行政院國家科學委員會專題研究計畫 成果報告

以鳥鳴聲為基礎之鳥種辨識系統架構之研究與開發(I)
研究成果報告(精簡版)

計 畫 類 別 ：個別型
計 畫 編 號 ：NSC 96-2221-E-216-031-
執 行 期 間 ：96 年 08 月 01 日至 97 年 07 月 31 日
執 行 單 位 ：中華大學資訊工程學系

計畫主持人：周智勳

計畫參與人員：碩士班研究生-兼任助理人員：丁泓承

報 告 附 件 ：出席國際會議研究心得報告及發表論文

處 理 方 式 ：本計畫涉及專利或其他智慧財產權，2 年後可公開查詢

中 華 民 國 97 年 10 月 23 日

On the Studies of Syllable Segmentation and Improving MFCCs for Automatic Birdsong Recognition

Chih-Hsun Chou*, Pang-Hsin Liu and Bingjing Cai+

Department of Computer Science and Information Engineering,
Chung Hua University, No.707, Sec.2, WuFu Rd.,
Hsinchu, 30067 Taiwan, R.O.C.
*E-mail: chc@chu.edu.tw

+School of Software, Yunnan University, Kunming, Yunnan,
P.R. China, 650091

Abstract

Birdsongs are typically divided into four hierarchical levels: note, syllable, phrase, and song, of which syllable plays an important role in bird species recognition. To improve the recognition rate of birdsongs, in this study an enhanced syllable segmentation method based on R-S endpoint detection method was presented. Furthermore, a decision based neural network with suitable reinforcement learning rule was developed as the classifier. The proposed methods combined with the well-known MFCCs feature vector form a birdsong recognition system that was applied to two recognition problems: one is the recognition of a set of arbitrary syllables and the other is the recognition of a section of a birdsong. Experimental results show the performances of the proposed methods.

Keywords: Birdsong recognition, syllable segmentation, Mel-frequency cepstral coefficients, decision based neural network.

1. Introduction

Birdsongs are typically divided into four hierarchical levels: note, syllable, phrase, and song [1]. Among them syllable was frequently used in bird species recognition [2-9], and the endpoint detection process is an important step for syllable segmentation. Endpoint detection detects the waveform boundaries of a signal so as to extract the voice parts and ignore the noise parts of the signal. The detection process can be accomplished via time domain approach and frequency domain approach. In time domain approach, detecting factors such as energy and zero crossing rates were widely applied [10-13]. Applying these two factors requires suitable methods for setting the decision threshold values. In addition to time domain, the energy computed in the frequency domain, called the sub-band energy, was also used as the segmentation index [14, 15]. Energies computed in both domains were applied in [16]. Other factors in the frequency domain such as spectrum, cepstrum and cepstrum entropy were used in the studies of [17-19].

A well known time domain technique named RS (Rabiner and Sambur) method [20] used signal energy to roughly search the locations of the start and end points of the voice parts and used the zero-crossing rates for fine adjusting. This method though useful in a silent environment is not robust in noisy environments [21]. On the other hand, due to the resistance to signal fading and echoing, the frequency domain analysis of the birdsong signal is better than the time domain approach [9]. It implies the necessity of frequency domain information of the birdsong for syllable segmentation. Syllable endpoint detection in the frequency domain is usually achieved by considering the magnitudes of the frequency bins of the signal. This approach, however, is uneasy to arrive at a satisfactory result if noticeable background noise is unavoidable. Because the respective shortcomings of both time-domain and frequency-domain approaches, the first subject of this study is to integrate the frequency-domain information for the RS method to design a syllable segmentation process called the improved RS method (IRS).

