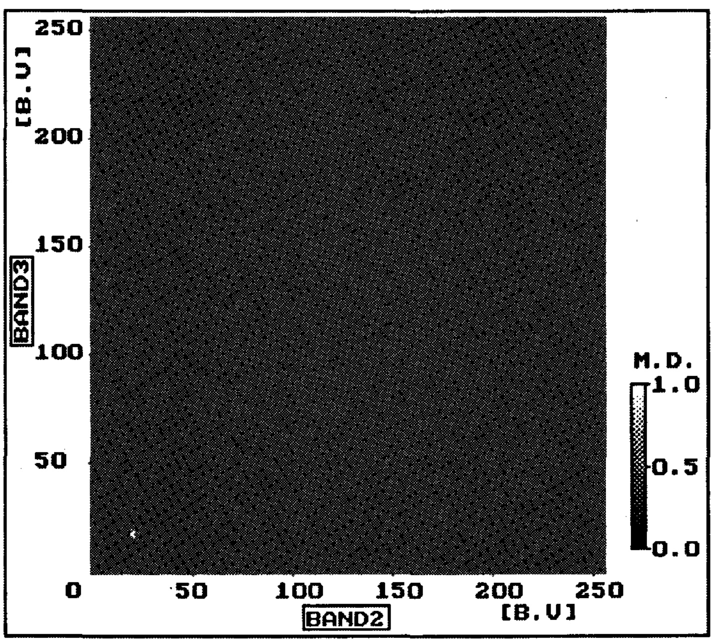

Fig. 6-a. Class "Water".

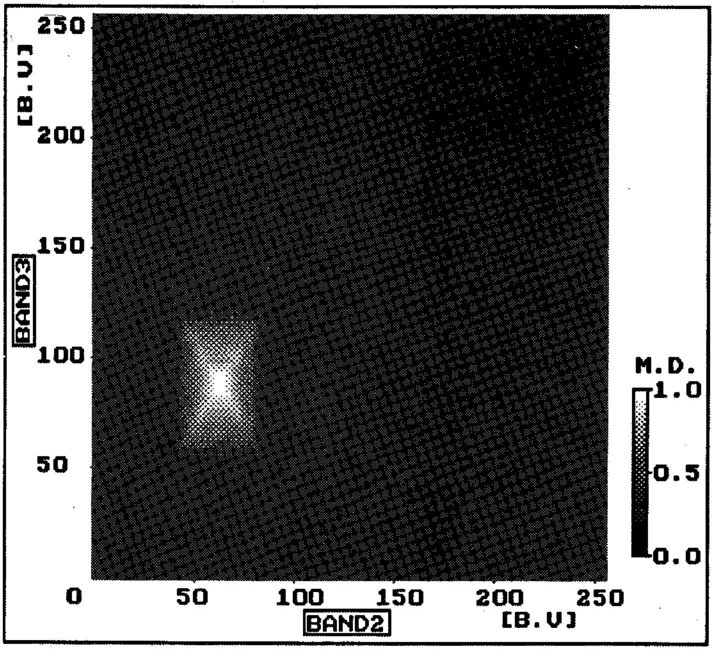

Fig. 6-b. Class "Village".

