

## A Segmentation Method for Continuous Speech Utilizing Hybrid Neuro-Fuzzy Network

CHING-TANG HSIEH, MU-CHUN SU, EUGENE LAI  
AND CHIH-HSU HSU\*

*Department of Electrical Engineering*

*\*Graduate Institute of Information Engineering  
Tamkang University*

*Taipei Hsien, Taiwan 251, R.O.C.*

This paper presents a hierarchical neuro-fuzzy network system for segmenting continuous speech into syllables. The system formulates speech segmentation as a two-phase procedure. In the first phase, the Hybrid Neuro-Fuzzy network (HNFN) is utilized to classify the speech signal into three different types. The hybrid model (HNFN) composed of a distributed representation of a fuzzy system (DRF) and a hyperrectangular composite neural network (HRCNN) is proposed and used to cluster frames. This special hybrid system may neutralize the disadvantages of each alternative. The parameters in the trained HNFN are utilized to extract both crispy and fuzzy classification rules. In the following phase, the self-tuning back-propagation neural network (STBNN) is utilized to solve the coarticulation effects of vowel-vowel (V-V) concatenation. In our experiments, a database containing continuous reading-rate Mandarin speech recorded from newscasts was utilized to test the performance of the proposed speaker-independent speech segmentation system. The effectiveness of the proposed segmentation system is confirmed by the experimental results.

**Keywords:** speech segmentation, hybrid neuro-fuzzy network, self-tuning back-propagation neural network, clustering analysis, neural network

### 1. INTRODUCTION

Speech segmentation plays an important role in speech recognition in reducing the requirement for large memory and in minimizing the computation complexity in large-vocabulary continuous speech recognition systems. In general, there are two kinds of segmentation. One is phonemic unit segmentation [1-3], which segments speech into phonemes based on the features of the phonemes. The other is syllabic unit segmentation [4, 5]. For both phonemic and syllabic unit segmentation, most of the approaches are based on the thresholds of the parameters used to segment the speech data. The thresholds of the segmentation features (e.g., the zero crossing rate and the average value of the first formant band for each frame) are usually set based on personal experience. Therefore, segmentation performance is not very reliable. This motivated us to formulate thresholds decision-making as a classification problem in order to automate the threshold setting procedure. Several approaches have been proposed for classification problems. Some works have focused on conventional probabilistic and deterministic classifiers [6-10]. One approach

---

Received February 14, 1997; revised July 10, 1998; accepted September 3, 1998.  
Communicated by Jhing-Fa Wang.

uses neural networks to classify patterns, but it is hard to analyze a network, and the learning process is relatively slow. In order to compensate for the above weak points, the Hybrid Neuro-Fuzzy Network (HNFN, a hybrid model composed of a distributed representation of a fuzzy (DRF) module and a hyperrectangular composite neural network (HRCNN module) is proposed. We apply HNFN to label speech data.

After we finish the frame-labeling phase, the next step is to correctly segment the speech signals into syllabic units. If we assume that every frame is classified correctly into one of three types: silence (0), consonant (1), or vowel (2), then the frames of the speech signals can be represented by the labeling sequence as

$$[00\dots 0][11\dots 1] 22\dots 2000\dots 0.$$

Therefore, it is easy to segment the sequence into syllabic units by detecting any transition in the labeling sequence. However, it is usually not possible to achieve a 100% recognition rate for any classification algorithm even if we use HNFN. In order to offset the non-perfect recognition performance of HNFN, the concept of transition states is applied. From a set of labeling sequences, several transition states can be obtained. Then, the segmentation errors carried in phase I can be remedied before executing phase II.

