Improving Remote Sensing Image Analysis through Fuzzy Information Representation

*Fangju Wang**
Department of Computer Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

ABSTRACT: This research reveals that an important factor which degrades current remote sensing image analysis lies in the loss of spectral information during the process of image classification, and the information loss is caused by the expressive inadequacy of the traditional method for representing geographical information.

This paper examines where and how the spectral information is lost in a conventional supervised classification process, proposes an alternative method to the traditional representation – the fuzzy representation which may reduce the information loss – describes a new classification algorithm which is based on the fuzzy representation, and presents the major achievements including higher overall classification accuracy and identification of types and proportions of component cover classes in mixed pixels.

INTRODUCTION

IN CLASSIFYING AND LABELING remote sensing images, information is currently represented in a one-pixel-one-class method. In practical situations, a pixel may represent the measurement of cover class mixture, intermediate conditions, or other complex cover patterns which cannot be properly described by a single class. The expressive limitation of this method has inevitably caused information loss which has in turn led to unsatisfactory image classification accuracy and poor extraction of information.

The current representation method is based on classical set theory which is ideally suitable for objects which can be precisely described by full membership in a set. Many geographical phenomena, however, cannot be accurately described by such membership. Impreciseness thus arises when the "one membership" method is used to represent geographical information. To improve remote sensing image analysis, an alternative representation method is required which allow for partial and multiple membership.

In this research, information loss in the image classification process is explored. A fuzzy remote sensing image classification technique has been developed. Differing from the previous work in applying fuzzy set theory to remote sensing image analysis (Cannon *et al.*, 1986; Jeansoulin *et al.*, 1981; Zenzo, 1987; Kent and Mardia, 1988) which employed the fuzzy methods in limited processing phases, this technique uses fuzzy sets for information representation throughout the entire process of image classification. Information loss can be largely reduced through this representation methods. Encouraging results have been achieved in experiments.

IMPRECISE REPRESENTATION METHOD FOR GEOGRAPHICAL INFORMATION

Geographical information is conventionally represented in *thematic maps*. A thematic map is a specific-purpose map which contains information about a single subject of *theme*. A map is a set of points, lines, and areas that are defined both by their location in space with reference to a coordinate system and by their non-spatial attributes about a theme (Burrough, 1986).

Currently, the linkages between the spatial entities and their attributes are implicitly based on the membership concept of classical set theory. According to the theory, a set has a precisely defined boundary and an element is either inside the set or outside it. When the attribute values are used to classify the spatial entities into classes (i.e., each class is characterized by an attribute value), an entity either completely belongs to a class or not at all. For example, in a land-cover map, a region can be described by only one cover type exclusively, such as grass *or* bare soil. Such a representation method has difficulties to deal with phenomena which cannot be described by membership in a single set. In the following, land-cover class mixture and within-class variability are used to reveal the expressive inadequacy of the current representation methods.

A region in a map corresponds to an area on the ground. Quite often, such a ground area contains a mixture of surface cover classes, for example, grass and underlying soil. In the case of a raster map (including remote sensing-derived maps), each pixel corresponds to a square cell on the ground. Mixture may also take place when the size of the cell is larger than the size of the features about which information is desired. Because currently only one cover class (usually the major one) can be assigned to a region (or a pixel), information about the other component classes and deviation of the assignment cannot be represented.

Different conditions may exist within a cover class in practical situations. For example, vegetation may be in different conditions which are caused by such factors as plant health, age, water content, and soil mineral content. However, these conditions cannot be differentiated in a thematic map unless more classes are introduced. It is clearly inaccurate to assign the same class to fresh grass and half dry grass without specifying their differences.

In the above situations, introducing more classes will lead to higher analysis costs and, no matter how finely the classes are defined, class mixture and within-class variability may exist. Using the existing method to represent geographical information inevitably leads to expressive limitations. In remote sensing image analysis, the expressive limitations lead to loss of valuable spectral information.

EXPRESSIVE LIMITATIONS AND INFORMATION LOSS

Probably because a primary objective of many remote sensing applications is the classification of images to generate thematic maps and the ground truth information required for classifier training is derived from thematic maps (which may not exist physically), information representation in remote sensing image analysis basically follows the conventional method. Each pixel can only be associated with one cover class. Such a method cannot properly represent class mixture and intermediate conditions which occur in most remote sensing images.

