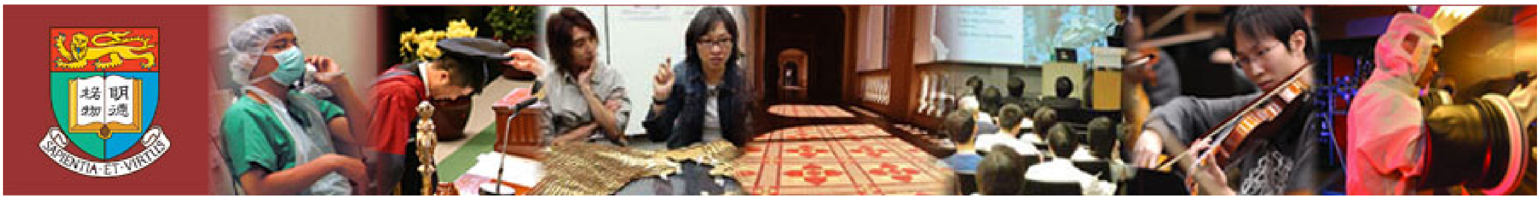

Title	A new fuzzy classifier with triangular membership functions
Author(s)	Yang, YS; Chan, FHY; Lam, FK; Hung, N
Citation	International Conference on Neural Networks Proceedings, Houston, TX, 9-12 June 1997, v. 1, p. 479-484
Issued Date	1997
URL	http://hdl.handle.net/10722/46013
Rights	©1997 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

A NEW FUZZY CLASSIFIER WITH TRIANGULAR MEMBERSHIP FUNCTIONS

[1]Yong-Sheng Yang [1]Francis H. Y. Chan [1]F. K. Lam [2]Hung Nguyen
[1]Dept. of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong
[2]School of Electrical Engineering, University of Technology, Sydney, NSW 2007, Australia
[1]ysyang@hkueee.hku.hk, [1]fhychan@hkueee.hku.hk, [1]fklam@hkueee.hku.hk, [2]htn@ee.uts.edu.au

ABSTRACT Fuzzy logic is widely applied in control and modeling for its robustness, simplicity and clarity. It is also applied in classifier design with rules directly generated from numerical data. Some available rule generation methods, however, are either too complicated to implement or impractical for high dimensions. In this paper, we propose a new fuzzy classifier architecture. At the very beginning the training data is clustered at the input space. Fuzzy sets are then defined based on these clusters with triangular membership function. The outputs in the rule conclusion are initially determined by the "normalized vote" in the corresponding cluster. Fuzzy sets and conclusions can be adjusted through training. The proposed fuzzy system is simple in structure, and can be fast trained and easily implemented. Its classification performance is generally better than artificial neural network.

I. INTRODUCTION

In recent years a lot of research was done on fuzzy modeling and fuzzy classifier [1]-[8] which generates rules from numerical data. Many methods cut the input space into fuzzy hyper-boxes by separating each input component so that it becomes impractical in case of too many inputs. Some methods are so complex in architecture [3] that they are only advantageous for large scale system.

In this paper, we propose a new fuzzy logic classifier which can be easily trained and implemented while keeping satisfactory classification performance. In input space the training data is clustered into hyper-spheres which are thought to be usually more appropriate than the hyper-boxes in reflecting the pattern distributions. Enlightened by [9], we then generate rules from each cluster. The rules can be tuned through training. It is an improved version of the fuzzy classifier presented in [14]. The improvement exists in, that a technology called "normalized vote" is originally introduced so that the classifier requires very little training before being adjusted to optimal parameter setting, and that triangular membership functions (MF) are applied in place of Gaussian-shaped MF so that the classifier is able to run faster and readily implemented on microprocessors or fuzzy chips.

Section II introduces the formation and training of the proposed fuzzy classifier. Section III presents two application examples verifying its better performance. Section IV contains some further discussion. And the system is then concluded in section IV.

