Fuzzy classification of Remote Sensing images
A pseudocolour representation of fuzzy partitions

Josep R. Casas *, Alain Hillion **, Christian Roux **, Luis Torres *, Antoni Gasull *

* Departament de Teoria del Senyal i Comunicacions
Universitat Politècnica de Catalunya, ETSET
Ap.30002, 08080 Barcelona, Spain

** Département de Mathématiques et Systèmes de Communication (GTI)
École Nationale Supérieure des Télécommunications de Bretagne
Technopôle de Brest-Iroise, BP 832, 29 285 Brest, France

ABSTRACT

In the field of Remote Sensing (RS) image classification, pattern indeterminacy due to inherent data variability is always present. Class mixture, too, is a serious handicap to conventional classifiers in order to settle proper class patterns. Fuzzy classification techniques improve the extraction of information yielded by conventional methods, i.e. statistical classification procedures, because both in the design of the classifier and when bringing out classification results, natural fuzziness present in real-world recognition processes is considered.

This paper presents first the application of a fuzzy classification algorithm from Kent and Mardia[1] to RS images, along with the analysis of the results and comparison against "hard" classifications. Secondly, we put forward one particular method to display these results (fuzzy partitions) by coding pixels' membership into a pseudocolour representation. This representation is intended to serve as an interface between fuzzy coefficients resulting from the classification process and a very natural way for humans to perceive information such as that of colour mixtures.

1. INTRODUCTION

Fuzzy set theory provides suitable tools to handle natural fuzziness present in decision processes when they are due to data variability rather than to randomness[2]. In classical systems, this kind of data variability is not considered by pattern recognition algorithms when the mathematical models are set up. Moreover, deciding on pixels' membership to one particular class ("hard classification") results in a loss of information always prone to classification errors that are difficult to recover from. This loss of information is smoothed by fuzzy partitions ("fuzzy classification"), where a group of membership grades -namely fuzzy coefficients- are attached to each pixel in order to indicate the extent to which the pixel belongs to certain classes[3]. Membership functions, or possibility distribution functions, provide additional information to overcome some limitations from usual classification systems and generalise the decision process. The classification through fuzzy coefficients may always be "hardened" in order to obtain a conventional "hard classification" result.

1.1 Fuzzy membership

In RS classification problems, the interpretation of fuzzy coefficients is twofold[1]. One can explain partial pixels' membership as if only a proportion of each pixel belongs to each class. Partial membership can also be seen as membership to mixture classes, obtained from mixture of classification prototypes. This mixing is particularly

This paper has been supported by the French-Spanish institutional grant n.172A

troublesome in some applications such as the mapping of soils or vegetation in sparsely vegetated arid areas[4], where fuzzy coefficients may be considered as straight estimates of class proportions in mixture pixels. Keeping in mind that the real on-the-earth surface of each pixel is 400 square meters, in the case of images from the French satellite SPOT, one can imagine either mixture of cover classes or surface proportions near to a limiting border between to different regions, in order to interpret partial membership.

1.2 Interface role

Fuzzy set theory may also play another role in pattern recognition, apart from that of the possibility-distribution interpretation of the concept of fuzziness. It serves as an interface between the linguistic variables seemingly preferred by humans and the quantitative characterisations appropriate for machines. The interface role[4] can be applied in two ways: first, to deal with input data -by defining a model that takes into account fuzziness of patterns- and second, to express in a reasonable format the results from fuzzy classification processes. Up to the authors' knowledge, fuzzy partitions are presented by a set of black and white images with the degrees of membership being coded in a grey level scale (figure 1). Section 4 presents a technique to code the degree of membership using a luminance scale but introducing, beside this, the coding of the membership to each class by a colour scale. This coding by colours benefits from the natural meaning of colour mixtures to human eyes. It allows to represent mixed classes accurately and includes parameters to threshold the level of fuzziness displayed in the picture representation by setting aside excessive information.

