

ePub^{WU} Institutional Repository

Manfred M. Fischer and Josef Benedikt

The Use of Fuzzy Set Theory in Remote Sensing Pattern Recognition

Paper

Original Citation:

Fischer, Manfred M. and Benedikt, Josef (1996) The Use of Fuzzy Set Theory in Remote Sensing Pattern Recognition. *Discussion Papers of the Institute for Economic Geography and GIScience*, 50/96. WU Vienna University of Economics and Business, Vienna.

This version is available at: <http://epub.wu.ac.at/4174/>

Available in ePub^{WU}: June 2014

ePub^{WU}, the institutional repository of the WU Vienna University of Economics and Business, is provided by the University Library and the IT-Services. The aim is to enable open access to the scholarly output of the WU.



WSG 50/96

**The Use of Fuzzy Set Theory
in Remote Sensing Pattern Recognition**

Manfred M. Fischer and Josef Benedikt

Institut für Wirtschafts-
und Sozialgeographie

**Wirtschaftsuniversität
Wien**

Department of Economic
and Social Geography

**Vienna University of
Economics and Business
Administration**

**Abteilung für Theoretische und Angewandte Wirtschafts- und Sozialgeographie
Institut für Wirtschafts- und Sozialgeographie
Wirtschaftsuniversität Wien**

**Vorstand: o.Univ.Prof. Dr. Manfred M. Fischer
A - 1090 Wien, Augasse 2-6, Tel. (0222) 313 36 - 4836**

Redaktion: Mag. Petra Staufer

WSG 50/96

**The Use of Fuzzy Set Theory
in Remote Sensing Pattern Recognition**

Manfred M. Fischer and Josef Benedikt

WSG-Discussion Paper 50

January 1996

ISBN 3 85037 055 0

ABSTRACT

Satellite images increasingly become a major data source for monitoring changes in the natural environment. A main task in the analysis of satellite images is concerned with the modelling of land use classes by reducing uncertainty during a classification process. In the approach presented in this paper uncertainty is perceived to be due to the vagueness of geographical categories caused by either the complexity of the category (like 'urban area') or by the use of the category in several application contexts. Two circumstances of use of an extended set theoretical concept (fuzzy sets) are discussed. First, two algorithms from the ISODATA class are used to determine the grades of membership to three a priori defined classes (woodland, suburban area, urban area) of a LANDSAT TM satellite image of Vienna, Austria. The results are visualized by a RGB image of the grades of membership to each category. Second, a measure of fuzziness is employed on the results. The measure relies on the concept of distance between a set/category and its complement. The so determined vagueness provide more information on the uncertainty of the different categories and may improve further information processing tasks.

Keywords: Fuzzy Sets, Remote Sensing, Pattern Recognition, Measure of Fuzziness

1. Introduction

Satellite remote sensing has become a valuable tool for gathering information about planet earth. Images received by satellite are useful not only for observing and monitoring changes in the natural environment. High-spatial-resolution images acquired by earth-orbiting sensors may also be used to monitor both the extent of urban areas and their composition in terms of land use. Unfortunately, many of the commonly used image processing techniques, such as the conventional parametric pixel by pixel classification algorithms, tend to perform poorly in this context. This is because urban areas comprise a complex spatial assemblage of disparate land cover types - including built structures, numerous vegetation types, bare soil and water bodies - each of which has different spectral reflectance characteristics. Thus, although a simple, direct relationship between land cover and the spectral response detected by a satellite sensor often exists in the natural environment, this is seldom the case for land cover in urban areas.

As a result, pattern recognition in urban areas is one of the most demanding issues in classifying satellite remote sensing data. This has generated interest in both neural networks and fuzzy systems. Neural networks deal with uncertainty as humans do, not by deliberate design, but as a byproduct of their parallel-distributed structure [see Gopal and Fischer 1994, Fischer et al. 1995]. In contrast, fuzzy systems directly build the basic insight that categories [classes] are not absolutely clear cut into an artificial system. There are significant differences between neural networks and fuzzy systems to pattern recognition. There are formal similarities as pointed out by Kosko [1992], but they vary greatly in detail. The noise and generalization abilities of neural networks grow out of the structure of networks, their dynamics, and their data presentations. Fuzzy systems start from highly formalized insights about the structure of categories found in the real world. The theory of fuzziness may be sometimes easier to use and simpler to apply to a particular remote sensing problem than neural networks may be. Whether to use one or another technology depends on the particular application and an engineering judgement.

Fuzzy set theory can be incorporated in the handling of uncertainties arising from deficiencies in the available information caused by incomplete, imprecise vague data and information in various stages of pattern recognition. In this paper we consider two circumstances where the concepts and techniques of fuzzy set theory may be helpful in the practice of remote sensing pattern recognition [Pao 1989]:

- ☐ The *first circumstance of use* is at the class-membership level. In the crisp case, classification consists of relegating a pixel to a membership in one of the a priori given land cover classes. In the fuzzy set approach, the class membership of a pixel itself is a fuzzy set, and the different class indices constitute the support for that fuzzy set. A pixel does not necessarily belong to just one of the classes. There is a certain degree of possibility that the pixel might belong to more than one class. The membership functions supply values for these various possibilities. Fuzzy clustering or the fuzzy ISODATA procedure is an instructive example of this first circumstance of use [see section 3].
- ☐ The *second circumstance of use* is where measures of fuzziness are used to make us aware of the vagueness of the land cover classification obtained. Vagueness in general is associated with the difficulty to delimit the land cover classes by sharp boundaries [see section 4].

