ISBN 978-981-07-8391-4
Proceedings of 2013 3rd International Conference on Computer Engineering and Bioinformatics (ICCEB 2013)
Bangkok, Thailand , November 23-24, 2013, pp. 20-25

Echo Cancellation Based on Multi-Layer Linear Neural Network

Li Shuangxun [1+], Wang Zhan [1], Lou Shengqiang [1] and Zhan Yonghong [2]

[1] College of Electronic Science and Engineering, National Univ. of Defense Technology, Changsha, China
[2] Research Department 3, Measurement Communication Research Institute, Beijing, China

Abstract. In this paper, the research work about the feature of the speech echo in phone lines and how to cancel this echo is presented. Generally, the LMS algorithm is often used in echo cancellation, because it is very easy to be implemented. But one of the disadvantages of the LMS algorithm is that the convergence speed is too slow, which is not suitable in fast time varying occasions. In order to solve this problem, an echo cancellation implement approach based on MLLNN (Multi-Layer Linear Neural Network) is proposed. The detail of the MLLNN-based echo cancellation workflows is presented in this paper, and the simulation experiment is put into practice. The simulation experiment result shows that the filter of MLLNN-based algorithm is more feasible and efficient than that of LMS-based algorithm in echo cancellation.

Keywords: Neural Network, Adaptive Filter, Echo Cancellation, LMS, BP

1. Introduction

With the development of communication technique, more and more communication devices provide the hands free system. The hands free system is convenient but the speech echo comes with it. In hands free devices, the speaker can amplify the speech signal, so high acoustic coupling is being generated between the microphone and speaker, and the sound wave echo comes out from the acoustic coupling. The reason of echo being generated is that, at the local end, the speech of the remote caller is played out through the speaker, and the sound wave generates several copies because of the multi-reflection when the sound wave broadcast in the room, which is called speech echo. Then the speech echo is received by the microphone at the local end, so the feedback loop is formed. Through the loop, the speech echo is transmitting to the caller, so the caller hears his own speech.

The speech echo transmitting in the feedback loop affects both the communication system stabilization and the speech quality, so we need to take some measures to get ride of the speech echo.

The basic principle[1-5] of the speech echo canceller is that an adaptive filter is being used to simulate the echo transmit path. Through some adaptive algorithm, the shock response of the adaptive filter is approximate equal to the shock response of the actual echo transmit path. Therefore we can get the echo forecast signal. Then the echo will be cancelled when we subtract the echo forecast signal from the microphone speech signal. If we want to have good communication speech quality, the echo cancellation is necessary for both the local end and the remote end. The diagram of the echo cancellation can be illustrated as the Fig. 1.

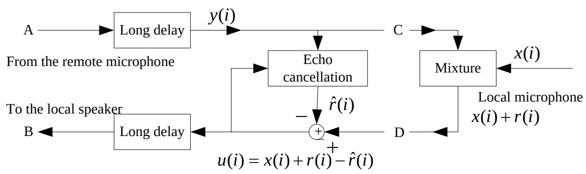

Fig. 1: The diagram of the echo cancellation algorithm.

[+] Corresponding author. Tel.: + 86 731 84573495; fax: +86 731 84575795.
E-mail address: lishuangxun@163.com.

In the Fig. 1, the $y(i)$ is the signal from the remote end A, the $x(i)$ is the signal of the local end, and the $r(i)$ is the echo signal. The $\hat{r}(i)$ is the simulated signal generated by the echo canceller, which is used to cancel the echo $r(i)$. The $u(i)$ is the transmitted signal to remote end B after echo cancellation. If the performance of the echo canceller is good enough, then $e(i)=r(i)-\hat{r}(i)$ should be approximate equal to zero, and the $u(i)=x(i)+e(i)$ should be approximate equal to $x(i)$. If the transfer function of the echo canceller is equal to the echo transmit path, then the estimated echo signal $\hat{r}(i)$ is equal to the actual echo signal $r(i)$. Because the process of the echo cancellation seems like a problem of system identification, so we try solving echo cancellation problem by utilizing the neural network method as being used in the system identification.

