International Workshop on Acoustic Echo and Noise Control (IWAENC2003), Sept. 2003, Kyoto, Japan

An Echo Canceller Based on the Structure of Dual-auxiliary Filters

Ou Xiongbing Chen Zhe Yin Fuliang

(School of Electronic and Information Engineering, Dalian Univ. of Tech. Dalian 116023 China)

Abstract The effect of double-talk detection directly affects the whole performance of the echo canceller, and a new structure of dual-auxiliary filters is presented to solve this problem in the paper. The new structure can exactly distinguish the state of double-talk from the echo path change. The simulations and the real time DSP-based realization prove the robustness of the method.

Key words adaptive filter; acoustic echo canceller; double-talk detection; dual-auxiliary filters

1. Introduction

There is the echo problem in many fields, such as telecommunication channels with large loop delay, teleconference system and mobile phone system. It's difficult to get high-speech-quality voice communication without proper echo control. To solve the problem, the echo canceller has been a very active research field in the recent years.

The kernel of the echo canceller is the adaptive filter, which is used to estimate the impulse response of the echo path by means of an adaptive algorithm. During double-talk, the near-end input signal contains not only the echo of the far-end input signal, but also the near-end talker's speech. In this case, the adaptation may be greatly disturbed because the far-end input signal doesn't correlate with the near-end talker's speech. The common approach for this is to use a double-talk detector and to enable or disable the adaptation according to the output of the double-talk detector. The double-talk problem hasn't yet been well solved for acoustic echo cancellation. The conventional method[2], which compares the receive-in and send-in signal levels, can't detect the double-talk precisely since the echo pass loss is unknown and variable in the acoustic echo circumstance. In the recent years, correlation method[3,4], which assumes that the receive-in signal doesn't correlate with the near-end talker's speech absolutely, has been proposed. But the experiments show that sometimes there is some correlation between these two signals. Furthermore, it's difficult to set an appropriate decision threshold that adapts to all kinds of noise circumstance. It's difficult to directly detect the double-talk because of the time-variant, delay and non-liner property of the echo path. So the echo canceller with a dual-filter structure[5] was proposed. The main idea of this method is to form a foreground and a background echo models. Only the background model is adapted and its tap weights are transferred to the foreground model when the residual echo produced by the background model is smaller. One problem of the method is that the mistransferring of the tap weights can't be inhibited especially during double-talk.

The echo canceller with a new structure of dual-auxiliary filters is proposed in this paper. The new structure absorbs the idea of the dual-filter structure, but overcomes its shortcoming. It can distinguish the echo path change from the double-talk precisely. In the paper, the new structure is introduced at first, and then the process of the proposed approach is described in detail. At last, the efficacy of the proposed approach is conformed through the computer simulation and the real time DSP-based realization.

2. New Dual-auxiliary Filters Structure And Application Method

The echo canceller based on the structure of dual-auxiliary filters is shown in Fig.1. Where $\mathbf{x}(n)$ denotes the receive-in signal vector, $d(n)$ denotes the send-in signal, $e(n)$ denotes the residual echo, $\mathbf{g}(n)$ denotes the impulse response of the echo path, and $v(n)$ contains the near-end talker speech and background noise. Here two auxiliary filters $\mathbf{h}_0$ and $\mathbf{h}_2$ are introduced beside the main adaptive filter $\mathbf{h}_1$. According to their function, $\mathbf{h}_0$ is called as the tentative filter and $\mathbf{h}_2$ is called as the backup filter.

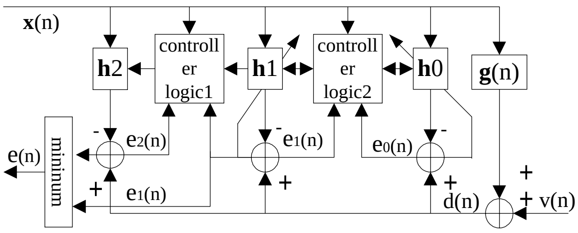

Fig. 1. Proposed echo canceller based on the structure of dual-auxiliary filters

If the effect after adaptation has been known, the appropriate control method can be adopted. That is, if the tap weights converge better after adaptation, then the adaptation is finished indeed. Otherwise, if the tap weights tend to diverge after adaptation, the current adaptation will be inhibited. Here the tentative filter $\mathbf{h}_0$ is proposed for tentative adaptive adjustment. At the same time, the residual echo signal without adaptation is also provided. To make the tendency of convergence or divergence more obvious, the data at the next period of time is also used to adapt. By comparing these two pairs of residual echo signals, whether to enable or disable the adaptation of the main adaptive filter will be decided. Moreover, to improve the robustness, the backup filter is introduced, whose function is similar to the foreground filter model in the dual-filter structure and used to backup the converged tap weights of the main filter.

