Dual Filter Acoustic Echo Canceller with Post-Filtering

G. Chicco† and J. Fouret‡

University College London †, Tecteon PLC ‡ .

Abstract: The proposed acoustic echo canceller (AEC) has two separate echo path models. The first one (back filter) is used for identifying echo path transfer characteristics. The second one (front filter) is used for cancelling the echo. In order to avoid any divergence, the parameters of the front filter are updated from the back filter according to a logical control. The residual error is post filtered to improve echo attenuation.

1 Introduction

Acoustic echo needs to be removed in order to provide good speech quality in hands free communication. Acoustic echo consists of acoustic coupling between the loudspeaker and the microphone directly as well as reflected via surrounding surfaces.

The challenges encountered in cancelling acoustic echo are the computational complexity, the influences of near-end speech and quick variations of echo path change. In hand free communications all these factors are significantly present. In almost all the AEC, residual echoes often remains at the output of adaptive filter. To achieve sufficient echo reduction, a post-filtering is needed. However, post-filtering technique needs to be handled with care, because distortion of the near-end speech could happen.

One of the biggest challenges in AEC is how to provide echo cancellation during double-talk. Double-talk periods occur when near-end speech is present within the room and added to the loudspeaker echo. During this period, the adaptive filter attempts to cancel both the echo and the near-end speech hence the adaptive filter diverges from the optimum solution. A passive solution is to freeze the adaptation of the filter. Due to echo path change, the frozen filter won't match the characteristics of the room impulse response and the echo won't be cancelled any longer. Another solution is to use robust algorithms, which estimate the room impulse function during double talk.

Section II introduces the proposed structure for acoustic echo canceller. The simulated acoustic echo canceller algorithm is first described, and then the use of the two echo path models for handling double-talk period is explained. Section III deals with post filtering technique. Section IV highlights the performance of the proposed system. Conclusions are given in Section V.

2. Proposed Structure

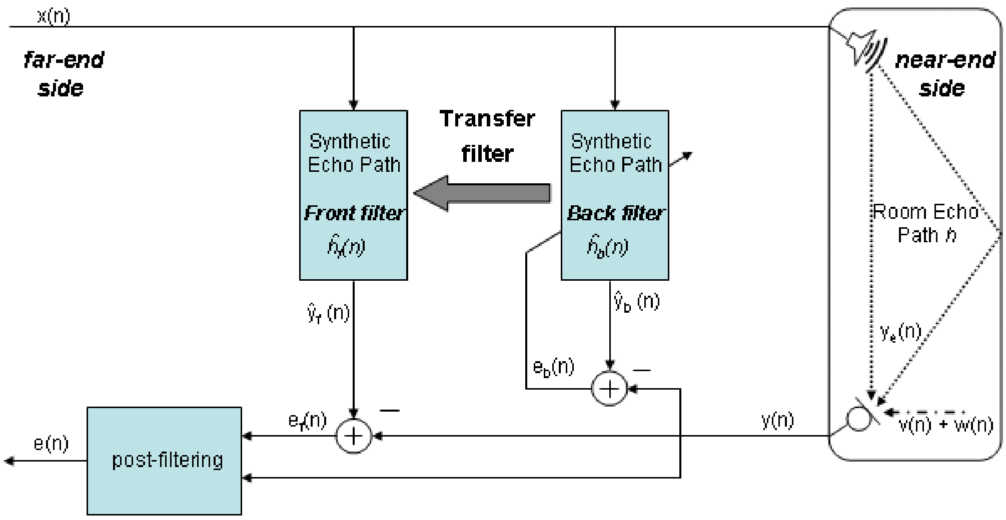

Figure 1: Structure of the acoustic echo canceller

In this section, an AEC is introduced as it is illustrated in Figure 1. The far-end signal x(n) is played in the room with echo path *h* which results the acoustic echo signal $y_e(n)$. The recorded signal y(n) by the microphone is combination of acoustic echo signal $y_e(n)$ plus the near-end noise w(n) and near-end speech v(n). The Finite Impulse Response (FIR) filter, $\hat{h}_b(n)$, estimates a replica of the echo path response. The estimated acoustic echo signal, $\hat{y}_b(n)$, is subtracted from the near-end signal of y(n) resulting in the residual error $e_b(n)$. The residual error $e_b(n)$ will be a free echo signal if the AEC has converged. Only the back filter is adapting with the help of its feedback error $e_b(n)$. The coefficients of $\hat{h}_b(n)$ are copied into $\hat{h}_f(n)$, whenever $\hat{h}_b(n)$ performs consistently better than $\hat{h}_f(n)$. The use of the two echo paths could slow down the convergence but it avoids any divergence during double talk. To reduce the remaining echo of $e_f(n)$ further, a post filtering filter is used. The aim is to shape the spectral properties of $e_f(n)$ in order to have the same spectral shape of the near-end y(n). Therefore the residual echo will be masked by the near-end signal.