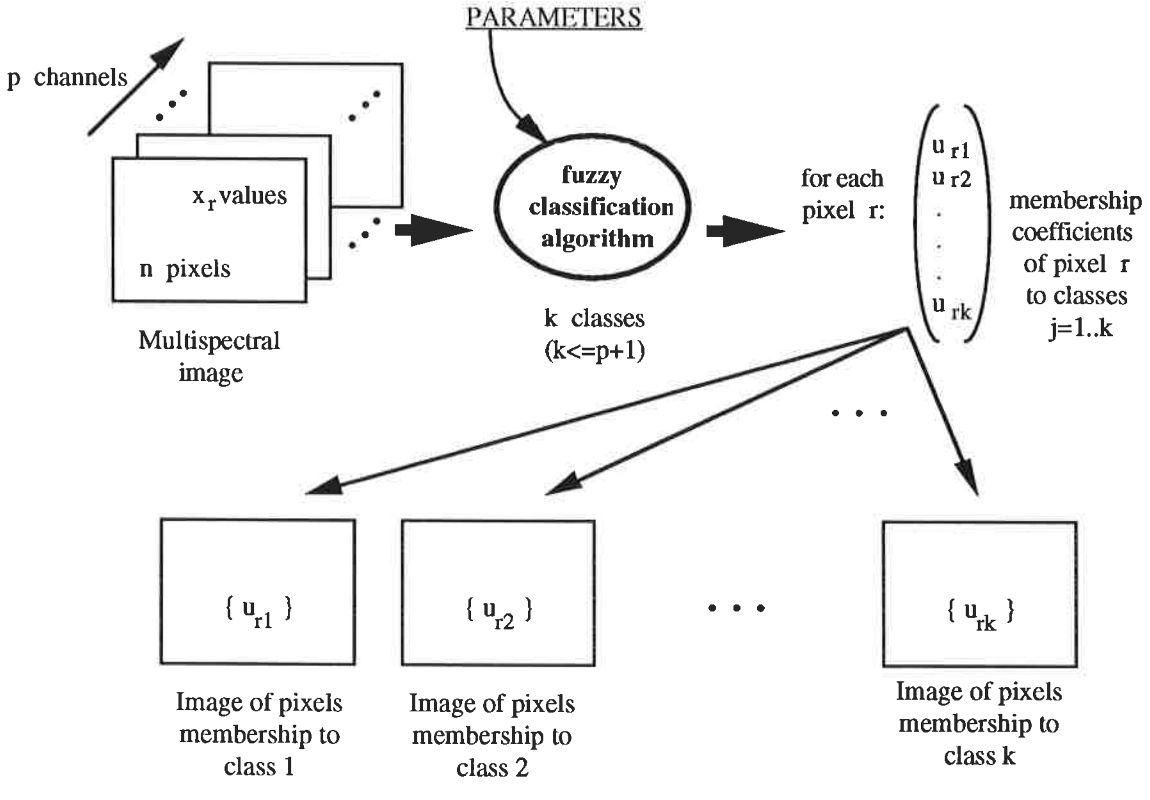


Figure 1. Fuzzy classification of a multispectral image (fuzzy partition).
Each image $\{ u_{rj} \}$ is the "fuzzy set" for class j .

2. THE FUZZY MEMBERSHIP MODEL

Kent and Mardia[1] have proposed a model for membership processes on spatial classification problems. The model is stated as follows:

Let $x_r \in R_p$ be a p-dimensional observation of a process X at sites $r \in D$, on the integer lattice Z^d. This process can be expressed in the form:

$$x_r \;=\; \Sigma_{j=1,k}\, u_{rj}\, \mu_j \;+\; e_r \;=\; \mu\, u_r \;+\; e_r \tag{1}$$

where $\{ u_r \}$ and $\{ e_r \}$ are two stochastic independent processes:

- μ is the matrix of prototypes, μ_j, for each class j, where j = 1, ... , k
- $U = \{ u_r \}$, the membership process, represents the fuzzy membership coefficients u_{rj}
- $E = \{ e_r \}$, the error process, is supposed to be Gaussian white noise

The target of the analysis is to estimate the fuzzy membership process U. At each point r, the coefficients u_{rj}, range in the interval [0 , 1], representing the proportion of each pixel r which belongs to class j or, also, the possibility or degree of membership of pixel r to class j . Thus,

$$\Sigma_{j=1,k}\, u_{rj} \;=\; 1, \quad \text{for each } r \in D \tag{2}$$

If the values of u_{rj} are restricted to "0" and "1", we have a conventional hard classification process, where u_{rj} is "1" for one class, j = n , and "0" for the rest, $j \neq n$, meaning that pixel r belongs to class n .

