STUDIES IN POSSIBILISIC INTERVAL RECOGNHION

Cao wenming
Nanjing city work university ,Zhongshan east-road No.54 Nanjing City,210005
Jiangsu,P.R.of China
Hu kedin Song wenzheng
Automation Res. Southeast University ,Nanjing 210018,Jiangsu P.R. of China

This paper introduces an algorithm for pattern recognition. The algorithm will classify a measured object as belonging to one of *N* known classes or none of the classes. The algorithm makes use of fuzzy techniques and *possibility interval* is used instead of probability. The algorithm was conceived with the idea of recognizing fast moving objects, but it is shown to be more general. Fuzzy ISODATA's use as a front end to the algorithm is shown. An expected value for a possibility interval distribution is also investigated. The algorithm actually proves to be adaptable to a wide variety of imprecise recognition problems. Some test results illustrate the use of the technique embodied in the algorithm and indicate its viability.

Keywords: Possibility interval , Fuzzy interval, Pattern recognition, Possibility interval distribution descriptors, Distributions expected value, Expert systems.

1. Introduction

In this paper a general pattern recognition technique is presented. It is in the form of an algorithm. The technique makes use of fuzzy techniques in place of binary and possibility in the place of probability interval. The calculus of fuzzy restrictions provides an underlying theory for the ideas present in this paper, however it is not used explicitly. The technique is intended to classify a measured object as belonging to one of *N* known classes. We will have enough information about the *N* classes to be able to fuzzily discriminate among objects in them.

Most pattern recognition problems deal with classes which have overlapping boundaries. They also deal with measurements in noisy environments and inexact measurements in general. They do not lend themselves to precise formulation and therefore a systematic method of dealing with imprecision is necessary. Possibility interval distributions which are based on fuzziness will provide our method.

As an example of the imprecision often encountered in pattern recognition problems consider the problem of recognizing handwritten characters. Many people write their B's so that they resemble 8's and vice-versa. The reader of a handwritten message can often discern the letter from the context of the text that surrounds it. In order for a pattern recognizer to recognize classify the handwritten B or 8 it must take into account the fuzzy boundaries of the characters. Even taking into account their fuzzy nature the context may still be necessary to simple example points out the pervasive fuzziness of pattern recognition problems.

The use of fuzzy techniques in the form of possibility interval distributions[1] will provide us with some advantages, over probabilistic methods, which we will describe. The calculation of a probabilistic interval expected value[2] requires a large number of observations be made. Our approach requires a possibility interval distribution descriptor to be described later. It is not analogous to the expected value and does not requested many observations. A traditional binary decision requires that an object either belong to a class or not. If we make a mistake we are 100% wrong. When we allow grades of membership in a class an algorithm will be more forgiving and less likely to be thrown off the track by an incorrect intermediate conclusion. Fuzzy techniques will lessen the severity of errors and lower the probability of misclassification.

2. Possibility interval [1,3]

In this section we will review and expand upon the necessary possibility concepts for pattern recognition. We will begin with a definition, in general terms, of a possibility distribution. Zadeb in [4] introduced the fuzzy concept of possibility and our definitions follow from that work

Definition 2.1 A possibility interval distribution is defined to be a triple

μ_A: $\{gm_{min}, gm_{max}, u\}$ for all u in U

in which A is a possibility interval set, gm_1, gm_2, is the minimum possible grade of membership of object u in U, and gm_2 is the maximum possible grade of membership of u in U

We further define a universe and a zero set

1:= (1,1, U) for all u in U,

0:=(0,0,U) for all u in U.

where in our application to an expert system U is the set of all possible states of pertinent working memory.