In this paper we describe the two above mentioned roles that fuzzy set theory might play in satellite remote sensing pattern recognition. First, we will introduce the basic concepts in section 2. Section 3 is concerned with the fuzzy extension of the clustering procedure ISODATA, while section 4 deals with the question of how to measure vagueness or fuzziness. The two circumstances of use of fuzzy set theory are discussed in view of a multispectral pixel-by-pixel classification task using a satellite image selected from a Landsat-TM scene from the city of Vienna and its northern surroundings. The spectral resolution of the five bands [TM2 to TM5, TM7] which were used was eight bits or 256 possible digital numbers. The geometric resolution of a pixel is 30 x 30 m². The purpose of the classification task was to distinguish between three broad land cover categories [classes]: woodland/public gardens; low density residential and industrial areas [suburban], and densely built up areas [urban]. The data base consisted of 256 x 256 pixels that are described by 5-dimensional feature vectors.

2. The Notion of Fuzzy Sets and Some Basic Concepts

According to the usual terminology, the term **conventional [hard or crisp] set** is used in this contribution for sets whose boundaries are sharp. A crisp set of all pertinent entities in any particular context is called a **universal set**. In the context of this paper, **X** may denote the set of pixels of the image of concern [i.e. the universal set].

Crisp sets contain objects [pixels] that satisfy precise properties required for membership. That is, each pixel is grouped unequivocally to a given set [land cover category in the application domain of this study]. In other words, the characteristic function of a crisp set assigns a value of either 1 or 0 to each object in the universal set, thereby discriminating between members and non-members of the crisp set under consideration. This requirement is a particularly harsh one for urban areas that contain mixtures of built structures. Most conventional classification [clustering] procedures have no natural mechanism for absorbing the effects of undistinctive or aberrant feature data. Accordingly, the notion of fuzzy set was introduced as a means for modifying the basic axioms underlying classification [clustering] models with the purpose of tackling the above mentioned issue.

In a way similar to conventional sets, **fuzzy sets**, are defined in each particular context within the relevant [conventional] universal set **X**. In contrast to hard sets, however, fuzzy sets involve uncertainty in determining whether an individual of **X** belongs to a given set or not. This uncertainty may be expressed in the following way [see Zadeh 1965]:

Given a universal set **X**, a **fuzzy set A** in **X** is defined as a set of ordered pairs

$$A := \{(x, \mu_A(x)) \mid x \in X\}, \quad (1)$$

where $\mu_A(x)$ is called the **membership function** defined by

$$\mu_A(x) : X \rightarrow [0,1]. \quad (2)$$

The membership function $\mu_A(x)$ maps **X** to the membership space $M = [0,1]$, the interval of real numbers from 0 to 1, inclusive. $\mu_A(x)$ expresses the grade of membership of **x** in **A**, i.e.

degree of compatibility of x with the concept represented by the fuzzy set A . If M contains only two points, 0 and 1, then A is not fuzzy. Thus, hard sets are a special type of fuzzy ones. This definition of a fuzzy set in terms of an ordered set of pairs allows to define subsets of X that do not have sharp boundaries. Fuzzy sets in $\mathcal{P}(X)$, the power set of all fuzzy subsets of X , can be operated upon a variety of operators of fuzzy complementation, intersection, union, etc. [see Yager 1979]. The original theory of fuzzy sets was formulated in terms of the following specific operators of set complement, union and intersection:

$$\mu_{A'}(x) = 1 - \mu_A(x) \quad \text{for all } x \in X \quad (3)$$

$$\mu_{A \cup B}(x) = \max[\mu_A(x), \mu_B(x)] \quad \text{for all } x \in X \quad (4)$$

$$\mu_{A \cap B}(x) = \min[\mu_A(x), \mu_B(x)] \quad \text{for all } x \in X \quad (5)$$

where A and B denote fuzzy sets of X , A' the complement of fuzzy set A ; $\mu_A(x)$ and $\mu_B(x)$ membership functions. When the membership space is restricted to the set $\{0,1\}$, these functions perform precisely as the corresponding operators for conventional sets, thus establishing them as clear generalizations of the latter. But it is important to note, that these above functions are not the only possible generalizations of conventional set operators. They may be called standard operations of fuzzy set theory.

In the context of this paper it is appropriate to follow Klir and Folger [1988] and to define the general class of fuzzy complements c , which possess appropriate axiomatic properties, in the following way: A general complement of a fuzzy set A is specified by a function [termed fuzzy complement function]

$$c: [0,1] \rightarrow [0,1] \quad (6)$$

which assigns a value $c(\mu_A(x))$ to each membership grade $\mu_A(x)$. This assigned value is viewed as the membership grade of x in the fuzzy set representing the negation of the concept represented by A . Any function c to be considered as a fuzzy complement has to satisfy the following four axiomatic requirements:

(i) **Boundary conditions** [i.e. c collapses into the ordinary complement for hard sets]

$$c(0) = 1 \text{ and } c(1) = 0$$

(ii) **c is a monotonic non-increasing function**

$$\text{if } a < b, \text{ then } c(a) \geq c(b), \text{ for all } a, b \in [0,1]$$

(iii) **c is a continuous function**

(iv) **c is involutive**, i.e. $c(c(a)) = a$, for all $a \in [0,1]$

where a and b represent degrees of some arbitrary elements of the universal set in a given fuzzy set [example: $a = \mu_A(x_i)$ and $b = \mu_A(x_j)$ for some $x_i, x_j \in X$ and some fuzzy set A].

