Design and Implementation Using Neural Networks and Its Application to Hearing Aid

Munehiro Namba and Yoshihisa Ishida
Department of Electronics and Communication, Meiji University
1-1-1, Higashi-mita, Tama-ku, Kawasaki, 214 Japan
email: ee12058@isc.meiji.ac.jp

ABSTRACT — Techniques for designing digital filters using neural networks, and an application example of the filter designed by this method to digital hearing aid are described. We apply neural networks to digital filters in order to get much more flexibilities on design than conventional adaptive digital filter techniques.

INTRODUCTION

The problem of designing digital filters is one of finding filter coefficients such that some aspect of the filter response approximates a desired behavior in a specified manner. The conventional adaptive digital filter techniqe, e.g., LMS or RLS algorithm, is currently one of most popular methods for designing because most of all computational problems can be resolved with computers. However, recent studies on neural networks have realized a digital filter whose structure seems to be a worthwhile subject to investigate. Digital filters using neural networks would have any behavior, i.e. arbitrary gain and phase responses, and futhermore, it is easy to compose both of finite impulse response (FIR) filter and infinite impulse response (IIR) filter. In this article, we will discuss techniques for design and implimentational aspects of the digital filter using neural networks. At first, the overall structure of the digital filter using neural networks will be discussed, and the learning argorithm will be described. Next, we will present an applicatioin example to digital hearing aid.

STRUCTURE OF DIGITAL FILTER USING NEURAL NETWORKS

Overall structure of the filter based on our method is shown in Fig.1(a) and Fig.1(b). Actual structure is as Fig.1(a), however it seems to be shown in Fig.1(b) in the learning process. In these figures, u_k and d_k represent the input signal and the desired signal to the filter, respectively as follows :

$$u_k = \sum_{p=0}^{P} \sin(2\pi f_p k) \tag{1}$$

$$d_k = \sum_{p=0}^{P} a_p \sin(2\pi f_p k + \theta_p) \tag{2}$$

where a_p is the amplitude and θ_p is the phase shift at that point. Then we can define the error by

$$e_k = d_k - y_k \tag{3}$$

The amplitude, a_p, and phase shift, θ_p, of the desired signal are picked up at each of P points in frequency domain, then the desired signal is given as Eq.(2). On the other hand, the input signal, u_k, is given in the same way, but its amplitude and phase shift are unit and zero, respectively. States of the neural networks are adjusted by back-propagating the error value given in Eq.(3) in time domain. These figures illustrate recursive system which can realize both of FIR and IIR filters, but non-recursive system can be also constructed. The recursive system consists of two neural networks, and each network has multi-layers whose details are discussed in next section.

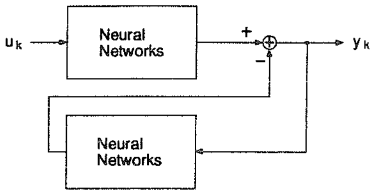

Fig.1(a) Actual structure of digital filter

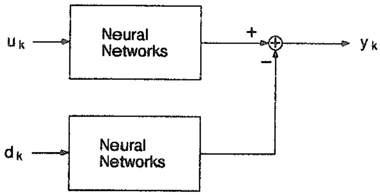

Fig.1(b) Structure of digital filter in the learning process

NETWORK LEARNING ARGORITHM

Fig.2(a) shows the structure of the neural networks. The hidden layer has three neurons, and the output layer has one neuron. All neurons are linear neurons due to the fact that nothing improvement has found in existance of neurons with sigmoidal functions in our method. The network is trained as follows :

Assuming that I_i is the input value of i-th node in the input layer, then the output value of the j-th node in the hidden layer, H_j, becomes

$$H_j = \sum_{i=0}^{M-1} W_{ij} I_i \tag{4}$$

where W_{ij} is the coupling coefficient between an input layer node and a hidden layer node. Then, the output value of the output layer node becomes

$$O = \sum_{j=0}^{N-1} H_j V_j \tag{5}$$

where V_j is the coupling coefficient between a hidden layer node and an output layer node.
Based on the principle of least square criterion, error function, E_k is defined by following equation :

$$E_k = \frac{1}{2}(d_k - y_k)^2 \tag{6}$$

We attempt to obtain a new value of V_j such that sum of the error would converge to minimum. Then, the following equation is derived.

