Journal of Theoretical and Applied Information Technology
10th February 2013. Vol. 48 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

APPLICATION OF FUZZY CLUSTERING NEURAL NETWORK IN CONJUNCTION SPEECH RECOGNITION

PEILING ZHANG, LINGFEI CHENG
School of Electrical Engineering and Automation, Henan Polytechnic University, Jiaozuo 454000, Henan China

ABSTRACT

In order to improve the approximation property of the past fuzzy clustering algorithms when identifying systems, a fuzzy clustering neural network (FCNN) is proposed and is applied to conjunction speech recognition system. Based on the fuzzy system model, FCNN presents every state as a fuzzy system and uses continuous frames as the system input. With improving fuzzy clustering identification algorithm, FCNN is acted as estimator of probability density function which could forecast output probability of the each state. This model not only can describe the inter-frames correlation information for speech signal efficiently, but overcome the deficiency of traditional hidden markov model which supposes each state's output is mixed Gauss distributing probability density function. Through the experiments of speaker-independent conjunction speech recognition, the effectiveness of FCNN could be verified.

Keywords: *Conjunction Speech Recognition, Fuzzy Clustering Neural network, Probability Density Function*

1. INTRODUCTION

Automatic recognition of speech by machine has become a hot research topic for more than four decades and has developed to the stage of large vocabulary, speaker independent and continuous speech recognition. In this process, hidden markov model (HMM) plays an important role. It is a powerful set of tools for providing a statistical model of both the static properties of sounds and the dynamical changes that occur across sounds. All most of excellent speech recognition systems are based on HMM.

However, traditional HMM has some inherent flaws, such as HMM is supposed that each state's output is mixture Gaussian probability density function and each state is independent. The convergence speed of basic Baum-Welch algorithm is fast, but system's recognition ratio is low. In order to overcome those deficiencies, some improved methods were put forward [1-7], but some were complex implementing or lack of theory base.

Therefore, a fuzzy clustering neural network (FCNN) model is proposed in this paper. Based on the fuzzy system model, every HMM state is regarded as a fuzzy system in this method. With continuous frames character vector of speech signal as the system's input, the model can forecast the probability density function of the system's output states by using improved fuzzy clustering identifying algorithm. Experimental results show the efficiency of the model in speech recognition system. Recognition ration on the basis of this model is high.

2. FUZZY CLUSTERING NEURAL NETWORK

2.1 Vector T-S fuzzy system model

The T-S fuzzy system model [8-10] proposed by Takagi and Sugeno is one efficient method to use for complex system recognition. It describes system's behavior through fuzzy rules, and then fuzzy approximates to them by a certain number of linear local models. Let $P(X,Y)$ be an identification system which is multi input single output system. Then, this system can be described using T-S model which is shown as following.

R^i : if x_{k1} is u_{i1}^k and x_{k2} is u_{i2}^k and... x_{kP} is u_{iP}^k

$$\text{then } \boldsymbol{y}_i^k = \boldsymbol{w}_i \boldsymbol{\theta}_k \tag{1}$$

where $i = 1,2,\cdots,c$ (c denotes the number of fuzzy rules), $k = 1,2,\cdots,N$ (N is the number of input samples), $x_k = (x_{k1}, x_{k2}, \cdots, x_{kP})^T \in X^P$ indicates the input vectors, $u_{ik} = (u_{i1}^k, u_{i2}^k, \cdots, u_{iP}^k)^T$ is subordinated vector for the k th input vectors which is relative to the i th class, $y_i^k = (y_{i1}^k, y_{i2}^k, \cdots, y_{iQ}^k)^T$ is the consequent part output vector of the i th rule,

Journal of Theoretical and Applied Information Technology
10th February 2013. Vol. 48 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$\boldsymbol{w}_i = (\boldsymbol{w}_{i1}, \boldsymbol{w}_{i2}, \cdots, \boldsymbol{w}_{iQ})^T$ denotes the consequent part parameter vector matrix of the i th rule, $\boldsymbol{\theta}_k = (1, x_{k1}, x_{k2}, \cdots x_{kP})^T$ is the consequent part input vector. Therefore, the output of system is shown as follows.

$$y_k = \sum_{i=1}^{c} u_i^k y_i^k \tag{2}$$

where $u_i^k = \prod_{j=1}^{P} u_{ij}^k$.

