Dealing with uncertainty is a common problem in pattern recognition. Rarely do object descriptions from different classes fall into totally disjoint regions of feature space. This uncertainty in class definition can be handled in several ways. In this paper we present several approaches to the incorporation of fuzzy set information into pattern recognition. We then introduce a new technique based on the fuzzy integral which combines objective evidence with the importance of that feature set for recognition purposes. In effect, the fuzzy integral performs a local feature selection, in that it attempts to use the strongest measurements first in the object classification. Algorithm performance is illustrated on real and synthetic data sets.

1. Introduction

The classification of objects from extracted features occupies a fundamental position in many areas of research interest and practical application. There are numerous approaches to the pattern recognition problem, [1]-[4] which can be organized by the different types of criterion functions which are used. Traditional approaches include decision theoretic methods, based on standard, or crisp, set theory, probabilistic techniques utilizing random models for uncertainty, and syntactic methods which rely on formal language theory. In the ideal case--well separated classes of tightly clustered objects--almost any recognition technique will work. However, this case is rarely found in real data. Hence, whatever approach is taken, knowledge about the structure and distributions of the data is extremely useful, if not crucial, to the successful classification of the objects of interest.

Fuzzy set theory was introduced by Zadeh in 1965 [5] as an alternative means of describing those situations where the defining characteristics of the sets themselves are vague or imprecise, and has been successfully used in many applications [3,4,7,10]. In this paper, we first present some of these methods for pattern recognition. We then present a new fuzzy pattern recognition algorithm which is based on the fuzzy integral with respect to a Sugeno measure.
2. Fuzzy Sets in Pattern Recognition

Given a universe $U$ of objects, a conventional "crisp" subset $A$ of $U$ is commonly defined by specifying the characteristic function of $A$, $\mu_A: U \rightarrow \{0,1\}$. Fuzzy sets are obtained by generalizing the concept of a characteristic function to a membership function $\mu: U \rightarrow [0,1]$. Most crisp set operations and set properties have analogs in fuzzy set theory.

Given a set of sample vectors $\{x_1, \ldots, x_n\}$, a fuzzy $c$ partition of these vectors specifies the degree of membership of each vector in each of $c$ classes. It is denoted by the $c \times n$ matrix $U$ where $u_{ik} = \mu_i(x_k)$ for $i = 1, \ldots, c$, and $k = 1, \ldots, n$ is the degree of membership of $x_k$ in class $i$. The following properties must be true for $U$ to be a fuzzy $c$ partition:

$$\sum_{i=1}^{c} u_{ik} = 1; \quad 0 < \sum_{k=1}^{n} u_{ik} < n; \quad u_{ik} \in [0, 1].$$

The most frequently used and documented fuzzy classification technique is the fuzzy $c$-means clustering algorithm [4]. It is an unsupervised approach which, like its crisp counterpart, looks for clusters in feature space by successively placing together all vectors which are "close to" the cluster center established for each class. In the crisp $c$-means, the vectors are assigned to respective clusters, while in the fuzzy $c$-means, their memberships in each class are updated during the iteration. The cluster centers then become a weighted average of the sample data [4]. Iterations continue until the process stabilizes.

In many pattern recognition problems, the classification of an input pattern is based on data where the respective sample sizes of each class are small and possibly not representative of the actual probability distributions, even if they are known. Under many circumstances, the K-nearest neighbor (K-NN) algorithm [6] is used to perform the classification. This decision rule provides a simple nonparametric procedure for the assignment of a class label to the input pattern based on the class labels represented by the $K$-closest neighbors of the vector.

One of the problems encountered in using the K-NN classifier is that normally each of the sample vectors is considered equally important in the assignment of the class label to the input vector. The fuzzy K-nearest neighbor algorithm [7] assigns class membership to a sample vector rather than assigning the vector to a particular class. The advantage is that no arbitrary total assignments are made by the algorithm. In addition, the vector's membership values provide a level of assurance to accompany the resultant classification. The basis of the algorithm is to assign membership as a function of the vector's distance from its K-nearest neighbors and those neighbors' memberships in the possible classes [7].

The labeled samples can be assigned class memberships in several ways. First, they can be given complete membership in their known class and nonmembership in all other classes. Other alternatives are to assign the samples' membership based on