*Department of Computing and Information Science, University of Guelph, Guelph, Ontario N1G 2W1, Canada

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 56, No. 8, August 1990, pp. 1163–1169

0099-1112/90/5608–1163$02.25/0

Pixel Spectral Characteristics

A ground cover class has spectral characteristics which depend upon the interaction mechanism between electromagnetic radiation and the material of that cover class. In a given radiometric band, a pixel value of a remote sensing image is a measure of the spectral characteristics of the corresponding ground cell. When the cell contains a single cover class, the pixel records spectral characteristics of that class, and when the cell contains a mixture of cover classes, the pixel value is a function of the reflectance from a mixture of component classes (Fung and Ulaby, 1983; Jensen, 1986; Thomas *et al.*, 1987). As the mixture proportions change from pixel to pixel, the reflectance value changes. Usually, the more of a cover class a pixel contains, the more spectral characteristics of that class it has. Thus, a mixed or heterogeneous pixel has its spectral characteristics which differ from those of a homogeneous pixel. Similarly, changes in conditions within a given cover class also cause variation in spectral characteristics.

Pixel measurement vectors are often considered as points in a spectral space. Pixels with similar spectral characteristics form groups which correspond to various ground-cover classes that the analyst defines. The groups of pixels are referred to as *spectral classes* while the cover classes are referred to as *information classes*. To classify pixels into groups, the spectral space is partitioned into regions, each of which corresponds to one of the information classes defined. As discussed before, traditionally the information classes are implicitly represented as classical sets. Thus, a partition of spectral space is based on the same principles. Decision surfaces are defined precisely by some decision rules to separate the regions. Pixels inside a region are associated with the corresponding information class. Such a partition is usually called a *hard* partition.

If X is a finite set $\{x_1, x_2, \ldots, x_n\}$, a hard k-partition of X is a family $\{A_i : 1 \le i \le k\} \subset P(X)$ which satisfies

$$\begin{aligned} &\cup_{i=1}^{c} A_i = X \\ &A_i \cap A_j = \phi,\ 1 \le i \ne j \le c \\ &\phi \subset A_i \subset X,\ \ 1 \le i \le c \end{aligned} \tag{1}$$

where $P(X)$ is the power set of X.

Formally, a hard k-partition space for X can be represented as

$$\begin{aligned} M_k = \{&\mathbf{U} \in \mathbf{V}_{kn} \mid u_{ij} \in \{0,1\}\ \ \forall i,j; \\ &\sum_{i=1}^{k} u_{ij} = 1\ \forall j;\ 0 < \sum_{j=1}^{n} u_{ij} < n\ \forall i\} \end{aligned} \tag{2}$$

where $\mathbf{V}_{kn}$ is a set of real $k \times n$ matrices, and $\mathbf{U} = [u_{ij}]$ is referred to as a *hard partition matrix*.

A serious drawback of a hard partition of spectral space is that a great quantity of spectral information is lost in determining and using pixel membership. In the following, the maximum likelihood classifier is taken as an example to illustrate this point.

Information Loss in Spectral Space Partition

The *maximum likelihood classifier* (MLC) is one of the most widely used methods for supervised classification. The decision rule applied to each unknown measurement vector, $\mathbf{x}$, is (Anuta, 1977): Decide $\mathbf{x}$ is in class c if and only if

$$P_c(\mathbf{X}) \ge P_i(\mathbf{X}) \tag{3}$$

where $i = 1,2,3,\ldots, m$ possible classes, and if the classes are Gaussianly distributed, then,

$$P_i(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\Sigma_i|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i)^T \Sigma_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i)} \tag{4}$$

where $\boldsymbol{\mu}_i$ is mean vector of class i, Σ_i is covariance matrix of class i, and N is dimension of the measurement vector $\mathbf{x}$. $P_i(\mathbf{x})$ is the conditional probability of the measurement vector $\mathbf{x}$ given that the class i is known.

By taking logarithm and assuming that the covariance matrices for all classes are equal,the decision rule is simplified to assigning the measurement vector $\mathbf{x}$ to class c, minimizing the square of *Mahalanobis distance*

$$(\mathbf{x} - \boldsymbol{\mu}_c)^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}_c) \tag{5}$$

between $\mathbf{x}$ and $\boldsymbol{\mu}_c$. Equation 5 precisely defines surfaces for a hard partiton.

In MLC, Mahalanobis distance measures the distance of a pixel vector from the mean position of a class which represents the spectral characteristics of that class. The shorter the Mahalanobis distance, the more similar the pixel is to a predefined class in terms of the spectral characteristics. The decison rule can thus be explained as classifying the unkown measurement vector $\mathbf{x}$ to class c whose mean is closest to $\mathbf{x}$ in the spectral space.