Fuzzy C-Means (FCM) algorithm is used to cluster input sample for above model, and then each of the class is corresponded one rule. The clustering rule is shown below.

$$\min\{J(U,V,M) = \sum_{k=1}^{N}\sum_{i=1}^{c}(u_{ik})^M (d_{ik})^2\} \tag{3}$$

where $u = [u_{ik}]_{c\times K}$ is membership matrix, $\sum_{i=1}^{c} u_{ik} = 1$, $u_{ik} \in [0,1]$, $k = 1,2,\cdots,N$, $d_{ik} = \|x_k - v_i\|$, $\|\cdot\|$ denotes Euclidean norm, $v_i = (v_{i1}, v_{i2}, \cdots, v_{iP})^T \in \boldsymbol{R}^P$ is the clustering centers of the i th class, $M \in [1,\infty]$ is fuzzy weighting exponent. But approximation ability of output subspace does not consider in formula (3) which causes a rather big errors for system identification.

2.2 Improved Fuzzy Clustering Identifying Algorithm

Supposing a sample space $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_N)$ ($\boldsymbol{x}_k = (x_{k1}, x_{k2}, \cdots, x_{kP})^T \in \boldsymbol{X}^P, k = 1,2,\cdots,N$) and a output space $\boldsymbol{Y} = (y_1, y_2, \cdots, y_N)$, $y_k = (y_{k1}, y_{k2}, \cdots, y_{kQ})^T \in \boldsymbol{Y}^Q$, $\boldsymbol{X}$ is divided into c range with mode clustering shown as follows: $\boldsymbol{K}_1, \boldsymbol{K}_2, \cdots, \boldsymbol{K}_c$ which satisfies $\boldsymbol{K}_i \cap \boldsymbol{K}_j = \varnothing$ and $\boldsymbol{X} = \cup \boldsymbol{K}_i$, therefore, for random $\boldsymbol{X}$, set of all probable fuzzy c divisions is expressed as following.

$$M_{fc} = \{U \in V_{c\times N} \,|\, u_{ik} \in [0,1], \forall i,k; \ \sum_{i=1}^{c} u_{ik} = 1, \forall k\} \tag{4}$$

where u_{ik} is subordinated vector for the k th input vectors which is relative to the i th class, $V_{c\times N}$ is the real matrix . The best fuzzy c division rule of X and Y is seek to the best small value of object function J . J must not only reflect cluster performance of the input space, but also reflect the approach performance of the output space. So a new object function J is introduced as a new fuzzy clustering rule which is defined below.

Definition: let $\boldsymbol{x}_k = (x_{k1}, x_{k2}, \cdots, x_{kP})^T \in X^P$ be input vector, $\boldsymbol{t}_k = (t_{k1}, t_{k2}, \cdots, t_{kQ})^T \in Y^Q$ is actual output vector of the system, $k = 1,2,\cdots,N$, N is the total number of samples; $\boldsymbol{v}_i = (v_{i1}, v_{i2}, \cdots, v_{iP})^T$ denote the clustering centers of the i th class, $i = 1,2,\cdots,c$, c is the cluster number; u_{ik} is fuzzy membership for the k *th* input vectors which is relative to the i *th* class; $\boldsymbol{w}_i = (\boldsymbol{w}_{i1}, \boldsymbol{w}_{i2}, \cdots, \boldsymbol{w}_{iQ})^T$ denotes the consequent part parameter vector matrix of the i *th* rule; $\boldsymbol{\theta}_k = (1, x_{k1}, x_{k2}, \cdots x_{kP})^T$ is the consequent part input vector; $\boldsymbol{y}_i^k = (y_{i1}^k, y_{i2}^k, \cdots, y_{iQ}^k)^T$ denotes consequent part output vector for the k *th* input vector of the i *th* rule; M is fuzzy weighting exponent. Then J is fuzzy clustering object function shown as follows.

$$J(U,V,M,W) = \sum_{k=1}^{N}\sum_{i=1}^{c}(u_{ik})^M \|x_k - v_i\|^2 + \sum_{k=1}^{N}\sum_{j=1}^{Q}\left|t_{kj} - \sum_{i=1}^{c} u_{ik} y_{ij}^k\right| \tag{5}$$

where $y_{ij}^k = w_{ij}\theta_k$, $\sum_{i=1}^{c} u_{ik} = 1$, $u_{ik} \in [0,1]$.

