Reprinted from the *Journal of the Institution of Electronics & Telecom. Engrs.*, 1977 Vol. 23, No. 3, pp 117-120

On Some Applications of Fuzzy Algorithm in Man-Machine Communication Research

D. DUTTA MAJUMDER* (Member) AND S. K. PAL*

(*Manuscript received on* 20 *February* 1976)

Present paper ennunciates some practical applications of fuzzy set theory in problems of man-machine communication. Problems are recognition of vowel speech sounds and identification of speakers from spoken words. Data used are derived from acoustic-phonetic and spectrographic analysis of large number of Hindi and Telugu (two of the major Indian Languages) speech sounds. It is explained that Fuzzy set theory provides a suitable algorithm which is substantively different from the conventional quantitative methods of system analysis yet presents an approximate but effective means of describing the behaviour of systems which are too ill-defined for precise mathematical analysis. In this paper, two decision algorithmic methods developed from fuzzy set theory and applied for machine recognition of vowels and identification of speakers with Telugu and Hindi speech sounds are presented along with the results of experiments.

Indexing Terms : Fuzzy Set, Formant Frequencies, Membership Function, Property Set, Weighting Co-efficient, Similarity Vector, Pattern Recognition, Recognition Score

PRESENT paper is a part of the work being conducted on man-machine communication research (Dutta Majumdar et al. 1968*a*, 1968*b*, 1969, 1975, 1976, Dutta 1974*a*, 1974*b*) at the Electronics and Communication Science Deptt. of Indian Statistical Institute. There are various approaches leading to the machine recognition of patterns as suggested by different authors (Duta Majumder et al. 1975, Dutta 1974*a*, Sebestyen 1962, Fu 1968). But most of the situations in the field of natural and social sciences are too ill-defined to be analysed. To study such complex behaved systems, the concept of fuzzy sets and its subsequent developments in decision processes (Zadeh 1965, 1973) could be applied to a reasonable extent. This algorithm is approximate but provides an effective and more flexible basis for analysis of systems which are not precisely defined.

This paper would first of all, present briefly an introduction to fuzzy sets with associated operations, then describe two methods for classificatory analysis on the basis of fuzzy algorithm and finally their implementation on some practical problems. First method is based on computation of weighted distance function and second method is based on computation of the similarity vectors of the property sets. Firstly, machine recognition of vowel speech sounds both for the Telugu and Hindi spoken words and secondly, speaker identification using the same data were experimented with. But the recognition algorithm for speaker identification could be further extended to any number varying in sex and age. All the data processing was conducted on the Electronics Digital Computer Honeywell 400.

*Deptt. of Electronics & Communication Science, Indian Statistical Institute, Calcutta 700 035.

FUZZY SET AND ITS OPERATIONS

A class of events $x_1, x_2, \ldots\ldots, x_n$ in the universe of discourse U is defined as a fuzzy set (A), if their transition from membership to non-membership is continuous rather than abrupt. $x_1, x_2, \ldots\ldots, x_n$ represent the supports of A at which the value of membership function $\mu_A(x)$ characterising the grade of membership of x in A is positive ranging between zero and one. A fuzzy set A with the finite number of supports could therefore be viewed as

$$A = \{(\mu_A(x_i), x_i)\}. \quad (1)$$

In summation form,

$$A = \mu_1/x_1 + \mu_2/x_2 + \ldots\ldots + \mu_n/x_n$$

$$\text{or, } A = \cup_i (\mu_i/x_i), \; i = 1, 2, \ldots\ldots, n \quad (2)$$

where $+$ sign stands for the union of supporting points, rather than arithmetic sum. μ_i being the grade of membrship of x_i in A denotes the degree to which an event x_i may be a member of A or belong to A, and as it approaches unity, the grade of membership of x_i in A becomes higher. For example, $\mu_A(x_i) = 1$ indicates strictly the containment of the event x_i in A and if x_i on the other hand does not belong to A, $\mu_A(x_i) = 0$. Any intermediate value would represent the degree upto which x_i could be a member of A.

If the support of a fuzzy set is only a single point $x_1 \in U$, then

$$A = \mu_1/x_1 \quad (3)$$

is called fuzzy singleton. Thus $A = 1/x_1$ for $\mu_1 = 1$ would obviously denote a non-fuzzy singleton. Basic operations related to fuzzy sets A and B having membership values $\mu_A(x)$ and $\mu_B(x)$ respectively are summarized here.