2. Multi-Layer Linear Neural Network

The diagram of the multi-layer linear neural network adaptive filter is illustrated as Fig.2. In the Fig.2, the $x(i)$ is the sampled discrete signal. If the number of the input data is p, the input signal vector is

$$x(n)=\left[x(n-1)\cdots x(n-p)\right]. \tag{1}$$

When the network is working, it will minimum the mean square error between the network output $y(n)$ and the desired output $d(n)$ through adjusting the weight value $h_{i,j,k}(n)$. The variable i is layer sequence number, and the variable j is the neural joint sequence number, the variable k is the path connection sequence number. The variable $S_{i,j}$ represents the output of the joint j from the layer i. The variable $h_{i,j,k}$ represents the weight value of the path connection k from the joint (i,j). The $N(i)$ is the number of joints from the layer i. The input signal of every joint can be presented as

$$S_{L,j}(n)=x(n-j). \tag{2}$$

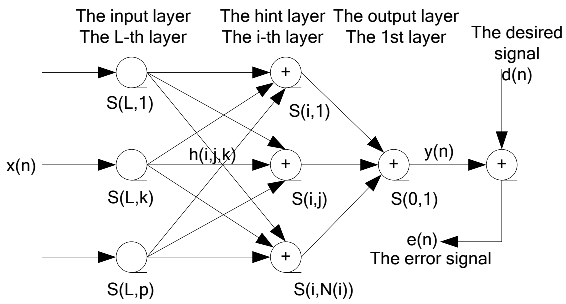

Fig. 2: The diagram of the multi-layer linear neural network.

So the output of the joint j from the layer i can be presented as

$$S_{i,j}(n)=\sum_{k=1}^{N(i+1)} h_{i,j,k}(n)S_{i+1,k}. \tag{3}$$

Therefore, the relationship between the output and the input of the network seems like a fir filter.Then the output of the FIR form can be formulated as

$$y(n)=S_{0,1}(n)=\sum_{k=1}^{p} h_k(n)x(n-k). \tag{4}$$

And the coefficient of the FIR is

$$h_k(n)=\sum_{k_1=1}^{N(1)}\sum_{k_2=1}^{N(2)}\cdots\sum_{k_{L-1}=1}^{N(L-1)}\prod_{i=0}^{L-1} h_{i,k_i,k_{i+1}}(n). \tag{5}$$

Here, we take the BP (Back Propagation) algorithm[6] as the weight training method, and minimum the mean square error $E\left(e^2(n)\right)$, where the $e(n)=d(n)-y(n)$. The iterative formula of the weight value is

$$h_{i,j,k}(n+1)=h_{i,j,k}(n)-\mu\delta_{i,j}(n)S_{i+1,k}. \tag{6}$$

For the output layer, the gradient error is

$$\delta_{0,1}(n)=2\left(y(n)-d(n)\right). \tag{7}$$

For the other layers, the gradient error is

$$\delta_{i,j}(n) = \sum_{m=1}^{N(i-1)} \delta_{i-1,m}(n) h_{i-1,m,j}(n). \tag{8}$$

If the $L = 1$, then the BP algorithm can be simplified to the LMS algorithm.

3. The Application of The BP Based Echo Cancellation

In this section, we will give an simulation experiment of echo cancellation by utilizing the BP based MLLNN. The experiment result shows that the BP based MLLNN algorithm is more feasible and efficient.

3.1. The simulation diagram of the echo cancellation

The simulation diagram of the echo cancellation is illustrated as Fig.3.

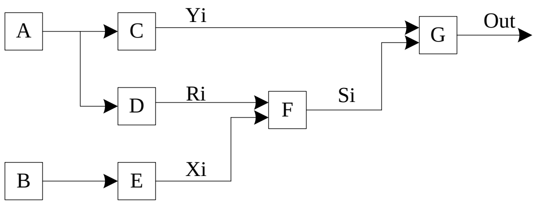

Fig. 3: The simulation diagram of the echo cancellation.

In Fig.3, the unit A and unit B are the speech input units, which are two periods of speech signals. Here, we call the two periods of speech signals as speech 1 and speech 2 separately. The unit C, unit D and unit E are speech delay units, which can attenuate and delay the speech signals from both the unit A and the unit B to simulate the echo signals. In the simulation process, the speech signal in the unit D is delayed, and the speech signals in the unit C and the unit E are not attenuated, in order to simulate the remote signal $y(i)$ and the local signal $x(i)$.

The unit F is a speech overlay unit, which is used to get overlay speech signal from the unit D and unit E. And the unit F is used to simulate the sum of the local signal and echo signal, which is $x(i)+r(i)$.