Definition 2.2 Let F be a fuzzy subset of a universe of discourse U which is characterized by its membership function with the grade of membership, μ interpreted as the compatibility of u with the concept labeled F

Let Y be a variable taking values in U and let F act as a fuzzy restriction, $R(Y)$, associated with Y. Then "Y is F" is a proposition which translates into

$$R(Y)=F \tag{2.1}$$

and associates a possibility distribution, Π_Y with Y which is postulated to be equal to $R(Y)$, i.e.

$$\Pi_Y=R(Y). \tag{2.2}$$

To go along with the above, the possibility interval distribution function associated with *Y* (which can be called the possibility distribution function of Π_Y) is denoted by Π_Y and is defined to be numerically equal to the membership function of *F*, i.e.,

$$\Pi_Y \cong \mu_F \qquad (2.3)$$

Thus, $\pi_Y(u)$, the possibility that $Y = u$, is postulated to be equal to $\mu_F(u)$.

From (2.2) the possibility interval distribution Π_Y may be regarded as an interpretation of the concept of a fuzzy restriction. As a consequence the mathematical apparatus of the theory of fuzzy sets and the calculus of fuzzy restrictions [5] provides a theoretical basis for the manipulation of possibility interval distributions. Calculus was not used in our manipulations which indicates their easily implementable nature.

It follows from the above definitions that our intuitive perception of the ways in when possibilities combine is in accord with the rules of combination for fuzzy restrictions. This assumption has, as far as is known, not been proven but was used in some of the thinking used to produce the algorithm presented in this paper.

It is of some value to touch on the relationship between possibility and probability. A high possibility does not imply a high probability and the lessening of the possibility of an event tends to lessen its probability but not vice-versa.

3. Possibility interval distribution operation

In this section we will show how possibility interval distributions arise in our pattern recognition technique. Some operators for manipulating and describing possibility distributions will be introduced.

The algorithm of this paper operates on features which are descriptors of the N classes to which an observed object may belong. These features are then the components of vectors, which are called feature vectors. The algorithm operates on a single feature vector which should be representative (a fuzzy average) of several different measurements of each feature of an object. The algorithm will attempt to classify the feature vector as belonging to one of the classes. The ideal or actual measurements for each of the *N* classes must be known.

Possibility comes into play when given a feature i we ask, "How well is the constraint 'feature i belongs to class M' satisfied when i is assigned to *Y*?". This question will be asked for each of the *N* classes, which will give us a possibility interval distribution for each feature i. Then some manipulations will be made to find the best M ($1 \leq M \leq N$) for the proposition "feature i belongs to class M". After this is done for all features a classification can be made based on criteria to be defined.

With possibility distributions comes the natural question of how to describe a distribution or set of distributions as an entity. The answer to that question is to define a Possibility Interval Distribution Descriptor (PIDD).

First let us define one type of PIDD which is interesting when we try to obtain some type of average value from a distribution. We will work with distributions in which impossibilities, possibilities of 0, are to be ignored. The PIDD does not need to be monotonously increasing with respect to its arguments and in fact is not. A high degree of possibility should influence our descriptive value more than the lower ones, since we have more confidence that it can occur. A possibility denotes more the ability of some action or object to occur than it does its probability. Therefore if we are unsure of even the occurrence of an action or object, we do not wish it to have a great effect upon our description of a distribution. A low degree of possibility certainly represents improbability.

Definition 3.1. Given a possibility interval distribution

$$\{X_{min1}/W_1+X_{min2}/W_2+\ldots+X_{minn}/W_n,\ X_{min1}/W_1+X_{min2}/W_2+\ldots+X_{minn}/W_n, U\}$$

the value of the possibility interval distribution descriptor is

$$PIDD=\{\ \Sigma_{i=1}^{n} X_{mini}^2/\Sigma X_{mini=1}^{n},\ \Sigma_{i=1}^{n} X_{maxi}^2/\Sigma X_{maxi=1}^{n}\}=[PIDD_{min}, PIDD_{max}].$$

This definition will give us a set number between 0 and 1 with the desire properties. We would like to be able to say we have a PIDD with an associate representative value, *R* for a distribution. We therefore need a way to find *R*.