If there exists a such that $c(a) = a$, then a is termed an **equilibrium** of c [termed e_c]. It is known that every continuous fuzzy complement has a unique equilibrium [Higashi and Klir, 1982].

The general fuzzy complement produced by function (6) with the four axiomatic requirements (i) - (iv) is denoted by $C(A)$ (or A^C), where

$$C: \mathcal{P}(X) \rightarrow \mathcal{P}(X) \quad (7)$$

is a function such that $c(\mu_A(x)) = \mu_{C(A)}(x)$ for all $x \in X$. $\mathcal{P}(X)$ denotes the set of all fuzzy subsets of X . We use the general class of fuzzy complements C in section 4 to define measures of fuzziness.

3. The First Circumstance of Use

The problem of the multispectral pixel-by-pixel classification task at hand is to assign pixels to three a priori given land cover types [classes, clusters] or in other words to partition the set X of pixels into $k=3$ subsets, where the elements of each set are as similar as possible to each other and at the same time, as different as possible from the other sets. Clustering algorithms attempt to partition X based on certain assumptions and/or criteria. Thus, the output partition [classification] is dependent on the criteria which are used to control the clustering algorithm [see Fischer 1982].

Many algorithms have been developed to obtain hard [crisp] classes from a given data set during the last two decades. Among these, the k -means algorithms and their generalizations, the ISODATA clustering procedures, are the most widely used. The performance of both cluster models is influenced by the choice of k , the initial cluster centres, the choice of the distance [similarity] measure and the order in which the samples are taken as input [in the case of the sequential versions]. In practice, the performance of any cluster activities depends more or less on extensive trial and error experiments and the experience of the user [see Bezdek and Pal, 1992].

Classical [crisp] clustering algorithms generate partitions such that each pixel is assigned to exactly one of the a priori given clusters. But, in practice, the separation of clusters is a fuzzy notion [especially in the case of images covering urban areas] and thus the concept of fuzzy sets offers special advantages over classical clustering by allowing algorithms to assign each pixel a partial or distributed membership to each of the k clusters. In this way fuzzy clustering procedures may yield more accurate representations of real data structures [see Pao 1989]. The following discussion of the fuzzy generalizations of the crisp ISODATA algorithm is based on Bezdek [1976,1980], Bezdek et al. [1984] and Pao [1989]. In the sequel we compare the partitioning of the set $X = \{x_1, \dots, x_n\}$ of n pixels into crisp versus fuzzy subsets [i.e. classes, clusters]. Each partition into crisp or fuzzy subsets can be described by a membership function $u_{q,p}$ and $\bar{u}_{q,p}$, respectively. These functions may be viewed to be mappings

$$u_{q,p} : X \rightarrow \{0,1\} \quad \text{for the hard case,} \quad (8)$$

$$u_{q,p} : X \rightarrow [0,1] \quad \text{for the fuzzy case.} \quad (9)$$

Allocation of the pixels to the different classes can be represented in terms of a k -partition (k,n) matrix which is crisp in the case of (8) and fuzzy in the case of (9). In the crisp case the element, $u_{q,p}$, of the matrix U denotes the degree of membership of pixel x_p in cluster q and can take on either the value 0 or 1. In the fuzzy case the elements of the matrix \underline{U} are $u_{q,p}$, which denote the degree of membership of pixel x_p to class q . Formally, the elements of U are subject to the following constraints:

$$u_{q,p} \in \{0,1\} \quad \text{for } 1 \leq q \leq k; 1 \leq p \leq n \quad (10)$$

$$\sum_{q=1}^k u_{qp} = 1 \quad \text{for } 1 \leq p \leq n \quad (11)$$

$$0 < \sum_{p=1}^n u_{qp} < n \quad \text{for } 1 \leq q \leq k \quad (12)$$

while the elements of \underline{U} satisfy the following three conditions:

$$u_{q,p} \in [0,1] \quad \text{for } 1 \leq q \leq k; 1 \leq p \leq n \quad (13)$$

$$\sum_{q=1}^k u_{qp} = 1 \quad \text{for } 1 \leq p \leq n \quad (14)$$

$$0 < \sum_{p=1}^n u_{qp} < n \quad \text{for } 1 \leq q \leq k \quad (15)$$

where q and q denote the cluster and p the pixel [pattern].

We shall sketch the way of getting the $u_{q,p}$ and $\underline{u}_{q,p}$, respectively, in following Bezdek [1976, 1980]. In the crisp ISODATA algorithm, the partition of the cluster centre v_q in the t -dimensional feature space is found to be the average of the positions of all the pixels in that class, i.e.

$$v_q = \frac{1}{|c_q|} \sum_{x_p \in c_q} x_p \quad (16)$$

where $|c_q|$ denotes the cardinality of the set of pixels in class C_q . This result is based on minimizing the sum of variances of all features j for each pixel in each cluster q . That is, the position of v_q is varied so to minimize

$$\sum_{q=1}^k \sum_{x_p \in c_q} \|x_p - v_q\|^2 = \sum_{q=1}^k \frac{1}{|c_q|} \sum_{x_p \in c_q} \sum_{j=1}^t (x_{pj} - v_{qj})^2. \quad (17)$$