For each pixel, more information could be derived by retaining the Mahalanobis distances for all classes instead of picking the maximally likely single class. First of all, the distances can provide information for differentiating between mixed pixels and homogeneous pixels: the vector of a homogeneous pixel must be very close to a class mean while the vector of a mixed pixel must be far from any class mean. Second, by comparing the distances, it is possible to identify cover-class components of a mixed pixel and estimate their propotions. Roughly, the more a pixel contains a cover class, the more similar is the pixel's measurement to the mean of that class in terms of spectral characteristics, and thus the smaller the distance to its mean in the spectral space. Third, the fact that a pixel measurement vector is half way between means of two classes in a spectral space may imply an intermediate condition. Furthermore, when a pixel is assigned a class in the conventional classification, the distances could be used to estimate the accuracy of the assignment. If the distance of the pixel to the mean of the class assigned is very small compared with the distances to the means of the other classes, the deviation of the assignment might be small with respect to spectral characteristics recorded. Although the above information is not always directly required by users, it might be significant for certain digital anaylses, especially for knowledge-based analysis.

However, the information contained in the Mahalanobis distance is not fully exploited in conventional classifiers. Let us examine how the information is lost in determining the pixel membership by using the simplified maximum likelihood classification. For each pixel, its Mahalanobis distances to all the class means are calculated and then compared. The closest class is assigned to the pixel. *Once membership of the pixel is decided, the distances are immediately discarded.* The assignment implies full membership in the class and no membership in other classes. The possibility that a pixel may partially belong to a class and simultaneously belong to more than one class is excluded. Final output of the classification is represented in a *one-pixel-one-class* image. A great deal of valuable information which has been derived when calculating the Mahalanobis distances has to be discarded because it is difficult to be represented in such a framework of classical set theory. This is an important reason for the current poor extraction of spectral information.

Information Loss in Classifier Training

In Gaussian cases, class means $\boldsymbol{\mu}$s and covariance matrixes Σs play a critical part in MLC decision-making. A mean and covariance matrix describe a classes: the mean is the representative element and the determinant of the covariance matrix describes the volume of the equidistance contour. In determining membership of a pixel measurement vector, the classifier decides in favor of the class with shorter distance to the pixel (i.e., the more likely class), and the class with smaller $|\Sigma|$ (i.e., the denser class). How accurate the classification is largely depends upon how accurate the estimates are.

In classifying remote sensing images, an important factor which reduces the estimate accuracy has not been paid deserved attention: the current methods for estimating the parameters cannot deal with class mixture and intermediate conditions properly.

μs and Σs are estimated from a set of sample pixels in classifier training. The following are conventional definitions of sample mean and sample covariance matrix for class c:

$$\mu_c = \frac{\sum_{i=1}^{n} \mathbf{x}_i}{n}, \tag{6}$$

$$\Sigma_c = \frac{\sum_{i=1}^{n} (\mathbf{x}_i - \mu_c)(\mathbf{x}_i - \mu_c)^T}{\mathrm{n}} \tag{7}$$

where n is the number of sample pixel vectors and $\mathbf{x}_i$ is a sample pixel vector ($1 \leq i \leq n$) which is known *a priori* in cover class c. Currently, the training data, i.e., the sample pixels, are represented in the conventional *one-pixel-one-class* method. Equations 6 and 7 imply that, for a sample pixel, only one value in one band in one training class can be used to yield the statistical parameters. No matter how much a pixel actually belongs to the class, it makes a full contribution. The spectral information from other possible cover types are all contributors to parameters of the class assigned to it. The existence of other cover classes is ignored. Although training sites are usually selected in relatively homogeneous areas, class mixture or intermediate conditions is inevitable. The statistical parameters estimated in this way must vary more or less from the "real" ones. A classification based on those parameters may thus contain mistakes.

Other existing classifiers, for example *Bayes classifier*, the *minimum distance classifier* for supervised classification, *hierarchical classifier*, and *K-means classifier* for unsupervised classification, are all based on the same information representation method and share the same problems with the maximum likelihood classifier. (Classifier training is not needed for unsupervised classification.) To improve the analysis of impercise data, a better representation method is required. The *fuzzy set* theory provides useful tools for dealing with the impreciseness.

FUZZY REPRESENTATION OF GEOGRAPHICAL INFORMATION

Let X be a universe of discourse, whose generic elements are denoted x: $X = \{x\}$. Membership in a classical set A of X is often viewed as a characteristic function χ_A from X to $\{0,1\}$ such that $\chi_A(x) = 1$ if and only if $x \in A$. A *fuzzy set* (Zadel, 1965) B in X is characterized by a *membership function* f_B which associates with each x a real number in $[0,1]$. $f_B(x)$ represents the "grade of membership" of x in B. The closer the value of $f_B(x)$ is to 1, the more x belongs to B. A fuzzy set does not have sharply defined boundaries and an element may have partial and multiple membership.

