01 Apr 2007

A Novel Normalized Cross-Correlation Based Echo-path Change Detector

M. A. Izbal

Steven L. Grant
Missouri University of Science and Technology, sgrant@mst.edu

A NOVEL NORMALIZED CROSS-CORRELATION BASED ECHO-PATH CHANGE DETECTOR

Mohammad Asif Iqbal and Steven L. Grant
University of Missouri-Rolla, Rolla, MO 65409. Email: {ammq2,sgrant}@umr.edu

***Abstract*— A double-talk detector is used to freeze acoustic echo canceller's (AEC) filter adaptation during periods of near-end speech. Increased sensitivity towards double-talk results in declaring echo-path changes as double-talk which adversely effects the performance of an AEC as we freeze adaptation when we really need to adapt. Thus, we need an efficient and simple echo-path change detector so as to differentiate any echo-path variations from double-talk condition. In this paper, we derive a novel test statistic for echo-path change detection. The proposed decision statistic detects any echo-path variations, is normalized properly and is computationally very efficient as compared to existing techniques. Simulations demonstrate the efficiency of the proposed algorithm.**

I. Introduction

Most teleconferencing conversations are conducted in the presence of acoustic echoes [1]; if the delay between the speech and its echo is more than a few tens of milliseconds, the echo is distinctly noticeable. An acoustic echo canceller (AEC) is used to remove the echo created due to the loudspeaker-microphone environment ($\mathbf{h}$) [3]. In an AEC the echo-path (loudspeaker-microphone path $\mathbf{h}$) is adaptively modelled using a filter ($\hat{\mathbf{h}}$), which is then used to synthesize a replica of the echo ($\hat{y}$). This synthesized replica of the echo is subtracted from the echo-corrupted microphone signal (m) to get an echo-free signal (e). When the near-end talker (v) is active or when the speech comes from both the far-end ($\mathbf{x}$) and near-end (v), the adaptive filter coefficients diverge from the true echo path impulse response if the adaptation is not halted. A double-talk detector is used to stop the AEC's filter adaptation during periods of near-end speech [3]. A double-talk detector should be able to detect a double-talk condition quickly and accurately so as to freeze adaptation as soon as possible; at the same time it should be able to track any echo-path changes and should be able to distinguish the double-talk from the echo-path variations [5]. Typically, better immunity towards double-talk results in declaring echo-path changes as double-talk which adversely effects the performance of an AEC as we freeze adaptation when we really need to adapt. Thus, we need an efficient and simple echo-path change detector so as to differentiate any echo-path variations from double-talk. An optimum decision variable for echo-path change detection should behave as follows:

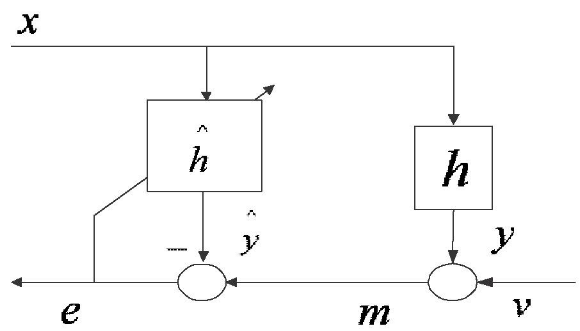

Fig. 1. Basic AEC Model

1) If no echo-path variations i.e. when the adaptive filter is converged $\xi_{EP} < T_{EP}$.
2) During echo-path variations i.e. when the adaptive filter is not converged $\xi_{EP} \geq T_{EP}$ and
3) ξ_{EP} is insensitive to near-end speech v.

Figure 1 shows the basic structure of the adaptive acoustic echo canceller. The far-end signal $\mathbf{x}$ is filtered through the room impulse response $\mathbf{h}$ to get the echo signal

$$y(n) = \mathbf{h}^{\mathbf{T}}\mathbf{x} \tag{1}$$

where

$$\mathbf{h} = [h_0 \quad h_1 \quad .... \quad , \quad h_{L-1}]^T,$$

$$\mathbf{x} = [x(n) \quad x(n-1) \quad .... \quad , \quad x(n-L+1)]^T,$$

and L is the length of the echo-path. This echo signal is added to the near-end speech signal v to get the microphone signal

$$m(n) = y(n) + v(n) \tag{2}$$

We define the error signal at time n as

$$e(n) = m(n) - \hat{\mathbf{h}}^{\mathbf{T}}\mathbf{x} \tag{3}$$

This error signal is used to adapt the L taps of the adaptive AEC filter $\hat{\mathbf{h}}$.

1-4244-1280-3/07/$25.00 ©2007 IEEE

This paper is structured as follows: In section 2 the proposed echo-path change detection statistic is derived. We do a comprehensive study on the proposed algorithm for echo-path change detection in section 3 which is followed by a summary and conclusion in section 4.