$$\frac{\partial E_k}{\partial V_j} = \frac{\partial E_k}{\partial y_k} \cdot \frac{\partial y_k}{\partial O} \cdot \frac{\partial O}{\partial V_j} \tag{7}$$

where

$$\frac{\partial E_k}{\partial y_k} = -(d_k - y_k), \quad \frac{\partial O}{\partial V_j} = H_j, \quad \frac{\partial y_k}{\partial O} = 1 \tag{8}$$

Therefore, we achive convergence toward improved value by taking incremental changes ΔV_j, that is,

$$\Delta V_j = -\eta_1 \frac{\partial E_k}{\partial V_j} \tag{9}$$

where η_1 is the coefficient of learning rate determing convergence speed.
In the same way, we have

$$\Delta W_{ij} = -\eta_2 \frac{\partial E_k}{\partial W_{ij}} \tag{10}$$

where

$$\frac{\partial E_k}{\partial W_{ij}} = \frac{\partial E_k}{\partial y_k} \cdot \frac{\partial y_k}{\partial O} \cdot \frac{\partial O}{\partial H_j} \cdot \frac{\partial H_j}{\partial W_{ij}} \tag{11}$$

and

$$\frac{\partial E_k}{\partial y_k} = -(d_k - y_k), \quad \frac{\partial O}{\partial H_j} = V_j, \quad \frac{\partial H_j}{\partial W_{ij}} = I_i, \quad \frac{\partial y_k}{\partial O} = 1 \tag{12}$$

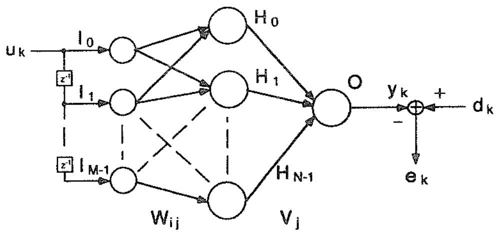

Fig.2(a) Neural Network System

A design example of IIR low-pass filter based on this learning algorithm is shown in Fig.2(b) and Fig2.(c). It is clear that the filter has an approximately linear phase response in the pass band.

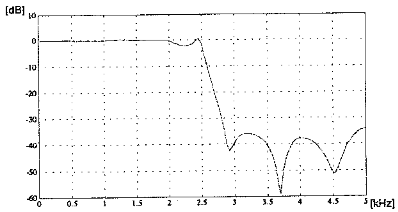

Fig.2(b) Gain Response

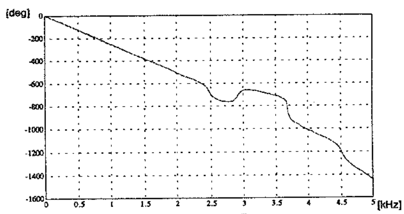

Fig.2(c) Phase Response

APPLICATION EXAMPLE TO DIGITAL HEARING AID

Next, we present an application example of the filter designed by our method to digital hearing aid. Digital filters can be composed that will make up for hearing loss which has been a common ailment in society. A generic view of hearing loss seems to be a sloping shape, that is, loss is incresing with respect to frequency. But we assume that a view of hearing loss is shown in Fig.3(a) for simplicity. To construct the filter that will make up for those loss, we determine the amplitude at each of 50 points in frequency domain and impose the linear phase constraint such that θ_p increases with regard to frequency constantly. Typicaly, the coefficient of learnig rate is constant, but if you focus attention on convergence speed, it must be varied with regard to time. However, we would like not to mention about it since the concepts of speeding up the learning rate are still under conditions of research by several persons, and it is the subject for a future study. The obtained results are shown in Fig.3(b) and Fig.3(c). As you see, we can get the filter with linear phase response.

$$\Delta W_{ij} = -\eta_2 \frac{\partial E_k}{\partial W_{ij}} \tag{10}$$

where

$$\frac{\partial E_k}{\partial W_{ij}} = \frac{\partial E_k}{\partial y_k} \cdot \frac{\partial y_k}{\partial O} \cdot \frac{\partial O}{\partial H_j} \cdot \frac{\partial H_j}{\partial W_{ij}} \tag{11}$$

and

$$\frac{\partial E_k}{\partial y_k} = -(d_k - y_k), \quad \frac{\partial O}{\partial H_j} = V_j, \quad \frac{\partial H_j}{\partial W_{ij}} = I_i, \quad \frac{\partial y_k}{\partial O} = 1 \tag{12}$$

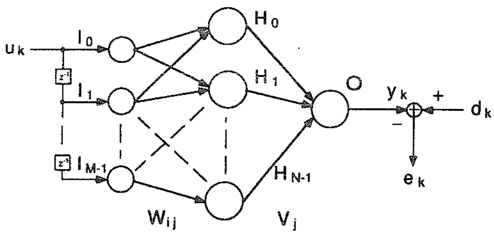

Fig.2(a) Neural Network System

A design example of IIR low-pass filter based on this learning algorithm is shown in Fig.2(b) and Fig2.(c). It is clear that the filter has an approximately linear phase response in the pass band.