In the design of a recognition system, a suitable classifier is usually required for manipulating specific feature. For birdsong recognition, the dynamic time warping (DTW) algorithm and the hidden Markov model (HMM) were used in [2, 22] for recognizing two bird species. The self-organizing map and supervised multilayer perceptron were applied in [23] for recognizing eight bird species. In [24], Gaussian mixed models (GMMs) were constructed for recognizing 11 bird species frequently appeared in the airport. The study in [4] combined the template-based technique and the time delay neural networks (TDNNs) to recognize the syllables of 16 bird species, while both Quadratic Discriminant Analysis and neural network were utilized in [7-9]. In [25], SVM (support vector machine) was applied to recognize two sets of birdsongs of six and eight bird species. In [6], Gaussian prototype was applied to model the histograms of birdsongs, and classification was achieved by comparing the similarity between the Gaussian prototypes. The second subject of this study is to develop a decision based neural network with suitable reinforcement learning rule for constructing an automatic bird species recognition system.

The remaining of this paper is as follows: Section 2 proposes the developed methods. Section 3 shows the experimental results. Section 4 is the conclusions.

2. The Proposed Birdsong Recognition System

The block diagram of the proposed system as shown in Fig. 2.1 containing preprocessing, feature extraction and recognition (species decision) are described in detail in the following.

2.1 *Preprocessing*

The procedure of preprocessing in this study is shown in Fig. 2.2, which contains four steps: improved syllable endpoint detection, normalization, pre-emphasis and segmentation.

2.1.1 Improved syllable endpoint detection

As stated in Section 2, both time-domain and frequency-domain information were employed in the proposed approach. This approach, called the improved RS method (IRS), is described in the following.

Step 1 Segment the input signal by using the RS method, and denote the result as $x(t)$.

Step 2 Compute the short time Fourier transform of $x(t)$ with frame size N = 512, and form the spectrogram of the signal.

Step 3 For each frame m, find the frequency Bin bin_m with the greatest magnitude.

Step 4 Initialize the syllable index j, $j = 1$.

Step 5 Compute the frame t at which the maximum magnitude occurs

$$t = \arg \max_{1 \le m \le M} (|X[bin_m]|), \tag{1}$$

and set the amplitude of syllable j as

$$A_j = 20 \cdot \log_{10} |X[bin_t]| (\mathrm{dB}), \tag{2}$$

in which M is the number of frames of $x(t)$, and $X[\cdot]$ denotes the spectrum of $x(t)$.

Step 6 Start from frame t and move backward and forward until frames h_j and t_j such that both $20 \cdot \log_{10} |X[bin_{h_j}]|$ and $20 \cdot \log_{10} |X[bin_{t_j}]|$ are smaller than $(A_j - 20)$ (dB).

Step 7 Start from frames h_j and t_j, find frames $h_j - \alpha$ and $t_j + \beta$ (α, β > 0) such that both $20 \cdot \log_{10} |X[bin_{h_j - \alpha - 1}]|$ and $20 \cdot \log_{10} |X[bin_{t_j + \beta + 1}]|$ are greater than $(A_j - 20)$. Then $h_j - \alpha$ and $t_j + \beta$ are called the head frame and tail frame of syllable j.

Step 8 Set $|X[bin_m]| = 0$,

$$m = h_j - \alpha, h_j - \alpha + 1, \cdots, t_j + \beta - 1, t_j + \beta. \tag{3}$$

Step 9 Let $j = j + 1$.

Step 10 Repeat Step 5 to Step 9 until $A_j < A_1 - 20$.

By using the above procedure, the boundaries of each syllable can be obtained for syllable segmentation. In Section 3, a comparison of birdsong recognition efficiencies for this method will be given.

2.1.2 Normalization and pre-emphasis

In this study, the amplitudes were linearly normalized to the region of [-1, 1]. Besides, since the amplitude of high frequency part is usually much smaller than the low frequency part, so pre-emphasis is applied to intensify the signal of high frequency part. The intensification is accomplished by a finite impulse response (FIR) filter with the following form

$$H(z) = 1 - a \cdot z^{-1}, \tag{4}$$

so that a signal $x[n]$ after the filtering process has the following property

$$\bar{x}[n] = x[n] - ax[n-1], \tag{5}$$

in which α is a scalar usually between 0.9 and 1, and was set as 0.95 in this study.