From the labeling sequence, we can not separate the vowel-vowel (V-V) concatenation cases without more information. We solve this problem by using the self-tuning back-propagation neural network (STBNN) with the energy and time derivative of the envelope of the unbiased log spectrum [12]. Once the network converges, it gets an optimal layer and node numbers, and the V-V concatenation can be separated. The effectiveness of this approach has been substantiated by segmentation experiments using samples of radio news continuously uttered in Mandarin by two females and two males. This paper is organized as follows. In section 2 we discuss the characteristics of Mandarin. Section 3 introduces the structure of HNFN and briefly describes the novel class of HRCNN and the DRF. Section 4 briefly introduces STBNN. A performance evaluation is given in section 5. Finally, concluding remarks are presented in section 6.

## 2. CHARACTERISTICS OF MANDARIN

The characteristics of mandarin speech are very different from those of other languages, such as English and French. Basically, a Mandarin word can be expressed as

$$[\text{Consonant,}] [\text{median,}] \text{Vowel} [\text{,tail}], \quad (1)$$

where [.] is optional. The “median” may be pronounced as /i/, /u/, or /iu/, and the “tail” may be produced as /n/ or /ng/. In most Mandarin speech recognition systems, it is usual to regard the combination of the “median”, “vowel”, and “tail” as a Vowel. Thus, a Mandarin word is usually represented as

$$[\text{Consonant,}] \text{Vowel}. \quad (2)$$

Since every Mandarin word may be expressed by Eq. (2), Mandarin speech can be partitioned into three different syllabic units: (1) silence, (2) consonant, and (3) vowel. Table 1 illustrates the relationships among phonemes and syllabic units.

**Table 1. Relationships among phonemes and classes units.**

Labels	Classes	Phonemes
0	silence	silence
1	consonants	f,d,t,g,k,h,j,b,p,y,m,n,ch,sh,t,z,ts,tz,s,r,l
2	vowels	i,u,a,o,e,io,ai,ei,ai,au,ou,el,ia,ie,iau,iou,ua,uo,uai,uei,ue,iue,iua,iu,n,ng

In our experiments, two features extracted from every frame, the average value of first formant band ( $V_i$ ) and the zero crossing rate (ZCR), were used for frame classification in the first phase. The first parameter was based on the spectral envelope of the speech signals obtained through unbiased log spectral estimation [12]. This parameter has been used in voiced/unvoiced detection, with good performance in [2]. It represents the degree of the periodic characteristic of vowels. The second parameter, usually used to decide the startpoints and endpoints of syllable [4], represents the degree of friction. These two parameters are used as the input variables to train HNFN.

### 3. HYBRID NEURO-FUZZY NETWORK

In Fig. 1, the phase I portion shows the structure of HNFN. HNFN consists of a DRF module, an HRCNN module, and a linear combiner. The input layers of the two modules simultaneously receive the same inputs. Therefore, the number of input units in the DRF module is equal to that in the HRCNN module. The linear combiner, which has one layer of linear units, receives the outputs of the two modules and gets an optimal output. In the following phase, STBNN is utilized to solve the coarticulation effects of V-V concatenation.

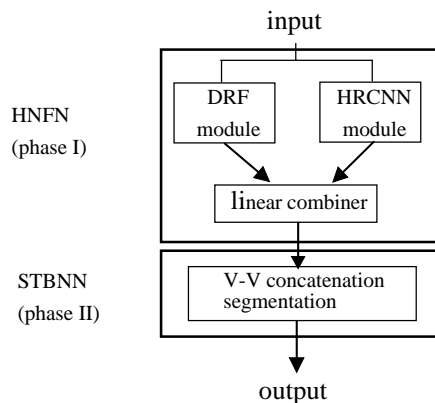


Fig. 1. Flow chart of syllabic segmentation.

