Double Talk Detection in Acoustic Echo Cancellation based on Variance Impulse Response

Sonika and Sanjeev Dhull

GJUS & T, Hisar, Haryana, India
Email: sonika.kadian@gmail.com

Abstract

In this paper we represent a new echo canceller using double talk detector that depends on variance impulse response. In this method detection is performed by calculating the variance of maximum value tap of recent tap in adaptive filter. Adaptive filter will stop updating its coefficients during double talk and it will resume updating after double talk. Simulation results by using real speech signals show the excellence of VIRE double talk detector.

Keywords: Double talk detector, echo canceller, adaptive filter.

Introduction

Nowadays, hands free conversation is popular in various fields of communication such as speaker phone system and teleconference systems. In acoustic echo canceller, however, the performance significantly degrades as double talk occurs i.e. near end and far end talkers coexist. Echoes in acoustic echo path degrade the communication quality. Therefore control of echoes is very important in communication. In this case the adaptive filter will stop updating its coefficients. Double talk deteriorates mainly the speech quality. In Fig. 1 we are having an acoustic echo canceller. The principle of acoustic echo canceller is to estimate the impulse response of an acoustic echo path by using an adaptive filter by generating a pseudo echo, and subtract it from the echo[1][2].
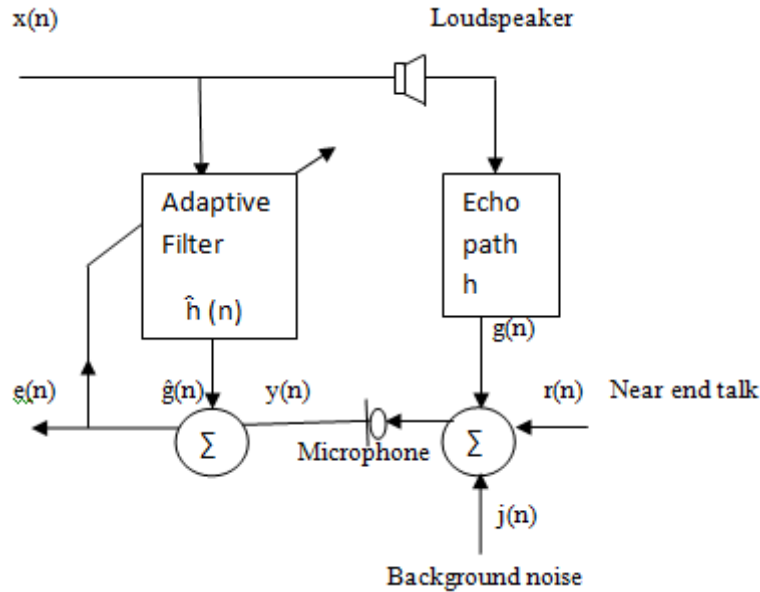
In this Fig. 1

$y(n) = g(n) + r(n) + j(n)$

$e(n) = y(n) - \hat{g}(n)$

$e(n)$ = error signal, $\hat{g}(n)$ = estimated echo signal by adaptive filter, $r(n)$ = near end speech

$j(t)$ = background noise.

Figure 1: Simple acoustic echo canceller.

Basic principle working of AEC
- A far end signal is delivered to the system.
- The far end signal is reproduced by the speaker in the room.
- A microphone also in room picks up the resulting direct path sound and consequent reverberant sound i.e..echo as a near end signal.
- The far end signal is subtracted from the pseudo far end replica produced by adaptive filter.
- The resulting signal represent sounds present in room excluding any direct or any reverberated sound produced by speaker i.e. known as error signal.

During double talk situations, path estimation of echo may be wrong as near end signal acts as interference to adaptive filter being used in AEC [3]. Therefore double talk detectors are needed in order to converge adaptive filters properly.

In this paper we proposed a double talk detector i.e. based on variance impulse response. Algorithm we are using for our coding is LMS (Least Mean Square).

The Proposed Double Talk Detector
Proposed DT detection algorithm uses the variance impulse response to solve the problem of double talk.

The Generic Doubletalk Detection Schemes:
Almost all types of doubletalk detectors operate in the same manner. Therefore, the general procedure for handling double talk is described by the following four steps.

- A VIRE variance β , is formed using available signals such as *x, g* and *e* and the estimated filter coefficients, ĥ.
- The VIRE variance, β , is compared to a preset threshold, T, (a constant), and double talk is declared if $\beta > \text{T}$.
- Once doubletalk is declared the detection is held for a minimum period of time T hold. While the detection is held the filter adaptation is disabled.
- If $\beta \leq \text{T}$ consecutively over a time T hold the filter resumes adaptation while the comparison of β to T continues until β > T again.