We can use the membership function values $u_{q,p}$ to remove the constraint $x_p \in c_q$ in (17) to arrive at a formulation of the problem as determining the minimum of the sum of the weighted variances, i.e.

$$\min J(U, v) = \sum_{q=1}^k \sum_{p=1}^n u_{qp} \|x_p - v_q\|^2. \quad (18)$$

One way to extend this problem to the fuzzy case is to draw an analogy between $u_{q,p}$ and $\underline{u}_{q,p}$, and to define the problem of finding fuzzy k-partitions as

$$\min J_m(\underline{U}, \underline{v}) = \sum_{q=1}^k \sum_{p=1}^n (\underline{u}_{qp})^m \|\underline{x}_p - \underline{v}_q\|^2 \quad (19)$$

where the squared distance between \underline{x}_p and \underline{v}_q shown in (19) is computed in the A - Norm

$$d_{qp}^2 = \|\underline{x}_p - \underline{v}_q\|_A^2 = (\underline{x}_p - \underline{v}_q)^T A (\underline{x}_p - \underline{v}_q) \quad (20)$$

where A is a positive definite (n,n) weight matrix which controls the shape that optimal classes assume. In practice, there are only three A-norms which enjoy widespread use: the euclidean norm, the diagonal norm and the Mahalanobis norm. The weight attached to each squared error is $(\underline{u}_{qp})^m$, the m-th power of \underline{x}_p 's membership in class q [$1 \leq m < \infty$]. As $m \rightarrow 1$, partitions that minimize J_m become increasingly hard. Increasing m tends to degrade membership towards the fuzziest state.

Differentiating the variance function (19) with respect to \underline{v}_q [for fixed \underline{u}] and the $\underline{u}_{q,p}$ [for fixed \underline{v}] and applying the condition (14), we obtain

$$\underline{v}_q = \frac{\sum_{p=1}^n (\underline{u}_{qp})^m \underline{x}_p}{\sum_{p=1}^n (\underline{u}_{qp})^m} \quad \text{for all } q = 1, \dots, k \quad (21)$$

and

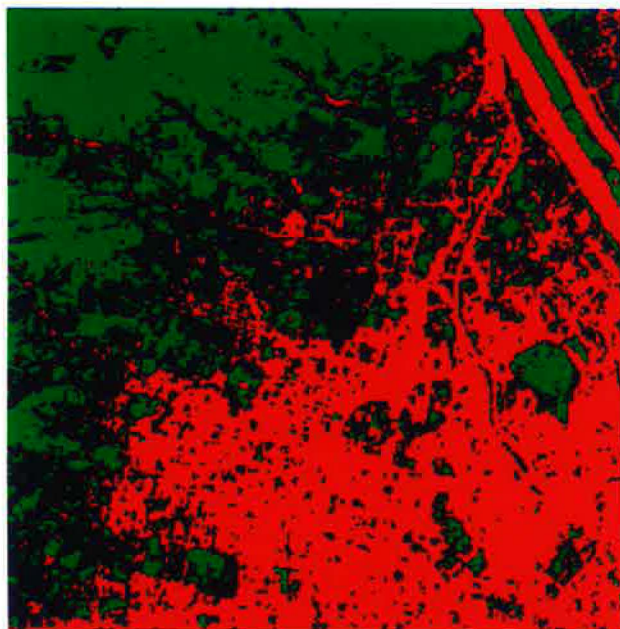
$$\underline{u}_{q,p} = \left[\sum_{i=1}^k \left(\frac{d_{q,p}}{d_{i,p}} \right)^{2/(m-1)} \right]^{-1} \quad \text{for } 1 \leq p \leq n; 1 \leq q \leq k \quad (22)$$

(21) und (22) provide necessary, but not sufficient conditions for optimizing J_m via iteration, by looping back and forth from (21) to (22) until the iterate sequence shows sufficiently small changes in successive entries of \underline{u} or \underline{v} .

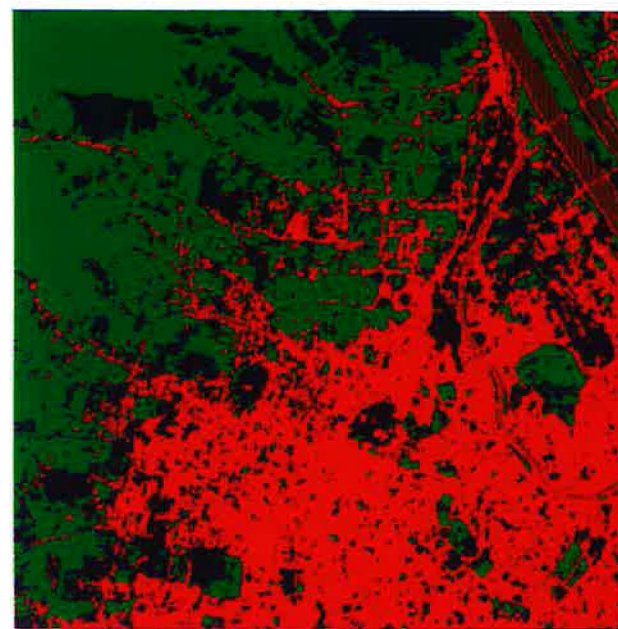
4. The Second Circumstance of Use

The fuzzy clustering procedure provides a richer and more flexible solution structure, one that classifies with a finer degree of detail at the class-membership level than the harshness that crisp procedures impose on the pixel-by-pixel classification problem [see the RGB-visualization displayed in Figure 1].

