THE PROBLEM OF MIXTURE
THE LINEAR MIXTURE MODEL VERSUS THE FUZZY MODEL

Vitor Haertel
Associate Professor
Jorge Silva Centeno, Nêstor A.Campana
Students
Federal University of Rio Grande do Sul,UFRGS,Brazil.

ABSTRACT

In natural scenes it often happens that more than one class is present in a pixel. The pixel spectral response in this case is not representative of either one of the component classes. One approach to solve this problem is the linear mixture model. More recently another model, implementing fuzzy mathematical concepts was introduced. Software compatible with the Brazilian SITIM-150 image processing system was developed in order to implement and compare both models. A test area located in the Emas National Park in Brazil is used in order to identify mixtures involving distinct proportions of vegetation.

KEY WORDS: Remote Sensing Application, Image Classification, Class Mixture.

1. INTRODUCTION

1.1 The Problem of Mixture

Current methods for image classification in Remote Sensing are applied on a pixel by pixel basis, implementing either probabilistic or deterministic algorithms (Tou & Gonzalez, 1974, Richards, 1986). Following these methods, image pixels are assigned to one of the existing information classes. Classification is based upon the radiance recorded by the satellite which is an integrated sum of the spectral responses of the materials within the instantaneous field of view (IFOV) of the sensor (Shimabukuro, 1987). In natural scenes, it often happens that pixels are constituted by more than one class, i.e.: are mixture pixels. In this case, the spectral response is not representative of any individual class, leading to incorrect classification results. Mixture pixels occur even in areas covered by homogeneous vegetation (eg.:agricultural fields).

Variations in canapy density across the area may cause varying degree of soil vegetation mixtures (Haertel, 1991) which, in turn, will cause variations in image spectral characteristics.

The current classification methods which assume that a pixel either belongs entirely to a class or does not belong to this class at all, are clearly not capable of dealing with the mixture problems. The only way possible would be to increase the number of information classes, to account for the mixture classes. This approach, however, would require an accurate knowledge of the mixture classes in the scene, which, in many cases, is not available and also lead to higher analysis costs.

The best way of dealing with the mixture problem in image classification consists in developing mathematical models capable of dealing with different proportions of informations classes in a single pixel.

1.2 Mathematical Models for the Mixture Problems

Two basic approaches were proposed to deal with the mixture problem: the Linear Mixture model and the Fuzzy Classification model.

The Linear Mixture model approach has been reported by several authors. A rather extensive review of this model is presented in Shimabukuro (1987) and is reviewed in section 2. The spectral contribution of each component class within a pixel is modelled in a linear relationship. The linear model, as applied to digital image classification,can be implemented in two different approaches. Given the component information classes ("pure" classes) one can estimate, for every pixel, its composition in terms of the component classes (Ranson,1975,Heimes,1977, Shimabukuro, 1987). A second approach consists in making use of the Linear Mixture model to estimate the mean vector and the covariance matrix of the "mixture classes" from the corresponding parameters associated with the component classes. This approach allows the user to "create" the mixture classes relevant to a particular situation,and is useful whenever the analyst is interested in some specific mixtures (eg.: vegetation and soil due to variations in vegetation density cover).

Image classification methods like the Gaussian Maximum Likelihood may then be applied to the entire scene.

A second mixture model, implementing Fuzzy mathematical concepts, was proposed by Wang (1990). The concept of class membership is introduced to account for multiple membership, i.e., for the mixture pixel. Any pixel may belong to more than one class. A membership function is introduced to estimate the degree to which a pixel belongs to each class. The membership degree is then associated to proportion of an information class in a pixel.

2. THE LINEAR MIXTURE MODEL

The linear mixture model assumes that in any spectral band 'i', the reflectance R_m associated with a mixture pixel can be equated to a linear function of the reflectances R_j associated with the component ("pure") classes. Each component class reflectance is weighted according to its proportion in a pixel:

$$R_{m,i} = \sum_{j=1}^{n} R_{j,i} \; X_j + \ell_i \qquad (1)$$

where:

$R_{m,i}$ = mean spectral reflectance associated with a mixture pixel for spectral band i.

$R_{j,i}$ = spectral reflectance associated with the component j for spectral band i.

X_j = proportion of component j in a pixel

j = 1,2,...,n (n=number of components).

i = 1,2,...,k (k=number of spectral bands)

ℓ_i = error term associated with spectral band 'i'.

To implement this model, the digital image must be converted from digital numbers available on CCTs into reflectances. This procedure is reported by some authors (eg.: Robinore, 1982; Markham and Barker, 1986).