Fuzzy set theory can provide a better representation for geographical phenomena, many of which cannot be described properly by a single attribute. In a fuzzy representation, land-cover classes can be defined as fuzzy sets and pixels as set elements. Each pixel has attached to it a group of membership grades to indicate the extent to which the pixel belongs to certain classes. Pixels with class mixture or in intermediate conditions can now be desbribed. For example, if a ground cell contains two cover types "soil" and "vegetation," it may have two membership grades indicating the extents to which it is associated with the two classes.

A CLASSIFICATION ALGORITHM BASED ON FUZZY REPRESENTATION

A fuzzy supervised classification algorithm has been developed based on the fuzzy information representation. This algorithm consists of two major steps: estimate of *fuzzy parameters* from fuzzy training data and a *fuzzy partition of spectral space*.

FUZZY PARTITION OF SPECTRAL SPACE

Fuzzy representation of geographical information enables a new method for spectral space partition. When information classes are represented as fuzzy sets, so can the corresponding spectral classes be represented. Thus, a spectral space is not partitioned by sharp surfaces. A pixel may belong to a class to some extent and meanwhile belong to another class to another extent. Such a partition is referred to as a *fuzzy partition of spectral space* (Wang, 1989b). Figure 1 illustrates membership grades of a pixel in a fuzzy partition.

Let $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ form a spectral space, a fuzzy k-partition space for X can be formally represented as

$$M_{fk} = \{\mathbf{U} \in \mathbf{V}_{kn} \mid u_{ij} \in [0, 1]\ \forall i, j; \sum_{i=1}^{j} u_{ij} = 1\ \forall j;\ 0 < \sum_{j=1}^{n} u_{ij} < n\ \forall\ i\} \tag{8}$$

where $\mathbf{V}_{cn}$, k, and n are as same as in Equation 2, and $\mathbf{U} \in M_{fk}$ is referred to as a *fuzzy partition matrix* which records a fuzzy k-partition.

A fuzzy k-partition of spectral space is a family of fuzzy set $F_1, F_2, \ldots, F_k$. u_{ij} (an element of $\mathbf{U}$) is the membership grade of $\mathbf{x}_j$ in fuzzy set F_i, i.e., $u_{ij} = f_{F_i}(\mathbf{x}_j)$. A hard partition matrix can be derived from the fuzzy partition matrix by changing the maximum value in each column into "1" and others into "0." A "hardened" classification image can be generated by assigning the label of the row with value "1" of each column to the corresponding pixel.

A fuzzy k-partition of spectral space conforms to real situations better. It allows more spectral information to be extracted and utilized in subsequent analysis. Another advantage of the fuzzy partition in cluster analysis is that stray pixels and pixels isolated between classes may be classified as such.

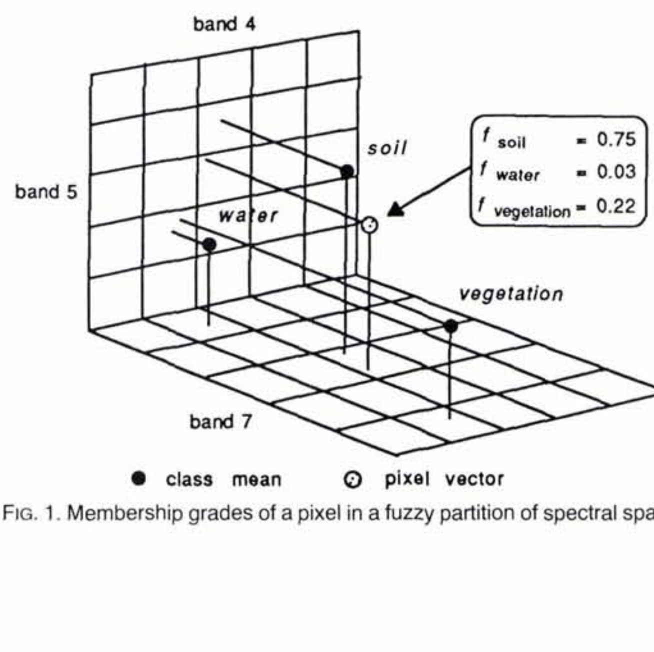

FIG. 1. Membership grades of a pixel in a fuzzy partition of spectral space.

Fuzzy Parameters for Image Classification

Fuzzy representation of geographical information makes it possible to calculate statistical parameters which are closer to the "real" ones. This can be achieved by means of the *probability measures of fuzzy events* (Zadeh, 1968).