Definition 3.2. Given a possibility interval distribution and a possibility interval distribution descriptor the *R* associated with the PIDD is

$$R=\{\Sigma X_{mini}W_{mini}/\Sigma X_{mini},\ \Sigma X_{mini}W_{mini}/\Sigma X_{mini}\}=[R_{min}, R_{max}]$$

The *R* is heavily influenced by the size of each possibility . It will tend to be gloss W_i's that are associated with the larger possibilities. The combination of the PIDD and *R*, written as PIDD/R, of a distribution will give an indication of what valus to expect from the distribution and how possible that value is (how good a indication it is).

For an example of the use of these definitions let's look at the case of a grocer who wants to know how many eggs to buy each day. He may have the follow in, distribution for his town:

{1/100+0.9/200+0.7/300+0.5/400+0.2/600+0.1/800, 1/100+1/200+0.8/300+0.6/400+0.3/600+0.2/800,U}

Then the PIDD for the distribution will be

$$PIDD_{min}=0.797, \qquad PIDD_{max}=0.805$$

R will be

$$R_{min}=260, \qquad R_{max}=287$$

We get, as a representation of our distribution, [0.797/260, 0.805/287] from Definitions 3.1 and 3.2. This can be interpreted as meaning that there is a 0.797 possibility that 260 eggs and 0.805 possibility that 287 eggs will be eaten in one day in the grocers town. The grocer may then decide to use this information to buy 290 eggs per day.

The techniques of Definitions 3.1 and 3.2 can also be used to describe several related possibility distributions. It will indicate what should be expected from them as a whole. The definitions must be extended a little to describe several distributions with one value.

The above definitions are applicable for the example shown. They are also applicable for the type of problems that

Zadeh [4] discusses. However, they do not accurately describe the kind of situation that pattern recognition problems generate. We will describe the properties a PIDD for pattern recognition should have and develop a useful PIDD in the following.

Consider the fact that a 0 possibility for a W_i will not change the PIDD of the distribution without the 0 possibility. In pattern recognition problems all possibilities or impossibilities must enter into the calculation of the PIDD.

When you have a feature vector and are comparing a feature against each of N classes a 0 possibility does give you some information. It indicates that the conceivability of the feature belonging to that class is negligible. That fact should be taken into account when the PIDD is computed.

The PIDD function must return a value between 0 and 1. We wish it to be monotonously increasing with respect to its arguments. The value of the function must be c when we operate on the set $\{x_i | x_i = c\}$ where $1 < i < N$ and $0 < c < 1$. The value of the PIDD for a distribution must be symmetric in the sense that the ordering of the x_i is not important. As an event becomes less possible we wish to take it into account less due to the nature of possibility, therefore we want the possibilities closer to one to sway the distribution more. The PIDD function should be convex for each variable separately or approximate this concept.

Let us introduce a PIDD which will have the properties discussed above. This function will be used in our algorithm.

Definition 3.3 Given a possibility interval distribution $X_1+X_2+\ldots+X_n$, possibility interval distribution descriptor will be given by

$$PIDD=\{ \Sigma_{i=1}^{n} X_{mini}^{2}+\Sigma X_{mini=1}^{n}/\Sigma X_{mini=1}^{n},\ \Sigma_{i=1}^{n} X_{maxi}^{2}+\Sigma X_{maxi=1}^{n}/\Sigma X_{maxi=1}^{n}\}=[PIDD_{min}, PIDD_{max}].$$

Theorem 3.1

$$(\Sigma_{i=1}^{n} A_i^2+\Sigma A_{i=1}^{n})/(\Sigma A_{i=1}^{n}+N) > (\Sigma_{i=1}^{n} a_i^2+\Sigma a_{i=1}^{n})/(\Sigma a_{i=1}^{n}+N)$$

with $A_i > a_i$ and $0<A_i<1$, $0<a_i<1$, $\forall i$.