2.2 *Feature extraction – the MFCCs*

Birdsongs are non-stationary signals requiring short-time analysis. The short time analysis for the signal obtained from Sec. 2.1 is accomplished by applying a Hamming window with size N equaling 512 samples. The equation of the applied Hamming window is

$$w[n] = \begin{cases} 0.54 - 0.46 \cos\left(\dfrac{2\pi n}{N-1}\right), 0 \le n \le N-1 \\ 0 \qquad\qquad\qquad\qquad\quad, \text{otherwise} \end{cases} \tag{6}$$

For checking the proposed endpoint detection method and the DBNN classifier, MFCCs based feature extraction for each syllable was used to construct the feature vector. This extraction process is described in the following

Step 1 Compute the fast Fourier transform (FFT) of each framed signal.

$$\tilde{x}[k] = \sum_{n=0}^{N-1} x[n] w[n] e^{-j2\pi nk/N}, 0 \le k < N, \tag{7}$$

Step 2 Compute the energy of each triangular filter band

$$E_j = \sum_{k=0}^{N/2-1} \phi_j[k] |\tilde{x}[k]|^2, 0 \le j < J, \tag{8}$$

where $\phi_j[k]$ denotes the amplitude(weight) of the j^{th} triangular filter at frequency bin k as shown in Fig. 2.3, E_j denotes the energy of j^{th} filter band, and J is the number of filters.

Step 3 Compute the MFCCs by Cosine transformation

$$c_i(m) = \sum_{j=0}^{J-1} \cos\left(m \frac{\pi}{J}(j + 0.5)\right) \log_{10}(E_j),$$

$$m = 1, 2, \ldots, 15, \tag{9}$$

where $c_i(m)$ denotes the m^{th} order MFCC of the i^{th} frame.

Step 4 The coefficients of the same order of all frames were averaged and then used to form a 15-dimensional syllable feature vector.

Such a feature vector was well-applied in many studies, so we used it for checking purpose. In the NN training process, the feature vectors of all training syllables were used as the inputs and the corresponding bird species as the desired outputs.

2.3 *Recognition by using the decision based neural network (DBNN)*

When applying the neural networks (NNs), one of the two learning types, supervised learning and unsupervised learning, is adopted. Supervised NNs can also be divided into two types named approximation-based formulation and decision-based formulation [26] depending on the arrangement of training data. The function of approximation-based formulation is to approximate the mapping between input and output data so as to minimize the mean square error between the network outputs and desired outputs. An example of such a type NN is the back-propagation NN with least mean squares learning rule. The decision-based formulation is usually used to decide which class the input data belongs to, means that it is more suitable for data classification [26]. The structure of the decision-based neural network (DBNN) applied in this study is shown in Fig. 2.4, in which the weight vectors $\mathbf{w}_j^i$ of the function ϕ_j^i were trained by using reinforcement type learning rules. The training process is described in the following.

Step 1 Initialize the weight vectors $\mathbf{w}_j^i$ of the function ϕ_j^i, $i = 1,2,...,C$, $j = 1,2,3$, where ϕ_j^i is a radius basis function of the form

$$\phi_j^i = \phi\left(\mathbf{x}, \mathbf{w}_j^i\right) = -\frac{\left\|\mathbf{x} - \mathbf{w}_j^i\right\|^2}{2}, \tag{10}$$

C is the number of classes and $\mathbf{x}$ is the input feature vector.

Step 2 Input the feature vector, $\mathbf{x}$, of the training birdsong whose bird species class is set as the *actual class*.

Step 3 Compute the value of each basis function ϕ_j^i.

Step 4 For each class network i, find the local winner $l = \arg\max_j(\phi_j^i)$ and $\phi_l^i = \max_j(\phi_j^i)$.

Step 5 For all local winners ϕ_l^i, $i = 1,2,...,C$, find the global winner, $g = \arg\max_i \phi_l^i$ and $\phi^g = \max_i \phi_l^i$.

Step 6 If g = *actual class*, then perform the reinforcement learning

$$\mathbf{w}_j^g = \mathbf{w}_j^g + \eta \nabla \phi(\mathbf{x}, \mathbf{w}_j^g), j = 1, 2, 3, \tag{11a}$$

else the anti-reinforcement learning

$$\mathbf{w}_j^g = \mathbf{w}_j^g - \eta \nabla \phi(\mathbf{x}, \mathbf{w}_j^g), j = 1, 2, 3, \tag{11b}$$

where $\mathbf{w}_j^g$ denotes the weight vector of ϕ_j^g and η is the learning rate.