2.1 Acoustic Echo Algorithm

To estimate the acoustic echo path response, the normalized least mean square (NLMS) algorithm can be used. This algorithm is described by the following set of equations [1]:

$$e_b(n) = y(n) - \hat{\mathbf{h}}_{\mathbf{b}}^{\mathbf{T}} \mathbf{x}(n) \qquad (1)$$

$$\hat{\mathbf{h}}_{\mathbf{b}}(n+1) = \hat{\mathbf{h}}_{\mathbf{b}}(n) + \frac{\mu}{[\mathbf{x}^{\mathbf{T}}(n)\mathbf{x}(n)] + \delta} \times e_b(n)\mathbf{x}(n) \qquad (2)$$

With an adaptive filter of length L, $\hat{\mathbf{h}}_{\mathbf{b}}(n) = \left[\hat{h}_b^0(n),...,\hat{h}_b^{L-1}(n)\right]^T$ is a weight vector of length L, $\mathbf{x}(n) = \left[x(n), x(n-1)..., x(n-L+1)\right]^T$ is an input vector of length L. $[.]^T$ denotes the matrix transpose, μ is a positive step size with value less than one to assure stability of the algorithm and δ is a very small positive number which prevents division by zero and stabilizes the solution.

This algorithm is a good compromise between tracking performance and computational load. However the depth of convergence is not sufficient and a post filtering is needed to remove the residual echo.

2.2 Handling double talk period

It is very difficult to create reliable double-talk detectors to freeze the adaptation and consistent double-talk robust algorithms. This study uses a two echo path model to avoid any divergence. The parameters of the back filter are transferred to the front filter, when the back filter is better than the front one. The following basic set of criterions [2] needed to be simultaneously verified:

$$P(e_b) < \beta \times P(e_f) \qquad (3)$$

$$P(e_b) < \sigma \times P(y) \qquad (4)$$

Where P(A) indicates a short time power of the signal A

$$P(y) < \varepsilon \times P(x) \qquad (5)$$

β, σ and ε are positive constant less than one. Condition (3) indicates that the back residual error signal power is lower than the front one. Condition (4) indicates that the power of the back residual error is less than the near-end power. No transfer is allowed by (5) if the near-end power is greater than the far-end one, because double-talk should be present.

3. Post filtering

The post filtering (Figure 2) is a small FIR filter transforming $e_f(n)$ into e(n). The values of this FIR are estimated by a NLMS adaptive filter, which takes as input signal z(n) a combination of the microphone signal y(n) and $e_f(n)$. The feedback signal is $e_h(n)$. The input signal z(n) is estimated as follow:

$$z(n) = \alpha(n) \times y(n) + [1 - \alpha(n)] \times e_f(n) \qquad \text{where } 0 \le \alpha(n) < \infty \qquad (6)$$

The factor α(n) is controlled in order to procure a minimum echo attenuation for single talk as well as introducing as little distortion as possible of the near-end during double-talk. A reliable and low computational estimation for α(n) is described in [3].

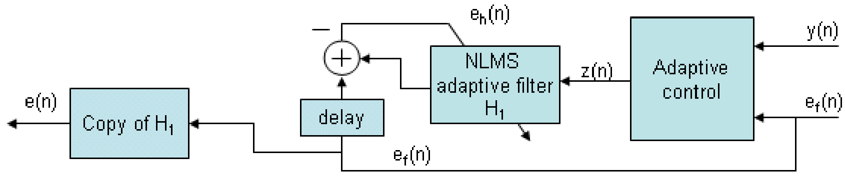

Figure 2: Block diagram of the post filtering block

4. Results

Using the image-source technique, a echo signal $y_e(n)$ was simulated for a room emulating the inside of a car. The echo path change inside the room has been simulated by moving the microphone position inside the room. The FIR filter $\hat{h}_f(n)$,and $\hat{h}_b(n)$ are 500 taps and the simulation are done at 8kHz. The post filtering filter did not have more than 20 taps. Female and male voices were present at the far-end and near-end respectively during double-talk simulation.