II. ALGORITHM

Our aim is to design a classifier which applies fuzzy inference so that it works like the way of human being. Fuzzy sets are to be defined on each dimension of the input data space. The classifier will work on many rules like:

IF $\mathbf{x}=(x_1, x_2,\ldots, x_n)\in F_k$, i.e., $x_1\in F_{k1}$ and $x_2\in F_{k2}$ and $\ldots$ and $x_n\in F_{kn}$ THEN $\mathbf{x}\in$ Class j

where $(x_1, x_2,\ldots, x_n)$ is the n-dimensional input, $F_k = F_{k1}\times F_{k2}\times\ldots\times F_{kn}$ is the fuzzy set in the premise of the k-th rule, and $j=1,2,\ldots,m$ (m is the total number of classes). To quantify the rule's conclusion, we define an output vector $w_k=(w_{k1},w_{k2},\ldots, w_{km})$, where w_{kj} is the degree of how much an input belongs to class j by a specific rule; $0\le w_{kj}\le 1$. In fact we have extended the crisp class to fuzzy class through the adoption of the output vector.

Based on this idea, the fuzzy set's centroid and width will be initialized by the results out of Basic ISO-DATA algorithm [10]. Suppose M clusters: C_1, C_2, ... , C_M, are generated from ISO-DATA with $(\bar{x}_{k1},\bar{x}_{k2},\ldots,\bar{x}_{kn})$ as C_k's centroid vector and $(\sigma_{k1},\sigma_{k1},\ldots,\sigma_{kn})$ as C_k's radius vector, where

$k=1,\ldots,M$, and radius vector being equal in each dimension. M fuzzy sets: $F_1, F_2, \ldots, F_M$, are then defined based on the M clusters respectively by the triangular-product membership function $\mu_{F_k}(\cdot)$ as: Given n-dimensional input $\mathbf{x}=(x_1, x_2, \ldots, x_n)$, the degree of $\mathbf{x}\in F_k$ is

$$\mu_{F_k}(x) = \prod_{i=1}^{n} \mu_{ki}(x_i)$$

where

$$\mu_{ki}(x_i) = \begin{cases} (x_i - \bar{x}_{ki} + 3\sigma_{ki})/(3\sigma_{ki}) & \text{for } \bar{x}_{ki} - 3\sigma_{ki} \le x_i \le \bar{x}_{ki} \\ (-x_i + \bar{x}_{ki} + 3\sigma_{ki})/(3\sigma_{ki}) & \text{for } \bar{x}_{ki} \le x_i \le \bar{x}_{ki} + 3\sigma_{ki} \\ 0 & \text{else} \end{cases} \quad (1)$$

and $i=1,\ldots,n$; $k=1,\ldots,M$. Fig. 1 illustrates an exemplar fuzzy set definition via clusters in two-dimensional input space. M rules are designed respectively corresponding to the M clusters as:

If $\mathbf{x}$ belongs to fuzzy set F_k, then outputs are $\mathbf{w}_k = (w_{k1}, w_{k2}, \ldots, w_{km})$

Where w_{kt} is yielded from the "normalized vote" in cluster k: if there is v_{kt} items in the cluster belongs to class t, then

$$w_{kt} = \frac{v_{kt}}{\sum_{s=1}^{m} v_{ks}}, \quad \text{for } t=1,\ldots,m;\ k=1,\ldots M.$$

$\mathbf{o}=(o_1,o_2,\ldots,o_m)$, the final outputs after the M product-inference rules and centroid average defuzzifier, are given by:

$$o_t = \frac{\sum_{k=1}^{M}\left[w_{kt}\prod_{i=1}^{n}\mu_{ki}(x_i)\right]}{\sum_{k=1}^{M}\left[\prod_{i=1}^{n}\mu_{ki}(x_i)\right]} \quad \text{for } t=1,2,\ldots,m \quad (2)$$

The sample is then assigned to the class with maximum output among the m values. The network representation of the fuzzy classifier is demonstrated in Fig. 2.