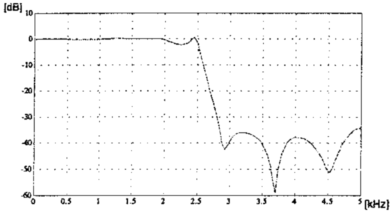

Fig.2(b) Gain Response

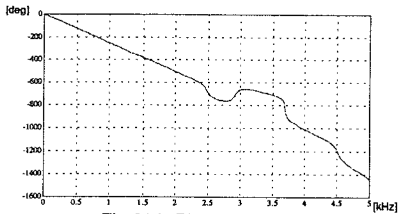

Fig.2(c) Phase Response

APPLICATION EXAMPLE TO DIGITAL HEARING AID

Next, we present an application example of the filter designed by our method to digital hearing aid. Digital filters can be composed that will make up for hearing loss which has been a common ailment in society. A generic view of hearing loss seems to be a sloping shape, that is, loss is incresing with respect to frequency. But we assume that a view of hearing loss is shown in Fig.3(a) for simplicity. To construct the filter that will make up for those loss, we determine the amplitude at each of 50 points in frequency domain and impose the linear phase constraint such that θ_p increases with regard to frequency constantly. Typicaly, the coefficient of learnig rate is constant, but if you focus attention on convergence speed, it must be varied with regard to time. However, we would like not to mention about it since the concepts of speeding up the learning rate are still under conditions of research by several persons, and it is the subject for a future study. The obtained results are shown in Fig.3(b) and Fig.3(c). As you see, we can get the filter with linear phase response.

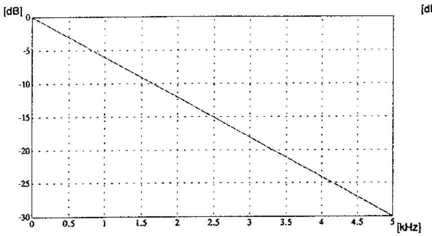

Fig.3(a) Hearing Loss

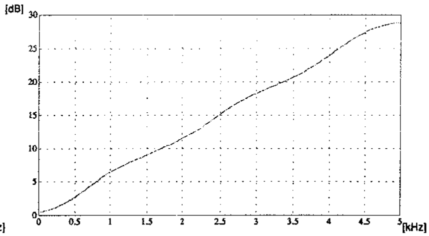

Fig.3(b) Obtained Gain Curve

If a wave form of the original speech is shown in Fig.4(a), hearing loss makes the amplitude of the speech signal to decay in high frequency domain as shown in Fig.4(b). We can recover the attenuated speech by our digital filter as shown in Fig.4(c). We think that a wave form of the speech signal could not be distorted since the obtained filter has a linear phase characteristic.