Restriction (2) allows to simplify the model, because the k coefficients are dependent. So, a new reduced space of differenced membership vectors is defined by mapping vectors u_r into vectors $v_r = (v_{r,1}, \dots v_{r,k-1})$ as follows:

$$v_{rj} \;=\; u_{rj} \;-\; u_{rk}, \quad \text{for each } j = 1, \dots, k-1 \tag{3}$$

The inverse mapping is a linear one:

$$u \;=\; A \cdot v \;+\; b \tag{4}$$

Two important restrictions more are forced on the model. First, if we want the choice of class k to be irrelevant in performing transformation (3) , i.e. the labelling of the classes to be exchangeable, it can be easily shown[1] that process $V = \{ v_r \}$ must have zero mean and a covariance matrix $\tau^2 G$ where τ^2 is a real positive number and $G = [g_{ij}]$ is the balanced equicorrelation matrix with $g_{ij} = 1/2$ if $i \neq j$ and $g_{ii} = 1$. The second restriction will arise later, in the fuzzy membership process estimation, and it forces the rank of matrix $\mu\ (p \times k)$ to be $k - 1$, and this implies that $k \leq p + 1$. So, the maximum number of classes to be identified will be as many as the components of the observed data increased in one.

The spatial distribution of vectors u_r (and so, v_r too) is assumed to be modelled by a Gaussian Markov random field or conditional autoregression (CAR)[5] satisfying

$$\begin{aligned} E\{ v_r \,/\, v_s ,\ s \neq r \} &= \Sigma_{s \in \{\beta_s\}} \left(\beta_{s-r}\, v_s \right) \\ E\{ v_r \cdot v_r^t \,/\, v_s ,\ s \neq r \} &= \Gamma \end{aligned} \tag{5}$$

where $\{ \beta_s \}$ is a set of weights for 4 or 8 pixels v_s in a finite symmetric neighbourhood of pixel v_r, and Γ is positive definite. Such a process is called QICAR in [1], for quasi-intrinsic Gaussian conditional autoregression.

To obtain a fuzzy classification, we estimate the process U given the data x_r using a Bayesian approach. From (1),

$$f(\{x_r\}/\{u_r\}) \quad \alpha \quad \exp\{ -1/2 \cdot \Sigma_{r \in D} \; [x_r - \mu u_r]^t \Sigma^{-1} [x_r - \mu u_r] \} \tag{6}$$

Posterior distribution of process V, given X, can also be given by a normal density:

$$f(\{v_r\}/\{x_r\}) \; \alpha \; f(\{x_r\}/\{v_r\}) \cdot f(\{v_r\}) \tag{7}$$

The maximisation of this posterior density, gives the following iterative solution to compute v_r, at each point $r \in D$:

$$v_r^{(\upsilon+1)} = [A^t \mu^t \Sigma^{-1} \mu A + \tau^{-2} G^{-1}]^{-1} \cdot [A^t \mu^t \Sigma^{-1} (x_r - \mu b) + \tau^{-2} G^{-1} \cdot \Sigma_{s \in \{\beta_s\}} (\beta_{s-r} v_s^{(\upsilon)})] \tag{8}$$

The parameters that should be given are:

- $\mu\,(p \times k)$ matrix of class prototypes
- $\Sigma\,(p \times p)$ covariance matrix of the error process E
- $\tau^2 > 0$, that weights the spatial variance of process V
- $\{ \beta_s \}$, the spatial neighbourhood considered (the simplest weighting is 1/4 or 1/8)

The initial estimate is given by setting $\tau^{-2} = 0$, to ignore the neighbourhood information.

3. APPLYING THE MODEL TO SPOT DATA

The model described in the previous section has been applied to the classification of multispectral images from the French satellite SPOT. The SPOT High Resolution Visible Imaging Instruments (HRV), based on charge coupled device detectors (CCD), provide images of the Earth surface for wavelengths into the following three spectral bands:

- 0,50 to 0,59 µm (green)
- 0,61 to 0,68 µm (red)
- 0,79 to 0,89 µm (near IR)