Because the echo path is usually long in the acoustic echo cancellation system, the computational complexity reduction is very critical. Adaptive filtering in subbands[7] is usually adopted to achieve some computational savings. This technique will also provide faster convergence speed. The proposed method is based on the data block, and let L denote the length of the data block, i the data block index, M the number of the subband, m the subband index. In subband m, let $\mathbf{x}_m(n)$ be the receive-in signal vector, $d_m(n)$ the send-in signal, $P_m(n)$ the level of the receive-in signal, $\mathbf{h}_{0m}(n)$, $\mathbf{h}_{1m}(n)$ and $\mathbf{h}_{2m}(n)$ correspond to the coefficient vector of the tentative filter, the main filter and the backup filter respectively. All the coefficient vectors are initialized to zero. Then the process of the algorithm is given as follows.

(1) In data block i-1, respectively from the receive-in signal and the send-in signal select the maximum energy subband j, k. The tap weights of the main filter in subband j and k are copied to the tentative filter: $\mathbf{h}_{0m}(n) = \mathbf{h}_{1m}(n)$, $m = j, k$. If j is equal to k, then only one subband needs to be considered. Then using the data at data block i-1, the tentative filter is adapted and $\mathbf{h}_{1m}(n)$ is used for filtering.

$$e_{0m}(n) = d_m(n) - \mathbf{h}_{0m}^T(n)\mathbf{x}_m(n) \tag{1}$$

$$\mathbf{h}_{0m}(n+1) = \mathbf{h}_{0m}(n) + \frac{\mu}{p_m(n)} e_{0m}(n)\mathbf{x}_m(n) \tag{2}$$

$$e_{1m}(n) = d_m(n) - \mathbf{h}_{1m}^T(n)\mathbf{x}_m(n) \tag{3}$$

where $m = j, k, n \in [(i-1)L,\ iL-1]$, μ is the step-size and $0 < \mu < 2$. Then, a pair of residual echo signals $e_{0m}(n)$ and $e_{1m}(n)$ are provided and their levels are computed.

$$P_0 = \sum_{n=(i-1)L}^{iL-1} e_{0j}^2(n) + \sum_{n=(i-1)L}^{iL-1} e_{0k}^2(n) \tag{4}$$

$$P_1 = \sum_{n=(i-1)L}^{iL-1} e_{1j}^2(n) + \sum_{n=(i-1)L}^{iL-1} e_{1k}^2(n) \tag{5}$$

(2) Save the tap weights of the tentative filter to the temp vector $\mathbf{h}_m(n)$: $\mathbf{h}_m(n) = \mathbf{h}_{0m}(n)$, $m = j, k$. To data block i, the tentative filter is adapted again, and $\mathbf{h}_m(n)$ is used for filtering.

$$e_{0m}(n) = d_m(n) - \mathbf{h}_{0m}^T(n)\mathbf{x}_m(n) \tag{6}$$

$$\mathbf{h}_{0m}(n+1) = \mathbf{h}_{0m}(n) + \frac{\mu}{p_m(n)} e_{0m}(n)\mathbf{x}_m(n) \tag{7}$$

$$e_{1m}(n) = d_m(n) - \mathbf{h}_m^T(n)\mathbf{x}_m(n) \tag{8}$$

where $m = j, k,\ \ n \in [iL,\ (i+1)L-1]$, then another pair of residual echo signals $e_{0m}(n)$ and $e_{1m}(n)$ are provided and their levels are computed.