II. Echo-path Change Detection Algorithm

In this section, we derive a novel normalized cross-correlation based echo-path change detector which detects any echo-path variations. Referring to figure 1, the cross-correlation between the microphone signal m, and the cancellation error e is given by:

$$\begin{aligned} r_{em} &= E[em^T] \\ &= E[(y+v-\hat{\mathbf{h}}^{\mathbf{T}}\mathbf{x})(y+v)^T] \\ &= E[(\mathbf{h}^{\mathbf{T}}\mathbf{x}-\hat{\mathbf{h}}^{\mathbf{T}}\mathbf{x}+v)(\mathbf{h}^{\mathbf{T}}\mathbf{x}+v)^T] \\ &= (\mathbf{h}-\hat{\mathbf{h}})^T E[\mathbf{x}\mathbf{x}^{\mathbf{T}}]\mathbf{h}+E[vv^T] \\ &= (\mathbf{h}-\hat{\mathbf{h}})^T R_{\mathbf{xx}}\mathbf{h}+\sigma_v^2 \end{aligned} \tag{4}$$

where σ_v^2 is the variance of the near-end speech, the far-end speech vector $\mathbf{x}$ the near-end signal v are independent and are assumed to be of zero mean. Variance of the microphone signal is given by:

$$\begin{aligned} \sigma_m^2 &= E[mm^T] = E[(y+v)(y+v)^T] \\ &= E[yy^T]+E[vv^T] = E[\mathbf{h}^{\mathbf{T}}\mathbf{x}(\mathbf{h}^{\mathbf{T}}\mathbf{x})^T]+\sigma_v^2 \\ &= \mathbf{h}^{\mathbf{T}}R_{\mathbf{xx}}\mathbf{h}+\sigma_v^2. \end{aligned} \tag{5}$$

and finally the variance of the cancellation error e is given by:

$$\begin{aligned} \sigma_e^2 &= E[ee^T] \\ &= E[((\mathbf{h}-\hat{\mathbf{h}})^T\mathbf{x}+v)((\mathbf{h}-\hat{\mathbf{h}})^T\mathbf{x}+v)^T] \\ &= (\mathbf{h}-\hat{\mathbf{h}})^T E[\mathbf{x}\mathbf{x}^T](\mathbf{h}-\hat{\mathbf{h}})+E[vv^T] \\ &= (\mathbf{h}-\hat{\mathbf{h}})^T R_{\mathbf{xx}}(\mathbf{h}-\hat{\mathbf{h}})+\sigma_v^2 \end{aligned} \tag{6}$$

We define our new normalized decision statistic to be

$$\xi_{AsifEPD} = \left| \frac{r_{em}-\sigma_e^2}{\sigma_m^2-r_{em}} \right| \tag{7}$$

substituting equations 4 , 5 and 6 in 7 we get:

$$\xi_{AsifEPD} = \left| \frac{(\mathbf{h}-\hat{\mathbf{h}})^T R_{\mathbf{xx}}\hat{\mathbf{h}}}{\mathbf{h}^{\mathbf{T}}R_{\mathbf{xx}}\hat{\mathbf{h}}} \right| \tag{8}$$

we observe from equation 8, for $\mathbf{h} \approx \hat{\mathbf{h}}$, $\xi_{AsifEPD} \approx 0$ and for $\mathbf{h} \neq \hat{\mathbf{h}}$, $\xi_{AsifEPD} > 0$. Thus the proposed echo-path change detector meets the needs of an optimal echo-path change detector.

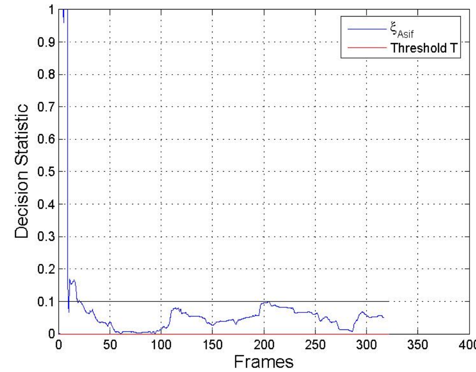

Fig. 2. ξ_{Asif} as function of time frames, Selecting detection threshold T.

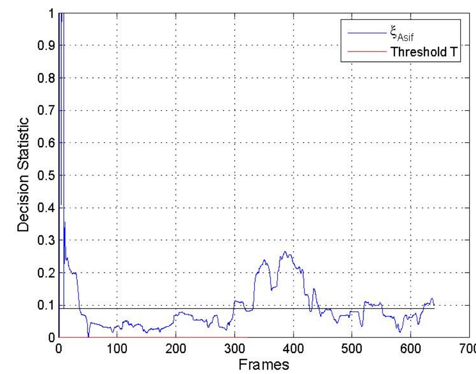

Fig. 3. ξ_{Asif} as function of time frames.