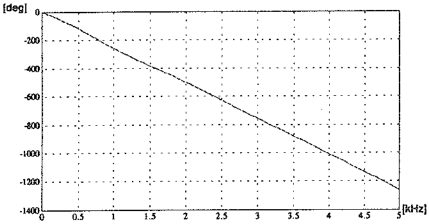

Fig.3(c) Obtained Phase Sift

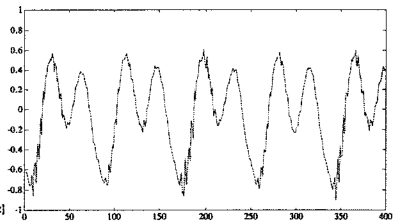

Fig.4(a) Original Speech

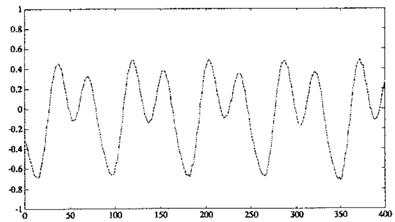

Fig.4(b) Speech Signal attenuated in high frequency domain

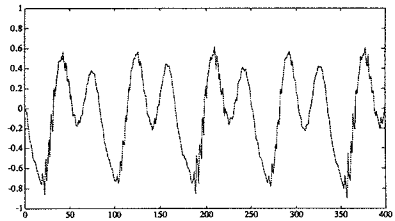

Fig.4(c) Recovered Speech

CONCLUSION

We have presented the techniques of designing digital filters using neural networks and an application example to hearing aid. By using neural networks, flexibilities on design are achieved so that one can compose a digital filter according to personal hearing loss. Clearly, these are able to be applied to any digital filter design and easily realized IIR filter with desired frequency responses which are difficult for conventional design techniques, so we think those techniques described above will receive wide application in the future. We will be glad if this article is any help to the people who concerned.

REFERENCES

Sen M.Kuo, and Steven Voepel, *Digital Hearing Aid with the Lapped Transform*, DIGITAL SIGNAL PROCESSING 3, pp.228 - 239, 1993.

Analysis and Synthesis of Human Voice Considering the Nonstationary Based on the Glottis Open and Close Characteristics

Hiroyuki KAMATA, Hiroyuki OKA and Yoshihisa ISHIDA

School of Science and Technology,
Meiji University.
1-1-1, Higashi-Mita, Tama-Ku,
Kawasaki-City, 214, Japan.

ABSTRACT - In this paper, we present a new structure of transfer function which is suitable to reconstruct the wave form of voiced speech. In case of voiced speech, there are two different states every pitch period based on the glottis open and close. Thus it is difficult to identify the voice generation system by one transfer function.
In this paper, two transfer functions that are connected as the parallel structure are used. When the compensation of the estimation error is requested, one transfer function is added to the other two transfer functions. By these three transfer functions, the voiced speech of human voice is perfectly reconstructed. These transfer functions are estimated by least mean square (LMS) method.
Furthermore, the method for reduction of voice information is also discussed in this paper.

INTRODUCTION

There are many methods that estimate the vocal tract transfer function of human voice[1]. Specifically, linear prediction (LP) is one of the methods that estimate the vocal tract transfer function of an auto-regressive (AR) model[2] and it is widely used in the area of speech processing.
On the other hand, an auto-regressive and moving-average (ARMA) model estimation is suitable for analyzing the voice that includes nasal and frictional consonants[3,4]. Furthermore, GARMA model estimation that uses the modeled glottis wave as the input signal is effective to reconstruct the voiced speech and it is valid for reduction in the voice information[5,6]. However, these methods often need the iterative calculation. Thus it is difficult to realize the real time processing.
Furthermore, most methods for estimating the transfer function do not consider about the difference of characteristics based on the glottis open and close.
In case of voiced speech, there are two different states every pitch period. In one state, the glottis is open, and the other state, it is closed. When the glottis is closed, the speech signal is generated by the resonance of the vocal tract or nasal tract, mainly. However, when the glottis is open, the influence of the lung is added to the characteristics of the speech signal.
In case of the conventional methods for analyzing the speech signal, the parameters of transfer function are calculated in the domain of the auto-correlation or the spectrum[1]. Thus the glottis open and close characteristics are combined, it becomes difficult to separate the vocal tract characteristic and the excitation from human voice. The GARMA method is the one that is calculated in the time domain, but it does not consider the glottis open and close characteristics.
In this paper, we present a new form of transfer function for estimating the voiced speech[7-9]. Generally, the cascade structure of transfer function is used in the field of speech analysis. In this paper, we use three transfer functions that are connected as the parallel structure. These transfer functions are requested for
(1): representing a part of glottis wave that have non-minimum phase characteristic.
(2): reconstruction the vocal and nasal tract characteristics.
(3): compensation of the estimation error.
These three transfer functions are calculated in the time domain simultaneously. As the method for estimating of transfer functions, the least mean square (LMS) method is applied. When the LMS method is used, the iterative calculation is not needed on this estimation.
Furthermore, we discuss about the voice information reduction. In the proposed method, moving-average (MA) model transfer functions are used for modeling of glottis wave and the estimation error respectively because these waves have non-minimum phase and time-varying characteristics, it is difficult to reconstruct these waves using ARMA transfer function. In this paper, we discuss about the difficulty of the information reduction.
Finally, we demonstrate experimental examples based on the analysis and synthesis of human voice, and show that our method is effective.