$$P_0^{'} = \sum_{n=iL}^{(i+1)L-1} e_{0j}^2(n) + \sum_{n=iL}^{(i+1)L-1} e_{0k}^2(n) \tag{9}$$

$$P_1^{'} = \sum_{n=iL}^{(i+1)L-1} e_{1j}^2(n) + \sum_{n=iL}^{(i+1)L-1} e_{1k}^2(n) \tag{10}$$

(3) Considering the inequalities:

$$P_0 \le \alpha \cdot P_1 \tag{11}$$

$$P_0^{'} \le \alpha^{'} \cdot P_1^{'} \tag{12}$$

where α and $\alpha^{'}$ are the proportion factor and are constrained to [0, 1]. Commonly, if it isn't in the double-talk state and is during the convergence process in data block i-1 and i, the inequality (11) and (12) above will be both satisfied. So if the inequalities are satisfied, it means that the adaptation tends to converge. Then the tap weights that have been saved in step (2) are copied to the corresponding subbands of the main filter: $\mathbf{h}_{1m}$(n) = $\mathbf{h}_m(n)$, $m = j, k$. In the other subbands, using the data at data block i-1, $m \in [0, M-1]$ and $m \neq j$, k, $n \in [(i-1)L, iL-1]$, the main filter is adapted.

$$e_{1m}(n) = d_m(n) - \mathbf{h}_{1m}^T(n)\mathbf{x}_m(n) \tag{13}$$

$$\mathbf{h}_{1m}(n+1) = \mathbf{h}_{1m}(n) + \frac{\mu}{p_m(n)} e_{1m}(n)\mathbf{x}_m(n) \tag{14}$$

To data block i, because the residual echo signals in subbands j and k have been provided in step (2), only in the other subbands, $m \in [0, M-1]$ and $m \neq j$, k, $n \in [iL, (i+1)L-1]$, the residual echo signals need to be computed utilizing the tap weights after adaptation.

$$e_{1m}(n) = d_m(n) - \mathbf{h}_{1m}^T(n)\mathbf{x}_m(n) \tag{15}$$

On the other hand, if the inequality (11) and (12) aren't both satisfied, it means that the adaptation may lead to diverge, so the adaptation is inhibited. Then to data block i, $m \in [0, M-1]$, $n \in [iL, (i+1)L-1]$, the old tap weights of the main filter is used for filtering.

$$e_{1m}(n) = d_m(n) - \mathbf{h}_{1m}^T(n)\mathbf{x}_m(n) \tag{16}$$

At the end of this step, the level P_s of the residual echo signal $e_{1m}(n)$ is computed.

$$p_s = \sum_{m=0}^{M-1} \sum_{n=iL}^{(i+1)L-1} e_{1m}^2(n) \tag{17}$$

(4) To data block i, $m \in [0, M-1]$, $n \in [iL, (i+1)L-1]$, the tap weights of the backup filter is used for filtering. Then the level P_b of the residual echo signal $e_{2m}(n)$ is computed.

$$e_{2m}(n) = d_m(n) - \mathbf{h}_{2m}^T(n)\mathbf{x}_m(n) \tag{18}$$

$$p_b = \sum_{m=0}^{M-1} \sum_{n=iL}^{(i+1)L-1} e_{2m}^2(n) \tag{19}$$

Considering the inequality:

$$\gamma \cdot P_s < P_b \tag{20}$$

where γ is the proportion factor. If the inequality (20) has been satisfied in the last N data blocks, the tap weights of the main filter need to be transferred to the backup filter: $\mathbf{h}_{2m}$(n) = $\mathbf{h}_{1m}$(n), $m \in [0, M-1]$. At last, the one with the smaller level between the two residual echo signals $e_{1m}(n)$ and $e_{2m}(n)$ are selected to synthesize the send-out signal.

In the algorithm above, to reduce the computational complexity while ensuring exact detection, only the subbands with the maximum level are selected in the tentative filter. If the computational amount permitted, the tentative adaptation can be done in all subbands. After step (1), whether to enable the adaptation of the main filter or not according to P_0 and P_1 is not decided immediately. Because even if the inequality (11) is satisfied, the misadjustment probably has occurred in the last several samples of data block i-1. In step (2), the data at data block i is used for adaptation and $P_0^{'}$, $P_1^{'}$ are provided. By considering the inequality (12), the adaptation effect using the data at data block i-1 will be further confirmed. To sum up, because two times of comparing are employed, the probability of misadjustment will be reduced effectively.

