

A New Fuzzy Approach for Pattern Recognition with Application to EMG Classification

¹Yong-Sheng Yang ¹F. K. Lam ¹Francis H. Y. Chan ²Yuan-Ting Zhang ³Philip. A. Parker

¹Dept. of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong,

²Dept. of Electrical Engineering, Chinese University of Hong Kong, Hong Kong

³Dept. of Electrical Engineering, University of New Brunswick, Fredericton, NB, Canada

¹ysyang@hkueee.hku.hk, ¹flam@hkueee.hku.hk, ¹fhychan@hkueee.hku.hk, ²ytzhang@ee.cuhk.hk, ³pap@unb.ca

ABSTRACT

A fuzzy logic system with center average defuzzifier, product-inference rule, nonsingleton fuzzifier and Gauss membership function is discussed in this paper. The fuzzy sets are initially defined by the cluster parameters from the Basic ISO-DATA algorithm on input space. The system is then trained via back error propagation algorithm so that the fuzzy sets are fine-tuned. The system is applied to functional EMG classification and compared with its ANN counterpart. It is superior to the latter in at least three points: higher recognition rate; insensitive to over-training; and more consistent outputs thus having higher reliability.

I. INTRODUCTION

The design of a pattern recognition system consists of several stages: data collection, formation of pattern classes, feature selection, specification of the classification algorithm, and estimation of the classification error. The crucial steps are feature selection and the classifier design. A good classifier should be able to learn and update from examples and to incorporate available knowledge. There are many methods for classifier design such as heuristic approach, deterministic approach, statistical approach, neural network approach and fuzzy approach. The heuristic approach and deterministic approach are of limited application in experience. The statistical approach requires mathematical model of the object, and the explicit model can only be approximate to the object. Neural network (ANN) approach learns from examples to generate complex but rigid boundary between classes, and is inscrutable to the designer and therefore causes some problems. Fuzzy logic tries to assimilate human being reasoning [9]. Classifier based on fuzzy logic can be more robust compared to ANN.

There are advantages of fuzzy logic systems in such applications as biomedical signal processing and classification. Biomedical signals are not always strictly repeatable, and may sometimes be even contradictory. One of the most useful properties of fuzzy logic system is that contradictions in the data can be tolerated. Furthermore, using the trainable fuzzy system, it is possible to discover patterns in data which are not easily detected. Finally, the experience of medical experts can be incorporated. It is possible to integrate this incomplete but valuable knowledge into the fuzzy logic system due to the system's similar reasoning style to the human being. This is a significant advantage over ANN[6].

In this paper, we propose a fuzzy logic system with center average defuzzifier, product-inference rule, nonsingleton fuzzifier and Gauss membership function. This fuzzy system is structurally analogous to radial basis-function neural network[7] so that most advantages of the latter are retained. The system is an adapted version from that proposed in [8] to solve pattern recognition problems. We apply the Basic ISO-DATA algorithm to cluster the training data set in the feature space. The cluster centers and RMS radii are then used to define the fuzzy sets of the system. Afterwards, all the system including the fuzzy sets is trained by the back error propagation algorithm. This fuzzy system is called ISO-FUZ. An application in EMG signal classification is presented.

Section II introduces the formation, training, implementation and performance of ISO-FUZ. Section III presents an application of ISO-FUZ in EMG classification, and the system performance evaluation in comparison to ANN approach. The system is concluded in section IV.

II. ISO-FUZ SYSTEM

A. Fuzzy system and its training

Many fuzzy logic systems in control problems use triangular membership function for the convenience of design and implementation. However, in general, the Gauss function is able to better reflect a fuzzy object, and the function is differentiable enabling membership function training by the steepest gradient method.