The unit G is the echo cancellation unit, which is used to cancel the echo of the overlay speech signal from the unit C and unit F. The output is the speech after echo cancelling.

The symbol meanings in the Fig.3 can be described as the followings.

Yi —— speech 1, simulating the remote reference signal $y(i)$.

Ri —— the attenuated value of the speech 1, simulating the echo signal $r(i)$.

Xi —— speech 2, simulating the local signal $x(i)$.

Si —— simulating the sum of the local signal and the echo signal, which is $x(i)+r(i)$.

Out ——simulating the output speech signal after echo cancelling, which is $u(i) = x(i) + r(i) - \hat{r}(i)$.

3.2. The algorithm unit

The speech delay unit is used to attenuate the amplitude of the speech 1, which is used to simulate the echo signal in the analog phone line. In mathematical, it can be formulated as

$$Output = Input \times 10^{-\frac{delayvalue}{20}}. \tag{9}$$

According to the protocol ITU-T G.165[7], we select 6dB as the delay value of the unit D, and select 0dB as the delay values of both the unit C and the unit E which means no attenuating.

The overlay unit is used to sum the speech 2 to the attenuated speech 1, which is used to simulate the overlay value between the local speech and the echo speech. In mathematical, it can be formulated as

$$S_i[i] = R_i[i] + X_i[i], i = 0,1,2,\cdots. \tag{10}$$

The echo cancellation unit is used to cancel the echo of the overlay speech signal from the unit C and unit F. The output is the speech after echo cancelling. The diagram of the echo cancellation unit is illustrated as Fig.4.

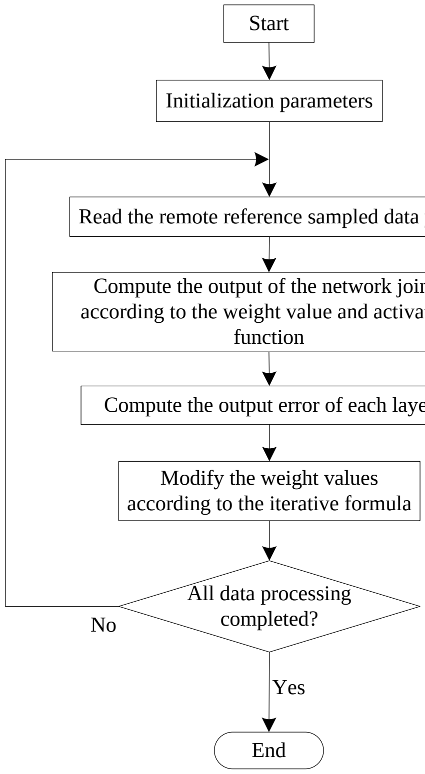

Fig. 4: The diagram of the echo cancellation unit.

3.3. The experiment result analysis

Fig.5 shows the experiment result analysis of echo cancellation for a period speech signal. The time domain waveform of speech 1 is showed in Fig.5a, which is used to simulate the remote reference signal. The time domain waveform of speech 2 is showed in Fig.5b, which is used to simulate the local speech signal. The time domain waveform of the overlay speech signal is showed in Fig.5c, which is used to simulate the sum of the local speech and echo speech signals. The time domain waveform of the output after echo cancellation is showed in Fig.5d.

From the experiment result, we can find that the speech waveform of Fig.5d and Fig.5b are very similar, and the correlation coefficients value of them is between 0.9 and 1. We can get the same experiment result when using other input sampling speech signal.

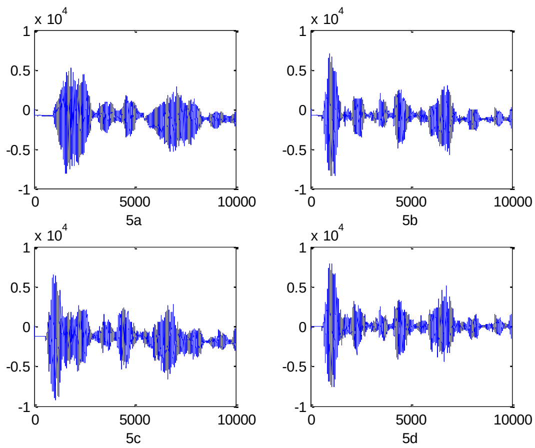

Fig. 5: The time domain waveform of input and output speech.