To make the fuzzy system work more precisely, we show how rules and membership functions are adjusted by the error back-propagation method. For N input-output pairs: $(\mathbf{x}^{(1)},\mathbf{y}^{(1)}), (\mathbf{x}^{(2)},\mathbf{y}^{(2)}),\ldots, (\mathbf{x}^{(N)},\mathbf{y}^{(N)})$, mean square error is defined as:

$$E = \frac{1}{2}\cdot\frac{1}{N}\sum_{p=1}^{N}\sum_{t=1}^{m}(o_t^{(p)} - y_t^{(p)})^2 \quad (3)$$

Calculating the partial differentiation of (3) with respect to $w_{kt}, \bar{x}_{ki}, \sigma_{ki}$ yields:

$$\frac{\partial E}{\partial w_{kt}} = \frac{1}{N}\sum_{p=1}^{N}\left\{(o_t^{(p)} - y_t^{(p)})\frac{A}{B}\right\} \quad (4)$$

$$\frac{\partial E}{\partial \bar{x}_{ki}} = \frac{1}{N}\sum_{p=1}^{N}\sum_{t=1}^{m}\left[(o_t^{(p)} - y_t^{(p)})\cdot\frac{A\cdot C}{B^2}\cdot\frac{\partial\mu_{ki}(x_i)}{\partial \bar{x}_{ki}}\right] \quad (5)$$

$$\frac{\partial E}{\partial \sigma_{ki}} = \frac{1}{N}\sum_{p=1}^{N}\sum_{t=1}^{m}\left[(o_t^{(p)} - y_t^{(p)})\cdot\frac{A\cdot C}{B^2}\cdot\frac{\partial\mu_{ki}(x_i)}{\partial \sigma_{ki}}\right] \quad (6)$$

where

$$\frac{\partial\mu_{ki}(x_i)}{\partial \bar{x}_{ki}} = \begin{cases} -1/(3\sigma_{ki}) & \text{for } \bar{x}_{ki} - 3\sigma_{ki} \le x_i \le \bar{x}_{ki} \\ 1/(3\sigma_{ki}) & \text{for } \bar{x}_{ki} \le x_i \le \bar{x}_{ki} + 3\sigma_{ki} \\ 0 & \text{else} \end{cases} \quad (7)$$

$$\frac{\partial\mu_{ki}(x_i)}{\partial \sigma_{ki}} = \begin{cases} -(x_i^{(p)} - \bar{x}_{ki})/(3\sigma_{ki}^2) & \text{for } \bar{x}_{ki} - 3\sigma_{ki} \le x_i \le \bar{x}_{ki} \\ (x_i^{(p)} - \bar{x}_{ki})/(3\sigma_{ki}^2) & \text{for } \bar{x}_{ki} \le x_i \le \bar{x}_{ki} + 3\sigma_{ki} \\ 0 & \text{else} \end{cases} \quad (8)$$

$$A = \prod_{\substack{j=1 \\ j\ne i}}^{n}\mu_{kj}(x_j^{(p)}), \quad B = \sum_{l=1}^{M}\left[\prod_{j=1}^{n}\mu_{lj}(x_j^{(p)})\right],$$

$$C = w_{kt}\cdot\sum_{l=1}^{M}\left[\prod_{j=1}^{n}\mu_{lj}(x_j^{(p)})\right] - \sum_{l=1}^{M}\left[w_{lt}\cdot\prod_{j=1}^{n}\mu_{lj}(x_j^{(p)})\right]$$

for $k=1,\ldots,M$; $i=1,\ldots,n$; $t=1,\ldots,m$. Hence $w_{kt}, \bar{x}_{ki}, \sigma_{ki}$ can be trained by the steepest-descent method to reduce E:

$$w_{kt}^{(new)} = w_{kt}^{(old)} - \alpha\cdot\frac{\partial E}{\partial w_{kt}} \quad (9)$$

$$\bar{x}_{ki}^{(new)} = \bar{x}_{ki}^{(old)} - \alpha\cdot\frac{\partial E}{\partial \bar{x}_{ki}} \quad (10)$$

$$\sigma_{ki}^{(new)} = \sigma_{ki}^{(old)} - \alpha\cdot\frac{\partial E}{\partial \sigma_{ki}} \quad (11)$$

where $\alpha=0.001\sim0.1$ is the step-size.

III. EXPERIMENTAL RESULTS

Initially, the performance of this fuzzy classifier is verified with the use of a set of synthesizes overlapping data. Then it is applied to the functional electromyogram (EMG) classification for prosthesis control [11]. For comparison purpose, a three-layer artificial neural network classifier [12] is also developed using the same data sets.