The reflectances $R_{j,i}$ of the component classes can then be estimated for the training sets available.

Usually the number of spectral bands utilized is larger than the number of component classes (K >n). In this case, the system (1) becomes overdetermined and a least squares procedure is then applied. Then, the numerical values for X_j (j=1,..., n) should be such that they minimize the suns of the squares of errors ℓ_i:

$$\sum_{i=1}^{K} \ell_i = \text{minimum} \qquad (2)$$

Also, two additional condition equations should be added in order to allow for a phisically meaningful solution to the proportions X_j:

$$X_j \in [0,1] \;\forall\, j \qquad (3)$$

$$\sum_{j=1}^{n} X_j = 1$$

The Linear Mixture problem can thus be written as:

$$\ell_i = R_{m,i} - \sum_{j=1}^{n} R_{j,i} \,.\, X_j$$

Minimize $\sum_{i=1}^{K} \ell_i^2$ as a function of the proportions X_j.

Subject to:

$$0 < X_j \qquad \forall j \qquad (4)$$

$$\sum_{j=1}^{n} X_j = 1$$

Numerical methods to solve this constrained least squares problem are presented in Shimabukuro (1987).

The Linear Mixture model, applied to each pixel individually, estimates the proportions of every component in it.

Another approach (Haertel, 1991) consists in using equation (1) to estimate the mean vector and covariance matrix for given mixture proportions (X_j). Mixture classes can then be selected and the entire image classified into the existing "pure" classes and the selected "mixture" classes, using a maximum likelihood classifier. This approach proved to be very useful when the mixture classes of interest are previously defined.

3. THE FUZZY MATHEMATICAL MODEL

Mathematical fuzzy techniques for Remote Sensing image classification were proposed by Wang (1990a) and Wang (1990b), by implementing the concept of partial and multiple membership, instead of the one-pixel-one-class conventional methods. The multiple membership concept can be understood as a result of multiple classes in a pixel i,e., the class membership grades measure the proportion of the component ("pure") classes in a pixel.

Wang (1990a) comments on the information loss that occurs in image classification methods such as the Gaussian Maximum Likelihood. Pixel probabilities of belonging to each of the information classes are estimated. The pixel is then assigned to the class associated with the largest probability. All the remaining probabilities are discarded, regardless of their magnitude, i.e., the possibility that a pixel may partially belong to more than one class is excluded.

The fuzzy classification method attempts to make use of the information contained in these discarded probabilities. This attempt is implemented via the concept of multiple partial membership. For each class a membership function is defined. These functions segment the multispectral space into fuzzy partitions rather than "hard" ones as in conventional classification, which means that a point (pixel) may partially belong to more than one class.

Wang (1990a) defines the membership function associated to class 'i' by:

$$f_i(X) = \frac{P_i^{*}(X)}{\sum_{j=1}^{n} P_j\,(X)} \qquad (5)$$

where:

X : vector associated to a pixel in the multispectral space.

P_i^* : Gaussian probability density function for class 'i', implementing the fuzzy mean and the fuzzy covariance matrix

Note that:

$0 \leq f_i(X) \leq 1 \qquad \forall\, i, X$

$\sum_{i=1}^{n} f_i(X) = 1$

'n' being the number of information classes.

It should be recalled that, in the conventional approach, the largest $f_i(X)$ ($i = 1,\ldots,n$) would be set equal to one and the remaining (n-1) to zero. In the fuzzy approach the 'n' membership functions associated to a pixel are saved and interpreted as a contribution of each information class to the pixel spectral response. Thus, the pixel components X_j are estimated by the membership functions.

4. EXPERIMENTS AND RESULTS

The Linear Mixture model and the Fuzzy Mixture model have been applied to Landsat -TM data. The study area is the Emas National Park in Brazil. The park is located in central-western Brazil at approximately 18° south latitude(figure 1).

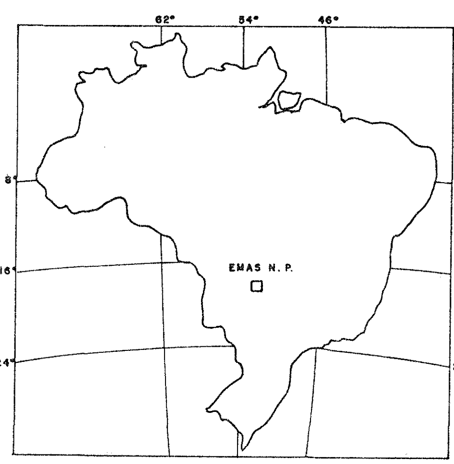

Fig.1 - Study Area Emas National Park

The climate is tropical. Winter is mild and corresponds to the dry season (June through September). Rains occur during the summer season (October through May).