### 3.1 HRCNN Module

The construction of an rule-based expert system involves the process of acquiring production rules. Production rules are often represented as “IF condition THEN act”. Back-propagation networks may not arrive at an inference structure close to this kind of high-level knowledge representation. Here, a novel class of HRCNN is presented. This class of HRCNN provides a new tool for machine learning. The classification knowledge is easily extracted from the weights of a trained HRCNN. The symbolic representation of a neural node with hyperrectangular neural-type junctions is shown in Fig. 2(a), and it is described by the following equations:

$$net_j(x) = \sum_{i=1}^n f((M_{ji} - x_i)(x_i - m_{ji})) - n, \quad (3)$$

and

$$out_j(x) = f(net_j(x)), \quad (4)$$

where

$$f(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}. \quad (5)$$

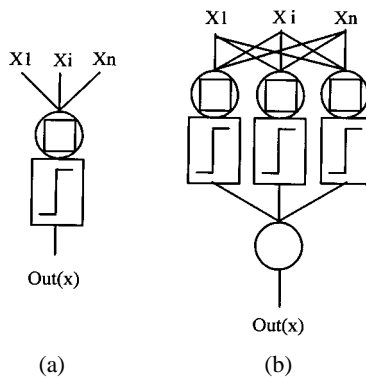


Fig. 2. (a) A neural node with “hyperrectangular neural-type” junctions. (b) A two-layer HRCNN.

$M_{ji}$  and  $m_{ji} \in R$  are adjustable weights of the  $j$ -th neural node,  $n$  is the dimensionality of the input variables, and  $out_j(x)$  is an output function of the  $j$ -th neural node with hyperrectangular neural-type junctions.

The supervised decision-directed learning (SDDL) algorithm generates a two-layer HRCNN as shown in Fig. 2(b) in a sequential manner by adding hidden nodes as needed. As long as there are no identical data in different classes, we can obtain a 100% recognition rate for training data.

#### 3.1.1 Extraction of crispy rules

According to Eqs. (3)-(5), we know that a neural node with hyperrectangular neural-type junctions outputs one only under the following condition:

$$\begin{cases} m_{j1} \leq x_1 \leq M_{j1} \\ m_{j2} \leq x_2 \leq M_{j2} \\ \dots \\ m_{jn} \leq x_n \leq M_{jn} \end{cases} \quad (6)$$

Therefore, the classification knowledge can be described in the form of a production rule:

$$\begin{aligned} &IF((m_{j1} \leq x_1 \leq M_{j1}) \cap \dots \cap (m_{jn} \leq x_n \leq M_{jn})) \\ &THEN out_j(x) = 1. \end{aligned} \quad (7)$$

The domain defined by the antecedent of Eq. (7) is an  $n$ -dimensional hyperrectangle. Owing to the characteristics of the data, such as the dispersion characteristic, the data may show the existence of many distinct clusters in input space. It is efficient to cope with this situation by using a set of hyperrectangles with different sizes and locations to fit the data. Each hyperrectangle corresponds to an “intermediate rule”. The final classification rule is the aggregation of the intermediate rules. This kind of expression is then mapped to a two-layer hyperrectangular composite neural network. In the composite neural network configuration shown in Fig. 2(b), input variables are assigned to input nodes, intermediate rules correspond to hidden nodes, and the final classification rule is assigned to an output node. Each hidden node is connected, with a weight of 1.0, to an output node. The output node implements the “OR” function.

### 3.1.2 Extraction of fuzzy rules

After having found a set of crispy if-then rules, we may fuzzify these crispy rules by using a reasonable fuzzy membership function. The membership function  $m_j(x)$  for the  $j$ -th hyperrectangle must measure the degree to which the input pattern  $x$  is close to the  $j$ -th hyperrectangle defined by  $[m_{j1}, M_{j1}] \dots [m_{jn}, M_{jn}]$ . As  $m_j(x)$  approaches 1, the patterns should be more contained by the hyperrectangle, with the value 1 which representing complete hyperrectangle containment. In addition, the measurement should reflect the degree of importance of the input feature on a dimension by dimension basis. The membership function that meets all these criteria is defined as

$$m_j(x) = \frac{V_j}{V_j + S_j^2 (Vol(x) - Vol_j)} \quad (8)$$

where

$$Vol_j = \prod_{i=1}^n (M_{ji} - m_{ji}) = \text{volume of the } j\text{-th hyperrectangular,} \quad (9)$$

$$Vol_j(x) = \prod_{i=1}^n \max(M_{ji} - m_{ji}, x_i - m_{ji}, M_{ji} - x_i), \quad (10)$$

and  $s_j$  is the sensitivity parameter that regulates how fast the membership values decreases as the distance between  $x$  and the  $j$ -th hyperrectangle increases. An example of this membership function for the two-dimensional case is shown in Fig. 3.