Figure 1: RGB - Visualization of the Three
Land Cover Classes:
Landsat TM Image of Vienna, Austria



(a) produced by ISODATA



(b) produced by fuzzy ISODATA

Woodland Suburb Urban

Woodland Suburb Urban

It is easy to see that the concept of a fuzzy set provides a basic mathematical framework for identifying and dealing with vagueness. Vagueness is associated with the difficulty of making sharp or precise distinctions between classes of pixels. Many measures of fuzziness have been proposed over the last two decades. Following Klir and Folger [1988] we consider a class of measures of fuzziness [or vagueness] which might be used to characterize classification results, received by using ISODATA and its fuzzy version, in terms of their degree of fuzziness and then apply these measures on the satellite image-based pattern classification problem [see section 3].

A measure of fuzziness is defined by a function

$$f: \mathcal{P}(X) \rightarrow \mathbb{R} \quad (23)$$

which satisfies certain requirements. $\mathcal{P}(X)$ denotes the power set of X [i.e. the set of all fuzzy subsets of X] and f assigns a value $f(A)$ to each fuzzy subset A of X which characterizes the degree of fuzziness of A . A meaningful measure of fuzziness has to fulfill the following three requirements [Klir 1987]:

- **First**, the degree of fuzziness must be zero for all hard subsets of $\mathcal{P}(X)$ and only for them, formally:

$$(i) \quad f(A) = 0 \quad \text{if and only if } A \text{ is a hard set.}$$

- **Second**, if - according to a particular meaning given to the concept of the degree of fuzziness or sharpness - set [class] A is considered as sharper [less fuzzy] than set [class] B , it is required that $f(A) \leq f(B)$, formally:

$$(ii) \quad \text{if } A < B \text{ then } f(A) \leq f(B)$$

where $A < B$ denotes that A is sharper than B .

- **Third**, the degree of fuzziness must be equal to the maximum value only for a fuzzy subset of $\mathcal{P}(X)$ that is perceived as maximally fuzzy, formally:

$$(iii) \quad f(A) \text{ assumes the maximum value if and only if } A \text{ is maximally fuzzy.}$$

Several measures of fuzziness have been suggested in the literature. But it was Yager [1979] who proposed the important idea to express the degree of fuzziness of a fuzzy set in terms of the lack of distinction between the set and its complement. The formulation of measures of fuzziness based upon this idea depends on the fuzzy complement employed. The measure of fuzziness utilized in this paper is employing Yager's approach within the general class of fuzzy complements as introduced in section 2, and is based on defining the sharpness relation $A < B$ in (ii) by

$$A < B \text{ if and only if } |\mu_A(x) - c(\mu_A(x))| \geq |\mu_B(x) - c(\mu_B(x))| \quad (24)$$

for all $x \in X$, and defining the term maximally fuzzy in (iii) by

$$\mu_A(x) = e_c \text{ for all } x \in X \quad (25)$$

provided that the complement employed has an equilibrium e_c .

Higashi and Klir [1982] showed that a general class of measures of fuzziness based upon Yager's approach [1979] is exactly the same as the class of measures of fuzziness in which the lack of distinction between the considered set and its complement is expressed in terms of a metric distance which is based on some form of aggregating the individual differences [Klir and Folger 1988]:

$$|\mu_A(x) - c(\mu_A(x))| = \delta_{c,A}(x) \text{ for all } x \in X \quad (26)$$

For example, using any metric distance from the class of Minkowski metrics

$$D_{c,r}(A, A^c) = [\sum_{x \in X} \delta_{c,A}^r(x)]^{1/r} \quad (27)$$

where A^c denotes the complement of A produced by function c and $r \in [1, \infty]$, the measure of fuzziness has the following form

$$f_{c,r}(A) = D_{c,r}(Z, Z^c) - D_{c,r}(A, A^c) \quad (28)$$

with Z denoting any arbitrary hard subset of X so that $D_{c,r}(Z, Z^c)$ is the largest possible distance in $\mathcal{P}(X)$ for a given c and r . The normalized version of this measure is given by

$$f_{c,r}(A) = 1 - [D_{c,r}(A, A^c) / D_{c,r}(Z, Z^c)] \quad (29)$$

so that

$$0 \leq f_{c,r}(A) \leq 1. \quad (30)$$

When the family of Minkowski metrics expressed by (25) is used, let $f_{c,r}$ denote the measure of fuzziness for the distance $d_{c,r}$. Thus

$$D_{c,r}(Z, Z^c) = |X|^{1/r} \quad (31)$$

Equation (28) becomes

$$f_{c,r}(A) = |X|^{1/r} - D_{c,r}(A, A^c) \quad (32)$$

and (29) becomes

$$f_{c,r}(A) = 1 - [D_{c,r}(A, A^c) / |X|^{1/r}] \quad (33)$$

It is important to note that measures of fuzziness defined in terms of different distance functions [i.e. r parameters] are based upon different measurement units. Although the choice of a unit is not a critical issue, it is often desirable to use a unit that is intuitively appealing in the sense that it has a simple interpretation in terms of some significant canonical situation [Klir and Folger 1988].