The vegatation covering this region corresponds to the "cerrado". Vegetation cover density varies across the Emas N.P.. Areas with a higher degree of soil moisture present a higher vegetation density cover. Areas presenting lower soil moisture contents show more sparsely distributed vegetation. During the dry season vegetation covering the drier area becomes more susceptible to fires. The fire that occurred in the park in 1988 was confined to this area.

Mixture pixels are clearly present. The "pure" information classes in this case are vegetation and soil. "Pure" vegetation pixels can be found in areas with higher soil moisture contents (along the drainage lines). Most of the park is then covered by mixture pixels. The different proportions of vegetation and soil can then be used to estimate conditions across the park with respect to general environmental conditions.

The classification process was performed considering five classes: two "pure" classes (vegetation and soil) and three mixture classes (0.75 vegetation and 0.25 soil, 0.5 vegetation and 0.5 soil, 0.25 vegetation and 0.75 soil). Computer programs were developed in order to implement the Linear Mixture model and the Fuzzy model for the Brazilian image processing system SITIM-150.

The Linear Mixture model was applied, to estimate the mean vector and covariance matrix of the mixture classes. The Gaussian maximum likelihood classifier was then applied to the scene. Results are presented in figure 2.

The Fuzzy model was also applied to the same scene. Equation 5 was used to estimate the membership grade of each pixel with respect to two information classes: vegetation and soil. A level slice criterion was then applied in order to define borders between the three mixture classes. The results are presented in figure 3.

The results of both methods are rather similar and in agreement with field data.

The Linear Mixture model,however,presented a more accurate classification result. This method applies a maximum likelihood classifier which is capable of better estimating borders between the mixture classes. The level slice method used to define borders in the Fuzzy model clearly presents a poorer performance.

5. CONCLUSION

The implementation of the mixture concept to natural scenes can be a valuable tool in natural resources. Variations in vegetation cover can be detected and quantified. Useful information can be obtained from this process. In the test area presented in this study, only the area with sparser vegetation, i.e., the area corresponding to dry vegetation was affected by a subsequent fire. This data can then be understood as a variation in soil and/or climate conditions.

REFERENCES

Haertel, V., Centeno, J. 1991. Utilização do conceito de pixel mistura no processo de classificação da cobertura vegetal em bacias hidrográficas. Revista Brasileira de Engenharia - Caderno de Recursos Hídricos. 9(2):87-100.

Heimes, J.F. 1977. Effects of scene proportions on spectral reflectance in lodgepole pine. M.Sc.thesis, Colorado State University, Fort Collins, Colorado.

Ranson, J.K. 1975. Computer assisted

classification of mixtures with simulated spectral signatures. M.Sc.thesis,Colorado State University, Fort Collins, Colorado.

Richards, J.A. 1986. Remote Sensing Digital Image Analysis: an Introduction. Springer-Verlag.

Shimabukuro, Y.E. 1987. Shade images derived from linear maxing models of multi -spectral measurements of forested areas. Ph.D.dissertation, Colorado State University, Fort Collins, Colorado.

Tou, J.T., Gonzalez, R.C. 1974. Pattern Recognition Principles. Addison-Wesley, Reading, Massachusetts.

Wang, F. 1990. Improving remote sensing image analysis through fyzzy information representation. Photogrammetric Engineering and Remote Sensing 56(8):1163-1169.

Wang, F. 1990. Fuzzy supervised classification of remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 28(2): 194-201.