A. Synthetic data classification

Two Gaussian distributed classes are generated in the two-dimensional space with same variances σ in each dimension. Their centers are located at (1,1) and (2,2) respectively. 200 samples are generated for each class. Half the samples are used for training, and the left for testing. We define C_{SEP}, the separation coefficient, to reflect the class overlap:

$$C_{SEP} = \frac{d}{3*(\sigma+\sigma)} = \frac{0.236}{\sigma}$$

where $d=\sqrt{2}$ is the class center distance.

Thirteen clusters are generated by ISO-DATA according to experiments, and hence there are thirteen rules. There are two outputs in each rule corresponding to the two classes. With relation to the separation coefficient, the classifiers' training mean square errors and classification error rates are listed in table I. Generally speaking, the fuzzy classifier works better than ANN with a little decrement in error rate. In addition it is noticed that fuzzy classifier needs very few training epochs.

B. EMG data classification

There are four classes of functional EMG data set [11] each of which containing 40 trials, half for training and half for testing. 24 features are extracted from these trials. Table II is the classification rates of four subjects. Slightly higher classification rate are acquired by the fuzzy classifier. In addition, as has been shown in [14], fuzzy classifier worked out more consistent outputs indicated by its smaller output variance than that of ANN classifier.

IV. DISCUSSION

A. Selection of rule number

The more rule number, the more partitions in the input space, and the more computations in inference and training. Within a certain limit, the performance of the system is improved as more rules are used because of more preciseness in matching the pattern distributions. However, with too many rules, the performance deteriorates as the system over-fits the training data. In order to identify the appropriate rule number, the system starts from small rule number which is usually triple the total number of classes. The rule number is then increased until the classification rate on test set does not improve any more.

B. About the training

In the fuzzy system, the fuzzy sets are initialized by the cluster parameters from ISO-DATA, and the outputs in the conclusions are initialized by the "normalized vote" among the classes within the cluster. The initialization guarantees good start point for the rule base demonstrated by the fact that the MSE and classification error rate are almost optimal even without training. After only a few epochs of training, the fuzzy system is expected to have more reliable generalization because of the fine tune of fuzzy sets and rule conclusion. As to ANN, the weights have to be initialized randomly and the training process is usually much longer as a consequence. A typical relation of classification error rate *vs.* training epoch number is demonstrated in fig. 3. The data set is EMG signal of above-mentioned-subject No. 1. It can be found that FUZ works out good results (MSE and error rate) even without any training at all.

In mathematical sense, the triangular membership function is not differentiable at the three sharp points. However, hardly does it make any side effect in training. The training by (9)-(11) never generates illegal membership functions.

C. Hardware implementation

Fuzzification, inference and defuzification of the system use only economical arithmetic computations. It can be implemented on general-purpose micro-controllers with low cost and fast running speed [13]. Some fuzzy logic co-processors can also realize the system conveniently.

V. CONCLUSION

We have proposed a new fuzzy classifier architecture. The training data set is clustered by Basic ISODATA algorithm. Fuzzy sets are then defined based on these clusters with triangular membership function. Each cluster will yield one rule. The outputs in the rule conclusion are initially determined by the "normalized vote" in the corresponding cluster. Fuzzy sets and conclusion outputs are adjusted by error back-propagation method afterwards.

The proposed fuzzy system is simple in structure, and can easily be trained and implemented. Moreover, its classification performance is generally better than ANN. Linguistic rules can also be integrated to the rule base.

VI. ACKNOWLEDGMENT

This work is supported in part by The University of Hong Kong Research Grants. The authors want to thank Dr. B. Hudgins and Mr. C. A. K. Englehart at the Institute of Biomedical Engineering, University of New Brunswick (UNB), Canada, for providing the EMG data set.