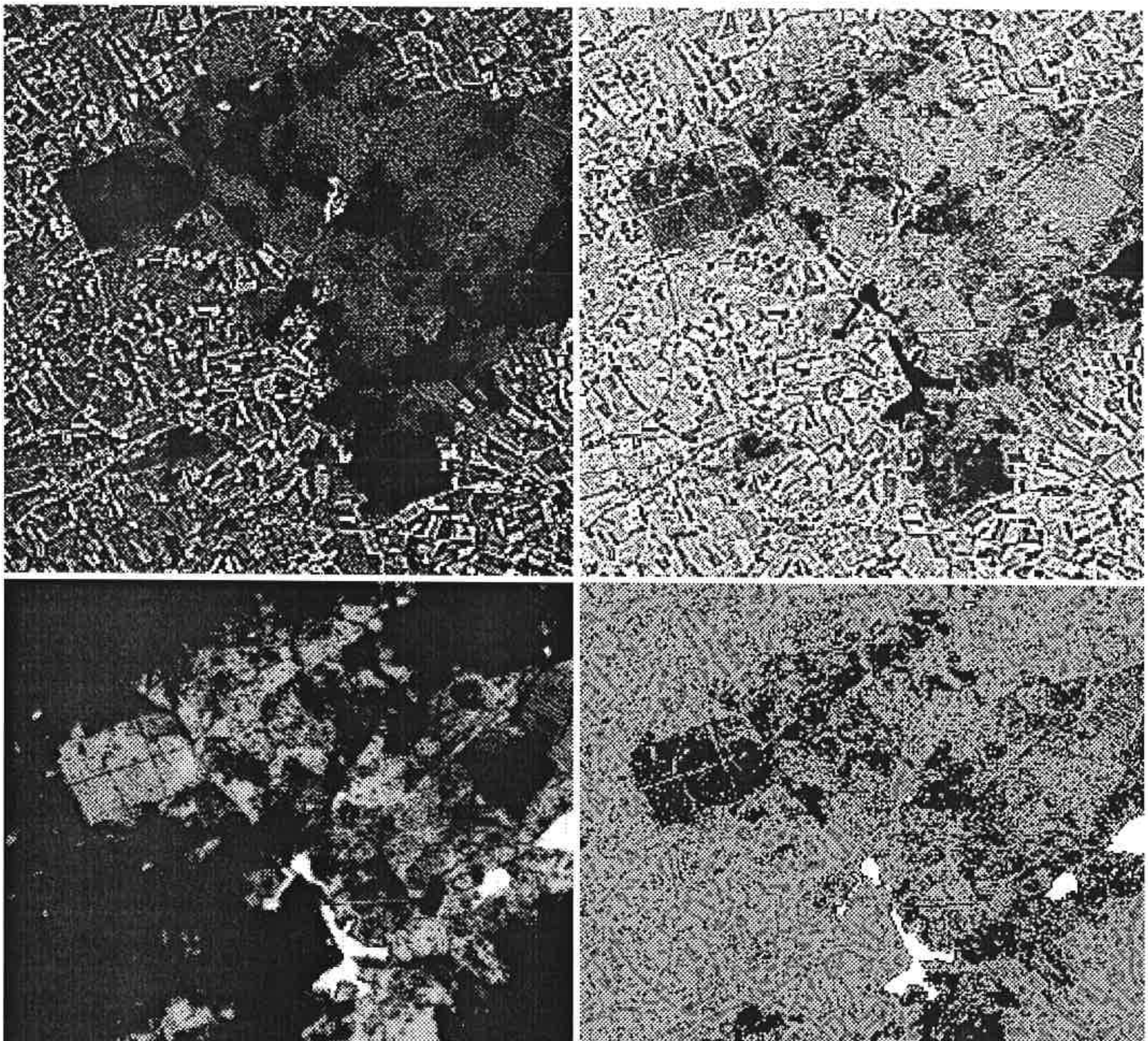
As noted previously, the restrictions imposed imply in particular that only four classes can be picked up from three dimensional observation data ($p = 3$). Thus, our aim was to separate some very general ground cover classes such as water, crops, urban or forests.