Theorem 3.1 indicates that whenever the possibility of every component of a distribution is greater than the possibility of the corresponding component of another distribution for a non unique ordering, the PIDD of the first distribution will be greater than that of the second. In fact it is shown in the appendix that only one $A_i > a_i$ (with the others equal) is necessary in order for this to be true. The PIDD function is also symmetric in the sense that its value will be the same for any ordering of the A_i.

For each feature of a feature vector a conditional possibility distribution will be generated. We need a way, given a conditional possibility distribution for a feature, to assign that feature to a class.

Definition 3.4. Given that x and w_i are non-interactive, but not necessarily independent,

$$\pi(w_i|x)=\begin{cases}\pi(w_i) & \text{for } \pi(w_i|x)<\pi(x), \quad (1)\\ [\pi(w_i), 1] & \text{for } \pi(w_i)>=\pi(x). \quad (2)\end{cases}$$

(2) will also be taken to be $\pi(w_i)$.

Note that $\pi(x)$ will be, for the purposes of this paper, the possibility that the measurement of feature x is good. It will be considered 1 for all x.

Given a set of w_i's and Definition 3.4 we need a way to assign x, which will be our feature, to one of the classes w_1. Analogous to the theory of probability the following definition will be used.

Definition3.5. Feature x will be assigned to the class i if

$\pi(w_i|x) > \pi(w_j|x)$, $j=1,2\ldots N$, $j\neq i$ where N is the number of classes.

Let a abs measure be defined by

$$abs([a,b]-[c.d])=[|a-c|,|b-d|], abs([a,b])=[|a|,|b|]$$

if $x=[a.b]$, $y=[c,d]\in P([0,1])$ then $abs(x-y)/abs(x)=([|a-c|/|b|,|b-d|/|a|])$

The function that will serve as a fuzzy restriction on the universe of fast moving aircraft, for the purposes of his paper, will be the following.

Definition 3.6 (a fuzzy interval restriction). Given a measured feature x and N classes for which the actual values of the object that x describes are known,

$$f(x, \text{class } i)=1-abs(x-\text{class } i)/abs(x)$$

where x, class $i\in P([0,1])$ and $P([0,1])$[2] be set of subset of $[0,1]$, $1=[1,1]|$, class i is defined to be the value of the object that x measures for class i. If $f(x, \text{class } i)\leq 0$ then $f(x, \text{class } i) = 0$.

The last statement of the definition implies that if the distance of class i from x greater than x, then the fuzzy interval of x being in class i is 0.

In words Definition 3.6 corresponds to the statement "Feature x is of class w_I". take a feature x whose degree of compatibility with the class w_I is given by 3.6. en it can be postulated that the proposition "Feature x is of class w_I", converts compatibility value from the degree of compatibility of x with the concept it is class w to the degree of fuzzy that x would be the measured value of the feature given the proposition "Feature x is of class w_i". he definitions and theorem of this section form the basis for the pattern recognition technique that will be presented in the next section.

4. The technique

Since the original thrust behind this paper was to use possibility in pattern recognition to recognize fast moving airplanes, most of the explanations and descriptions from here on will involve that problem .Let us assume that s

features that describe an airplane have been found and that these s features do a good job of differentiating between airplanes. As an airplane travels along several different measurements will be taken of that airplane's S features. Say K measurements are made, then we will have K feature vectors each of s features. what we need is one representative feature vector, which can then be classified as belonging to one of the N classes of airplanes we are looking for. In order to be consistent with the ideas behind this paper we would also like to use fuzziness in the process of finding a representative feature vector of our data.

In order to find a representative feature vector the fuzzy ISODATA algorithm (with $m = 1.4$) as similar presented by Bezdek and Castelaz[6] is used in the manner described by the following.

First a description of the fuzzy ISODATA algorithms is in order. The algorithm requires that a predetermined number of classes (c) greater than or equal to 2 be searched for. The algorithm takes n feature vectors, each of S features, and finds C duster centers for the data. It gives each of the n feature vectors a grade of membership for each of the C cluster centers. Two cluster centers are searched for. The membership value of each feature to cluster center should be within some epsilon. If this is not the case we consider the data to be inconsistent and do not proceed further. If the data is consistent we take the average of the two cluster centers and operate our algorithm on the resulting representative feature vector of s elements.

