ACOUSTIC ECHO CANCELLATION USING MULTIPLE SUB-FILTERS

R.N. Sharma , A.K. Chaturvedi and G. Sharma
Department of Electrical Engineering
IIT Kanpur
Kanpur, UP, INDIA
0512-259-7897
nath@iitk.ac.in
*

Abstract

We present the modeling of the acoustic echo path based on process segmentation approach. A new algorithm is proposed using the concept of decomposing the long adaptive filter into low order multiple sub-filters. Simulation results show that the decomposed LMS adaptive algorithm significantly improves the convergence rate while keeping the steady state error almost same as that of the original long adaptive filter.

1 Introduction

Acoustic echo originates due to the coupling of the loudspeaker and microphone in hands-free telephony and teleconferencing [1]. A simplified model for acoustic echo path is developed based on the idea that the propagation delay is caused due to the speed of the sound wave, while reflections experience attenuation of high frequency components and some energy loss. The response of the acoustic echo path is broken into frames according to the reflections received at the microphone using process segmentation approach for modeling of non-stationary processes [2]. Usually acoustic echo cancellers are realized by adaptive FIR filters, requiring thousands of coefficients to accurately model the echo return path. This leads to excessive burden of computation and slower convergence rate. One of the way to mitigate this slowly convergent and computationally intensive long adaptive filter problem is to use decomposition. This idea is based on distributing the load of adjusting a long adaptive filter to low order multiple sub-filters updated individually by a separate adaptive algorithm. It is generally found that adaptive LMS algorithm with lower order has faster convergence [3]. In most of the cases, the eigen-value spread of the auto correlation decreases as the order of the filter decreases except for white input [4].

*0-7803-7651-X/03/$17.00 ©2003 IEEE

2 Modeling of Acoustic Echo Path

It has been observed that the reflected sound signal experiences attenuation, propagation delay and energy loss. The model cannot be static unless there is no change or movement of the person and objects in the environment. Keeping in view the characteristics of the reflected sound signal, the model for multiple reflections together with the direct path from loudspeaker to microphone can be obtained. A simple case is considered where only three reflections are significant as shown in fig.1.

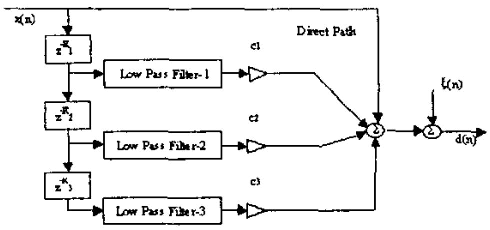

Figure 1: Model for Acoustic Echo Channel

In this model there is a pass-through connection, which represents the direct path and other three branches include the acoustic reflections. The delays involved in different reflections are obtained by transmitting a random sequence. The received signal at microphone exhibits high correlation from a copy of the original sequence. This is due to the direct transmission path and echoes of the original signal. From the correlation function the peaks will determine delays in reflections, accordingly the received signal is broken into a sequence of delay length "frames". Physically the delay is caused by the speed of sound (360 mtr/sec.), while the Low Pass Filters (LPF) model the attenuation experienced by reflected sound wave to high frequency components and acoustic environment being passive, i.e. echoes decay with time. This model can be viewed as an acoustic transfer function between loudspeaker and microphone.

3 Adaptive Algorithms Based on Decomposition Approach

Here we will discuss only the adaptive filter portion of the Acoustic Echo Canceller (AEC) rather than discussing the other components like 'Double Talk Detector' (DTD) and residual echo suppressor. The idea of decomposing the input signal vector and the weight vector into sub-vectors was first presented in [5]. Here the decomposition is to partition the long single adaptive filter into smaller multiple sub-filters. Each sub-filter is updated by an individual adaptive algorithm. Adaptive algorithms are constructed depending upon how the error signal is generated. The error signal can be obtained at each stage of the sub-filter for its updation or it can be a common error obtained at the last stage. Both the arrangements are shown in fig.2.

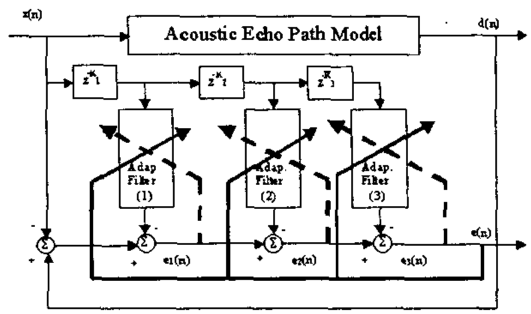

Figure 2: Configuration Based on Decomposition Approach

The adaptation factor is separately chosen for each sub-filter. The common error mode has the advantage that it introduces parallelism that will help in achieving fast convergence. In the different error mode although algorithm appears to be fast at the beginning but the steady state error is high.