The fuzzy logic systems with center average defuzzifier, product-inference rule, nonsingleton fuzzifier and Gauss membership function can be described by the following function [8]:

input: $\underline{x} = (x_1, x_2, \dots, x_n)^T \Rightarrow$ output: $f(\underline{x})$ by

$$f(\underline{x}) = \frac{\sum_{l=1}^M y_l^{-1} \left[\prod_{i=1}^n a_i^l \exp\left(-\left(\frac{x_i - \bar{x}_i^{-1}}{\sigma_i^l}\right)^2\right) \right]}{\sum_{l=1}^M \left[\prod_{i=1}^n a_i^l \exp\left(-\left(\frac{x_i - \bar{x}_i^{-1}}{\sigma_i^l}\right)^2\right) \right]} \quad (1)$$

where $\bar{x}_i^{-1}, \sigma_i^l, a_i^l$ are the centroid, width, maximum membership of fuzzy set F_i^l respectively; y_l^{-1} is the centroid of fuzzy set G^l ; the corresponding rule base consists of a collection of fuzzy IF-THEN rules in the following form:

$$R^l: \text{IF } x_1 \text{ is } F_1^l \text{ and } \dots \text{ and } x_n \text{ is } F_n^l, \text{ THEN } y \text{ is } G^l \quad (l=1,2,\dots,M)$$

The renowned back-propagation algorithm can be used to train the fuzzy system. Suppose that we are given an input-output pair (\underline{x}, d) , $\underline{x} \in U \subset \mathbb{R}^n$, $d \in V \subset \mathbb{R}$ (\mathbb{R} is real number set, \mathbb{R}^n is n tuples of real number); our task is to determine a fuzzy logic system $f(\underline{x})$ as expressed in (1) such that

$$e = \frac{1}{2} [f(\underline{x}) - d]^2 \quad (2)$$

is minimized, where d is the desired output corresponding to input \underline{x} . Supposing $a_i^l = 1$ (so that membership achieves maximum value 1 at some point), by performing the partial differentiation on equation (2) with respect to y_l^{-1} we get:

$$\bar{y}_l^{-1}(k+1) = \bar{y}_l^{-1}(k) - \alpha_1 \left. \frac{\partial e}{\partial y_l^{-1}} \right|_k = \bar{y}_l^{-1}(k) - \alpha_1 \frac{f-d}{b} z^l \quad (3)$$

where $l=1,2,\dots,M$, $k=0,1,2,\dots$, and α_1 is a constant learning stepsize, and

$$f = a/b, \quad a = \sum_{l=1}^M (y_l^{-1} z^l), \quad b = \sum_{l=1}^M z^l, \quad z^l = \prod_{i=1}^n \exp\left(-\left(\frac{x_i - \bar{x}_i^{-1}}{\sigma_i^l}\right)^2\right)$$

respectively. Analogously,

$$\bar{x}_i^{-1}(k+1) = \bar{x}_i^{-1}(k) - \alpha_2 \frac{f-d}{b} (y_l^{-1} - f) z^l \frac{2(x_i - \bar{x}_i^{-1}(k))}{\sigma_i^l{}^2(k)} \quad (4)$$

$$\sigma_i^l{}^2(k+1) = \sigma_i^l{}^2(k) - \alpha_3 \frac{f-d}{b} (y_l^{-1} - f) z^l \frac{2(x_i - \bar{x}_i^{-1}(k))^2}{\sigma_i^l{}^3(k)} \quad (5)$$

where $i=1,2,\dots,n$; $l=1,2,\dots,M$; $k=0,1,2,\dots$; α_1 and α_2 are learning stepsizes.

The operations by equations (3)–(5) with the presentations of all input-output pairs iteratively will gradually minimize the summation of error in equation (2), and the fuzzy system is trained. A preset mean square error (MSE), is used to terminate training:

$$\text{MSE} = \sqrt{\frac{1}{N} \sum_{(\underline{x}, d)} [f(\underline{x}) - d]^2} \quad (6)$$

where N is the number of training input-output pairs.

B. Convergence problem and initialization of fuzzy sets and performance assessment

For the fuzzy system with Gauss-shaping membership function, no quick or good convergence is guaranteed without appropriate initialization due to the “local nature”.