REFERENCES

[1] S. Abe, M.-S. Lan, "A method for fuzzy rules extraction directly from numerical data and its application to pattern classification", *IEEE Trans. on Fuzzy Systems*, FEB 1995, VOL. 3, NO. 1, pp. 18-28

[2] S. Horikawa, T. Furuhashi, Y. Uchikawa, "On fuzzy modeling using fuzzy neural networks with the back-propagation algorithm", *IEEE Trans. on Neural Networks*, SEP 1992, VOL. 3, NO. 5, pp. 801-806

[3] N. Imasaki, J. Kiji, M. Arai, "Framework for fuzzy neural networks", *Science of Artificial Neural Networks,* 1992, Dennis W. Ruck, Editor, Proc. SPIE 1710, pp. 535-546

[4] S. Mitra, S. K. Pal, "Fuzzy multi-layer perceptron, inference and rule generation", *IEEE Trans. on Neural Networks*, JAN 1995, VOL. 6, NO. 1, pp. 51-63

[5] C.-T. Sun, "Rule-base structure identification in an adaptive-network-based fuzzy inference system", *IEEE Trans. on Fuzzy Systems*, FEB 1994, VOL. 2, NO. 1, pp. 64-73

[6] Y. J. Lee, V. Cherkassky, J. R. Slagle, "Adaptive fuzzy-rule-based classifier", World Congress on Neural Networks, JUN 1994, San Diego, VOL. I, pp. 699-704

[7] L.-X. Wang, J. M. Mendel, "Generating fuzzy rules by learning from examples", *IEEE Trans. SMC*, NOV/DEC 1992, VOL. 22, NO. 6, pp. 1414-1427

[8] R. Katayama, Y. Kajitani, Y. Nishida, "A self generating and tuning method for fuzzy modeling using interior penalty method", Proc. of the 2nd international conference on Fuzzy Logic and Neural Networks, JUL 1992, Japan, pp. 357-360

[9] B. Kosko, "Fuzzy systems as universe approximators", *IEEE Trans. on Computers*, NOV 1994, VOL. 43, NO. 11, pp. 1329-1333

[10] R. O. Duda, P. R. Hart, "Pattern classification and Scene analysis", John Wiley & Sons, USA, 1973

[11] B. Hudgins, P. A. Parker, "A new strategy for multifunction myoelectric control", *IEEE Trans. BME*, JAN 1993, VOL. 40, NO. 1, pp. 82-94

[12] R. P. Lippmann, "An introduction to computing with neural nets", *IEEE ASSP Magzine*, APR 1987, VOL. 4, NO. 2, pp. 4-22

[13] J. Yan, M. Ryan, J. Power, "Using fuzzy logic: towards intelligent systems", Prentice Hall, New York, 1994

[14] Y.-S. Yang, F. K. Lam, F. H. Y. Chan, Y.-T. Zhang, P. A. Parker, "A new fuzzy approach for pattern recognition with application to EMG classification", *Proceedings of International Conference on Neural Network (ICNN'96)*, Vol II, pp 1109-1114, June 1996, Washington D.C.

[15] Y.-S. Yang, "Fuzzy classification of biomedical signals", Master of philosophy thesis, The University of Hong Kong, Oct. 1996, Hong Kong

[16] Y.-S. Yang, H. Nguyen, F. H. Y. Chan, F. K. Lam, "Fuzzy classification of biomedical signals", *Proceedings of International Conference on Biomedical Engieering (BME'96)*, pp 28-30, June 1996, Hong Kong

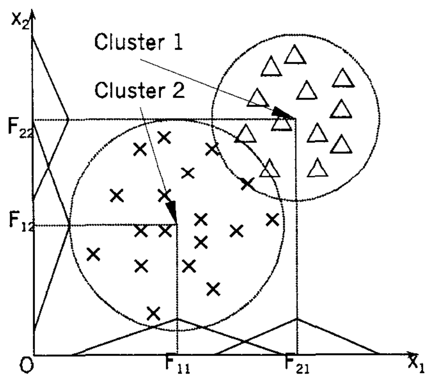

Fig. 1 An exemplar fuzzy set definition via clusters in two-dimensional input space