3.1 Convergence Behavior

The input time series $\{x(n)\}$ is assumed stationary Gaussian zero mean. The microphone output is described as

$$d(n) = \sum_{j=0}^{M} H_j^T X_j(n) + \xi(n); \qquad (1)$$

where $H_j = [h_{j,1}, h_{j,2}, \cdots, h_{j,Lj}]^T$ is the Wiener solution of the jth sub-filter,
$X_j(n) = [x(n-(k_1+k_2+...+k_j)), x(n-(k_1+k_2+\ldots+k_j+1)), \ldots,$
$\ldots, x(n-(k_1+k_2+...+k_j+L_j-1))]^T$ input sequence

M is the number of reflections taken into account and L_j is the length of LPF corresponding to each reflection. Also $H_0 = 1.0, X_0(n) = x(n)$ and $\xi(n)$ is the ambient noise assumed independent of sequence $X_j(n)$ with a variance $\varepsilon_{min} = E\left[\xi^2(n)\right]$.

The order of each adaptive sub-filter is considered to be same as that of the individual reflection path coefficient length in exact modeling. Therefore the output of the adaptive sub-filter is

$$y(n) = \sum_{j=0}^{M} W_j^T(n) X_j(n); \qquad (2)$$

where $W_j(n) = [w_{j,1}(n), w_{j,2}(n), \ldots, w_{j,Lj}(n)]^T$ and

$$e_i(n) = e_{i-1}(n) - W_i^T(n) X_i(n); \quad i = 1, 2, \ldots, M \quad (3)$$

with $e_0 = d(n) - x(n)$ and $W_0 = 1.0$

The LMS adaptation of each sub-filter $W_i(n)$ is given as

$$W_i(n+1) = W_i(n) + \mu_i X_i(n) e_i(n); \quad i = 1, 2, \ldots, M \quad (4)$$

In common error mode each $e_i(n) = e(n)$ $\forall \; i = 1, 2, \ldots, M$ and given as

$$e(n) = d(n) - \sum_{j=0}^{M} W_j^T(n) X_j(n) \quad (5)$$

we define the weight error vector $V_i = W_i(n) - H_i$

Then the Mean Square Error (MSE) under well known assumption is given as

$$\varepsilon(n) = \varepsilon_{min} + \sum_{i=0}^{M} \sum_{\substack{j=0 \\ i \neq j}}^{M} E\left[V_i(n)\right] R_{i,j} \; E\left[V_j(n)\right] + \sum_{j=0}^{M} tr\{R_{i,i} \; E\left[V_i(n) V_i^T(n)\right]\} \quad (6)$$

where $R_{i,i} = E\left[X_i(n) X_i^T(n)\right]$ and $tr(*)$ represents the trace operation. For the MSE to converge, it is necessary that $E\left[V_i(n)\right]$ and $tr\{R_{i,i} E\left[V_i(n) V_i^T(n)\right]\}$ converges.

To check the convergence subtracting H_i on both sides in (4), substituting $e(n)$ and taking expectation.

$$E\left[V_i(n+1)\right] = (I - \mu_i R_{i,i}) E\left[V_i(n)\right] - \mu_i \sum_{\substack{j=0 \\ j \neq i}}^{M} R_{i,j} E\left[V_j(n)\right] \quad (7)$$

$$\Rightarrow 0 < \mu_i < 2/tr(R_{i,i}) \quad \forall \; i = 1, 2, \ldots, M \quad (8)$$

This gives the essential condition for convergence but still tighter bound for the step size can be obtained to confirm stability and convergence of sub-filters. Therefore we obtain the correlation matrix of weight error vector by post multiplying (7) by its transpose and taking expectation. Applying the gaussian fourth moment factoring theorem and defining the rotation vector $\acute{V}_i(n) = Q_i^T V_i(n)$ where Q_i is eigen vector matrix of $R_{i,i}$. The MSE convergence and stability is ensured [6] if the step-size μ_i is chosen such that