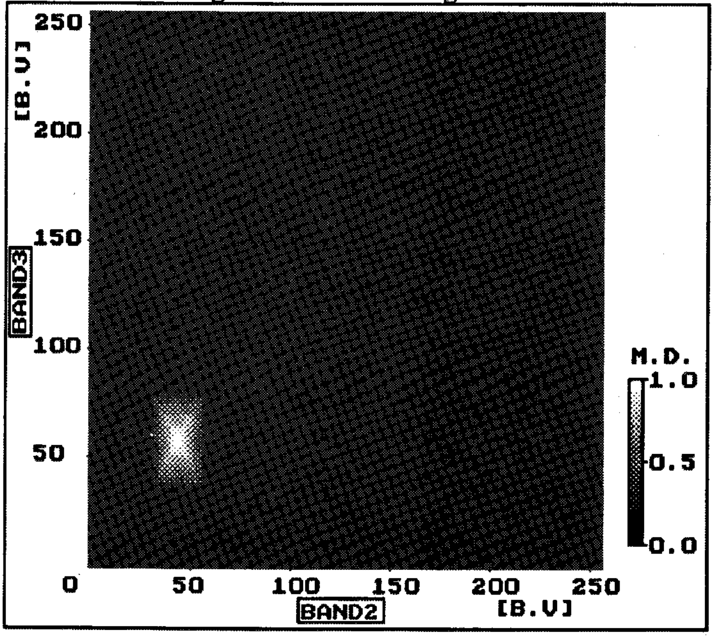

Fig. 6-c. Class "Agric.".

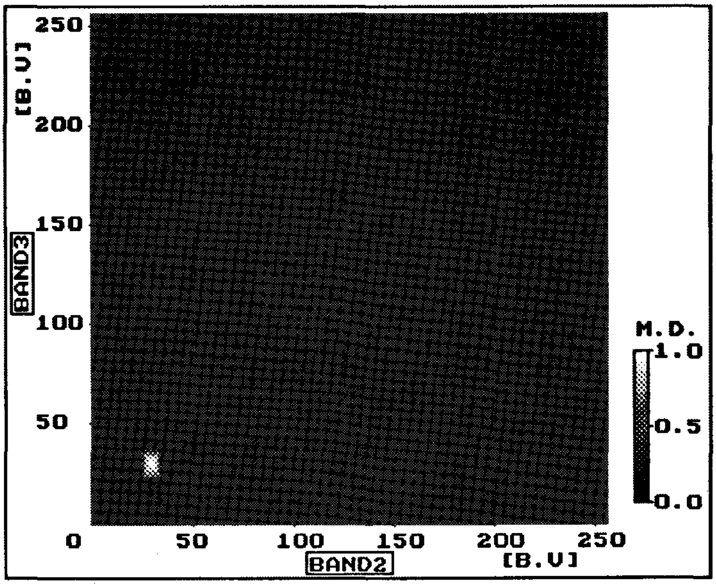

Fig. 6-d. Class "Forest 1".

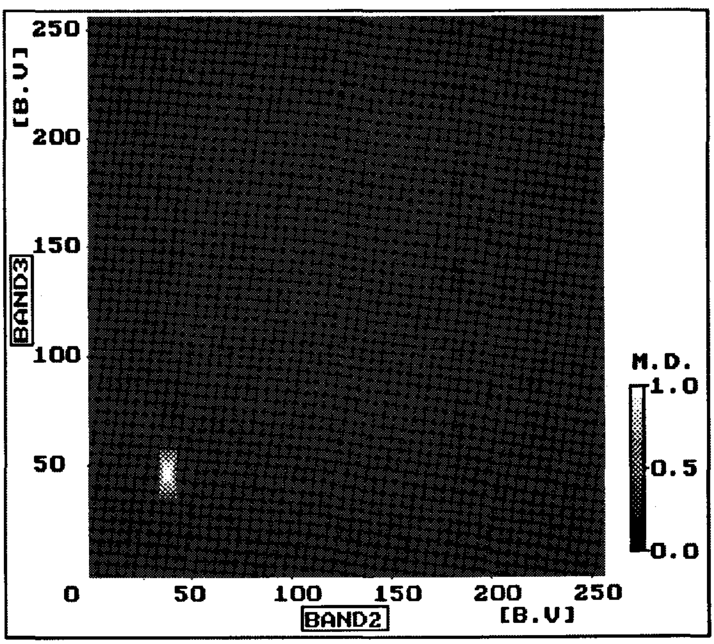

Fig. 6-e. Class "Forest 2".

Fig. 6. Two Dimensional Fuzzy Partition (Band 3 (B.V: Brightness Value) versus Band 2 (B.V) versus Membership Degree (M.D.)).