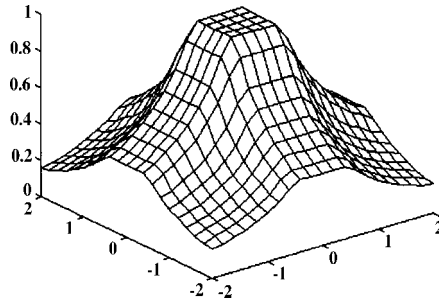


Fig. 3. An example of  $m_j(x)$  when  $s_j$  is small for the two-dimensional case.

The connection between the  $j$ -th hidden node and the output node represents the fraction of positive examples that fall within the  $j$ -th hyperrectangle. The “maximum-membership defuzzification scheme” is utilized in the last stage of a fuzzy system implemented using HRCNNs [13]. This scheme is motivated by the popular probabilistic methods of maximum-likelihood and maximum-a-posteriori parameter estimation. The other approach to finding  $w_j$  is to use the LMS or the back-propagation algorithm.

### 3.2 DRF Module

In [14], the concept of distributed representation of fuzzy rules was introduced. This approach was applied to extract fuzzy rules for classification patterns. In this approach, fuzzy rules corresponding to various fuzzy partitions are simultaneously utilized in the fuzzy inference. Assume there are  $L$  partitions in each input feature. All the fuzzy rules corresponding to the fuzzy partitions from 2 to  $L$  are simultaneously used in fuzzy inference. Therefore, the total number of distributed fuzzy rules is  $2^2 + 3^2 + \dots + L^2$ . Fig. 4 illustrates the difference between the ordinary approach and the distributed approach.

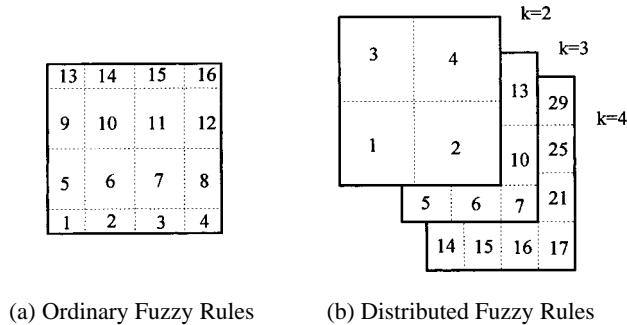


Fig. 4. Representation of different fuzzy rules ( $L=4$ ).

### 3.3 Linear Combiner of HRCNN and DRF Modules

The linear combiner, which has one layer of a linear unit, receives the outputs of the two modules and gets an optimal output. Given input vector  $\mathbf{x}$ , the  $i$ -th output unit produced by the DRF module is  $Y_i^{(D)}(x)$ , and the  $i$ -th output unit produced by the HRCNN module is  $Y_i^{(H)}(x)$ . Then, the  $i$ -th unit in the linear combiner receives these outputs through the weights  $f^{(D)}(a, b)$  and  $f^{(H)}(a, b)$ , respectively. Thus, the output  $Y_i(x)$  of the  $i$ -th unit in the linear combiner is given by

$$Y_i(x) = Y_i^{(D)}(x) \bullet f^{(D)}(Y_i^{(D)}(x)Y_i^{(H)}(x)) + Y_i^{(H)}(x) \bullet f^{(H)}(Y_i^{(D)}(x)Y_i^{(H)}(x)), \quad (11)$$

$$f^{(D)}(a, b) = \begin{cases} 1 & \text{if } a^2 - b \geq THD \\ 0 & \text{if } a^2 - b < THD, \end{cases} \quad (12)$$

$$f^{(H)}(a, b) = \begin{cases} 0 & \text{if } a^2 - b \geq THD \\ 1 & \text{if } a^2 - b < THD. \end{cases} \quad (13)$$

Using the two weighting factors, the method for selecting a module is dependent on the performance of the two modules in the silence-consonant-vowel classification. The threshold THD is fixed because these modules are speaker-independent, speaking-rate independent, and gender-independent [7, 8].