$$0 < \mu_i < 2/3tr\{R_{i,i}\} \quad \forall \; i = 1, 2, \ldots, M \quad (9)$$

3.2 Experimental Results

The reflection energy losses are taken as $c_1 = 0.9, c_2 = 0.7, c_3 = 0.5$ respectively. These attenuation constants depends upon the size of the room and surface from which reflections occur. Echo path impulse response measured at a distance of 1.0 ft using pc loudspeaker and unidynamic microphone is shown in fig.3. The sampling frequency of the experiment is 16,000 Hz. The delay is measured by transmitting a random sequence and obtaining its correlation with microphone output. The output of the microphone is then segmented according to the peaks observed in the correlation function. Each segmented portion of the output is then modelled with FIR filter. Here three significant reflections are detected having delays 92,135 and 172 respectively apart from direct path. The order of the FIR filter to model each segment is 12. Under this data if a single adaptive filter is to be used then 184 coefficients are needed to synthesize the acoustic echo, whereas only 36 coefficients are enough to compensate the echo effect using multiple sub-filters for sufficient order case. The step size for decomposed adaptive filters are chosen

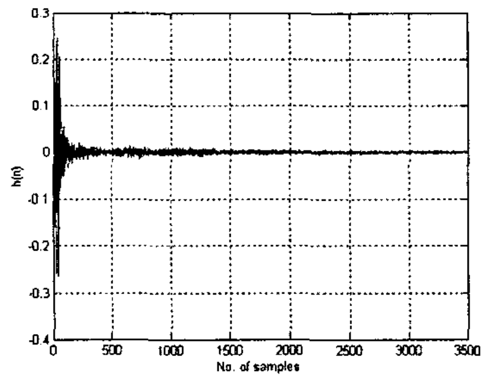

Figure 3: Impulse Response of Acoustic Echo Path

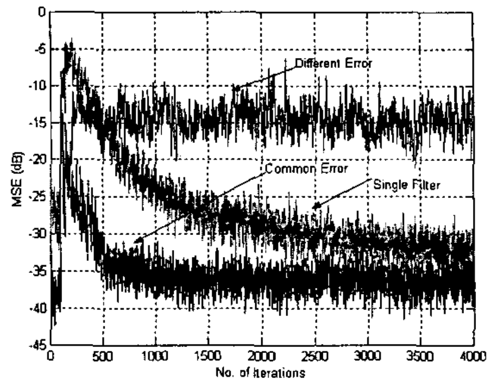

Figure 5: Plots of MSE with Colored Input

$\mu_1 = \mu_2 = \mu_3 = 0.3$ with white input and $\mu = 1.0$ in single adaptive filter. Simulation results shown in fig.4 for SNR 45dB have been obtained after averaging 100 independent runs. The same configuration has been tested for colored signal generated after passing white input through a ARMA process having Transfer Function $\{(z-0.5)/(z-0.6)\}^{16}$ with step size $\mu_1 = \mu_2 = \mu_3 = 0.1$. Results for this are shown in fig.5.

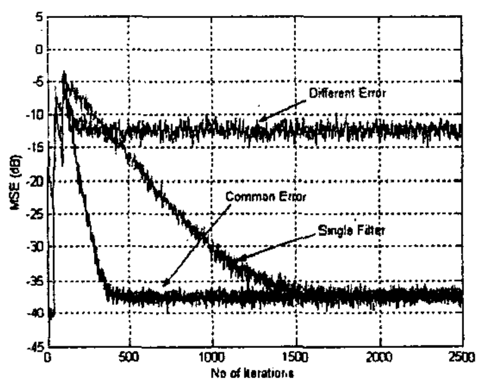

Figure 4: Plots of MSE for Different Approaches

References

[1] J. Benesty, T. Gansler, D.R. Morgan, M.M. Sondhi, and S.L. Gay, *Advances in Network and Acoustic Echo Cancellation*, Springer-Verlag, 2001.

[2] M. Niedzwiecki, *Identification of Time - Varying Processes*, John Wiley and Sons, Ltd. 2000.

[3] Z. Pritzker, and A. Fuer, "Variable length Stochastic Gradient Algorithm," *IEEE Trans.Signal Processing*, **39**, 997–1001, 1991.

[4] J. Homer, R.R. Bitmead, and I. Mareels, "Quantifying the Effects of Dimension on the Convergence Rate of the LMS Adaptive FIR Estimator," *IEEE Trans. Signal Processing*, **46**, 2611–2615, 1998.

[5] A.K. Chaturvedi, and G. Sharma, "A New Family of Concurrent Algorithms for Adaptive Volterra and Linear Filters," *IEEE Trans. Signal Processing*, **47**, 2547–2551, 1999.

[6] A. Feuer,and E. Weinstein, "Convergence Analysis of LMS Filter with Uncorrelated Gaussion Data," *IEEE Trans. Acoust.Speech,Signal Processing*, **33**, 222–230, 1985.