Our algorithm is computationally very efficient, we only need to do 3 multiplications, 3 additions, 2 subtractions and a division to compute the decision statistic at each sample (i.e. 9 operations per sample) as compared to $2l$ + 2 multiplications, l +1 divisions and $3l+1$ additions (i.e. $6l$+4 operations) per sample are required for the decision statistic proposed in [5]. Further, our decision statistic is normalized appropriately i.e. it is approximately zero in the absence of echo-path variations and is greater than zero during echo-path variations.

III. Experiments and Results

The values of r_{em}, σ_m^2 and σ_e^2 in 7 are exact and not available in practice. As a result, the final decision statistic is given by:

$$\xi_{Asif} = \left| \frac{\hat{r}_{em}-\hat{\sigma}_e^2}{\hat{\sigma}_m^2-\hat{r}_{em}} \right| \tag{9}$$

1-4244-1280-3/07/$25.00 ©2007 IEEE 2007 IEEE Region 5 Technical Conference, April 20-21, Fayetteville, AR

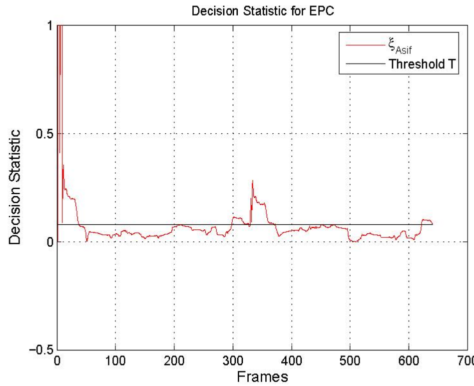

Fig. 4. ξ_{Asif} as function of time frames.

where the estimates denoted by a hat are obtained using the exponential recursive weighting algorithm, [4] [2]:

$$\begin{aligned}\hat{r}_{em}(t) &= \lambda\hat{r}_{em}(t-1)+(1-\lambda)e(t)m^T(t)\\ \hat{\sigma}_m^2(t) &= \lambda\hat{\sigma}_m^2(t-1)+(1-\lambda)m(t)m^T(t)\\ \hat{\sigma}_e^2(t) &= \lambda\hat{\sigma}_e^2(t-1)+(1-\lambda)e(t)e^T(t)\end{aligned}$$

The echo-path change detector works as follows:

1) When $\xi_{Asif} > T$ (T is a properly chosen detection threshold), we declare that the echo-canceller has not converged i.e. the echo-path has changed, the adaptation is enabled even if the double-talk detector declares a double-talk.
2) Whenever $\xi_{Asif} < T$, the detector decides that the echo-canceller has converged i.e. there are no echo-path variations.

Detection threshold T is chosen to be slightly greater than the steady state value (the value of ξ_{Asif} in the absence of any echo-path variations) as shown in figure 2. The recorded digital speech sampled at 16 KHz is used as far-end speech $\mathbf{x}$ and near-end speech v and a measured $L = 8000$ sample (500 ms) room impulse response of a $10' \times 10' \times 8'$ room is used as the loudspeaker-microphone environment $\mathbf{h}$. The room response was collected using a stereo system.

To create echo-path variations we moved from the collected left channel impulse response to the right channel response after 320 frames, as can be seen in figure 3 these changes in echo-path were detected. The echo-path change statistic goes above the detection threshold T as observed, and hence the variations in echo-path are detected. Next, we increased the echo-path gain by 2 dB after 320 frames. Simulations demonstrate the efficiency of the algorithm, even variations in filter coefficients by 2dB are detected as shown in figure 4.

IV. Conclusions

We have proposed a novel normalized sample by sample echo-path change detector. To summarize we list the major advantages of the proposed echo-path change detector:

- Detects any echo-path variations and is normalized appropriately i.e. the detection statistic is greater than zero only for echo-path variations.
- Low added complexity when implemented with any adaptive algorithm.
- Independent of the near-end speech/doubletalk.

The proposed echo-path change detector can also serve as a good download test for a two-path AEC, and when used with a good double-talk detector makes the complete system (AEC) very robust.

References

[1] Simon Haykin. *Adaptive Filter Theory*. Prentice Hall, Englewood Cliffs, NJ, 1991.

[2] Y. Hua. *Adaptive Filter Theory and applications*. PhD thesis, South-East University, Tiangsu, China, March 1989.

[3] J.Benesty, T. Gansler, D.R. Morgan, M.M. Sondhi, and S.L. Gay. *Advances in Network and Acoustic Echo Cancellation*. Springer, Inc., New York, 2001.

[4] B. Porat. Second-order equvalence of rectangular and exponential windows in least-squares estimation of autoregressive processes. *IEEE Transactions on Acoustics,Speech and Signal Processing*, 1985.

[5] Hua Ye and Bo-Xiu Wu. A new double-talk detection algorithm based on the orthogonality theorem. *IEEE Transactions on Communications*, 39:1542–1545, November 1991.