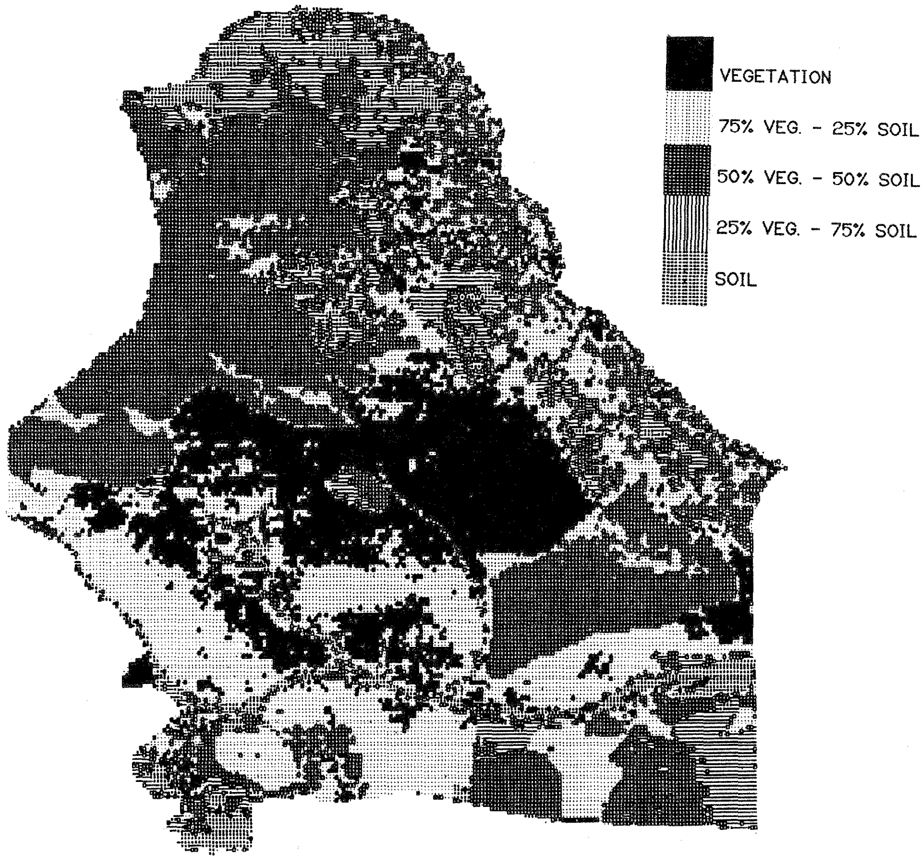

Fig.2 - linear Mixture model classification

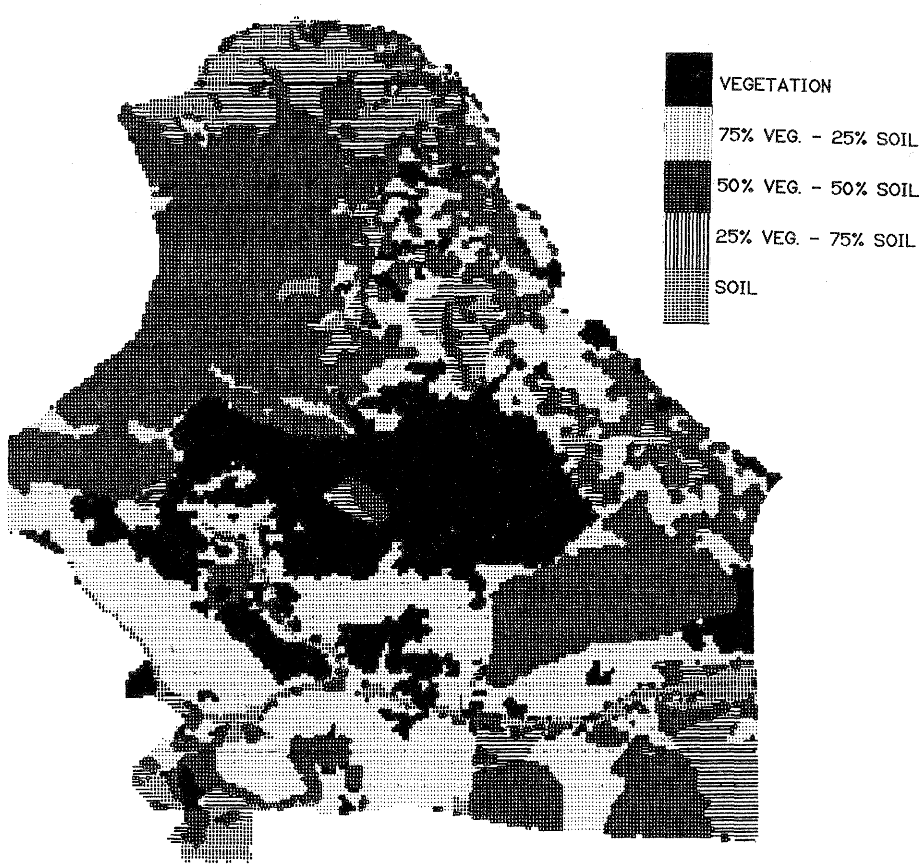

Fig.3 - Fuzzy model classification