No matter how powerful they are, classification methods work well only on effective features, which tend to gather the elements belonging to the same class into one or more clusters in the feature space, such that they tend to

disclose their similarities quantitatively. Basic Isodata algorithm [2] is able to cluster the items automatically. If the parameters of these clusters are used to initialize the fuzzy sets, each fuzzy set will be able to represent all or part of the items in the same class at the very beginning, so that blindness is avoided and convergence is accelerated in the following training stage. Based on this idea, the fuzzy set's centroid and width can be initialized by the results out of the following Basic Isodata algorithm:

1. Choose randomly c samples from the training data set as the initial centers of clusters, where c is the desired cluster total to be generated.
2. Assign each training sample to the cluster whose center is closest to the sample.
3. Update the cluster centers by averaging the current samples in the clusters.
4. If any cluster center changed position, go to step 2 to cluster again; otherwise, go to step 5 to define fuzzy sets with these clusters.
5. Assign the centers of each cluster as the centroid of fuzzy sets; compute averaged RMS radii of each clusters and assign these values to the widths of fuzzy sets respectively.

Such an initialization proved to be a very effective way to ensure good convergence in the training process with training epoch number much fewer than ANN with comparable size. The ISO-DATA algorithm costs trivial time compared with the subsequent training process.

C. Implementation considerations and performance assessment

The fuzzy system is designed to solve pattern recognition problems. When there are more than two pattern classes, we may compose a multi-output fuzzy system by paralleling several above-stated single output fuzzy systems, each output corresponding to the activation level of one class. The ideal output of the selected class is usually set to 0.9, and others to 0.1.

With the input and output components normalized between 0 and 1, we usually choose $\alpha_1 = 0.001 \sim 0.010$, and $\alpha_2 = \alpha_3 = 10\alpha_1$, since $\frac{-1}{x_1}$ and σ_1 are insensitive to training. The cluster number c is often chosen to be 2~4 times of the pattern classes, and rule number is equal to cluster number: $M=c$. The stopping training error MSE can be as high as 0.2~0.4 when satisfactory classification rate is achieved. In comparison, the MSE in ANN has to be about 0.1 in order to get similar classification performance.

Compared to ANN, the fuzzy system takes much fewer training epochs and often has higher classification rate. The fuzzy system does not require low training MSE to acquire good recognition rate. When trained to the same MSE as ANN, the fuzzy system can generate more consistent outputs from test set, thus more reliable result. The threshold is more flexible than ANN. In addition, there is little over-training side effect for fuzzy system, while ANN can easily suffer from over-training. When over-trained, SIG-ANN can yield such complex boundaries (over-fitting) between classes that the performance of generalization is undermined. Fuzzy systems approximates the ideal classifier (if it exists) by many patches (B. Kosko, 1992) [5]. It is the "local" nature of these patches that makes the fuzzy system robust and insensitive to over-training.

III. EMG CLASSIFICATION VIA FUZZY APPROACH

A. Feature selection

In [3], EMG signal was acquired using a bipolar surface electrode pair placed over each of the biceps and triceps brachii muscle groups of the subject. The signal were then divided into several time segments to preserve pattern structure, and features were extracted from these segments. The features included mean absolute value (MAV), mean absolute value slope (MAVSLP), zero-crossing (ZC), slope sign changes (SSC), and wave length or wave complexity (WC). In our fuzzy approach classification, most of the features are the same, except that SSC was not adopted, as it proved to be insignificant. In this paper, we use 6 segments (i.e. 240 ms) instead of 5 to collect more information. The reproduction of the ANN approach is also based on 6 segments EMG for comparison purposes.

B. Classification implementation

Fig. 1 illustrates the fuzzy approach configuration. In the figure, the block "Fuzzy Logic System" is composed of 4 parallel sub-system generating OUT_EX, OUT_FL, OUT_PR and OUT_SP, the activation level of each function respectively. The output logic generates the final classification result.