The Echo Return Loss Enhancement (ERLE) is used to measure the performance of the AEC during single talk. The ERLE is defined as the ratio of the power of the residual error signal and near-end power immediately after the cancellation. The following formula has been used:

$$ERLE(n) = 10\log\left(E\left[e^2(n)\right] \middle/ E\left[y^2(n)\right] \right) \qquad (7)$$

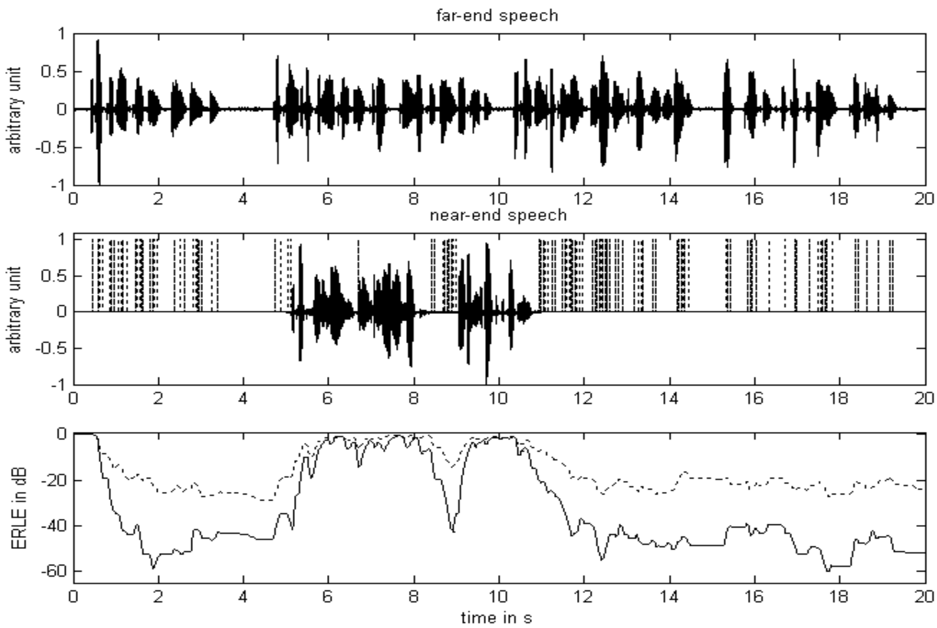

Figure 3: Performance of the AEC during double talk and single talk periods

The AEC structure is tested with single talk and double talk periods. The far-end and near-end are plotted on the first two graphs (Figure 3). The last graph shows the ERLE with post-filtering (solid line) and without (dashed line). On the second graph, the dashed lines indicate when the coefficients were transferred from the back filter to the front one.

The simulation results (Figure 3) show that the post-filtering reduces the echo by further 20dB attenuation for single talk periods. Moreover the transfer is never done during high level of near-end which avoids divergence.

To evaluate the behaviour of the post-filtering and the reliability of the transfer logic, different types and speech levels have been simulated. Table 1 summarises the behaviour of the post-filtering. The performance of the logic transfer is gathered in Table 2.

Echo to noise ratio (ENR) in dB	30			10		
Length of the back filter	300	400	500	300	400	500
Further attenuation brought by post-filtering (in dB)	10	16	20	9	11	13

Table 1: Performance of the post-filtering for different ENR and length of adaptive back filter

Note that the post-filter echo reduction improvement is proportional to the base AEC echo reduction. Moreover, this post filtering is sensitive to near-end noise affecting its overall performance.

Near-end to far end speech ratio (NFR) in dB	5	0	-5
Probability of missing good transfer	0.25	0.3	0.4
Probability of bad transfer	0.1	0.1	0.2

Table 2: Performance of the logic transfer for different NFR

The logic transfer in Table 2 illustrates low probability of bad transfer "in overall performance" which avoids divergence during double talk periods. However, the probability of missing a good transfer is rather high. So this logic transfer is too conservative and slows down the attenuation of the AEC.

5. Conclusions

This paper has presented an AEC, where the performance of the post filtering has been underlined and a reliable method to handle double talk has been demonstrated. The simulation results has also shown that even with the post filtering, the near-end audio is still highly intelligible during double talk. Further work needs to be done to carry out a reliable estimation during double talk in order to reduce the probability of missing a good transfer, without impacting or increasing the probability of bad transfer. As a consequence, AEC attenuation will be increased leading to even better performance of the post-filter.

Acknowledgments

The authors wish to thank M. Ahmadi, D. Cross, A. Karshenas and H. Sepehr for valuable comments.

References

[1] **J.Benesty T.Gansler,D.R Morgan, M.M.Sondhi, and S.L.Gay**, *Advances in Network and acoustic echo Cancellation*, in Digital signal processing, Springer, ISBN 3-540-41721-4.

[2] **Kazuo Ochiai, Takashi Araseki, and Takashi Ogihara**, *Echo Canceler with Two Echo Path Models,*IEEE Transactions on Communications, Vol.COM-25, No.6, June 1977.

[3] **Rainer Martin , and Stefan Gustafsson**, *The echo shaping approach to acoustic echo control*, Institute of Communication Systems and Data Processing, Aachen University of Technology, Elsevier Science, Speech Communication 20 (1996) 181-190.