In formula (5), the first item of definition requires the distance of input vector and cluster center must be the smallest, which reflects the input space's fuzzy division; The second item requires the estimated value approaches the actual value, which reflects the smallest error between the input space's fuzzy division mapped to output space and the actual output, and then obtains the best fuzzy clustering for the fuzzy model. U (the best fuzzy division) can be obtained by the following theorem.

Theorem 1: Given the clustering center $\boldsymbol{v}_i$, the fuzzy weighting exponent M and the rule number c , J will be the smallest when the above parameter satisfies the formula (6) shown as follows.

$$u_{ik} = \frac{[\sum_{j=1}^{Q}\left|y_{ij}^k\right| / \|x_k - v_i\|^2]^{\frac{1}{M-1}}}{\sum_{I=1}^{c}[\sum_{j=1}^{Q}\left|y_{ij}^k\right| / \|x_k - v_i\|^2]^{\frac{1}{M-1}}} \tag{6}$$

2.3 FCNN Network Structure

Based on the above improved fuzzy clustering identification algorithm, a new kind of neural

Journal of Theoretical and Applied Information Technology
10th February 2013. Vol. 48 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

network model—fuzzy clustering neural network (FCNN) is proposed. Figure 1 shows its structure. It is made up of two parts: the first part is fuzzy classifier, which is constituted by three layer networks shown in figure 2. It has *P* input nodes which correspond to *P* components for input vector. Meanwhile, it has *C* hide nodes and the *i th* hide node denotes deviation between input vector and the *i th* clustering centers. Their transfer function is shown below.

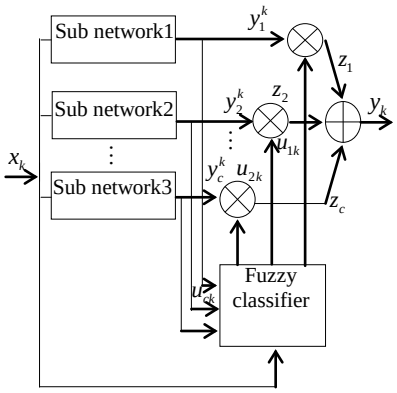

Figure1: FNCC Network Structure

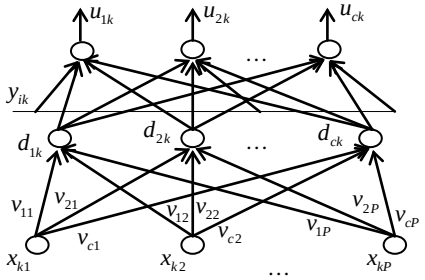

Figure 2: Fuzzy Sorter Structure

$$d_{ik} = \|x_k - v_i\|^2 = \sum_{j=1}^{P}(x_{kj} - v_{ij})^2 \qquad (7)$$

The output layer has *C* nodes whose output denotes the membership of input vector to one type. Their transfer function is obtained by formula (6). The connection weight between input nodes and hide nodes denotes the clustering center for one type which needs to be optimized by study algorithm. There is not connection weight between the hidden nodes and the output nodes, which is united with each subnet to form the input of the third layer.

The second part is shown in figure 3 which is made up of *C* subnets, and every subnet is a two-layer net. $\boldsymbol{w}_i = (\boldsymbol{w}_{i1}, \boldsymbol{w}_{i2}, \cdots, \boldsymbol{w}_{iQ})^T$ is connection weight matrix (where $\boldsymbol{w}_{ij} = (w_{j0}^{(i)}, w_{j1}^{(i)}, w_{j2}^{(i)}, \cdots, w_{jP}^{(i)})$), the input vector is $\boldsymbol{\theta}_k = (1, x_{k1}, x_{k2}, \cdots x_{kP})^T$,and the *i th* subnet's output is $\boldsymbol{y}_i^k = \boldsymbol{w}_i \boldsymbol{\theta}_k$, which denotes consequent part output vector for the *k th* input vector of the *i th* rule. Therefore, the whole system output is shown below:

$$y_k = \sum_{i=1}^{c} u_{ik} y_i^k = \sum_{i=1}^{c} u_{ik} \boldsymbol{w}_i \boldsymbol{\theta}_k \qquad (8)$$