3. Computer Simulations

To compare the performance, the simulations based on the dual-filter method and the new structure of dual-auxiliary filters are both implemented. The speech data used to simulate is recorded from the actual room environment and the sampling frequency is 16KHz. The distance between the microphone and the loudspeaker is three meters. The related parameters are set as M = 16, L = 512, $\alpha = \alpha' = 1.0$, N= 8, $\gamma = 4.0$. And in both methods the step size $\mu = 0.5$ is used for the NLMS algorithm. The number of tap weights in each suband from 0kHz to 4kHz is 300 for 300ms acoustic echo path length and in each suband from 4kHz to 8kHz is 150 for 150ms acoustic echo path length.

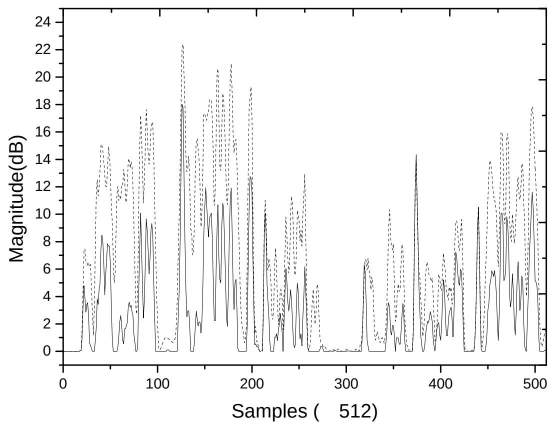

Fig. 2. ERLE comparison of the new method (dashed line) and the dual-filter method (solid line)

The ERLE performance of the proposed and the dual-filter method is shown in Fig.2. As is shown in Fig.2, double-talk occurs at about the $1.0*10^5$th (200*512) sample, and ends at about the $2.4*10^5$th (470*512) sample. It can be seen that the ERLE using the proposed method is more than 4dB up to 8dB over that using the dual-filter method when it is not during double-talk, and the former can also provide higher cancellation amount than the later during double-talk. One reasonable explanation to this is that not only more adjustment is hoped, but also the validity of every time adjustment needs to be ensured. In the proposed method, the tentative filter can ensure the validity of the main filter's adaptive adjustment.

4. Conclusion

This paper presents a new kind of echo canceller based on the structure of dual-auxiliary filters. In the new structure, the advantage of the dual-filter structure is retained by means of the backup filter, and its shortcoming is overcome by means of the tentative filter. The computer simulations have shown that the new structure can exactly distinguish the state of double talk from the echo path change. Based on TI's TMS320C549, an acoustic echo canceller using the proposed method has been test in a real acoustic field. The results of the real time experiment are consistent with the computer simulations.

REFERENCES

[1] B. Widrow and S.D. Stearns. Adaptive Signal Processing. N.J., Prentice-Hall, Englewood Cliffs, 1985.

[2] D.L. Duttweilr. A twelve-channel digital echo canceller. IEEE Trans. on Communications, 1978, Vol. 26(5): 647-653.

[3] H.Ye, B.-X. Wu. A new double-talk detection algorithm based on the orthogonality theorem. IEEE Trans. on Communications, 1991, Vol. 39(11): 1542-1545.

[4] J. Benesty, D. R. Morgan, J. H. Cho. A Family of Doubletalk Detectors Based on Cross-Correlation. International Workshop on Acoustic Echo and Noise Control. Pocona Manor, Pennsylvania, USA, 1999. 108-111.

[5] K. Ochiai, et al. Echo canceller with two echo path models. IEEE Trans. on Communications, 1977, Vol. COM-25(6): 589-595.

[6] Jianfeng Liu. A novel adaptation scheme in the NLMS algorithm for echo cancellation. IEEE Signal Processing Letters, 2001, Vol. 8(1): 20-22.

[7] A. Gilloire, M. Vetterli. Adaptive filtering in subbands with critical sampling: analysis, experiments, and application to acoustic echo cancellation. IEEE Trans. on Signal Processing, 1992, Vol. 40(8): 1862-1875.