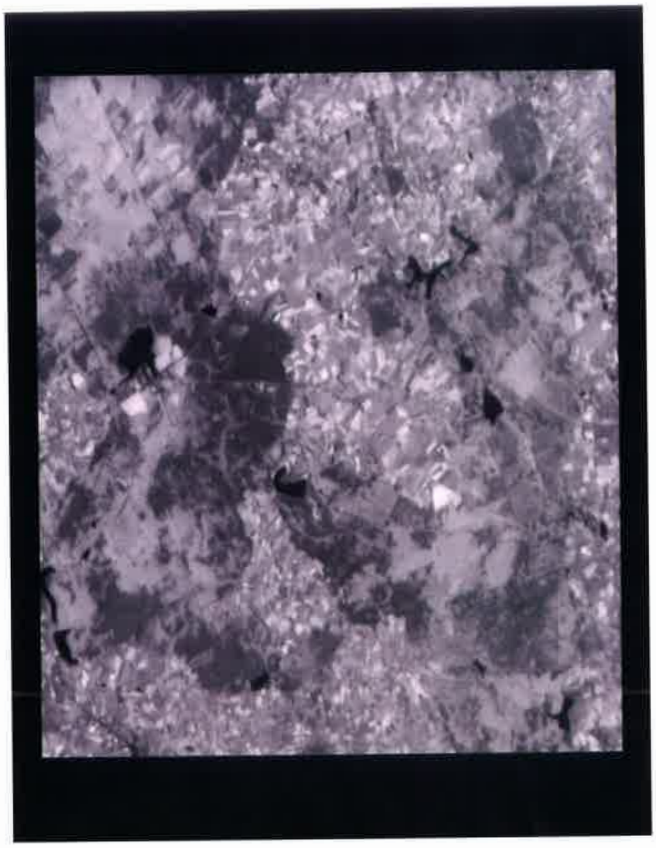
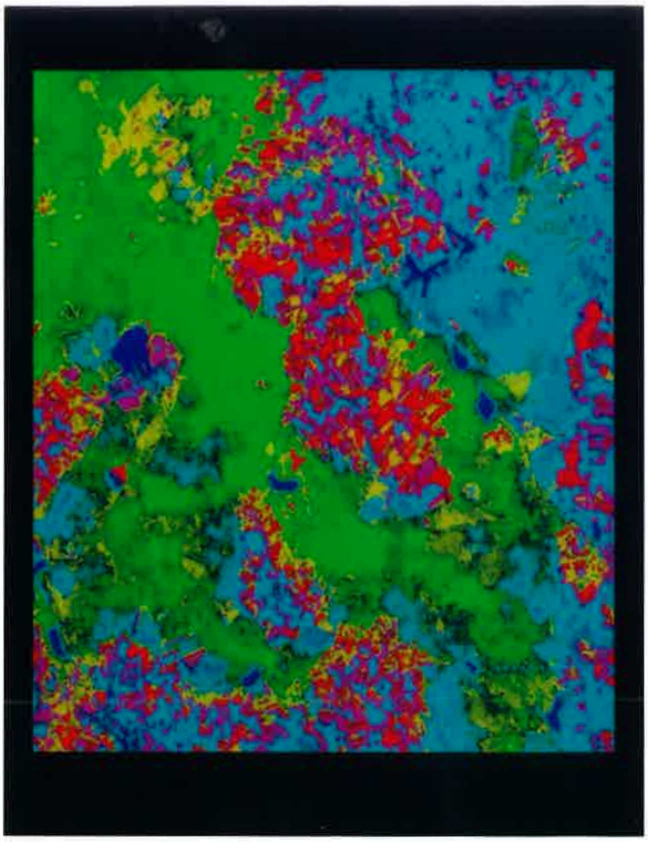
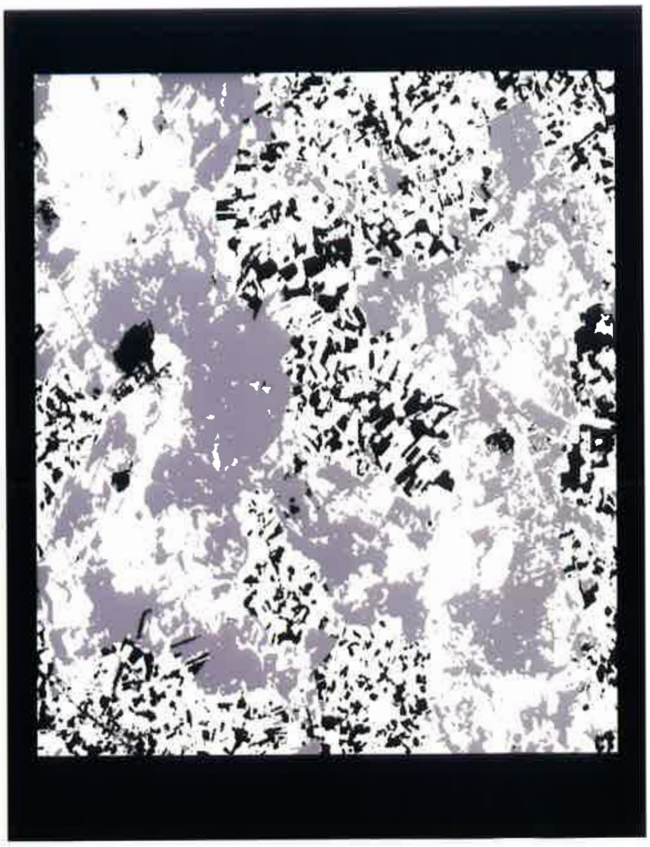
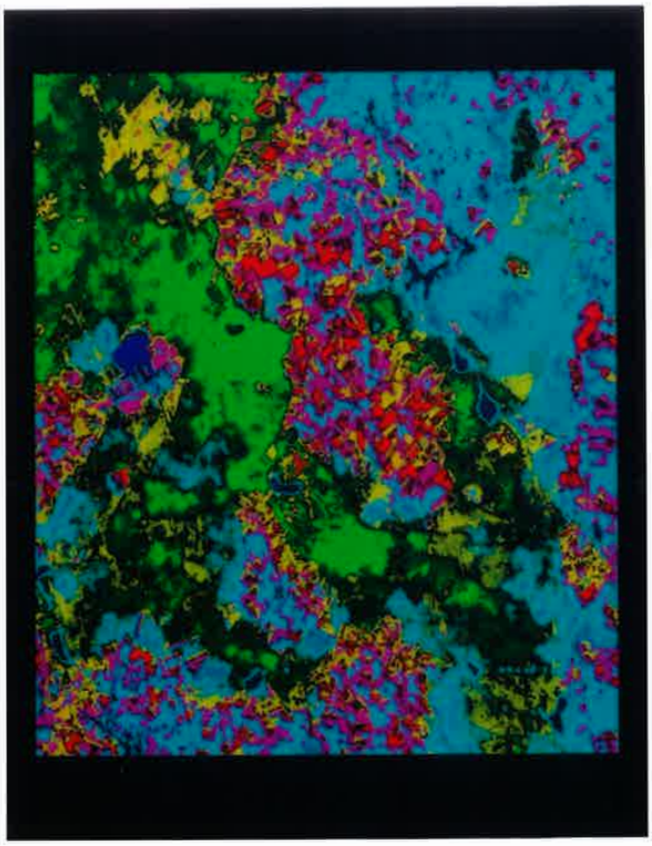
A sample of a fuzzy partition which has been obtained by this procedure is shown in figures 2-a, b and c. It can be seen, for instance, how some pixels on the upper left corner representing a road going through a wood are attached with the corresponding membership coefficients. The road is hardly one pixel wide (20 meters), so mixture "road" pixels are more likely than pure "road" pixels. Fuzzy classification algorithms perform better in such cases than conventional statistical classifiers, because these observations are not considered as noisy pixels but as points with few possibilities of membership to neighbouring dominant classes.