A fuzzy classification (fig. 7) of the whole scene has been carried out to represent the extent of each class throughout the study area. Figures 7.c and 7.e reveal a non negligible mixture between the Agricultural land and the Forest 2 classes over the image with a preference for the agricultural class (higher possibility degree). It was expected because of the biophysical nature of these classes which involve a possibility of mixture between them. However, the fuzzy classification shows a quasi hard distribution of the remaining classes (Water, Village and Forest 1) excepting some sporadic pixels.

Concerning the hard classification through which the comparison will be done, very good results are obtained for the training and test pixel classification by the Maximum Likelihood and Explicit Fuzzy methods (Tables 3 and 4) with an accuracy at the neighborhood of 90%. We note that the Average Accuracy is the mean of the accuracies obtained for each class, whereas the Overall Accuracy represents the ratio of the true pixel number to the total pixel number for the whole of the classes.

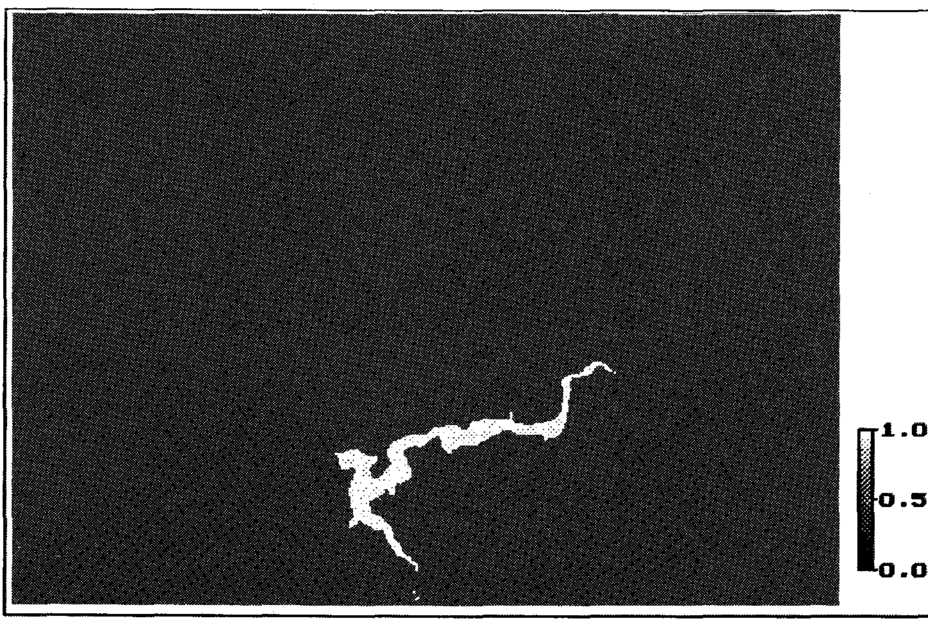

Fig. 7a. Fuzzy Classification Result (class "Village").