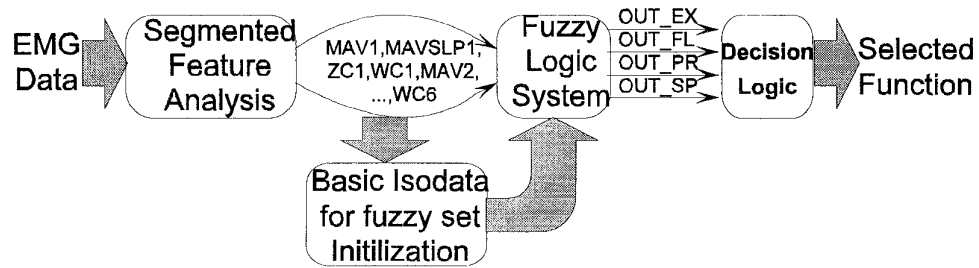


Fig. 1. Schematic diagram of fuzzy approach for EMG classification

The block “Basic Isodata for Initialization” initializes the system’s fuzzy sets before the training phase. The desired output of OUT_XXs are set to 0.9 for selected function and 0.1 for others at the training phase. The four sub-systems are trained simultaneously with the presentations of the whole training set. When the mean square error (*MSE*) is lower than a preset threshold (usually 0.1~0.3), the training ends. The block “Decision Logic” finds the maximum value among OUT_XXs. If this value is greater than 0.3, that function is selected; otherwise, no function is activated. This threshold was optimally determined over many experiments. It should be noted that the OUT_XX’s have relative meaning rather than absolute, and they should not be regarded or interpreted as probabilities.

C. Results and discussion

We denote the EMG Neural Network classification method with sigmoid activation function proposed in [3] as **SIG-ANN**, and the EMG fuzzy system classification method proposed in this paper as **ISO-FUZ**.

Data sets from four subjects were used in the experiments. They were the same set as used in [3]. The training vs. test set size was 20 vs. 20 for each function per subject so that at least 160 trials consists of four function categories were required in each subject’s data set. Both SIG-ANN method and ISO-FUZ method were applied on these data sets. The (correct) recognition rates of the two approaches were measured and compared in Table I. The rates were close, with ISO-FUZ being slightly better.

Table I. Comparison of EMG recognition rates by SIG-ANN and ISO-FUZ

Subject Id.	Rate of SIG-ANN	Rate of ISO-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. 2 demonstrates the outputs from ISO-FUZ system compared with SIG-ANN outputs on each test trial of the same subject’s data set. The figure suggests that, ISO-FUZ tends to give more consistent and hence more “stable” results than SIG-ANN does, since the former output variations over the same class were obviously smaller than the latter.

When the decision threshold in the output logic is increased, the error rate of classification also increases due to more rejected trials. On the other hand, if the threshold is arbitrarily decreased, say to less than 0.3, it will not make much sense since the small threshold is too close to the inactivate level (=0.1). We investigated the relation between error rate and decision threshold. As an example, the threshold of ISO-FUZ can be varied from 0.3 to 0.5 without increasing the error rate, while only from 0.3 to 0.4 for SIG-ANN. The example demonstrates ISO-FUZ’s higher threshold tolerance than SIG-ANN.

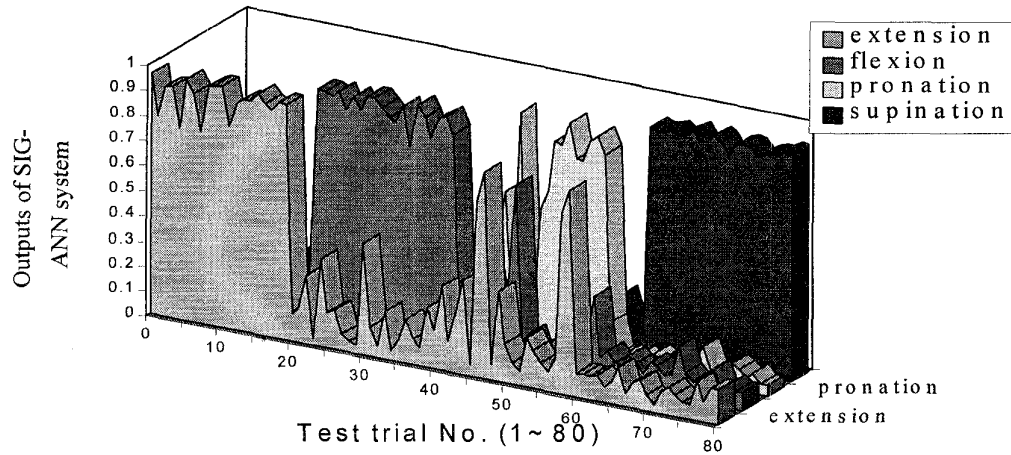
Like the neural network, after too many epochs of training, the classification performance of the fuzzy system will deteriorate with an increased error rate. This over-training problem was investigated by experiments, with results illustrated in Fig. 3. It should be noted that, due to their different structure, the training stepsizes are different, and they are optimally decided by experimentation respectively. It was found that ISO-FUZ is less sensitive to over-training compared with SIG-ANN. In both methods, the mean square error (*MSE*) over the training set decreased as the training epoch increased, but the classification rate over the test set did not decrease monotonically, with the rate