Table 1: The Effect of the Complement and Distance (value of r) Employed on the Measure of Fuzziness defined by (28) and (29)**(a) Fuzzy Complement $c(a) = 1 - a$**

Category	Clustering Procedure	$D_{c,2}(A, A^c)$	$f_{c,2}(A, A^c)$	$f_{c,2}(A, A^c)$	$D_{c,3}(A, A^c)$	$f_{c,3}(A, A^c)$	$f_{c,3}(A, A^c)$	$D_{c,5}(A, A^c)$	$f_{c,5}(A, A^c)$	$f_{c,5}(A, A^c)$
Woodland	ISODATA	192.8	63.2	0.25	32.4	7.9	0.19	7.8	1.3	0.14
	fuzzy ISODATA	209.5	46.5	0.18	34.3	6.1	0.15	8.1	1.1	0.12
Suburban	ISODATA	142.6	113.4	0.44	25.6	14.7	0.36	6.7	2.5	0.27
	fuzzy ISODATA	205.2	50.8	0.19	33.7	6.6	0.16	8.0	1.2	0.13
Urban	ISODATA	150.9	105.1	0.41	19.2	21.1	0.52	6.8	2.3	0.26
	fuzzy ISODATA	165.6	90.4	0.35	28.1	12.2	0.30	6.9	2.2	0.24

(b) Fuzzy Complement $c(a) = (1 - a^2)^{1/2}$

Category	Clustering Procedure	$D_{c,2}(A, A^c)$	$f_{c,2}(A, A^c)$	$f_{c,2}(A, A^c)$	$D_{c,3}(A, A^c)$	$f_{c,3}(A, A^c)$	$f_{c,3}(A, A^c)$	$D_{c,5}(A, A^c)$	$f_{c,5}(A, A^c)$	$f_{c,5}(A, A^c)$
Woodland	ISODATA	200.9	55.0	0.21	33.4	6.8	0.17	8.1	1.1	0.12
	fuzzy ISODATA	200.5	55.5	0.21	33.2	7.1	0.18	7.9	1.2	0.13
Suburban	ISODATA	158.6	97.4	0.38	27.6	12.6	0.31	7.0	2.2	0.24
	fuzzy ISODATA	210.4	45.6	0.18	34.5	5.8	0.14	8.2	0.9	0.11
Urban	ISODATA	160.9	95.0	0.37	27.9	12.4	0.30	7.1	2.1	0.23
	fuzzy ISODATA	199.0	56.9	0.22	32.8	7.5	0.18	7.8	1.3	0.15

Figure 2a: The Three Land Cover Classes
and Two Complements of the Yager-Class
at the Class Membership Level:
Produced by the ISODATA Algorithm