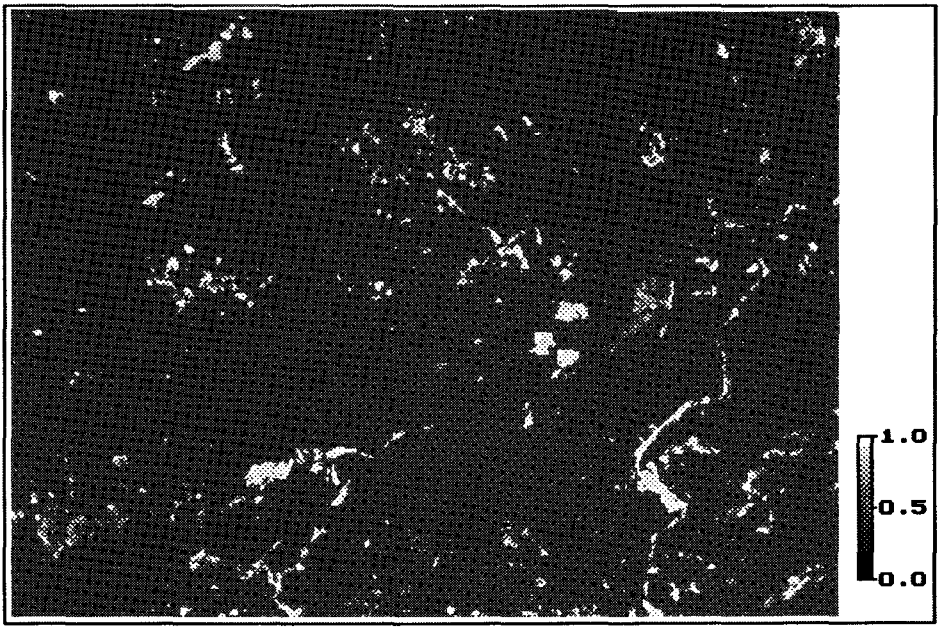

Fig. 7-b. Fuzzy Classification Result (class "Village").

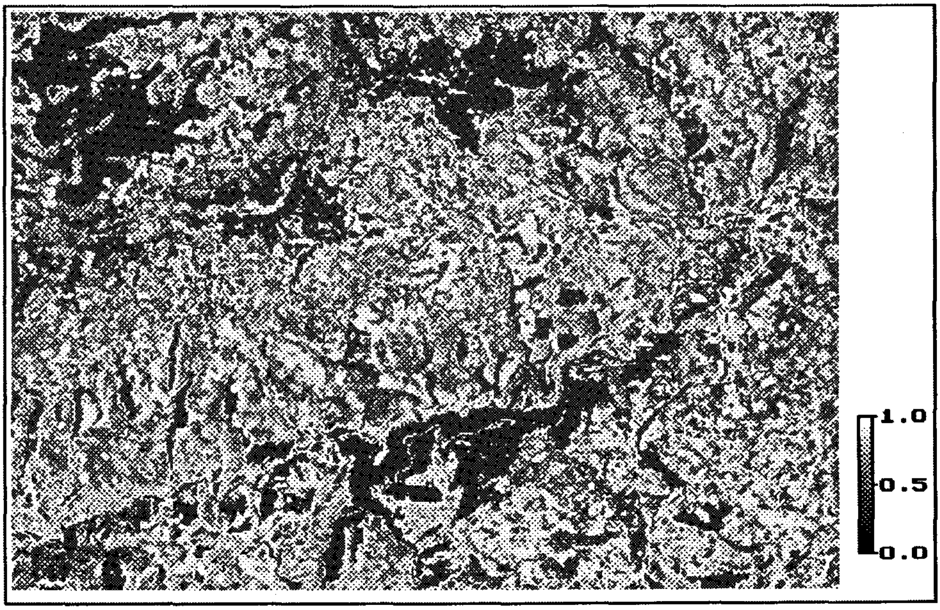

Fig. 7c. Fuzzy Classification Result (class "Agric.").

Fig. 7-d. Fuzzy Classification Result (class "Forest 1").