3.1 Key parameters

It has been proven that parameter τ^2 acts as a smoothing parameter[1]. As it decreases, the covariance values of the QICAR process are smaller, and more importance is given to the neighbouring information. Then, the resulting partitions tend to present uniform and large areas. Moreover, this parameter also affects the convergence of the algorithm -as can be deduced from the iterative solution applied (8)- and large values of τ^{-2} may lead the algorithm to divergence. For the particular case of the image shown in figure 1, we have obtained some good classification results with τ^2 about 0.05, and convergence problems have arisen with smaller values.

Figure 2. (a,b,c) Fuzzy partition of a SPOT multispectral image from an area near to Rennes, France. (d) Hard classification obtained by assigning the class with the largest fuzzy coefficient to each pixel.



Initial estimates of means and variances used in the algorithm, i.e. μ and Σ, hardly represent an approximation to reality. This is because they are computed with the assumption that they are data coming from a hard membership process with random variability, instead of 'fuzzy data' with different degrees of membership to different classes. Actually, covariance matrices obtained from this data show great disparity due to this assumption. Such a problem has taken us to the point of considering the mean and covariance values again, after the application of the algorithm, and computing new 'fuzzy' mean and covariance values in the form[3]:

- fuzzy mean: $$\bar{\mu}_j = (\Sigma_{r \in D} \, u_{rj} \cdot x_r) / \Sigma_{r \in D} \, u_{rj} \tag{9}$$

- fuzzy covariance: $$\bar{\Sigma}_j = (\Sigma_{r \in D} \, u_{rj} \, (x_r - \bar{\mu}_j) / \Sigma_{r \in D} \, u_{rj} \tag{10}$$

These estimates perform better if applied again to the iterative solution (8). And if computed again after a new fuzzy partition, then rapidly converge to target values.