bouncing up a little after reaching its minimum point. We define B , the error rate's bouncing amplitude, as the measure of over-training side effect:

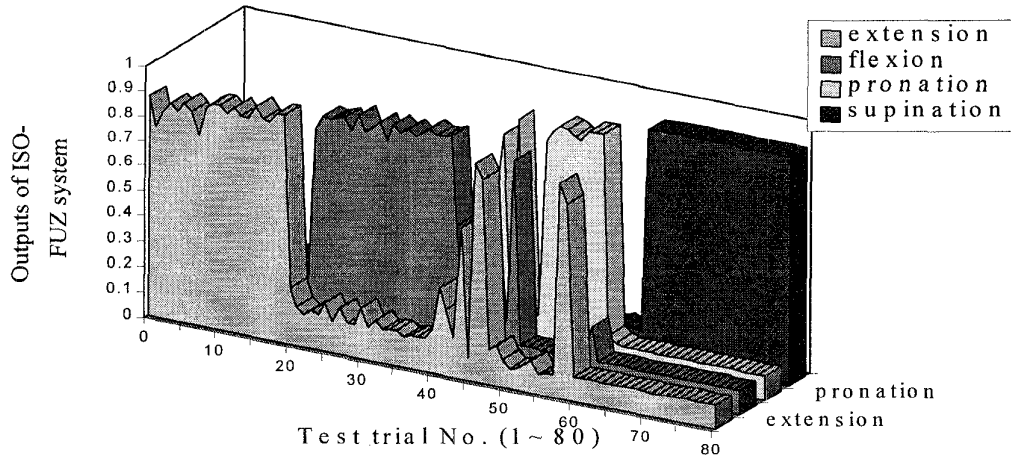
$$B = \frac{\text{MAX}}{R_E(n) \text{ is Minimum}} \{R_E(3n) - R_E(n)\}$$

where $R_E(n)$ is the test error rate after n training epochs. Greater B implies more serious over-training effect. In the example of Fig. 3, B was 0.025 for SIG-ANN, and 0.013 for ISO-FUZ, i.e., the over-training side effect of SIG-ANN is double that of ISO-FUZ in this example. ISO-FUZ takes less over-training risk.

Their training speeds are about at the same order. One of the exemplary training times is 23 seconds for SIG-ANN and 12 seconds for ISO-FUZ with our software running on Pentium PC. Fig. 3 indicates that ISO-FUZ achieves optimal recognition rate after only several training epochs, while SIG-ANN takes dozens of epochs. Although each epoch of ISO-FUZ is longer than SIG-ANN, the whole training is still quicker. It is noted that ISO-FUZ does not require very low MSE to reach low error rate.