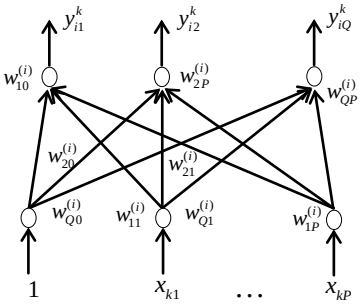

Figure 3: Sub-Network Structure

2.4 The Training Algorithm Of FCNN

The training algorithm of FCNN adopts gradient-drop method. It defines an error function as follows:

$$E_k = 1/2 \sum_{i=1}^{Q}(t_{ki} - y_{ki})^2 \qquad (9)$$

Then the net algorithm in figure 1 can be trained by the following three steps:

(1)Initialize weight value $w_{ij}^{(t)}(0)$ 、 clustering center $v_{ij}(0)$ 、 study rate $\eta(0)$, choose iterative error ε and the maximum iterating times T .

(2)With $k = 1,2,\cdots,N$, $i = 1,2,\cdots,Q$, $j = 0,1,2,\cdots,P$, $t = 1,2,\cdots,C$, the v_{ij} and $w_{ij}^{(t)}$ are updated according to the following algorithm.

$$v_{ij}(n+1) = v_{ij}(n) - \eta(n)\frac{\partial E_k}{\partial v_{ij}} \qquad (10)$$

$$w_{ij}^{(t)}(n+1) = w_{ij}^{(t)}(n) - \eta(n)\frac{\partial E_k}{\partial w_{ij}^{(t)}} \qquad (11)$$

$$\frac{\partial E_k}{\partial w_{ij}^{(t)}} = \frac{\partial E_k}{\partial y_{ki}}\frac{\partial y_{ki}}{\partial z_{ti}}\frac{\partial z_{ti}}{\partial y_{ti}^{(k)}}\frac{\partial y_{ti}^{(k)}}{\partial w_{ij}^{(t)}} = -(t_{ki} - y_{ki})u_{tk}x_{kj} \qquad (12)$$

$$\frac{\partial E_k}{\partial v_{tj}} = -\sum_{i=1}^{Q}(t_{ki} - y_{ki})\sum_{m=1}^{c}\frac{\partial y_{ki}}{\partial z_{mi}}\frac{\partial z_{mi}}{\partial u_{mk}}\frac{\partial u_{mk}}{\partial d_{tk}}\frac{\partial d_{tk}}{\partial v_{tj}}$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$= -2\sum_{i=1}^{Q}(t_{ki} - y_{ki})\sum_{m=1}^{c} y_{mi}^{k}(v_{tj} - x_{kj})\frac{\partial u_{mk}}{\partial d_{tk}}$$

$$= -\frac{2u_{tk}(x_{kj} - v_{tj})\sum_{i=1}^{Q}(t_{ki} - y_{ki})(\sum_{m=1}^{c} u_{mk} y_{mi}^{k} - y_{ti}^{k})}{(M-1)d_{tk}} \quad (13)$$

(3)Calculate total error $E = \sum_{k=1}^{N} E_k$, if $E < \varepsilon$ or $n > T$, the program is end. Otherwise, let $n = n+1$, $\eta(n) = \eta(0)[1 - n/T]$, the program goes to step (1).

3. ACOUSTICAL MODEL OF SPEECH RECOGNITION SYSTEM

There are the nonlinear characteristics in speech signal, and then it must need nonlinear method to process them. Therefore, FCNN is acted as the speech recognition model.

Supposing the input characteristic vector of speech signal $\boldsymbol{Y} = \boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_N$, $\boldsymbol{X} = \boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_L$ ($\boldsymbol{x}_n \in S_1 \, S_L$) is HMM's state sequence and S_j is HMM's status, so the output probability of speech recognition system based HMM is shown as follows.

$$P(\boldsymbol{Y}|\lambda) = \sum_{X} P(\boldsymbol{Y}, \boldsymbol{X}|\lambda)$$

$$= \sum_{\boldsymbol{X}} \{P(\boldsymbol{X})[\prod_{n=1}^{N} P_{x_n = S_l}(\boldsymbol{y}_n|\lambda)]\}$$

$$= \sum_{\boldsymbol{X}} \{P(\boldsymbol{X})[\prod_{n=1}^{N} P_{x_n = S_l}(\boldsymbol{y}_n|\boldsymbol{y}_{n-1}, \boldsymbol{y}_{n-2}, \cdots, \boldsymbol{y}_{n-p})]\} \quad (14)$$

Formula (14) denotes conditional probability of the first p characteristic vector which is different from traditional HMM's output probability. It effectively depicts speech signal's correlation between frames.