3.2 'Hardening' the classification

Last picture in figure 2-d presents a conventional hard classification obtained by "hardening" the fuzzy partition from 2-a, b and c. Each pixel has been assigned to the class with the largest fuzzy membership value. Of course, this results in an important loss of information, but having considered the fuzzy nature of the observations during the classification procedure produces better results than simply applying conventional hard models, because the loosy step is then delayed as much as possible.

Fuzzy partitions can also be used as reliability measures of pixels' membership. In the hard version of the fuzzy process estimate in figure 2, for instance, we have considered pixels with membership values smaller than 0.6 as unclassified pixels (in black at 2-d).

4. PSEUDOCOLOUR REPRESENTATION OF FUZZY PARTITIONS

It would be desirable to find a representation of fuzzy data so that the extra information about class mixture could be displayed in only one picture. That should be an intermediate mode of displaying fuzzy partitions where the information loss would be smaller than if converted to hard classifications. Pseudocolour representations are often provided for displaying multispectral image classifications when the number of classes is high, but usually do not benefit from the information that colour mixtures are able to convey. Conversely, the colours are chosen to be the most distinguishable ones, in order to easily identify different cover classes.

4.1 Mapping fuzzy coefficients in a colour scale

We propose a colour representation of fuzzy partitions starting from the idea of false colour compositions (figure 3). Membership coefficients from different classes are taken as colour signals R, G and B, with appropriate mapping of the fuzzy interval, [0, 1], into colour component ranges. Such a colour coding should use different hues for different class types and luminance values proportional to the reliability of the classification of each pixel in one or two main classes. In this way, additional information provided by fuzzy partitions can be displayed as colour mixtures. Pixels with partial membership to several classes will melt into composite colours with smaller intensity values, while pure class type pixels will bright with more saturated colours. So, less reliable pixel assignments appear darker, and limiting frames between two classes fade from a pure class colour to the neighbouring pure class one.

4.2 Parameters to threshold the level of fuzziness

Some restrictions on mapping membership values must be provided in order to control the level of fuzziness displayed on the composite colour image. Such mapping, for a fuzzy coefficient at a single pixel, should not be independent of the rest of them at the same pixel. Thresholding the colour assignment at special values will ensure that certain situations will not happen. Some numerical examples can illustrate this point:

Fuzzy coefficients			*RGB values*			
u_{r1}	u_{r2}	u_{r3}	*R*	*G*	*B*	
1	*0*	*0*	*255*	*0*	*0*	*pure class (red)*
1 to .5	*< .3*	*< .3*	*255-90*	*0*	*0*	*dominant class (red)*
.5 to 0	*1 to .3*	*-*	*170-100*	*170-70*	*0*	*hybrid (red-green: yellow)*
0.5	*0.5*	*0*	*170*	*170*	*0*	*hybrid (light yellow)*
0 to .3	*0 to .3*	*0 to .3*	*0*	*0*	*0*	*small reliability (black)*

Obviously, if these thresholds are increased, we would reach a conventional hard classification, where pixels would be displayed in pure colours for each class. The threshold in the last sample line, would control the number of unclassified pixels.

We manage the mapping of the fuzzy coefficients at two levels. First, checking its relative values in order to assign a particular grey level index to each pixel. Second, several look up tables (LUT) specially designed can be used to display a pseudocoloured image. At first sight, this way to proceed seems to be a "false colour" technique, but we have called it "pseudocolour representation" because of the use a unique gray level scale and LUT files. The intermediate step of assigning suitable grey level values to each set of fuzzy coefficients, before applying the look up tables, is the key to the representation.