In this study, the function ϕ_j^i is defined as a radius basis function as shown in (10), so the gradient operation result in (11) is

$$\nabla \phi(\mathbf{x}, \mathbf{w}_j^g) = \mathbf{x} - \mathbf{w}_j^g. \tag{12}$$

In the recognition process, the feature vector of a test birdsong was obtained by the same process as the training part. After inputting the feature vector to the DBNN, the global winner of the network (i.e. the network output) indicates the species class the test birdsong belongs to. Because back-propagation NN (BPNN) was widely applied in many studies, both BPNN and DBNN were applied in this experiment for comparison.

3. Experimental results

The bird species vocalization database used in this study was obtained from a commercial CD [27] containing both birdcall and birdsong files of 420 bird species recorded in the field. Each file contains vocalizations of the same bird species. The database of 420 bird species makes it much larger than any used in previous studies. Meanwhile, recordings in the field are usually in noisy environment, incomplete and interrupted. Because of this, sometimes, only some syllables of a birdsong, rather than a complete birdsong, could be used in the recognition process. So two recognition problems were applied in the experiments, one is the recognition of a set of arbitrary syllables of a bird species and the other is the recognition of a section of a birdsong. The sampling rate of these vocalization signals is 44.1 kHz with 16-bit resolution and a monotone type PCM format. In the experiment, the frame size was set as 512 samples with three-fourths frame overlapping.

3.1 *Recognition of a set of arbitrary syllables from the song of a bird species*

The purpose of the first experiment is to examine the efficiencies of the improved RS endpoint detection method and the using of DBNN. For comparison, both RS and IRS methods were used to segment the syllables of each birdsong file. After segmentation, half syllables of each birdsong file were randomly selected for training and the remaining for testing. The same feature extraction process was applied to both training and testing syllables. The extracted feature vector of a test syllable was used as the input of the NN whose output indicates the bird species of the test syllable. By comparing the network output and the actual species for each test syllable, the recognition rates (*RR*s) of all test syllables can be obtained. Table 3.1 shows the *RR*s of using both segmentation methods and both types of NN. It shows that the DBNN improved the *RR* of about 5% to

7% on the BPNN for both segmentation methods. Meanwhile, the IRS segmentation method also improved the *RR* of about 3% to 5% on the RS method. In addition, the combination of IRS segmentation method with DBNN achieved a *RR* of 64.93% evidently higher than the result of using BPNN and the RS segmentation method (54.27%).

3.2 *Recognition of a section of a birdsong*

The second experiment is the recognition of a section of a birdsong containing several consecutive syllables. Meanwhile, both segmentation methods and both types of NN were also used for comparison. All the syllables in the birdsong section were recognized individually, and the species with the largest number of NN outputs was recognized as the species of the test birdsong. So in this case, the recognition rate *RR* was defined as

$$RR(\%) = \frac{\text{number of birdsongs recognized correctly}}{\text{number of all birdsongs}} \cdot 100\% \quad (13)$$

The *RRs* of using all the four structures are shown in Table 3.2. It can be found that when the NN is fixed, the IRS segmentation method improved the *RRs* of 3.64% and 0.31% on RS method. When the segmentation method is fixed, the DBNN improved the *RRs* of 10.68% and 7.35% on BPNN. The above comparisons exhibit the efficiencies of the proposed methods.

3.3 *Recognition with threatened birdsongs taken into account*

The territoriality instinct is innate for some bird species so that they vocalize threaten voices when infringed. The threaten voice is used to warn the companies and frighten the invaders. Such a type of vocalization is usually distinct from the regular birdsong requiring additional categorization. That is, two types of birdsongs are considered if the bird species also vocalizes threatened voices. In this case, feature extraction process is apply for both regular birdsongs and threatened voices so that a bird species was represented by two feature vectors. In the recognition process, the extracted feature vector of the test syllable (birdsong) was matched to both feature vectors of all species to determine the most likely species.