(*i*) A is equal to B ($A = B$)

$$\Rightarrow \mu_A(x) = \mu_B(x), \text{ for } \forall x \quad (4)$$

(*ii*) A is complement of B ($A = \neg B$)

$$\Rightarrow (i - \mu_B(x)) = \mu_A(x), \text{ for } \forall x \quad (5)$$

(*iii*) A is subset of B ($A \subset B$)

$$\Rightarrow \mu_A(x) \leqslant \mu_B(x), \text{ for } \forall x \quad (6)$$

(*iv*) Union of A and B ($A \cup B$)

$$\Rightarrow \mu_{A\cup B}(x) = \vee [\mu_A(x), \mu_B(x)], \text{ for } \forall x \quad (7)$$

$\vee$ denotes maximum

(*v*) Intersection of A and B ($A \cap B$)

$$\Rightarrow \mu_{A\cap B}(x) = \wedge [\mu_A(x), \mu_B(x)], \text{ for } \forall x. \quad (8)$$

$\wedge$ denotes minimum

(*vi*) Product of A and B (AB)

$$\Rightarrow \mu_{AB}(x) = \mu_A(x)\, \mu_B(x), \text{ for } \forall x. \quad (9)$$

In addition to the above mentioned operations, there are other operations applied on fuzzy set A. These are,

(*i*) Concentration of A ($CON(A)$)

$$\Rightarrow \mu_{CON(A)}(x) = [\mu_A(x)]^2, \text{ for } \forall x \quad (10)$$

(*ii*) Dilation of A ($DIL(A)$)

$$\Rightarrow \mu_{DIL(A)}(x) = [\mu_A(x)]^{1/2}, \text{ for } \forall x \quad (11)$$

iii) Contrast Intensification of A ($INT(A)$)

$$\Rightarrow \mu_{INT(A)}(x) = \begin{cases} 2[\mu_A(x)]^2 & , 0 \leqslant \mu_A(x) \leqslant 0.5 \\ [1-2(1-\mu_A(x))^2], & 0.5 \leqslant \mu_A(x) \leqslant 1.0. \end{cases} \quad (12)$$

All these operations have the effect of altering the fuzziness of a set. Effect of $DIL(A)$ is opposite to that of concentration which reduces the magnitude of $\mu_A(x)$ in relatively smaller amount for those x having higher membership value in A compared to those with low μ_A value. Contrast intensification as its name implies, reduces the fuzziness of A by increasing the value of $\mu_A(x)$ which are above 0.5 and decreasing those which are below it.

DECISION ALGORITHM

Consider an N-dimensional feature vector space Ω_x containing m ill-defined pattern classes $C_1, C_2, \ldots C_j, \ldots C_m$ to be recognized with defined set of N-dimensional prototypes $R_1, R_2, \ldots R_j, \ldots R_m$ such that

$R_j^{(l)} \epsilon R_j$

$l = 1, 2, \ldots h_j$; h_j is the number of reference vectors in set R_j.

Present classification model aims to assign an unknown event with N number of features

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \\ \vdots \\ x_N \end{bmatrix}$$

in Ω_x, to its appropriate class C_j for which its μ-value $\mu_j(x)$ or magnitude of the similarity vector $S_j(x)$ possesses maximum value where x_n represents the nth measured feature the event has.

Method I

If W_{jn} (having magnitude less than or equal to unity) and R_{jn} associated with jth class represent the magnitude of weighting co-efficient and reference feature respectively for nth component and x_{ijn} is the numerical value of nth component for ith token in C_j containing M no. of events such that

$$W_{jn}^2 = \frac{1}{\frac{1}{M}\sum_{i=1}^{M} x_{ijn}^2 - R_{jn}^2} \quad (13)$$

$$R_{jn} = \frac{1}{M}\sum_{i=1}^{M} x_{ijn}. \quad (14)$$

Then membership function of an unknown pattern X corresponding to jth class could be expressed as

$$\mu_j(x) = \left[1 + \left[\frac{d(x, R_j)}{E}\right]^F\right]^{-1} \quad (15)$$