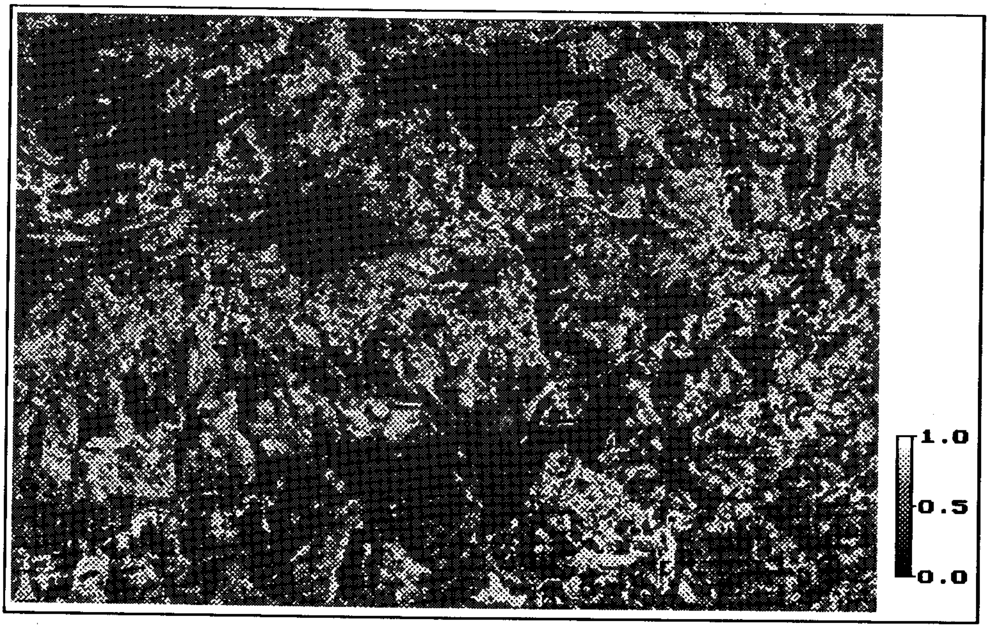

Fig. 7-e. Fuzzy Classification Result (class "Forest 2").

Table 3-a. Confusion Matrix of the Training Pixels for the Maximum Likelihood Method.

TRUE CLASS	CLASSIFIED AS					ACCURACY [%]
	WATER	VILLAGE	AGRIC.	FOREST1	FOREST2	
WATER	75	0	0	0	0	100.00
VILLAGE	0	164	7	0	3	94.25
AGRIC.	0	18	176	0	36	76.52
FOREST1	0	0	0	104	0	100.00
FOREST2	0	1	18	0	113	85.61
Average Accuracy [%]= 91.28			Overall Accuracy [%]= 88.39			

Table 3-b. Confusion Matrix of the Test Pixels for the Maximum Likelihood Method.

TRUE CLASS	CLASSIFIED AS					ACCURACY [%]
	WATER	VILLAGE	AGRIC.	FOREST1	FOREST2	
WATER	34	0	0	0	0	100.00
VILLAGE	0	69	18	0	0	79.31
AGRIC.	0	1	88	0	26	76.52
FOREST1	0	0	0	52	0	100.00
FOREST2	0	0	9	0	63	87.50
Average Accuracy [%]= 88.67			Overall Accuracy [%]= 85.00			

Table 4-a. Confusion Matrix of the Training Pixels for the Explicit Fuzzy Method.

TRUE CLASS	CLASSIFIED AS					ACCURACY [%]
	WATER	VILLAGE	AGRIC.	FOREST1	FOREST2	
WATER	75	0	0	0	0	100.00
VILLAGE	0	163	10	0	1	93.68
AGRIC.	0	20	194	0	16	84.35
FOREST1	0	0	3	101	0	97.12
FOREST2	0	0	39	0	93	70.45
Average Accuracy [%]= 89.12			Overall Accuracy [%]= 87.55			

Table 4-b. Confusion Matrix of the Test Pixels for the Explicit Fuzzy Method.

TRUE CLASS	CLASSIFIED AS					ACCURACY [%]
	WATER	VILLAGE	AGRIC.	FOREST1	FOREST2	
WATER	34	0	0	0	0	100.00
VILLAGE	0	74	13	0	0	85.06
AGRIC.	0	3	103	0	9	89.57
FOREST1	0	0	0	52	0	100.00
FOREST2	0	0	18	0	54	75.00
Average Accuracy [%]= 89.92			Overall Accuracy [%]= 88.06			

The Artificial Neural Network based method, whose characteristics are given in Table 5, shows weak results compared to the other methods proving the Gaussian methods have a behavior which increases positively with

1. The band number, and
2. The class signature homogeneity

Table 5a. Confusion Matrix of the Training Pixels for the Neural Network Method.