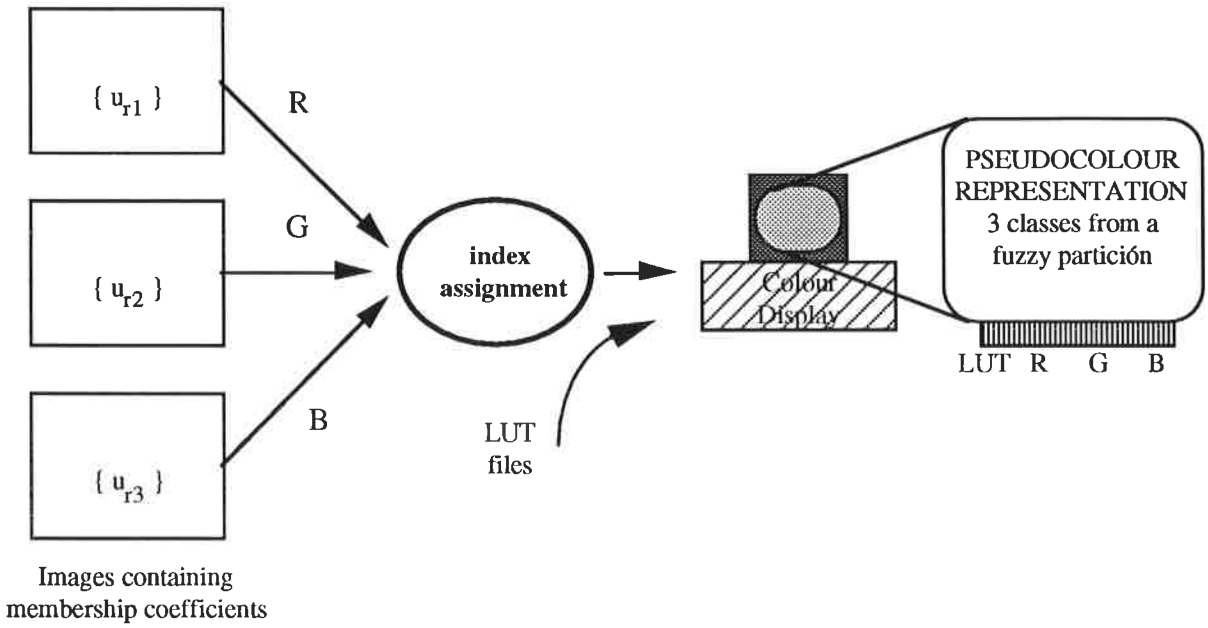


Figure 3. False colour representation technique for a fuzzy partition.

5. SUMMARY

In Remote Sensing image classification, innumerable factors may cause data varying within class clusters and overlapping between different clusters. Thus, analysis of spatial information, that is, spatial texture, has become necessary to improve the results of the simplest pixel-by-pixel Bayesian classifiers. But such techniques are not enough to perform accurate classifications. Class mixture may be encountered not only at the level of neighbouring pixels, but also within the spectral data values observed at every single pixel.

Better than considering multispectral data as hard membership samples of spatial processes contaminated with noise is to treat them as fuzzy membership processes. Fuzzy models for spatial classification are, in this framework, quite well suited to analyse RS data. We have proven a fuzzy algorithm to classify SPOT images, and the results confirm the former ideas.

We have also shown one technique to code the results of these "fuzzy classifiers". Instead of presenting several images with pixels' membership coded in a grey level scale, the classification is displayed in a colour image. Such technique allows to represent partial membership accurately if hue and luminance parameters of colour displays are managed in the right way.

6. REFERENCES

1. J.T. Kent and K.V. Mardia, *"Spatial Classification Using Fuzzy Membership Models".* IEEE Trans. on PAMI , Vol 10 No.5, pp.659-671 , Sept. 1988.
2. Yoh-Han Pao, *"Adaptive pattern recognition and neural networks".* Adison-Wesley, 1989.
3. F. Wang. *"Fuzzy Supervised Classification of Remote Sensing Images".* IEEE Trans. on GRS, Vol 28 No.2, pp.194-201 March 1990.
4 R.A. Showengerdt, *"Techniques for Image Processing and Classification in Remote Sensing".* Academic Press, Orlando, 1983.
5. J.W. Woods, *"Two-Dimensional Discrete Markovian Fields".* IEEE Trans. on IT , Vol 18 No.2, pp.232-240 , March 1972.
6. D.K.D. Majumder, S.K. Pal. *"Fuzzy Mathematical Approach to Pattern Recognition".* J. Wiley. New York, 1983.