In probability theory, an *event*, A, is a member of a σ-field, $\mathcal{A}$, of subsets of a sample space Ω. When A is a fuzzy event, i.e., a fuzzy set of points in Ω, a probability measure of A can be defined in the Lebesque-Stieltjes integral as

$$P(A) = \int_{\Omega} f_A(x)\, dP \tag{9}$$

where f_A is the membership function of $A(0 \leq f_A(x) \leq 1)$. Equation 9 is a generalization of the conventional definition of probability measure of A. Partial membership of a point x in A can be taken into consideration in calculating $P(A)$.

Similarly, mean and variance of fuzzy event A relative to a probability measure P can be expressed as

$$\mu_A^* = \frac{1}{P(A)} \int_{\Omega} x\, f_A(x)\, dP \tag{10}$$

and

$$\sigma_A^{*2} = \frac{1}{P(A)} \int_{\Omega} (x - \mu_A^*)^2\, f_A(x)\, dP. \tag{11}$$

The mean and variance calculated in this way are called a *fuzzy mean* and *fuzzy variance*. From Equation 10 and 11, the discrete type of sample mean and sample covariance matrix, the multivariate analog of variance, can be obtained for a land-cover class c which is represented as a fuzzy set. The fuzzy mean can be expressed as

$$\boldsymbol{\mu}_c^* = \frac{\sum_{i=1}^{n} f_c(\mathbf{x}_i)\mathbf{x}_i}{\sum_{i=1}^{n} f_c(\mathbf{x}_i)} \tag{12}$$

where n is total number of sample pixel vectors, f_c is membership function of class c, and $\mathbf{x}_i$ is a sample pixel measurement vector $(1 \leq i \leq n)$.

The fuzzy covariance matrix can be similarly expressed as

$$\Sigma_c^* = \frac{\sum_{i=1}^{n} f_c(\mathbf{x}_i)(\mathbf{x}_i - \boldsymbol{\mu}_c^*)(\mathbf{x}_i - \boldsymbol{\mu}_c^*)^T}{\sum_{i=1}^{n} f_c(\mathbf{x}_i)}. \tag{13}$$

The fuzzy mean and fuzzy covariance matrix can be considered as extensions of the conventional mean and covariance matrix. When $f_c(\mathbf{x}) = 0$ or 1, Equations 12 and 13 become definitions of the conventional mean and covariance matrix. The fuzzy representation permits that how much a pixel belongs to a class determines how much it contributes to the mean and covariance matrix of that class. When training data are presented in a fuzzy partition matrix, Equations 12 and 13 are applied to each row to generate a fuzzy mean and fuzzy covariance matrix for each class.

EXPERIMENTS AND RESULTS

The fuzzy supervised classification algorithm has been applied to Landsat MSS and TM data. Encouraging results have been achieved in identifying types and proportions of component land covers in mixed pixels, and improving overall classification accuracy. In this section, results of classifying an MSS image are presented.

The study area is southwest of Hamilton City, Ontario, Canada. The image was taken in July, 1978. The image for the experiments is 300 by 300 pixels in size. Seven cover classes are defined for the classification. They are water body, industrial or commercial area, residential area, forests or woods, grass or crop land, pasture or other vegetation, and bare soil. Seven membership grades are assigned to each pixel. Results of the classification are recorded in a 300 by 300 by 7 fuzzy partition matrix.

The Membership Function

A fuzzy set is characterized by its membership function. To perform a fuzzy partition on a spectral space, a membership function must be defined for each class. In this work, the membership functions are defined based on maximum likelihood classification algorithm with fuzzy mean $\boldsymbol{\mu}^*$ and fuzzy covariance matrix Σ^* replacing the conventional mean and covariance matrix. The following is the definition of membership function for cover class c:

$$f_c(\mathbf{x}) = \frac{P_c^*(\mathbf{x})}{\sum_{i=1}^{m} P_i^*(\mathbf{x})} \tag{14}$$

where

$$P_i^*(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}|\Sigma_i^*|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_i^*)^T \Sigma_i^{*-1}(\mathbf{x}-\boldsymbol{\mu}_i^*)}, \tag{15}$$

N is the pixel vector dimension, m is number of predefined classes, and $1 \leq i \leq m$.

The membership grades of a pixel vector $\mathbf{x}$ depend on $\mathbf{x}$'s position in the spectral space. $f_c(\mathbf{x})$ increases exponentially with the decrease of $(\mathbf{x}-\boldsymbol{\mu}_c^*)^T\Sigma_c^{*-1}(\mathbf{x}-\boldsymbol{\mu}_c^*)$, i.e., the distance between $\mathbf{x}$ and $\boldsymbol{\mu}_c^*$. $\sum_{i=1}^{m} P_i^*(\mathbf{x})$ serves as a normalizing factor.

When the membership function of class c is applied to a pixel vector, the membership grade of the pixel in class c is calculated for each pixel, seven membership grades can be obtained which form a column in the fuzzy partition matrix.