Properly used fuzzy ISODATA will provide a representative feature vector which the possibility interval distribution based pattern recognition algorithm, that will now be presented, will operate on.

The algorithm presented here consists of five main steps. They will be listed and then explained. It is necessary, a priori, to know the actual value of each feature in the feature vector for every class that an object may be put in.An algorithm for possibility based pattern recognition using fuzzy techniques is as follows:

Algorithm 4.1. (1) Given n feature vectors of s features each that are taken from measurements of an object; use fuzzy ISODATA to get a representative feature vector X of the data. It is also necessary that for each of the S features an actual value is known for each of the j classes to be used for classification purposes.

(2) Get the possibility interval distribution associated with each x_i (the ith feature of the representative feature vector):

$$Pidist_{ij} = f_i(x_i, y_{ij})$$

where y_{ij} is the actual value of feature i for class j. Also $Pidist_{ij} \leq 0 \Rightarrow Pidist_{ij} = 0$. Each function f_i corresponds to the fuzzy restriction for feature i and may be distinct.

(3)This step is obtained from definitions 3.4 and 3.5.

(A) Classify each x_i as belonging to class I if $Pidistj > Pidist_K$, $j \neq k$

(B) Find the class that x_i belongs to after class i. That is, find the class that has the second highest possibility (after j) of' x_i belong to it. The second most possible class is A if $Pidist_{iL_i} > Pidist_{iK}$, $L_i \neq K$, and $m \neq L_i$. Note that a different class can be chosen for each row i.

(4). (A) If all the x_i belong to class i then get the possibility interval distribution descriptor $Pidist_{ij}$ call it PIDDl. Also get the possibility interval distribution descriptor gotten from step 3(B), call it. PIDD2. PIDD2 will be the PIDD1 of the compendium of the number 2 classes.

$$PIDD1 = \{ (\Sigma_i^s{}_{=1} Pidist_{ij}^2 + \Sigma_{i=1}^s Pidist_{ij}) / (\Sigma_{i=1}^s Pidist_{ij} + s), \ (\Sigma_i^n{}_{=1} Pidist_{ij}^2 + \Sigma_{i=1}^s Pidist_{ij}) / (\Sigma Pidist_{iji=1}^n + s) \} = [PIDD_{min}, PIDD_{ma}]$$

$$PIDD1 = \{ (\Sigma_i^s{}_{=1} Pidist_{iL_i}^2 + \Sigma_{i=1}^s Pidist_{iL_i}) / (\Sigma_{i=1}^s Pidist_{iL_i} + s), (\Sigma_i^n{}_{=1} Pidist_{iL_i}^2 + \Sigma_{i=1}^s Pidist_{iL_i}) / (\Sigma Pidist_{iL_i i=1}^n + s) \} = [PIDD_{min}, PIDD_{max}]$$

where A is the class gotten from step 3(B), for each i.

(B)If all the x_i do not belong to class i, stop and declare that impossible and a new set of data is needed.

(5) If 4(A) was satisfied, then if PIDD1 > T and PIDD1 - PIDD2 > D, classify the object described by feature vector X as belonging to class i. Otherwise no decision will be made and new data or different features could be tried. The particular T and D used will depend on the application classification is. This algorithm is flexible enough to allow multiple applications due to the multiplicty of functions available in step two. We will use the same function for all features in our example. The function is the one defined in 3.6.

Note that step 4(B) indicates either some bad measurements were made or the feature selection was not good enough. It is a built in error check.

The requirement of step 5, that PIDDl be greater than T is used in case an object has been measured that does not belong to one of the N classes that are being used for classification purposes. The requirement that the difference between PIDDI and PIDD2 be greater than D tends to guard against the conceivability of misclassification. In many of our tests the values $T = 0.7$ and $D = 0.1$ were used.