## 4. LABELING SEQUENCE DIVIDED INTO SYLLABIC UNITS BY STBNN

### 4.1 Preprocessing of the Labeling Sequence into Syllabic Units

After frames have been classified, the speech signal can be represented by a labeling sequence, e.g.,

$$[02...10],[1012...1].2212022. \quad (14)$$

As mentioned before, the achieved classification performance is usually not perfect (i.e., 100% correct). In phase I, because we use frames as the input for classification and do not consider the relations between pre-frames and post-frames, if we segment the labeling sequence into syllabic units according to a simple strategy, where a transition in the labeling sequence represents the existence of a new syllabic unit, then a lot of unnecessary syllabic units will be added. Therefore, we need more complicated control strategies to limit the range of segmentation points and to label possibilities. Here, we propose a method to offset the imperfect classification performance. From observations of some labeling sequences, we have found the following regularities:

- (1) The continuation of label 0 will not exceed 5 frames for consonants. This is the difference between consonants and silence.
  - (2) The consonants will usually start with label 1. Sometimes, a syllable, such as /r/, /b/, /l/, /g/, /d/, /m/, and /N/, will begin with label 2. These properties can indicate the start-point of a syllable.
  - (3) Then, there will exist at least 5 continuous frames labeled 2 so as to end the syllabic unit.
- Table 2 illustrates some labeling sequences.

Based on these observations, we can represent these three regularities with a finite state network (FSN) or nondeterministic finite automata (NFA) as shown in Fig. 5. By using the NFA, we can remedy the labeling errors which occur in the first phase [7].

**Table 2. Some possible forms of syllables.**

the forms of syllables	
consonants	vowels
110	
112	
11...1	
11...10	
11...101000	
11..1000000	2222...2
1112001	
12	
122	
22211...1	
20210	

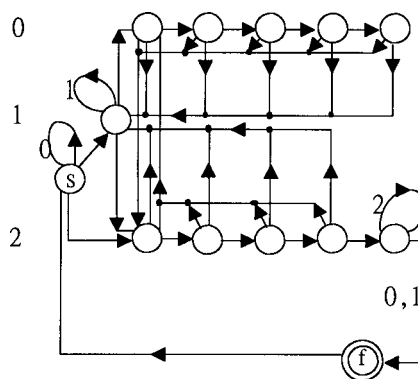


Fig. 5. The nondeterministic finite automata used to offset the labeling errors which occur in phase I.

Owing to different speaking rates and characteristics of Mandarin words, a long sequence of 2's will exist. This will result in the "V-V concatenation problem". That is, only one syllabic unit will be detected even if several vowels are pronounced sequentially. For

the second phase, we use STBNN with the energy and time derivative of the envelope of the unbiased log spectrum [12] to solve this problem within the continuation of label 2 sequence [11].

### 4.2 The Structure of STBNN

The mapping capability of an artificial neural network (ANNs) depends on its structure, including the numbers of layers and hidden nodes. A small network with few hidden layers and processing elements might not be adequate for a complex problem, and using a large network with too many layers and processing elements might lead to poor generalization (Alpaydion, 1991). In this section, we will describe the use of a dynamic learning algorithm (Nabhan, 1994) for multilayered feedforward ANNs which is called STBNN, to solve the V-V concatenation problem.

We use a dynamic algorithm to modify the network structure, and the structure can be modified by adding or deleting neurons/layers as shown in Fig. 6.