Improvement in Overall Classification Accuracy

Improvement in overall classification accuracy has been achieved using the fuzzy mean and fuzzy covariance matrix. For comparison, the conventional maximum likelihood classification was also performed on the same data sets. The fuzzy training data were "hardened" for training the conventional classification algorithm. The statistical parameters generated by the fuzzy and conventional algorithms are somewhat different (Wang, 1989a).

Overall accuracy of the fuzzy classification was assessed using the traditional error matrix method with reference to aerial photographs taken in the same year. For accuracy assessment, a hard partition matrix was derived by hardening the fuzzy partition matrix and then a "hardened" (one-pixel-one-value) classification image was generated from the hard matrix. Six hundred and sixty sample pixels were selected by a random number generator. The labels of the sample pixels in both the "hardened" image and the image generated by the conventional algorithm were identified. The identification results were recorded in two error matrixes (see Table 1). The values along the diagonal represent the numbers of correctly classified pixels of each class and the values along a given row indicate how misclassified pixels are distributed among the classes.

From the error matrixes, a 91.21 percent overall accuracy was estimated for the fuzzy classification, and an 86.06 percent overall accuracy for the conventional classification. It can be asserted with probability 0.99 that the magnitude of the error of the estimate is less than 3.5 percent (Freund, 1988). An improvement

TABLE 1. ERROR MATRIXES FOR THE IMAGE IN FIGURE 2.

	Water	Industrial	Residential	Forest	Crop	Soil	Pasture
Water	32	2	0	0	0	0	0
Industrial	3	62	3	0	0	2	1
Residential	0	7	163	1	1	0	2
Forest	0	0	1	70	6	0	0
Crop	0	0	1	6	159	0	4
Soil	0	4	0	0	0	42	2
Pasture	0	0	2	2	6	2	74

Sum of diagonals = 602. Sum of entries = 660. Overall accuracy = 602/660 = 0.9121.

	Water	Industrial	Residential	Forest	Crop	Soil	Pasture
Water	30	4	0	0	0	0	0
Industrial	5	57	6	0	0	2	1
Residential	0	8	158	1	3	0	4
Forest	0	0	0	65	12	0	0
Crop	0	0	2	10	151	0	7
Soil	0	6	0	0	0	39	4
Pasture	0	0	2	2	9	5	68

Sum of diagonals = 568. Sum of entries = 660. Overall accuracy = 568/660 = 0.8606.

of 5.15 percent has been achieved. Similar accuracy levels and improvements were measured from the results of classifying 1974, 1981, and 1984 MSS images of the same study area and TM images of another study area.

IDENTIFICATION OF MIXED PIXELS AND COMPONENT CLASSES

More information about land cover has been derived by using the fuzzy classification. It has been verified that a membership grade calculated from Equation 14 is proportional to the percentage to which the pixel contains a given type of land cover. Figure 2 shows a subimage in which eight pixels are selected to illustrate this fact. Figure 3 shows an aerial photograph of the same area which was taken 20 days later than the satellite image. In the photograph, approximate locations of the pixels selected are indicated by labeled arrows. Membership grades of the eight pixels are recorded in the fuzzy partition matrix in Table 2.

Homogeneous and mixed pixels can be differentiated by analyzing the membership grades. From Table 2, *C*, *F*, *G*, and *H* can be identified as mixed pixels because each of them contains more than one cover class and each of the component classes is not negligible, while *A*, *B*, *D*, and *E* as relatively homogeneous pixels.

Proportions of component cover classes in a pixel can be estimated from the membership grades. For instance, it can be estimated that *A* contains almost 100 percent of forests/woods, *B* contains about three fourths forests/woods and one fourth grass/crop land, *C* contains about one half forests/woods and one half grass/crop land, and *D* contains about 90 percent of grass/crop land. This estimation conforms to the real land covers well. In this way, intermediate conditions between the two classes have been well identified and represented by the fuzzy method. Figure 4 illustrates approximate positions of the pixel vectors of *A*, *B*, *C*, and *D*, as well as class means of forests/woods and grass/crop land in a spectral space of bands 4, 5, and 7. It can be observed that the percentage of a cover type in a pixel is roughly inversely proportional to the distance between the pixel vector and the mean of that class. The fuzzy classification allows more information contained in the positions to be extracted and utilized in the subsequent analysis, whereas a conventional classification simply classifies *A*, *B*, and *C* into forests/woods and *D* into grass/crop land and is unable to specify the differences between *A*, *B*, and *C*.