This experiment also dealt with the recognition of a section of a birdsong. The *RRs* of using all the four structures with the above distinction are shown in Table 3.3, in which titles A and B represent the recognition without and with the distinction of threatened voices. It can be found that the *RRs* of case B exceeded case A of about 3.87% to 7.51%.

4. Conclusion

Recordings in the field are usually in noise environment, incomplete or interrupted making the birdsong recognition much harder. To overcome these problems, the processes for syllable segmentation and classification were studied. In the syllable segmentation, an improvement of the RS method that simultaneously considers the time domain and frequency domain characteristics of the syllable was proposed. For the study of classifier, a DBNN with reinforcement learning rule was presented and compared to the using of BPNN. Two types of recognition problems were considered in the experiments, one is the recognition of a set of arbitrary syllables and the other is the recognition of a section of a birdsong. Experimental results show that the proposed methods evidently improved the *RRs* regardless of the consideration of threatened voices.

References

[1] C.K. Catchpole and P.J.B. Slater, *Bird Song*: *Biological Themes and Variations*, Cambridge University Press, 1995.

[2] S.E. Anderson, A.S. Dave and D. Margoliash, "Template-based automatic recognition of birdsong syllables from continuous recordings," *Journal of the Acoustical Society of America*, vol. 100, no. 2, pp. 1209-1219, 1996.

[3] A. Härmä and P. Somervuo, "Classification of the harmonic structure in bird vocalization," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. V-701-4, 2004.

[4] S.A. Selouani, M. Kardouchi, E. Hervet and D. Roy, "Automatic birdsong recognition based on autoregressive time-delay neural networks," in *Proceedings of the ICSC Congress on Computational Intelligence Methods and Applications*, pp. 1-6, 2005.

[5] A. Härmä, "Automatic identification of bird species based on sinusoidal modeling of syllables," in *Proceedings of the IEEE International Conference on Acoustics, Speech, Signal Processing*, vol. 5, pp. 545-548, 2003.

[6] P. Somervuo, "A. Härmä, Bird song recognition based on syllable pair histograms," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. V-825-8, 2004.

[7] A.L. McIlraith and H.C. Card, "Bird song identification using artificial neural networks and statistical analysis," in *Proceedings of the Canadian Conference on Electrical and Computer Engineering*, vol. 1, pp. 63-66, 1997.

[8] A.L. McIlraith and H.C. Card, "A comparison of backpropagation and statistical classifiers for bird identification," in *Proceedings of IEEE International Conference on Neural Networks*, vol. 1, pp. 100-104, 1997.

[9] A.L. McIlraith and H.C. Card, "Birdsong recognition using backpropagation and multivariate statistics," *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2740-2748, 1997.

[10] J.G. Wilpon, L.R. Rabiner and T. Martin, "An improved word-detection algorithm for telephone-quality speech incorporating both syntactic and semantic constraints," *AT&T Bell labs. Tech. J.*, vol, 63, pp. 479-498, Mar. 1984.

[11] L.R. Rabiner and M.R. Sambur, "An algorithm for determining the endpoints of isolated utterances", *Bell Syst. Tech. J.*, vol.54, pp. 297-315, Feb. 1975.

[12] J.C. Junqua, B. Reaves and B. Mark, "A study of endpoint detection algorithms in adverse conditions: Incidence on a DTW and HMM recognize," *Proceedings of Eurospeech*, pp. 1371-1374, 1991.

[13] S.N He, and J.B Yu, "A novel Chinese continuous speech endpoint detection method based on time domain features of the word structure," *IEEE International Conference on Communications*, *Circuits and Systems and West Sino Expositions*, vol. 2, pp. 992-996, July 2002.

[14] W.J. Zhang, and J.Y. Xie, "Endpoint detection based on MDL using subband speech satisfied auditory model," *IEEE International Conference on Neural Networks and Signal Processing*, vol. 2, pp. 892-895, Dec. 2003.

[15] S.N. He and J.B. Yu, "The performance analysis of Chinese speech endpoint detection based on continuous multi sub-band spectral features," *Proceedings of the IEEE International Conference on Communications, Circuits and Systems and West Sino Expositions,* vol. 2, pp. 997-1002, July 2002.