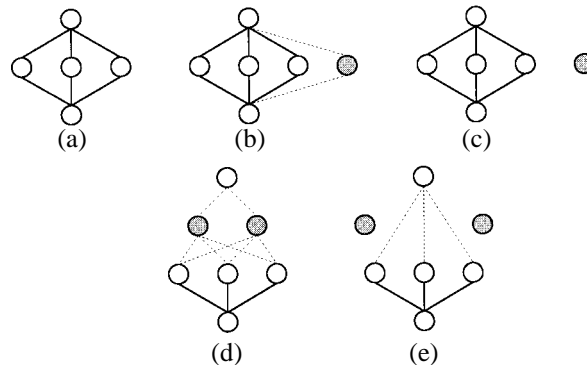


Fig. 6. Structure modification. (a) The original structure. (b) Neuron is added. (c) The added neuron is deleted. (d) Layer is added. (e) The added layer is deleted.

The number of neurons in the new layer  $N_{new}$  is given by

$$N_{new} = integer\left(\frac{N_0 + N_h}{2}\right), \tag{15}$$

where  $N_0$  is the number of neurons in the output layer.  $N_h$  is the number of neurons in the last hidden layer. For a three-layer network structure, the number  $N$  of neurons in the hidden layer is given as follows:

$$N = integer\left(\frac{N_0 + N_i}{2}\right), \tag{16}$$

where  $N_i$  is the number of neurons in the input layer. Fig. 7 shows the iterations among the different models of the STBNN system.

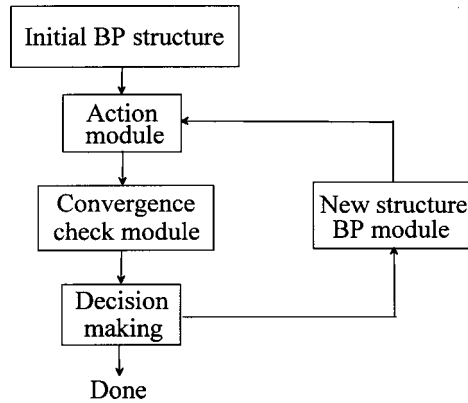


Fig. 7. The structure of STBNN.

## 5. PERFORMANCE EVALUATION

### 5.1 HNFN, HRCNN and DRF Performance Comparison

The overall experimental data were taken from news uttered in Mandarin by two females and two males. The speech signal was sampled at a rate of 10k Hz. The frame length was 25.6ms, and its period was 10ms. There were in total 22590 frames in our database. First, we used 2000 training frames for HNFN, HRCNN and DRF to extract distributed fuzzy classification rules and both crispy and fuzzy classification rules, respectively. The results are shown in Fig. 8 (a-c), respectively. From Fig. 8, we find that the decision boundaries are nonlinear, unlike the result of conventional approach based on

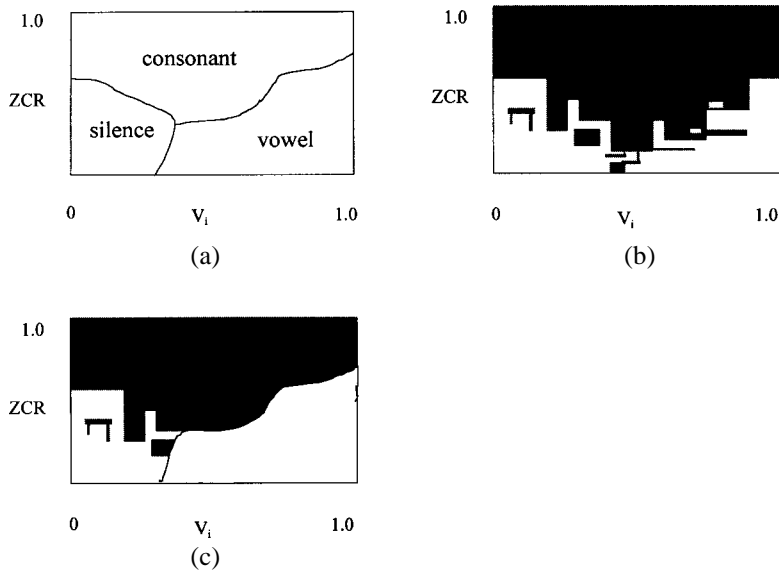


Fig. 8. The decision regions of the (a) DRF, (b) HRCNN, (c) HNFN.