Similarly, the land cover of *E*, *F*, and *G* can be estimated from their membership grades: *E* contains almost 100 percent of

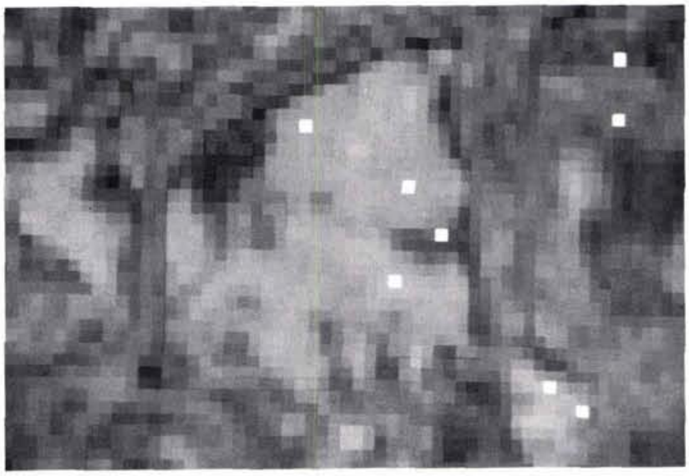

FIG. 2. A subimage in the study area.

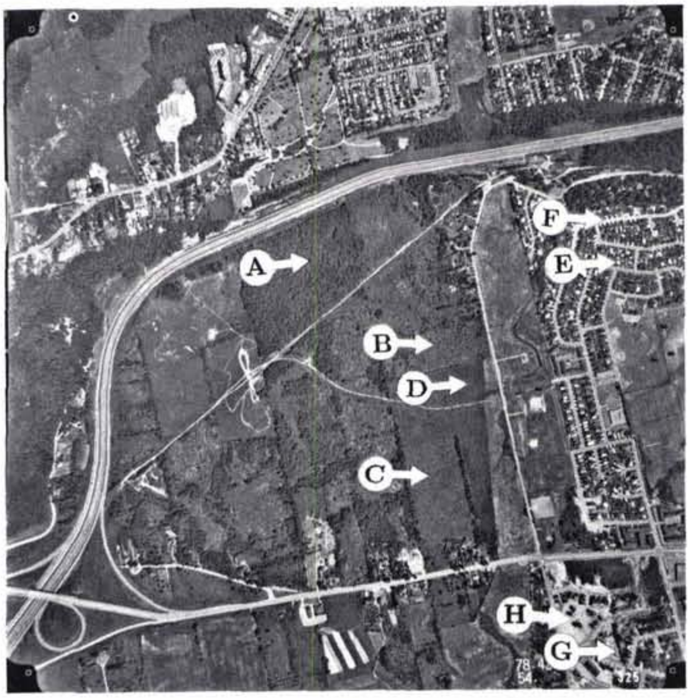

FIG. 3. An aerial photograph of the same area as in Figure 2.

TABLE 2. MEMBERSHIP GRADES OF SELECTED PIXELS.

pixel	A	B	C	D	E	F	G	H
water	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
industrial	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
residential	0.00	0.00	0.00	0.00	0.99	0.64	0.48	0.24
forest	0.99	0.77	0.54	0.00	0.00	0.13	0.00	0.00
grass	0.00	0.23	0.45	0.87	0.00	0.22	0.17	0.00
pasture	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14
bare soil	0.00	0.00	0.00	0.12	0.00	0.00	0.35	0.62

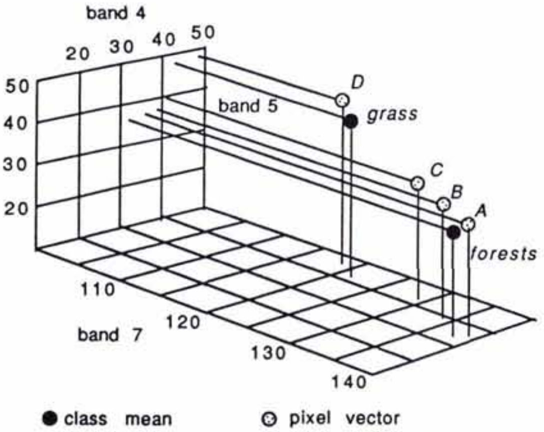

FIG. 4. A spectral space of bands 4, 5, and 7.

residential area, *F* contains two thirds of residential area and one third vegetation, and *G* contains about one half residential area and one third bare soil. The major cover type of the three pixels is residential area. However, the three pixels are in different situations. *E* is in a completed residential area, and has no other cover type. *F* is at a fringe of a residential area, thus containing some vegetation. *G* is under construction and, thus, contains bare soil and other cover types. Their differences can be well understood by analyzing the membership grades.