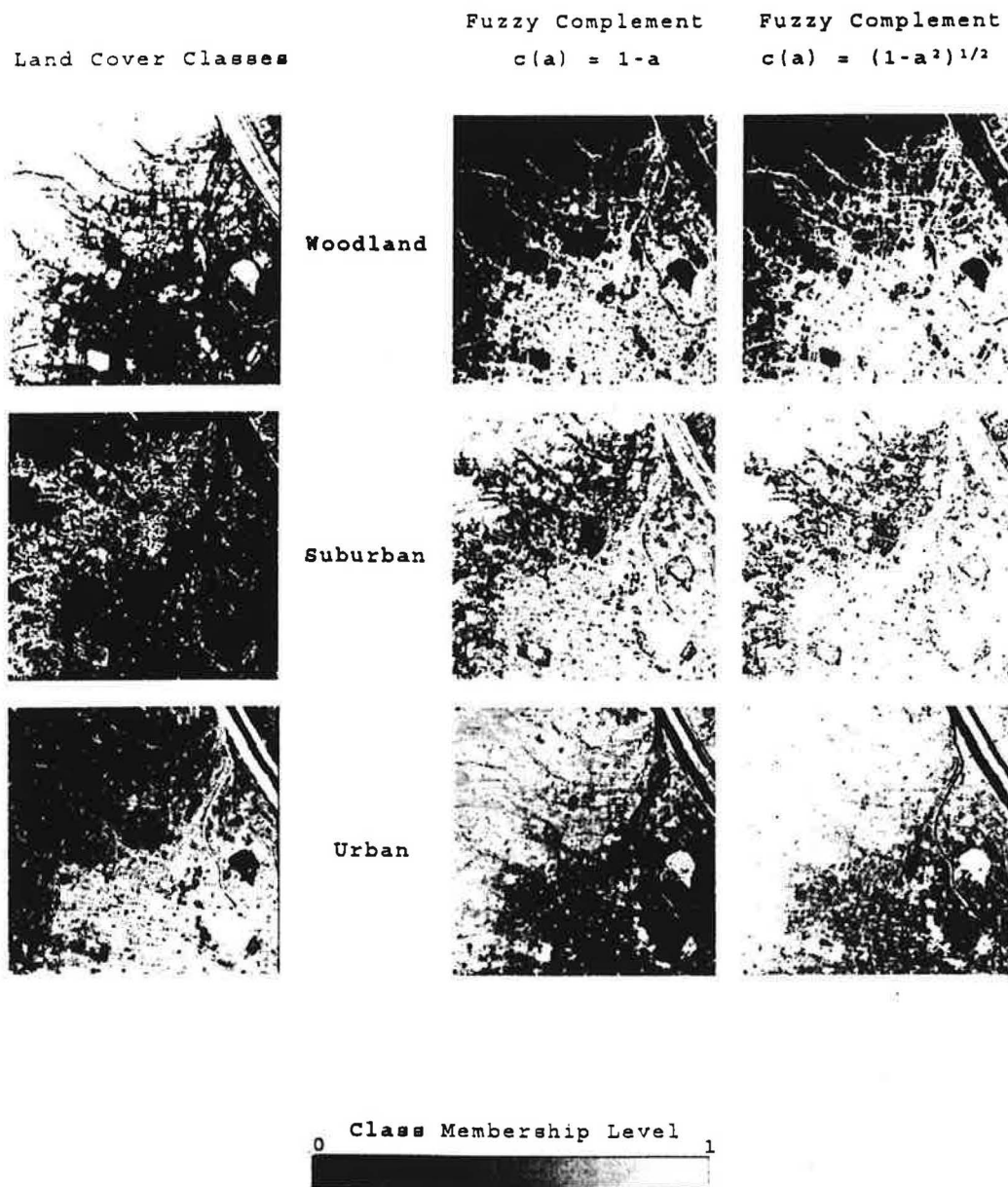


Figure 2b: The Three Land Cover Classes
 And Two Complements of the Yager-Class
 at the Class Membership Level:
 Produced by the Fuzzy ISODATA Algorithm