[16] S.E. Bou-Ghazale and K. Assaleh, "A robust endpoint detection of speech for noisy environments with application to automatic speech recognition," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, pp. 3808-3811, May 2002.

[17] J.A. Haigh, and J.S. Mason , "Robust voice activity detection using cepstral features," *Proceedings of the IEEE TENCON* , pp. 321-324, 1993.

[18] B.F. Wu, and K.C. Wang, "Robust Endpoint Detection Algorithm Based on the Adaptive Band-Partitioning Spectral Entropy in Adverse Environments," *IEEE Transactions on Speech and Audio Processing*, pp. 762-775, Sept. 2005.

[19] H.G. Xu, H.S. Li, J. Liu and R.S. Liu, "Endpoint detection algorithm for Mandarin digit recognition using DSP," in *Proceedings of the 6th International Conference on Signal Processing*, vol. 1, pp. 548-551, Aug. 2002.

[20] 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, pp. 297-315, 1975.

[21] K. Kitayama, M. Goto, K. Itou and T. Kobayashi, "Speech Stgarter: Noise-robust endpoint detection by using filled pauses," *Proceeding of the 8th European Conference on Speech Communication and Technology* (*Eurospeech 2003*), pp. 1237-1240, Setp. 2003.

[22] J.A. Kogan and D. Margoliash, "Automated recognition of bird song elements from continuous recordings using dynamic time warping and hidden Markov models: A comparative study," *Journal of the Acoustical Society of America*, vol. 103, no. 4, pp. 2185-2195, 1998.

[23] A. Selin, J. Turunen and J.T. Tanttu, "Wavelets in recognition of bird sounds," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 51806, 9 pages.

[24] C. Kwan et al., "An automated acoustic system to monitor and classify birds," *EURASIP Journal on Applied Signal Processing*, vol. 2006, Article ID 96706, pp. 1-19.

[25] S. Fagerlund, "Bird species recognition using support vector machines," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, Article ID 38637, 8 pages.

[26] S.Y. Kung, *Digital Neural Networks*, Prentice-Hall, Inc., 1993.

[27] T. Kabaya and M. Matsuda, *The Songs & Calls of 420 Birds in Japan*, SHOGAKUKAN Inc., Tokyo, 2001.