(a)



(b)

Fig. 2. (a) Output of SIG-ANN system (learning rate: 0.25; training MSE=0.1)
 (b) Output of ISO-FUZ system (learning rate: 0.015; training MSE=0.1)
 For both: Trial 1~20 are extension; 21~40 are flexion; 41~60 are pronation;
 61~80 are supination. Subject: No. 1; Correct classification rate:
 91.3% (7 trials misclassified or rejected in the category of pronation)

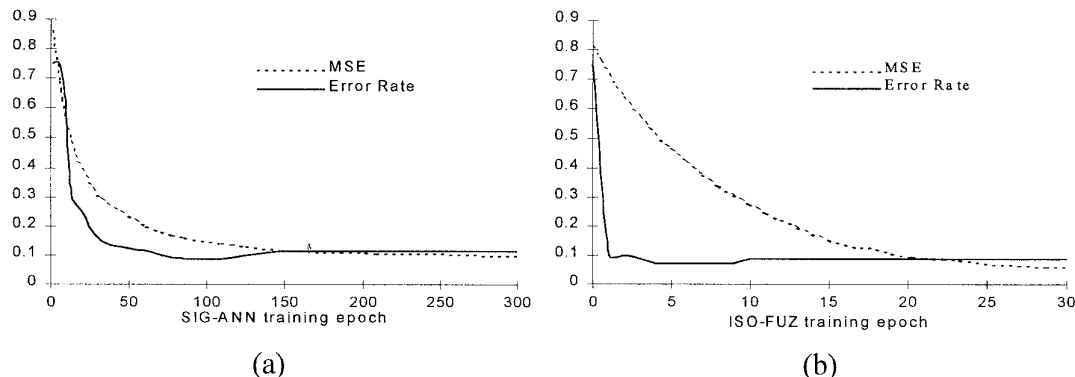


Fig. 3. Mean square error over training set and classification error rate over test set vs. training epochs. (a) SIG-ANN system, training step size = 0.25; (b) ISO-FUZ system, training step size = 0.015, Isodata cluster number = 12, rule = 12×4 .

IV. CONCLUSION

This paper has proposed a fuzzy system (ISO-FUZ) whose fuzzy sets are initialized by clusters from Basic ISO-DATA algorithm. The system with the fuzzy sets are then trained by the training data set. The training process and generalization capability turned out to be superior to the neural network (SIG-ANN), primarily in that the fuzzy system gives more consistent classification results and is insensitive to over-training. When applied to pattern recognition problem, the threshold is more flexible than SIG-ANN, demonstrating its robustness. More favorably, it can incorporate expert experience easily by adding fuzzy units in parallel.

The system has been applied to the functional EMG classification for prosthesis control purpose. The recognition rate is better than ANN approach.

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 and for their valuable comments on this project. We appreciate the editorial suggestions by Dr. M.H. Du of our group.

REFERENCES

- [1] C. H. Chen, "On the relationships between statistical pattern recognition and artificial Neural networks", in *Neural networks in pattern recognition and their application*, C. H. Chen, World Scientific, 1991, pp 153-158
- [2] R. O. Duda, P. R. Hart, *Pattern classification and Scene analysis*, John Wiley & Sons, 1973, chapter 6, pp 201-202
- [3] B. Hudgins, P. A. Parker, "A new strategy for multifunction myoelectric control", *IEEE Trans BME*, Vol 40, No. 1, pp 82-94, JAN 1993
- [4] H.M. Kim, J.M. Mendel, "Fuzzy basis functions: comparisons with other basis functions", *IEEE Trans. on Fuzzy Systems*, VOL. 3, NO. 2, pp 158-168, MAY 1995
- [5] B. Kosko, "Fuzzy systems as universe approximators", *proc. of 1992 IEEE International Conference on Fuzzy systems*, San Diego, pp 1409-1418, 1992
- [6] R.P. Lippmann, "An introduction to computing with neural nets", *IEEE ASSP Magazine*, Vol. 4, No. 2, pp.4-22, APR 1987
- [7] J. Moody and C.J. Darken, "Fast learning networks of locally-tuned processing units", *Neural Computation*, Vol 1, No. 2, pp 281-294, 1989
- [8] L.X. Wang, *Adaptive fuzzy systems and control: design and stability analysis*, Prentice Hall, 1994, chapter 3, pp 29-31
- [9] L.A. Zadeh, "Fuzzy sets", *Informat. Control*, No. 8, pp 338-353, 1965