TRUE CLASS	CLASSIFIED AS					ACCURACY [%]
	WATER	VILLAGE	AGRIC.	FOREST1	FOREST2	
WATER	75	0	0	0	0	100.00
VILLAGE	0	160	7	0	7	91.95
AGRIC.	0	17	115	0	98	50.00
FOREST1	0	0	0	104	0	100.00
FOREST2	0	0	3	0	129	97.73
Average Accuracy [%]= 87.94			Overall Accuracy [%]= 81.54			

Table 5b. Confusion Matrix of the Test Pixels for the Neural Network Method.

TRUE CLASS	CLASSIFIED AS					ACCURACY [%]
	WATER	VILLAGE	AGRIC.	FOREST1	FOREST2	
WATER	34	0	0	0	0	100.00
VILLAGE	0	64	23	0	0	73.56
AGRIC.	0	2	48	0	65	41.74
FOREST1	0	0	0	52	0	100.00
FOREST2	0	0	0	0	72	100.00
Average Accuracy [%]= 83.06			Overall Accuracy [%]= 75.00			

The results of the training samples reveal that the Maximum Likelihood method has the advantage to produce an efficient adaptation to the training sites due to the use of second order statistics (covariance). If we consider that the mean accuracy is the mean of the Average and Overall accuracies for the training and test pixels, the Explicit Fuzzy method shows the best results with a mean accuracy of 88.66% compared to the Maximum Likelihood and the Neural Network methods with respectively 88.33% and 81.88%. The comparison of the classification accuracies between the training and test samples reveals a decrease of 3% for the Maximum Likelihood method and of nearly 6% for the Neural Network method. However, the Explicit Fuzzy approach provides an increase of 0.66% proving a better generalization ability because of the use of first order statistics making the method less dependent on the training samples. In a graphic language (Fig. 6-b, c and d), the hard classification put under light the fact that the main conflict in the class separation is between the Agricultural land and Forest 2 classes, as expected in the fuzzy classification, because the three classifiers provide different results concerning this conflict.

1. The Explicit Fuzzy method favors the Agricultural land class in accordance with the ground truth.
2. The Neural Net approach prefers the Forest 2 class.
3. The Maximum Likelihood classifier places itself in the middle.

We note the great precision of the Neural Network method for detecting the track leading to the village at the east part of the image and the small river at the west-south side of the scene.

Another important factor in the comparison is certainly the Total Classification Time of the 132252 pixels of the study area which labels the Explicit Fuzzy method as a very fast classifier. As shown in Table 6, this last needed a total classification time of 26.42 seconds nearly 52 times and 3 times shorter than respectively the Maximum Likelihood and the Neural Network methods. As we can see, the Explicit Fuzzy method provides an encouraging classification accuracy and a far shorter classification time, which is one of the major handicaps of the Maximum Likelihood classifier because of its necessary complex matrix computations.

Table 6. Total Classification Time.

CLASSIFIER	TOTAL TIME
Maximum Likelihood	1372.81 [s]
Explicit Fuzzy	26.42 [s]
Neural Network	78.21 [s]

5. CONCLUSIONS

We have proposed in this work a new method for Fuzzy Supervised Classification which reveals promising performance, and particularly, a fast classification time, which becomes a factor of great importance with the advent of new generations of sensors providing broad collections of data (for example, HIRIS with 192 spectral bands). The Maximum Likelihood classifier has the advantage to adapt itself very well to the training sites because of the use of second order statistics. However, the test sites make always its weakness in the generalization ability appear, which is not the case for the Explicit Fuzzy classifier proving the reliability of its classification results. This problem is more acute for the Neural Network approach because the more the network learns the more it passes from a generalized memory (classifier) state to a pure memory state. Certainly, it has the advantage to be "Distribution Free" but they do not guarantee to solve the problem of signature homogeneity because they depend strongly on the training conditions (hidden node number, weight initialization, convergence process behavior ...) and the learning time can become very long if the training pixel number is important. Furthermore, the Gaussian distribution model known for its efficiency, robustness and reliability provides better performances as the band number increases, a property which is not always true for the Artificial Neural Network classifier.