4. The Comparison between the LMS Algorithm

Generally, we use the LMS algorithm in echo cancellation processing. In order to verify the stability and fastness of the BP based MLLNN algorithm, we make a convergence property comparison between the LMS algorithm and the BP based MLLNN algorithm.

Firstly, we check the convergence property of the LMS algorithm. When selecting 32-order FIR filter and using different learning value μ, we can find the following results in estimating the mean square error (MSE) $E\left(e^2(n)\right)$: The max value of μ is $\mu_{\max} = 0.054$; Considering both the fast convergence property and the minimum MSE, the value of μ is $\mu = 0.042$.

Under the same conditions, we use the BP based MLLNN algorithm in echo cancellation. Here, we select 4 layers as the processing neural network. The number of input layer neurons is $N(4) = 32$. The number of the first hint layer neurons is $N(3) = 4$. The number of the second hint layer neurons is $N(2) = 4$. The number of the third hint layer neurons is $N(1) = 1$. We select the same μ value as the LMS algorithm and estimate the mean square error (MSE). Then we can get the experiment result as showed in Fig.6. By comparing the result of the two algorithms, we can draw a conclusion that the convergence speed of the BP based MLLNN algorithm is faster than that of the LMS algorithm.

From the experiment, we also find that the convergence speed is faster when the μ value is larger for the BP based MLLNN algorithm, and the convergence speed is faster when the number of the neural network layers is larger for the BP based MLLNN algorithm. But if the same number of neural network layers is not changed, only increasing the number of hint layer joint, the processing complexity is increased, while the convergence speed is being improved little.

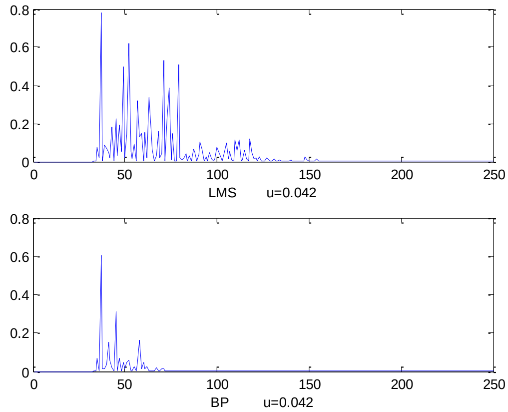

Fig. 6: The μ value is 0.042.

5. Conclusions

In this paper, the basic principle of echo cancellation is introduced. An echo cancellation approach is proposed by utilizing multi-layer linear neural network based on BP algorithm. From the relationship of the input and output, the adaptive filters based on the LMS algorithm and the BP based MLLNN algorithm are both linear systems. And the two algorithms are both based on gradient descent method. But due to the weight value of the BP based MLLNN algorithm being nonlinear, the convergence property of BP based MLLNN algorithm is better than the LMS algorithm. Considering the computation and the realization, the ASIC (Application Specific Integrated Circuit) of neural network or linear filtering device can be used in the echo cancellation processing.

6. References

[1] Duttweiler. D L. A twelve-channel digital echo canceller. *IEEE Trans. Common.*, 1978, 26 (5): 647-653.

[2] Lu. Bin, Li. Lemin, and Li. Zhengmao. Performance simulation of adaptive filters based on multi-layer linear neural networks. *Journal of UEST of China.* 1996, 25(12): 461-465.

[3] Liu. Jianfeng. Discussion and analysis of the adaptive echo cancellation technology. *Journal of Digital Communication.* March 2000, pp. 8-9.

[4] Lei. Ming, Tang. Kun, Cui. Huijuan, and Du. Wen. Improved acoustic echo cancellation. *Journal of Tsinghua University (Science and Technology).* 2001, 41(1): 37-40.

[5] Zhou. Jingli, Yang. Shutang, Han. Qi, and Yu. Shengsheng. Research on algorithm of echo cancellation within video conference system. *Mini-Micro Systems.* 1998, 19(6): 37-42.

[6] Xu. Li-na, Neural Networks Control, 1st ed. *Electronic Industry Press.* China: Beijing, 2009, Ch. 2, pp. 18-21.

[7] Zhou. Weixiang, and Sun. Debao. The simulation research on phone line echo cancellation based on ITU-TG.165 protocol. *Journal of Computer and Information Technology*. May 1999, pp. 23-27.