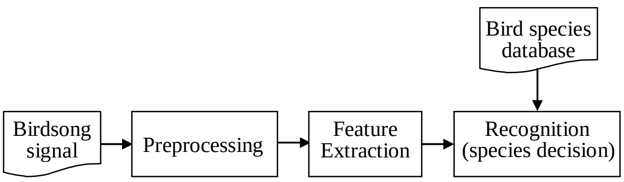

Figure 2.1 Block diagram of the proposed system

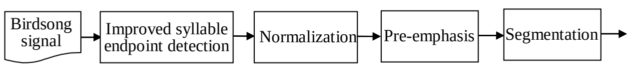

Fig. 2.2 Block diagram of the preprocessing

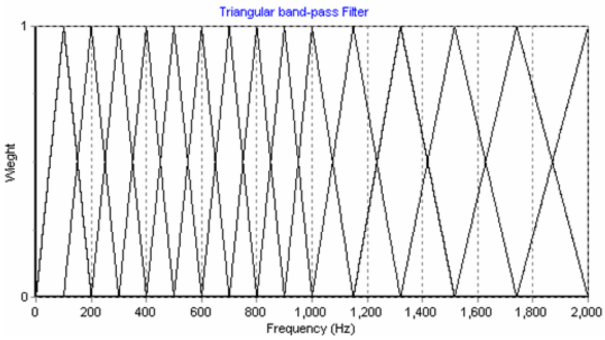

Figure 2.3 Applied triangular filters for computing the MFCCs

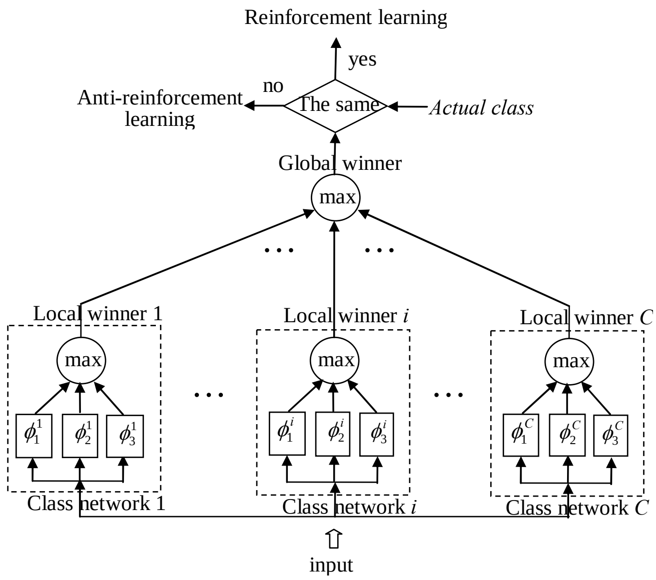

Figure 2.4 Structure of the applied DBNN.

Table 3.1 *RR*s of using different segmentation methods and different types of NN

RS segmentation		**IRS segmentation**	
MFCC+BP	MFCC+DBNN	MFCC+BP	MFCC+DBNN
54.27%	61.63%	59.76%	64.93%

Table 3.2 *RR*s of using various structures

Structure	***RRs***
RS+BPNN	54.69%
IRS+BPNN	58.33%
RS+DBNN	65.37%
IRS+DBNN	65.68%

Table 3.3 *RR*s of using various structures

Structure	**A**	**B**
RS+BPNN	54.69%	59.16%
IRS+BPNN	58.33%	64.42%
RS+DBNN	65.37%	69.24%
IRS+DBNN	65.68%	73.19%

計畫成果自評：

1. 研究內容與原計畫相符程度、達成預期目標情況：

本計劃完成音節切割的改良技巧以及開發新型式分類器，對鳥鳴聲辨識率有所貢獻。

2. 研究成果之學術或應用價值、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等

本計劃研究成果之一部份投稿至APSCC2008國際研討會(EI)並被接受，預計於12/10會中發表。另外研究成果亦已投稿至國際期刊IEICE(SCI)審查中。

出席國際學術會議心得報告

計畫編號	NSC96-2221-E-216-031
計畫名稱	以鳥鳴聲為基礎之鳥種辨識系統架構之研究與開發
出國人員姓名 服務機關及職稱	周智勳，中華大學資訊工程系副教授
會議時間地點	September 5-7, 2007 Kumamoto City International Center, Kumamoto, Japan.
會議名稱	The Second International Conference on Innovative Computing, Information and Control (ICICIC2007)
發表論文題目	Bird Species Recognition by Comparing the HMMs of the Syllables

一、 參加會議經過

這次會議於 9/4 出發，當晚先住在福岡，隔天一早搭 JR 至熊本轉電車至會場。此次會議會場分三個場地，不是很集中，感覺比較缺乏整體性。我報告所屬的 session 為 Intelligent Informatics (II) (A10-06, ICICIC-2007-1474), September 5, 2007, 13:30-15:00，該 session 大部分為日本及大陸的學者，報告方面發現大家的英文似乎都不是很好。

這次會議，地緣關係，台灣來的學者相當多，還遇到不少熟人。用餐地點有兩處設於別棟大樓，有點距離，而且似乎刻意把同一國家的學者擺一起。熊本城就在會場附近，第一天報告完論文，去該城參觀了一趟。由於是買華航自由行行程參加會議，因此三天的會議都是住福岡通車熊本，有那麼點小累。

圖一 於會場看板前留影

圖二 於會場報到處留影，報到處還特別設置台灣報到處，可見台灣來的學者相當多

二、與會心得

1. 會場設於某商用及辦公大樓之 5，6 樓，獨立性較差。
2. 半日遊我沒參加，但藉由半日遊行銷觀光是不錯的活動。
3. 個人報告的流暢性還可加強，但欠缺他人的建議。
4. 本 session 報告出席率相當完整，聆聽的人數也不少，報告者雖然英文能力普通，但也都相當盡力的發表內容。