The strength of the proposed method is certainly its simplicity and its optimal extraction of the data through the explicit fuzzyfication, exploited efficiently by the MIN fuzzy reasoning rule. Another advantage of the Explicit Fuzzyfication is its modular architecture involving a high flexibility for easily inserting new bands or removing bands without disturbing the remaining parts of the classifier. Furthermore, the implementation of the Explicit Fuzzy method on Artificial Neural Networks would not pose any difficulty.

ACKNOWLEDGMENTS

We would like to thank gratefully the Royal Jordanian Geographic Center for providing the raw images.

REFERENCES

1. **P. D. Heermann and N. Khazenie, 1992,** " Classification of multispectral remote sensing data using a back-propagation neural network", IEEE Trans. Geosci. Remote Sensing, vol. 30, pp. 81-88.

2. **H. Bischof, W. Schneider, and A. J. Pinz, 1992** " Multispectral classification of Landsat-Images using neural networks", IEEE Trans. Geosci. Remote Sensing, vol. 30, pp. 482-490.

3. **T. Yoshida and S. Omatu, 1994,** " Neural network approach to land cover mapping", IEEE Trans. Geosci. Remote Sensing, vol. 32, pp. 1103-1109.

4. **S. B. Serpico and F. Roli, 1995** " Classification of multisensor remote-sensing images by structured neural networks", IEEE Trans. Geosci. Remote Sensing, vol. 33, pp. 562-578.

5. **J. A. Benediktsson, P. H. Swain, and O. K. Ersoy ,1990,** " Neural network approaches versus statistical methods in classification of multisource remote sensing data", IEEE Trans. Geosci. Remote Sensing, vol. 28, pp. 540-552.

6. **L. A. Zadeh,** " Fuzzy sets, 1965." Inform. Contr., vol. 8, pp. 338-353.

7. **B. Kosko, 1992**, Neural Networks and Fuzzy Systems. Prentice Hall.

8. **J. C. Bezdek, R. Ehrlich and W. FullL, 1984,** " FCM: The fuzzy c-means clustering algorithm", Computers and Geoscience, vol. 10, pp. 191-203.

9. **R.L. Cannon, J. V. Dave, J. C. Bezdek and M. M. Trivedi, 1986** " Segmentation of a Thematic Mapper Image using the fuzzy c-means clustering algorithm", IEEE Trans. Geosci. Remote Sensing, vol. 24, pp. 400-408.

10. **F. Wang, 1990** " Fuzzy supervised classification of remote sensing images", IEEE Trans. Geosci. Remote Sensing, vol. 28, pp. 194-201.

11. **G. M. Foody, N. A. Campbell, N. M. Trodd and T.F. Wood, 1992,** " Derivation and applications of probabilistic measures of class membership from the maximum likelihood classification", Photogram. Eng. and Remote Sensing, vol. 58, pp. 1335-1341.

12. **J. B. Adams, M. O. Smith** and **A. R. Gillespie, 1989,** "Simple models for complex natural surfaces : A strategy for the hyperspectral era of remote sensing", Proceedings of the IEEE International Geosciences and Remote Sensing Symposium, pp. 16-21.

13. **G. M. Foody** and **D. P. Cox, 1994** " Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions", International Journal of Remote Sensing, vol. 15, pp. 619-631.

14. **G. M. Foody,1996,** " Relating the land-cover composition of mixed pixels to artificial neural network classification output", Photogram. Eng. and Remote Sensing, vol. 62, pp. 491-499.

15. **T. A. Warner** and **M. Shank,1997,** " An evaluation of the potential for fuzzy classification of multispectral data using artificial neural networks", Photogram. Eng. and Remote Sensing, vol. 63, pp. 1285-1294.

16. **K. Hornik, M. Stichcombe** and **H. White, 1989,** " Multilayer feedforward networks are universal approximators", Neural Networks, vol. 2, pp. 359-366.

17. **J. D. Paola** and **R. A. Schowengerdt, 1995** " A detailed comparison of Backpropagation Neural Network and Maximum-Likelihood classifiers for urban land use classification", IEEE Trans. Geosci. Remote Sensing, vol. 33, pp. 981-996.