This algorithm will work for any application in which knowledge of the N classes that an object to be measured can belong to is known. The possibility distribution descriptor function or fuzzy restriction function will need to be changed to be consistent with the application desired.

If one had many feature vectors and wished to see how they clustered up, this algorithm would not be useful. Instead an algorithm of the fuzzy ISODATA type should be used. This algorithm should be particularly useful for the recognition of fast moving objects that may belong to one of N classes.

An important part of any algorithm for pattern recognition is the selection of the features that make up the feature vector. How many features should there be and what features should be used? The number of features should be as few as possible for an accurate classification to occur. The type of features used should be those that best differentiate between classes.

Note that if the features do not differentiate well between classes, the requirement that PIDDI - PIDD2 >0.1 may have to be relaxed. This will depend on the actual application.

Most of Algorithm 4.1 is taken from previous definitions and its veracity is evident. However, the classification done in steps 3(A) and 3(B) is not completely taken from previous definitions. Intuitively it seems that the steps would be true. fl' is in fact easy to prove them true by induction and this has been done.

Algorithm 4.1 is the primary result of this section. It will classify a feature vector as belonging to one of *N* classes or none of them, which may indicate an inconsistent feature vector has been used. It uses fuzzy techniques in the form of possibility distributions which more accurately describe actual situations than binary and probabilistic techniques may. Therefore this classifier should be at least as accurate as a standard classifier. The algorithm presented in this section lends itself very easily to the form of a computer program. A program in C has been written to implement the algorithm. A fuzzy ISODATA program, also in C, has been written as a front end, step 1. The following section of this paper will show an example of the use of the algorithm and the results. Some other tests will be discussed and conclusions drawn.

5. Analysis and conclusions

let us make use of an example problem to clarity the algorithm. Our example will make use of some actual data that describes Russian and American airplanes. A feature vector of three features has been chosen. The features are chosen for their simplicity, not optimality. The fact that simple features are often not optimal points out the need for a good method of selecting features. Step 5 of Algorithm 4.1 will be relaxed so that PIDD1 - PIDD2> 0.04 will suffice for classification to occur.

Example 5.1 The data used in this example is from [6]. The two classes that an object is to be classified in are: (1) American fighter airplanes with top speeds of almost 1300 mph. (2) Russian fighter airplanes with top speeds of about 1300mph.

Three features will be used for classification, they are:

(1) airplane length,
(2) airplane height,
(3) wing span.

The known values for a Russian airplane of this type are:

(1) 15.7m~15.78m, (2) 4.5m~4.6m, (3) 7.16m~7.20m.

The known values for an American airplane of this type are:

(1) 17.00m~17.20m, (2) 4.00m~4.10m, (3) 6.5m~6.60m.

let's assume that an American F- 104 star fighter flies by and the following measurements are obtained in its representative feature vector after the operation of fuzzy ISODATA. The values for features 1, 2, and 3 are

(1)16.69m~16.72m, (2) 4.15m~4.21m, (3) 6.68m~6.70m.

Then from step 2 of the algorithm the Pidist matrix will look like

$$\text{Pidist}=\begin{matrix}[0.9814\ ,\ 0.9816] & [0.9436,0.9466] \\ [0.9638\ ,\ 0.9639] & [0.9156,\ 0.9187] \\ [0.9731,\ 0.9742] & [0.9281,0.9292]\end{matrix}$$

Clearly each element of column one is greater than each element of column two for each row i. Therefore step 3 will classify each feature 'ci as belonging to class 1 and the number two class is then class 2. Since step 4(A) is satisfied PIDD1 and PIDD2 will be calculated:

PIDD1 =[0.97279,0.97614] PIDD2 =[0.92916,0.9301].

From step 5, PIDDl - PIDD2 =[0.97279 - 0.92916,0.97614-0.9301]=[0.04363,0.04604] Since [0.04363,0.04604]> 0.04 the feature vector will be classified as belonging to class 1.