where E is an arbitrary positive constant and F is any integer. These constants are included only to affect the fuzziness of a set (Dutta Majumder and Pal 1976). Here $d(X, R_j)$, the distance between X and prototype R_j is defined† as the smallest of the distances between X and each reference vector $R_j^{(l)}$ in R_j. Thus

$$d(X, R_j) = \wedge_l \left| X - R_j^{(l)} \right|$$

$$\text{and } \left| X - R_j^{(l)} \right| = \left[\sum_{n=1}^{N} [W_{jn}(X_n - R_{jn}^{(l)})]^2 \right]^{1/2} \quad (16)$$

where $l = 1, 2, \ldots\ldots\ldots\ldots, h_j$

$j = 1, 2, \ldots\ldots\ldots\ldots, m.$

Equation (15) defines the grade of membership of X in C_j and its μ-value would be 1 for $d(X, R_j) = 0$. Thus an unknown pattern X could be identified to be a member of kth class if $\mu_k(X)$ is maximum i.e.

$$\text{if } \mu_k(x) = \vee [\mu_j(x)], \text{ decide } X \epsilon C_k \quad (17)$$

j and k may have any integer value from 1 to m.

Method II

Let $p_1, p_2, \ldots\ldots\ldots, p_n \ldots\ldots\ldots\ldots, p_N$ be the N properties each of which represents some aspects of the unknown pattern X and has value only in the interval [0, 1] such that

$$P_n = \left[1 + \frac{|\bar{x}_n - x_n|^F}{E}\right]^{-1} \quad (18)$$

†This is one of the most accepted definitions of the distance function used in statistical classificatory analysis (Rao 1952, Sebestyen 1962).

$$\text{where } \bar{x}_n = \vee_j \frac{1}{h} \sum_{l=1}^{h} x_n^{(l)} \tag{19}$$

h = no. of reference vectors in a class. Constants E and F have the same effects as in the previous method in affecting the fuzziness of a set.

There are h no. of reference prototypes in a class C_j and each prototype may be represented as

$$R_j^{(l)} = \left\{ p_{1j}^{(l)}, p_{2j}^{(l)} \ldots\ldots p_{nj}^{(l)} \ldots\ldots p_{Nj}^{(l)} \right\} \tag{20}$$

$$l = 1, 2, \ldots\ldots\ldots h$$
$$j = 1, 2, \ldots\ldots\ldots m$$

$P_{nj}(l)$ denotes the property p_n of the lth prototype in C_j. Then the similarity vector $S_j(X)$ for the pattern X with respect to jth class is

$$S_j(x) = \left\{ s_{1j}, s_{2j}, \ldots\ldots s_{nj} \ldots\ldots, s_{Nj} \right\} \tag{21}$$

where numerical value of s_{nj} denoting the degree upto which nth property of X resembles with that of nth class is given by

$$\text{where } s_{nj} = \frac{1}{h} \sum_{l=1}^{h} s_{nj}^{(l)} \tag{22}$$

where

$$s_{nj}^{(l)} = \left[1 + A \left| \frac{p_{nj}^{(l)} - p_n}{p_{nj}^{(l)}} \right| \right]^{-1}. \tag{23}$$

A is any positive constant. Thus, the grade of membership of an unknown pattern $X = \{p_1, p_2 \ldots p_n \ldots\ldots p_N\}$ with reference to all the classes are found from the magnitudes of their corresponding similarity vectors and decide it to belong to C_k if.

$$\vee_j S_{nj} = S_{nk}$$
$$n = 1, 2, \ldots\ldots\ldots\ldots, N \tag{24}$$
$$k, j = 1, 2, \ldots\ldots\ldots\ldots\ldots m.$$

IMPLEMENTATION OF THE THEORY AND EXPERIMENTAL RESULTS

To demonstrate the application of the theory to the problem of vowel speech sound and speaker recognition, vowel formant frequencies F_1, F_2, and F_3 were extracted from selected Telugu and Hindi speech samples in CNC combination (Dutta Majumder et al. 1974, 1973). There are 10 sets of Telugu vowels which were transformed to 6 classes which are phonetically non-identical namely. /ə/, /a:/, /I/, /U/, /E/ and /o/ such that

$$C_I \supset C_i^{1-}, C_i :$$
$$C_U \supset C_u^{-1}, C_u :$$
$$C_E \supset C_e^{1-}, C_e :$$
and $$C_0 \supset C_0^{-1}, C_0 :$$

where shorter and longer subsets being phonetically alike are treated in the same class.