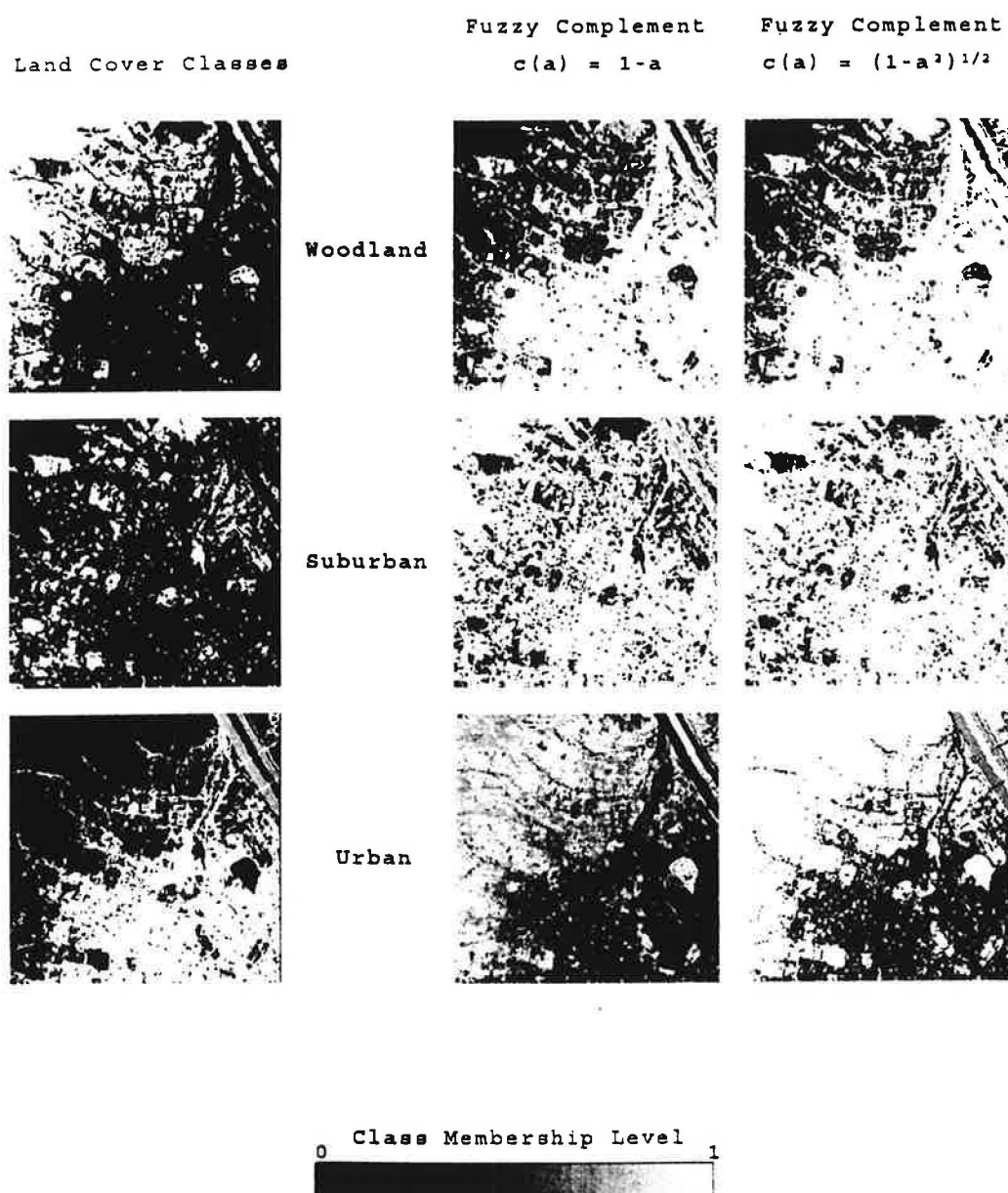


Table 1 presents the results of the calculations of the measure of fuzziness for each of the three land cover classes produced by ISODATA and its fuzzy version. Note that the results are obtained by using the image that is obtained next to the last iteration so we are able to compare the uncertainty or vagueness of the two algorithms used. To illustrate the effect of the complement and distance metric employed, the table shows calculations of the measure of fuzziness [regular and normalized] for two complements of the Yager class and for three distances of the family of the Minkowski metrics [$r = 2, 3, 5$]. For each case, we first calculate the local differences $\delta_{c,A}(x)$ for all pixels belonging to the corresponding land cover category. Then, using equation (28), we calculate for each case the distance $D_{c,r}(A, A^c)$ between the given land cover category A and its complement A^c , as well as the distance $D_{c,r}(Z, Z^c)$ between an arbitrary crisp set on X and its complement. By equation (31), we have $D_{c,r}(Z, Z^c) = |X|^{1/r}$, which is independent of the complement employed. Finally using (32) and (33) we compute the measure of fuzziness $f_{c,r}(A)$ and its normalized version $\tilde{f}_{c,r}(A)$, respectively. It is important to note that measure of fuzziness [both regular and normalized] decreases with increasing r . This is true for both complements used in this paper. Moreover, Table 1 clearly shows that the classification result produced by the fuzzy version of ISODATA is generally less vague than that generated by the conventional ISODATA. This result is visualized in Figure 2a and 2b.

5. Outlook

Fuzzy sets are a generalization of conventional set theory that were introduced by Zadeh [1965] as a mathematical framework to represent and deal with vagueness. Over the last 10 years fuzzy models have supplanted more conventional scientific applications and engineering systems, especially in control systems and pattern recognition. Fuzzy set theory might be incorporated in various stages of pattern recognition. In this paper attention focused on two circumstances, in which the concepts and techniques of fuzzy set theory are uniquely helpful in the practice of pattern recognition: the class-membership level and the output level.

A satellite image-based pattern classification problem has been chosen to illustrate the use of fuzzy set theory. The satellite image consists of 256×256 pixels selected from a LANDSAT-5 TM scene from the city of Vienna and its northern surroundings. The purpose of the multispectral classification was to distinguish three broad land cover classes on a pixel-by-pixel basis. We were able to illustrate that a crisp classification provides no hint of the details of the situation that pixels often represent a complex spatial assemblage of disparate land cover types especially in urban areas. In contrast, fuzzy clustering procedures like the fuzzy version of ISODATA used in this study can yield more accurate representations of real data structures. Moreover, we discussed general measures of fuzziness which are powerful tools to express the degree of fuzziness of the resulting classes of the pattern recognition process.

Finally, it is our expectation and contention that the synthesis between fuzzy and neural approaches to pattern recognition is one of the top issues in the current research agendas for the near future.

References

- Bezdek, J.C. (1976): *A physical interpretation of fuzzy ISODATA*, **IEEE Transactions on Systems, Man and Cybernetics** 6, pp.387-389.
- Bezdek, J.C. (1980): *A convergence theorem for the fuzzy ISODATA clustering algorithms*, **IEEE Transactions on Pattern Analysis and Machine Intelligence** 2, pp.1-8.
- Bezdek, J.C. and Pal, S.K. (1992) (eds.): **Fuzzy Models For Pattern Recognition. Methods that Search for Structures in Data**. IEEE Press, New York.
- Bezdek, J.C., Ehrlich, R. and Full, W. (1984): *FCM: The fuzzy c-means clustering algorithm*, **Computers and Geosciences**, vol. 10, no.2-3, pp.191-203.
- Dubois, D. and Prade, H. (1980): **Fuzzy Sets and Systems: Theory and Applications**. Academic Press, New York.
- Fischer, M.M. (1982): **Eine Methodologie der Regionaltaxonomie**. University of Bremen Press, Bremen [= Bremer Beiträge zur Geographie und Raumplanung, vol.3].
- Fischer, M.M., Gopal, S., Stauffer, P. and Steinnocher, K. (1995): *Evaluation of neural pattern classifiers for a remote sensing application*, submitted to **Geographical Systems**.
- Gopal, S. and Fischer, M.M. (1994): *The application of artificial neural networks in remote sensing: Theoretical and methodological issues*. In: Ernste, H. (ed.): **Pathways to Human Ecology**, pp.17-36. Steiner, Wiesbaden.
- Higashi, M. and Klir, G.J. (1982): *On measures of fuzziness and complements*. **International Journal of General Systems** 8, pp.169-180.
- Klir, G.J. (1987): *Where do we stand on measures of uncertainty, ambiguity, fuzziness, and the like?* **Fuzzy Sets and Systems** 24, pp.141-160.
- Klir, G.J. and Folger, T.A. (1988): **Fuzzy Sets, Uncertainty and Information**. Prentice Hall, Englewood Cliffs.
- Kosko, B. (1992): **Neural Networks and Fuzzy Systems**. Prentice Hall, Englewood Cliffs.
- Pao, Y.-H. (1989): **Adaptive Pattern Recognition and Neural Networks**. Addison Wesley, Reading (Ma).
- Yager, R.R. (1979): *On the measures of fuzziness and negation. Part I: Membership in the unit interval*, **International Journal of General Systems** 5, pp.221-229.
- Zadeh, L.A. (1965): *Fuzzy Sets*. **Information and Control** 8, pp. 338-353.

Acknowledgement

The authors gratefully acknowledge Professor Karl Kraus (Department of Photogrammetric Engineering and Remote Sensing, Vienna Technical University) for his assistance in supplying the image data used in this study. This work is supported by a grant from the Austrian Fonds zur Förderung der wissenschaftlichen Forschung (P-09972-TEC)