Class 1 is the class of American airplanes that have top speeds of =1300 mph. The. F-104, which was measured, is an American airplane with top speed 1300 mph; therefore it does belong to class 1. For this simple example Algorithm 4.1 has made a correct classification.

The algorithm has been tested on the problem of determining to which constellation a star belongs. In these tests the algorithm classified a star as belonging to the proper constellation or not at all. There were no misclassifications, which is one characteristic that we feel is very important. In the few cases that a classification was not made the circumstances seemed reasonable.

In the following we will suggest some extensions to the algorithm via mechanisms, which may be implementable as part of a fuzzy expert system .

One mechanism would determine when restrictions should be placed on the use of a possibility interval distribution obtained from a feature in the proceeding calculations. For instance if the top three interval possibilities in a distribution are within 0.05 of each other that distribution could be ignored in the proceeding calculations. If the feature proved consistently poor on a specific data set then the possibility distribution induced from it need not be calculated. If the features were good differentiators between classes nothing of this son would need to be done, but it may not always be possible to find good features.

The mechanism could also perform a weighting of features. The features could be weighted by some criteria according to the distributions obtained from them. For instance if one feature belonged to one class very~strongly ([0.9,1]) and to the other classes with a possibility of 0 the feature could be weighted 1 and other features weighted could be weighted in a similar manner. The classification could be made on a small number of excellent features.

The weighting of features and the disregarding of those that provide little information would allow us to effectively make use of more features. We would also be able to allow the relaxation of the restriction that the highest interval possibility for each feature must belong to the same class, which may be too restrictive for some applications. We could also discontinue using a feature if it continually provided no discriminatory information.

In order to obtain a representative feature vector from several measurements an algorithm such as the one suggested by Definitions 3.1 and 3.2 may be used. It has been used and has provided a useful feature vector in tests. The PIDD's associated with the representative values can be thought of as weights. The suggested weights could be used in some, as yet undetermined, fashion.

The extensions to the algorithm that we have suggested will need to be tuned to specific applications. There appears to be a place for a fuzzy expert system which could determine weighting factors and the worth of distributions among other things. It would enable less tuning and experimentation to be done by the users of the algorithm.

The use of possibility interval distributions has enabled us to develop an algorithmic framework for a viable pattern recognition technique. It is easily implemented as a computer program and will not require exorbitant CPU time. Since the resources were unavailable, we were unable to compare our algorithm against others with some consistent data. Since fuzziness tends to minimize the severity of errors and provide flexibility we feel that our algorithm performs at a high level. The tests that have been made have given us encouraging results. less training on new data should be required and this is an advantage. Since we do not attempt to perform crisp matches, noise and imprecise measurements will not have as great an effect as they may on other recognizers.

The use of possibility interval, based on fuzzy interval sets, provides a method for a viable pattern recognition technique. Its ability to deal with the imprecision inherent in many pattern recognition problems will enable it to perform as well or better than conventional recognizers.

References

[1] J. Buckley and W. Siler Fuzzy operators for possibility interval sets , Fuzzy Sets and System 22 (1987) 215-227

[2] Cao Wenmin ,Hu kedin, Song Wenzhong Weighted fuzzy interval logic and its application, Journal of southeast university Vol.25 No.5A sept. 1995

[3] Cao Wenmin, Hu kedin, Song Wenzhong Application for Fuzzy interval expect, Journal of southeast university .Vol 26 No.5A Sept. 1996

[4] R.E. Bellman ,R Kalaba and L.A Zedeh Abstraction and pattern classification, J. Math. Anal Appl. 13(1966)1-7

[5] R.E Bellman and L. A Zedeh Local and fuzzy logic in :D.Epstein ,Ed., Modern Uses of Multiple-valued logics,(Reidel ,Dordrecht, 1977).

[6] Cchant ,Editor ,Military Aircraft of the World (Presidio Press, Novato, CA 1981)