thresholding features, which will result in linear decision boundaries. Then an open test was implemented on both systems. Table 3 shows the results. From Table 3, we find that the HNFN approach is better than the HRCNN and DRF approaches based on comparison of (1) the number of fuzzy rules and (2) the correct recognition performance.

**Table 3. Labeling performance achieved by DRF, HRCNN, HNFN: training set (8.85%) and testing set (91.14%).**

System	No. of Rules	Class			Average
		silence	consonant	vowel	
DRF	1,239	89.1%	74.5%	99.4%	87.7%
HRCNN (Fuzzy)	132	94.2%	89.7%	94.7%	91.2%
HNFN (Fuzzy)	134	96.1%	85.1%	98.2%	93.1%

After frames have been classified, the speech signal can be represented by a labeling sequence, e.g., [00...0], [11...1].22...200..0. Then, preprocessing of the labeling sequence into syllabic units will be done to offset the imperfect classification performance using the transition state.

But owing to different speaking rates and the characteristics of Mandarin words, a long sequence of 2's will exist. This will result in the "concatenating-syllable problem". That is, only one syllabic unit will be detected even if several vowels are pronounced sequentially. In phase II, we solve this problem using the STBNN method.

## 5.2 System Evaluation

In order to simulate the circumstances surrounding V-V concatenation, we used "V-X-V" type syllables as the experimental data, where "V" corresponds to a vowel and "X" corresponds to a group of vowels. "X" consists of the following syllables: /ai/, /au/, /an/, /ang/, /in/, /ing/, /ei/, /en/, /ou/, /ong/, /e'n/, /e'ng/, /iun/, /ia/, /ie/, /iau/, /iou/, /ua/, /uo/, /uai/, /uei/, and /ue/. For each training frame "V-X-V", the input parameters to the STBNN system are the energy and time derivative of the envelope of the unbiased log spectrum [12]. There were in total 528 speech data of "V-X-V", which were uttered six times in Mandarin by the original four speakers mentioned in the above section. Table 4 shows the result. There was an error rate of approximately 5% such that the boundary of V-V concatenation could not be detected. The correct boundary of segmentation was decided by observing the waveform of the speech, energy and time variation of the envelope of the spectrum, and by listening to the test speech.

**Table 4. The result of "V-X-V" syllable segmentation with STBNN.**

Speaker	Male 1	Male 2	Female 1	Female 2
Insertion	8	12	15	16
Deletion	35	38	32	38
Total error	43	50	47	54
Boundary	1,056	1,056	1,056	1,056