Thus in the present problem, $N = 3$, j has maximum value of 6 and h equals 1 for $C_ə$ and C_a: and 2 for C_I, C_U, C_E and C_0.

Table 1 shows typical values of formants for each vowel. Let us consider for example, the sample of class [I] whose $d(X, R_j)$ values are

For Class	[ə]	=	4.93
	[a:]	=	11.67
	[I]	=	$\wedge$ [1.08, 1.80] = 1.08
	[U]	=	$\wedge$ [8.64, 13.57] = 8.64
	[E]	=	$\wedge$ [2.67, 1.57] = 1.57
and	[O]	=	$\wedge$ [8.11, 18.80] = 8.11

Corresponding μ-values with $E=F=1$ are found from eqn. (15) to be 0.17, 0.08, 0.48, 0.10, 0.39 and 0.11 respectively and verify its containment in C_I. Besides, these values defining the degree to which it resembles the membership of the class to which it belongs other possible order of classes is indicated to be /E/, /ə/, /U/,/O/ and /a :/. This serialized result is similar to their phonetical order.

Again membership values computed on formants of vowel /ə/ (Table 1) come to be $\mu_ə = 0.24$, $\mu_{a:} = 0.17$, $\mu_I = 0.13$, $\mu_U = 0.28$, $\mu_E = 0.23$ and $\mu_0 = 0.54$ and decide the classes in order of possibility as /O/, /U/, /E/, /ə/, /a:/ and /I/ which is almost correct from the same point of view.

There are about 900 such samples whose proper classes were assigned exactly in same way and results are tabulated in Table 2. Maximum number of misrecognized samples for vowels /I/, /E/, /ə/, /a:/, /U/ and /O/ as expected correspond to classes /E/, /I/, /a:/, /ə/, /O/ and /U/ respectively.

TABLE 1 — TYPICAL VALUES OF FORMANT FREQUENCIES

Vowel class	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
/ə/	550	1550	2400
/a:/	700	1200	2550
/I/	400	2250	2600
/U/	350	0850	2600
/E/	500	1900	2650
/O/	500	0950	2600

TABLE 2—RECOGNITION SCORE FOR TELUGU VOWELS

Classified as	Actual Vowel Class					
	/ə/	/a:/	/I/	/U/	/E/	/O/
/ə/	52	17		1	15	7
/a:/	8	70				
/I/			155		25	
/U/	1			144		34
/E/	9		17		154	1
/O/	2	2		6	13	138
Recognition Score (%)	72.2	78.65	90.1	95.36	74.4	76.66

TABLE 3—RECOGNITION SCORE FOR HINDI VOWELS

Classified as	Actual Vowel Class /ə/	/a:/	/I/	/U/	/E/	/O/
/ə/	9					
/a:/	1	10				
/I/			20			
/U/				10		
/E/					10	
/O/						20
Recognition Score (%)	90.0	100.0	100.0	100.0	100.0	100.0

TABLE 4 — TYPICAL VALUES OF FORMANT FREQUENCIES FOR HINDI VOWEL [i:] UTTERED BY THREE SPEAKERS AND THEIR RECOGNITION RATE

Speaker	F_1(Hz)	F_2(Hz)	F_3(Hz)	No. of samples	Correct Rate (%)
X	350	2350	2800	38	89.47
Y	350	2250	2700	66	90.90
Z	300	2350	3000	66	100.00

Hindi vowels consist of 8 classes (/ə/, /a:/, /i¹-/, /i:/, /u⁻¹/, /u:/, /E/ and /O/) and number of samples were only 10 for each class. As in the Telugu vowels longer and shorter subsets were pooled together in the same category and recognition rate obtained by this method is shown in Table 3.

Second recognition scheme described in the last section was applied to the problem of speaker identification. Typical formant frequencies for vowel [i:] uttered by 3 informants X, Y and Z say, in the age group of 28/30 years are tabulated in Table 4.