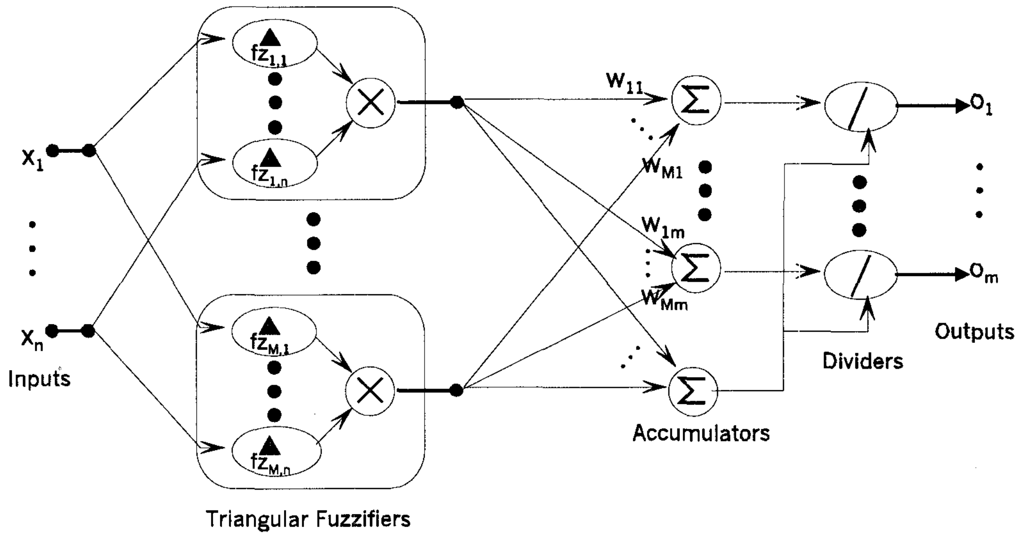

Fig. 2 Network representation of fuzzy classifier with triangular membership functions

Table I. Comparison of ANN and Fuzzy classifier on synthetic overlapping data

C_{SEP}	ANN-MSE	FUZ-MSE	ANN ErrorRate	FUZ ErrorRate
2.00	0.43	0.00	0.0%	0.0%
1.00	0.33	0.02	0.5%	0.5%
0.75	0.20	0.10	1.5%	1.5%
0.65	0.20	0.17	1.5%	1.0%
0.55	0.22	0.25	5.5%	4.5%
0.45	0.30	0.38	11.5%	10.0%
0.35	0.34	0.43	12.5%	12.0%
0.25	0.42	0.47	19.5%	20.0%

Table II. Comparison of EMG recognition rates by ANN and Fuzzy classifier

Subject Identification	Classification Rate by ANN	Classification Rate by FUZ
No. 1	91.3%	92.5%
No. 2	78.8%	80.0%
No. 3	96.3%	97.5%
No. 4	83.8%	88.8%

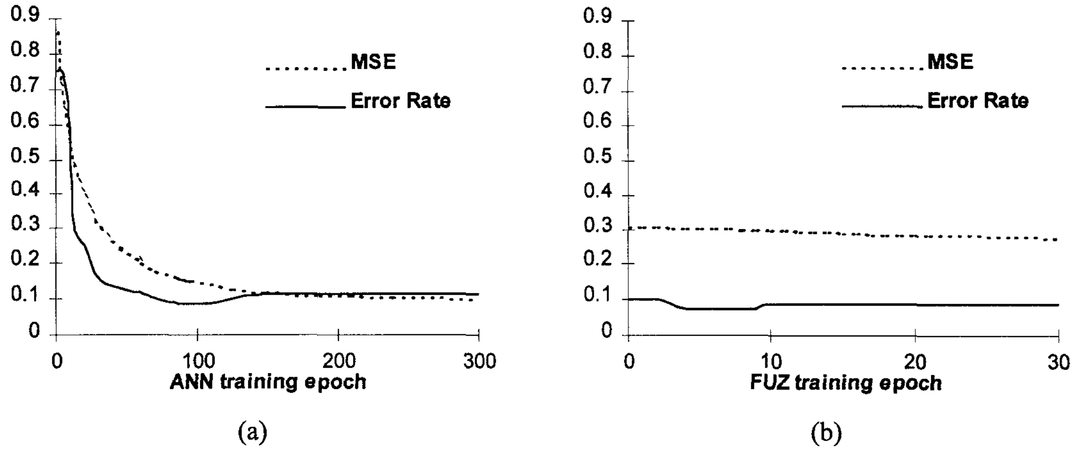

Fig. 3. Mean square error over training set and classification error rate over test set *vs.* training epochs. (a) ANN system, training step size = 0.25; (b) FUZ system, training step size = 0.04, Isodata cluster number M = 13, rule = 13 × 4.