The identification of pixel component cover classes has been successfully applied in a land-cover change detection expert system (Wang, 1989b) to facilitate automated image analysis. When determining a change type, if only one cover class is available for each date, it is sometimes difficult for an expert system to identify the type of change which took place between the two dates. However, more cover classes may provide important information. For example, assuming cover type of *H* at the first date was vegetation, if only one cover class, "bare soil," is known for the second date, it is difficult to determine that this change is a crop rotation, an urban development, or something else. However, the identification of the residential area component for the second date, even though in a small percentage, makes the decision-making much easier.

In addition, if a hard partition (conventional classification output) is required from the fuzzy parition, the membership grades enable assessing classification accuracy for *individual* pixels. For example, if class "forests/woods" is assigned to pixels *A*, *B*, and *C* when the fuzzy partition is hardened, it can be inferred from their membership grades of 0.99, 0.77, and 0.54 in that class that deviation with the assignment to pixel *A* is very small, the deviation with pixel *B* is bigger, and the deviation with pixel *C* is even bigger. Accuracy assessment for individual pixels may contribute to the integration of remote sensing image analysis systems and geographical information systems. The current accuracy assessment methods used in remote sensing cannot provide local accuracy levels of a data set which are usually required by the error models of geographical information systems. The incompatibility between the accuracy assessment methods has been considered as one of the major hindrances to incorporating remote sensing inputs into geographical information systems. The fuzzy techniques may help bridge this gap (Wang, 1989a).

CONCLUSION

Good initial results have been obtained in improving remote sensing image analysis through the fuzzy representation method. The major achievements include increasing overall classification accuracy, differentiating mixed and homogeneous pixels, identifying types and proportions of component cover classes in mixed pixels, and assessing classification accuracy for individual pixels. This research has broken up a pixel which is considered as the atomic unit in the conventional anaylsis. Much more information has been extracted; thus, we can make fuller use of the valuable remotely sensed data. The techniques developed in this research may also largely facilitate knowledge-based remote sensing image analysis and integration of remote sensing and geographical information systems.

ACKNOWLEDGMENT

This work was made possible by the financial support from the International Development and Research Center (IDRC) Canada, which enabled the author to undertake graduate studies in Canada.

REFERENCES

Anuta, P. E., 1977. Computer Assisted analysis Techniques for Remote Sensing Data Interpretation, *Geophysics*, Vol. 42, No. 3, pp. 468–481.

Burrough, P. A., 1986. *Principles of Geographical Information Systems for Land Resources Assessment*, Clarenden Press, Oxford.

Cannon, R. L., 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 Transactions on Geoscience and Remote Sensing*, Vol. 24, No. 3, pp. 400–408.

Freund, J. E., 1988. *Modern Elementary Statistics* Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Fung, A. D., and F. T. Ulaby, 1983. Matter-Energy interaction in the microwave region. *Manual of Remote Sensing (Second Edition)*, (R. N. Colwell, ed.) American Society of Photogrammetry, Falls Church, Virginia.

Jeansoulin, R., Yves Fonyaine, and Werner Frei, 1981. Multitemporal Segmentation by Means of Fuzzy Sets *1981 Machine Processing of Remotely Sensed Data Symposium*, pp. 336–339.

Jensen, J. R., 1986. *Introductory Digital Image Processing: A Remote Sensing Perspective* Prentice-Hall, New Jersey.

Kent, J. T., and Mardia K. V., 1988. Spatial Classification Using Fuzzy Membership Models, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, No. 5, pp. 659–671.

Thomas, I. L., V. M. Benning, and N. P. Ching, 1987. *Classification of Remotely Sensed Images*, IOP Publishing limited, Bristol, England.

Wang, F., 1989a. *Integrating Expert Vision Systems and Spatial Databases by unifying Knowledge Represenation Schemes: Development in Remote Sensing Image Analysis Expert Systems and Geographical Information Systems*, Ph.D. dissertation, University of Waterloo, Ontario, Canada.

———, 1989b. "A Fuzzy Expert System for Remote Sensing Image Analysis" *Proceedings of IEEE IGARSS'89*, Vancouver, Canada, Vol. 2, pp. 848–851.

Zadeh, L. A., 1965. Fuzzy Sets, *Information and Control*, Vol. 8, pp. 338–353.

———, 1968. Probability Measures of Fuzzy Events, *Journal of Mathematical Analysis and Application*, Vol. 10, pp. 421–427.

Zenzo, S. D., R. Bernstein, S. D. Degloria, and H. G. Kolsky, 1987. Gaussian Maximum Likelihood and Contextual Classification Algorithms for Multicrop Classification, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. GE-25, No. 6, pp. 805–814.

(Received 5 September 1989; revised and accepted 3 January 1990)