As the third formant F_3 is more significant compared to others for bearing information about sex and age of an informant, the recognition parameters choosen are (F_3/F_1), (F_3/F_2) and F_3. In this problem, $N = 3$, $m = 3$ and for reference prototypes in each class $h = 5$. For example, consider the feature of informant Y whose associated properties (eqn. 18) are

$$p_1 = 0.9994626$$
$$p_2 = 0.9999990$$
$$p_3 = 0.0588235$$

where $\bar{x}_1 = 10.033$, $\bar{x}_2 = 1.3$, $\bar{x}_3 = 3100$, $E = 100$ and $F = 2$.

Using equations (22) and (23) and taking constant A to be inverse of the standard deviation of the prototype features, the magnitudes of the similarity vectors for this pattern with reference to three classes are found to be 2.50, 2.986 and 2.965 respectively which indicate its appropriate containment. Present procedure was repeated entirely in the same fashion for the other events and net rate of identifications are mentioned in Table 4.

CONCLUSION

Use of fuzzy set to the classification problem of imprecisely defined patterns is experimented for computer recognition of speakers and spoken vowels. Knowledge of weighting co-efficients and reference vectors used in this paper could be available also from any size of samples selected within a class. For Telugu speech sounds, since vowel recognition is done irrespective of speakers, its net recognition rate is lowered as compared to Hindi vowels.

In the problem of speaker identification, the machine was trained for each informant by 5 samples which can be changed to any size to have improved score. Further work on this algorithm and classification of a number of male and female speakers in the different age groups is in progress.

ACKNOWLEDGEMENT

Authors wish to express their thanks to Prof. C. R. Rao, FRS, Director of Indian Statistical Institute for his kind interest in the work and also to the Council of Scientific and Industrial Research for Awarding fellowship to Mr. S. K. Pal, one of the authors.

REFERENCES

1. Dutta Majumder (D) and Dutta (AK). *a.* Some Studies on Automatic Speech Coding and Recognition Procedure. Ind. Jour. Phys. **42**, 1968; 425-443.
2. Dutta Majumder (D) and Dutta (AK). *b.* A Model for Spoken word Recognition. Proc. Automazione Estrumentazione. Milan, Italy, 1968; 249-259.
3. Dutta Majumder (D) and Dutta (AK). An Analyser Coder for Machine Recognition of Human Speech. JITE. **15**, 4; 1969; 233-243.
4. Dutta Majumder (D), Dutta (AK) and Pal (SK). Computer Recognition of Vowel Speech Sounds Using Three-dimensional Weighted Discriminant Function. J. CSI. **7**, 1; 1976.
5. Dutta Majumder (D) and Pal (SK). The Concepts of Fuzzy Sets and its application in Pattern Recognition Problems. Proc. Eleventh Annual Symp. of the Computer Soc. of India. Hyderabad, 1976.
6. Dutta (AK). *a.* An Experimental Procedure for Handwritten Character Recognition. IEEE Trans. **C-23**, 5; 1974; 536-545.
7. Dutta (AK). *b.* Computer Recognition of Handwritten Roman Cursive Script. Proc. IEEE International Symp. on Information Theory. Notre Dame, Indiana, USA, 1974.
8. Sebestyen (GS). Decision Making Procedures in Pattern Recognition. 1962. The Macmillan Co., NY. P 8–82.
9. Fu (KS). Sequential Methods in Pattern Recognition and Machine Learning. 1968. Academic Press, London. P3–18.
10. Zadeh (LA). Fuzzy Sets. Inform. Control. **8**; 1965; 338-353.
11. Zadeh (LA). Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. IEEE Trans. **SMC-3**, 1; 1973; 28-44.
12. Dutta Majumder (D), Dutta (AK), Ganguly (NR), Mukherjee (B) and Sarkar (R). Studies on Acoustic-Phonetic Features of Telugu Speech Sounds. ECSL Series on Phonological Studies. Mahalanobis Memorial Issue, Nov. 1974.
13. Dutta Majumder (D), Dutta (AK) and Ganguly (NR). Some Studies on Acoustic Features of Human Speech in relation to Hindi Speech Sounds. Ind. Jour. Phys. **47**, 1973; 598-613.
14. Rao (CR). Advanced Statistical Methods in Research. 1952. J. Wiley & Sons. NY.