Basic HMM is assumed that each state's output is mixture Gaussian probability-density function and used Maximum Likelihood (ML) criterion to train parameter of HMM. Experimental results show the result of this assumption is suboptimum. So FCNN is acted as estimator of probability density function, which could forecast output of the each state. Supposing the input vector of FCNN is constituted by first p characteristic vector of speech signal, that is to say $I_t = (y_{t-1}, y_{t-2}, \cdots, y_{t-p})$, the output vector of FCNN is y_t . Let $O_t = y_t$. However, the FCNN's actual output is K_t , the output probability density function of speech recognition system based FCNN is shown below.

$$b_j(O_t) = P(O_t|I_t) = P_{X_t = S_j}(y_t|y_{t-1}, y_{t-2}, \cdots, y_{t-p})$$

$$= (2\pi)^{-Q/2} \exp[-\frac{1}{2}\sum_{t=1}^{Q}(O_{ti} - K_{ti})^2] \quad (15)$$

where Q is dimension of characteristic vector.

From formula (15), system's state output probability not only is determined by its state, but also the first p frame characteristic vector. So the frame's correlation of speech signal is effectively used. Meanwhile, system's state output probability is determined by the FCNN approaching the degree to the actual voice system. That is to say, the FCNN more approaches the real system which output probability is bigger; otherwise, the output probability is smaller.

4. EXPERIMENTAL RESULTS

There are 50 male talkers' voices in data base including ten digits from 0 to 9. Each male talker speaks ten mandarin digit strings and length of each mandarin digit string is different. Where, voices of 30 talkers are used to train, others are used to test.

The spectral analysis uses the MFCC with the following characteristics: sampling rate-16KHz; analysis window size-320 samples (20 msec); analysis window shift-160 samples (10 msec); MFCC dimension-12; the MFCC's derivative dimension-12. Therefore, the feature parameters is a 26 dimension feature vector consisting of 12 dimension MFCC, 12 dimension MFCC's derivative, their normalized energy and their normalized energy derivative.

The type of CHMM is left-to-right models which have 7 states and each state is mixture Gaussian probability density function, where the number of per-state's function is 5.

Fuzzy clustering neural network which is based on fuzzy system model, is regarded each HMM state as a fuzzy system, letting continuous frames character vector of speech signal as the system's input, applying improved FC recognition algorithm to form a new kind of FCNN to forecast the system's output, so that it can realize the estimate of output probability density function. Based on some experiment analyses and comparison, the parameters of FCNN are input vector frames $p = 4$, states number $L = 5$, weighting exponent $M = 2$, and clustering sort number $c = 5$.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Two conjunction speech recognition systems which based on basic CHMM and FCNN are accomplished using MATLAB. Types of error recognition include replace error (S), insertion error (I) and deleting error (D). S denotes one digit is wrongly recognized as the other digit. I means inserting a new digit. D means that one digit is deleted from mandarin digit string. So, accurate ratio, correct ratio and error ratio are defined as following.

$$Acc = \frac{N-(S+I+D)}{N} \times 100\% \quad (16)$$

$$Corr = \frac{N-(S+D)}{N} \times 100\% \quad (17)$$

$$E = \frac{S+I+D}{N} \times 100\% \quad (18)$$

The two speech recognition system's results are shown in table 1, table 2. Table 3 and table 4 are shown recognition results without replace error that 8 are wrongly recognized as 2.

Table 1 Recognition Results Of Speaker-Dependent System

Parameter	CHMM(7 States, 5 Gauss)	FCNN (5 States)
Acc	91.729%	89.923%
Corr	92.481%	97.287%
E	8.271%	13.561%

Table 2 Recognition Results Of Speaker-Independent System

Parameter	CHMM(7 States, 5 Gauss)	FCNN (5 States)
Acc	78.212%	84.522%
Corr	85.475%	95.347%
E	21.788%	15.034%

Table 3 Recognition Results Of Speaker-Dependent System

Parameter	CHMM(7 States, 5 Gauss)	FCNN (5 States)
Acc	93.985%	87.816%
Corr	94.737%	98.844%
E	6.015%	12.139%

Table 4 Recognition Results Of Speaker-Independent System

Parameter	CHMM(7 States, 5 Gauss)	FCNN (5 States)
Acc	84.637%	86.518%
Corr	91.899%	97.468%
E	15.363%	13.429%

From the statistical results, some conclusions can be drawn as following:

(1) FCNN has higher recognition ratio than traditional HMM, especially in the correct ratio. It further shows that the new recognition model effectively introduces the frame's correlation of speech signal. System's state output probability is determined by the FCNN approaching the degree to the actual voice system. It overcomes the assumption that each state's output is mixture Gaussian probability density function. Meanwhile, as the un-linearity characteristic for speech signal, the new recognition model uses un-linearity tools (neural network) to process speech signal. Experimental results show that the model is efficiency.