## REFERENCES

1. C. Furuichi and S. Imai, "Phonemic units segmentation in various phonetic environments," *Transactions on Institute of Electronics, Information and Communication Engineers*, Vol. J72-D-II, No. 8, 1989, pp.1221-1227.
2. C. T. Hsieh and J. T. Chien, "Segmentation of continuous speech into phonemic units," in *Proceedings of International Conference on Information and Systems*, 1991, pp. 420-424.
3. H. K. Liao, "Phonemic units segmentation and recognition for Mandarin Chinese in various phonetic environments," M.S. Thesis, Institute of Computer Science, Tamkang University, 1993.
4. L. R. Rabiner and M. R. Sambur, "An algorithm for determining the endpoints of isolated utterances," *The Bell System Technical Journal*, Vol. 54, No. 2, 1975, pp. 297-315.
5. R. G. Chen, "Automatic segmentation techniques for Mandarin speech recognition," M. S. Thesis, Department of Electrical Engineering, National Taiwan University, 1978.
6. N. DeClaris and M. C. Su, "A novel class of neural networks for knowledge representation from experimental data," *International Symposium on Artificial Neural Networks*, 1993, pp. SC1-01-07 (tutorial notes).
7. C. T. Hsieh and S. C. Chien, "Speech segmentation and clustering problem based on fuzzy rules and transition states," *Twelfth International Association of Science and Technology for Development International Conference on Applied Informations*, 1994, pp. 291-294.
8. C. T. Hsieh, M. C. Su and S. C. Chien, "Use of a self-learning neuro-fuzzy system for syllabic labeling of continuous speech," in *Proceedings of the International Joint Conference of the Fourth IEEE International Conference on Fuzzy System and the Second Fuzzy Engineering Symposium*, Yokohama Japan, 1995, pp. 1727-1734.
9. T. M. Nabhan and A. Y. Zomaya, "Toward generating neural network structures for function approximation," *Neural Network*, Vol. 7, No. 1, 1994, pp. 89-99.
10. M. C. Su, C. T. Hsieh and L. J. Inh, "Use of the hybrid neuro-fuzzy system for syllabic speech segmentation," *International Association of Science and Technology for Development International Conference on Modelling and Simulation*, Pittsburgh, 1995, pp. 293-296.
11. C. T. Hsieh, M. C. Su, E. Lai, Y. H. Lee and L. J. Inh, "A segmentation method for continuous speech utilizing neural network," *International Association of Science and Technology for Development International Conference on Signal and Image Processing*, Las Vegas, 1995, pp. 159-163.
12. S. Imai and C. Furuichi, "Unbiased estimator of log spectrum and its application to speech signal processing," *Transactions on Institute of Electronics, Information and Communication Engineers*, Vol. J70-A, No. 3, 1987, pp. 471-480.
13. B. Kosko, *Neural Networks and Fuzzy System: A Dynamical Systems Approach to Machine Intelligence*, Prentice-Hall, NJ, 1992.
14. H. Ishibuchi, K. Nozaki and H. Tanaka, "Distributed representation of fuzzy rules and its application to pattern classification," *Fuzzy Sets and System*, Vol. 52, 1992, pp. 21-31.



**Ching-Tang Hsieh** ( 謝堂 ) is an associate professor of Electrical Engineering at Tamkang University, Taiwan, Republic of China. He received the B. S. degree in electronics engineering in 1976 from Tamkang University and the M. S. and Ph.D. degrees in 1985 and 1988, respectively, from the Tokyo Institute of Technology, Japan. From 1990 to 1996, he acted as the Chairman of the Department of Electrical Engineering. His current research interests include speech analysis, speech recognition, natural language processing, image processing, neural networks, and fuzzy systems.



**Mu-Chun Su** ( 蘇文坤 ) received the B. S. degree in electronics engineering from National Chiao Tung University, Taiwan, Republic of China, in 1986, and the M. S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, MD, in 1990 and 1993, respectively. He was the 1992 IEEE Franklin V. Talor Award recipient for the most outstanding paper, co-authored with Dr. N. DeClaris. He is an associate professor of Electrical Engineering at Tamkang University, Taiwan. His current interests include neural networks, fuzzy systems, computer-aided medical systems, pattern recognition and speech processing.



**Eugene Lai** ( 賴文 ) was born in April 5, 1941, and received B. S. degree in National Taiwan University, Republic of China, in 1963, and Ph.D. degree in Iowa State University in 1971. He is a professor of Electrical Engineering at Tamkang University, Taiwan. His current interests include electromagnetics and semiconductor physics.



**Chih-Hsu Hsu** ( 蘇智 ) received the B. S. degree in electrical engineering from Fengchia University, Taiwan, Republic of China, in 1990, and the M. S. degree in information engineering from Tamkang University, Taiwan, in 1993. Currently, he is working toward the Ph.D. degree in information engineering of Tamkang University and a lecture of information engineering at Dahan Junior College of Engineering and Business, Taiwan. His research interests include speech synthesis and recognition, signal processing, fuzzy system, and neural networks.