(2) The replace error (S) takes place mainly that 8 are wrongly recognized as 2, which in traditional HMM accounts for 46% and in FCNN accounts for 76%. By comparing the results of statistics, it found that the two recognition systems of accurate ratio and correct ratio are greatly improved without replace error that 8 are wrongly recognized as 2.

5. CONCLUSION

This paper presents a speech recognition model based on fuzzy clustering neural network. It not only effectively describes speech signal's correlation of frame, but also overcomes the assumption that each state's output is mixture Gaussian probability density function. This model is based on fuzzy system model, taking each HMM state as a fuzzy system, letting continuous frames character vector as the system's input, applying improved fuzzy clustering recognition algorithm form a new kind of FCNN to forecast the system's output. Therefore, it can realize the estimate for every state output probability density function. The experimental results have demonstrated the efficiency of the new recognition model proposed in this paper.

ACKNOWLEDGEMENTS

This work was supported by the Young Scientists Fund of the National Natural Science Foundation of China (11105042); Henan province university science and technology innovation talent support plan (2011HASTIT013) and the higher school control engineering key disciplines open foundation of Henan province, China (KG2011-18).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFERENCES:

[1] Y. Shao, C. H. Chang, "Bayesian separation with sparsity promotion in perceptual wavelet domain for speech enhancement and hybrid speech recognition", *IEEE Transaction on Systems, Man and Cybernetics, Part A: Systems and Humans*, Vol. 41, No. 2, 2011, pp. 284-293.

[2] E. Erzin, "Improving throat microphone speech recognition by joint analysis of throat and acoustic microphone recordings", *IEEE Transaction on Audio, Speech and Language Processing*, Vol. 17, No. 7, 2009, pp. 1316-1324.

[3] S. Windmann, R. Haeb-Umbach, "Approaches to iterative speech feature enhancement and recognition", *IEEE Transaction on Audio, Speech, and Language Processing*, Vol. 17, No. 5, 2009, pp. 974-984.

[4] V. Mitra, Nam Hosung, C. Y. Espy-Wilson, E. Saltzman, L. Goldstein, "Gesture-based Dynamic Bayesian Network for noise robust speech recognition", Proceedings of International Conference on Acoustics, Speech and Signal Processing, IEEE Conference Publishing Services, May 22-27, 2011, pp. 5172 - 5175

[5] B. H. Juang, W. Chou, C. H. Lee, "Minimum classification error rate methods for speech recognition", *IEEE Trans. Speech and Audio Processing*, Vol. 5, No. 3, 1997, pp. 257-265.

[6] V. Gupta, M. Lennig, P. Mermelstein, "Integration of acoustic information in a large vocabulary word recognizer", Proceedings of International Conference on Acoustics, Speech and Signal Processing, IEEE Conference Publishing Services, Apr. 1987, pp. 697-700.

[7] L. Deng, M. Aksmanoric, X. Sun, C. F. J. Wu, "Speech recognition using hidden Markov models with polynomial regression function as stationary states", *IEEE Trans. Speech and Audio Processing*, Vol. 2, No. 4, 1994, pp. 507-520.

[8] T. Takagi, M. Sugeno, "Fuzzy Identification of systems and its application to modeling and control", *IEEE Transaction on Systems, Man and Cybemetics*, Vol. 15, No. 1, 1985, pp. 116-132.

[9] A.F. Gomez-Skarmeta, M. Delgado, M.A. Vila, "About the use of fuzzy clustering techniques for fuzzy model identification", *Fuzzy Sets and Systems*, Vol. 106, No.2, 1999, pp. 179-188.

[10] M. Sugeno, Y. Takahiro, "A fuzzy-logic-based approach to qualitative modeling", *IEEE transactions on Fuzzy Systems*, Vol.1, No.1, 1993, pp. 7-31.