The hold time, T hold, in steps 3 and 4 is essential to suppress detection dropouts due to the noisy behavior of the detection statistic [4]. Although there are some possible variations most of the DTD algorithms keep this basic form and only differ in how they form the VIRE variance. A nonlinear processor, (NLP), is a signal processing circuit or algorithm that is placed in the speech path after echo cancellation in order to provide further attenuation or removal of residual echo signals that cannot be removed completely by an echo canceller. A non-linearity, a distortion, or an added noise signal is examples of signals that cannot be fully cancelled by an echo canceller. Therefore, these signals are typically removed or attenuated by a nonlinear processor. A comfort noise generator (CNG) is used to generate background noise for voice communications during periods of silence that occur during the course of conversation.
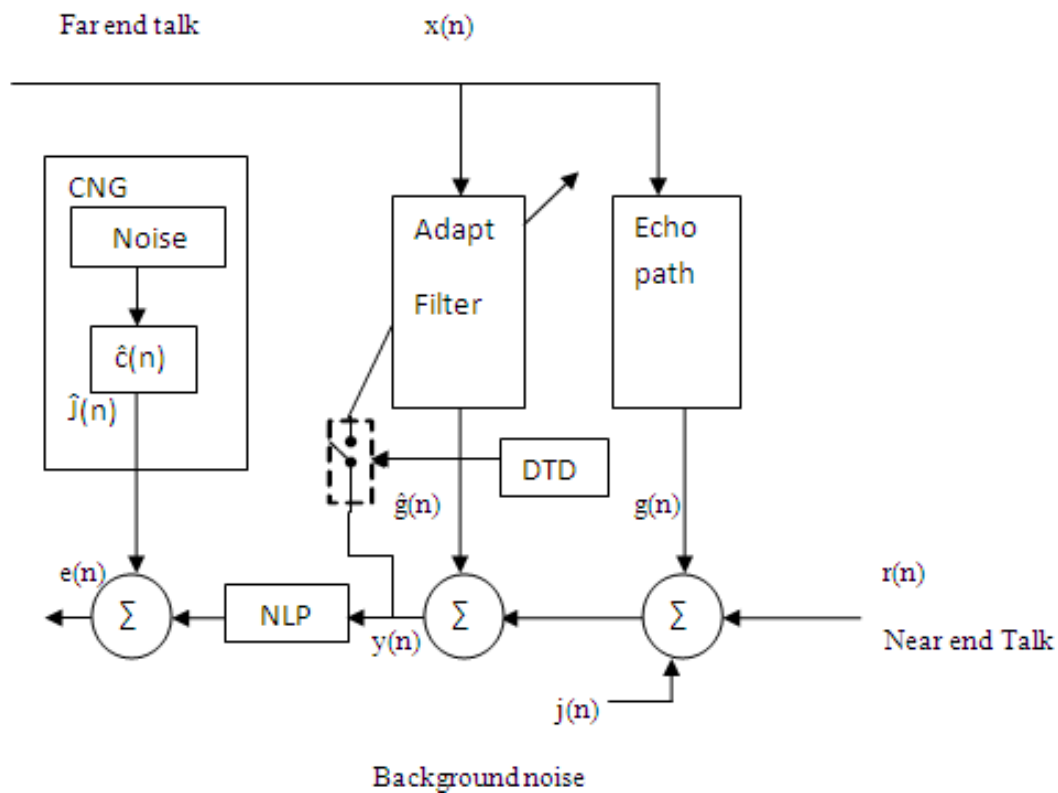

Figure 2: Echo Canceller using DTD.

The following notations are used:

ĉ(t) = adaptive coloring filter /comfort noise, NLP = non linear processor, e(t) = error signal

g(t) = estimated echoic signal, ĝ (t) = replica of estimated signal, r(t) = near end talk, j(t) = background noise

Here
e(t) = d(t) − ŷ(t)
= d(t) + s(t) + n(t) − ŷ(t)

In this system CNG (comfort noise generator) and NLP (non linear processor) are used in order to improve the further output.

There are mainly two methods for sorting out the problem of dtd:

Geigel method

This method is having very simple Geigel algorithm with low complexity. It is based on the assumption that, when we receive the signal, far end talk has lower power than that of the near end. Talk detection can be done with threshold basis [5].

VIRE method

In this paper we are dealing with variance impulse response algorithm. Variance Impulse Response algorithm calculates the variance of maximum value of the recent taps in adaptive filters. If the variance exceeds some threshold, which could be varied over time, we have double talk. In other words, if estimated room impulse response changes a lot, we assume that it is not the room has changed, but some other sources of sound has occurred[6].

We have evaluated out the performance of our proposed double talk detection algorithm through computer simulation. We have used the least mean square algorithm for the adaptive filter and acts step size parameter is set to 1. Adaptive filter length is taken to be 1000, double talk memory assumed to be 1000 and forgetting factor is equals to.97 but it can be changed and also the white Gaussian noise strength is statically assigned to be -27.

In Fig. 3 we have shown different signals, Fig. 4 shows the Vire Variance, Fig. 5 & 6 shows the comparison of ERLEs.
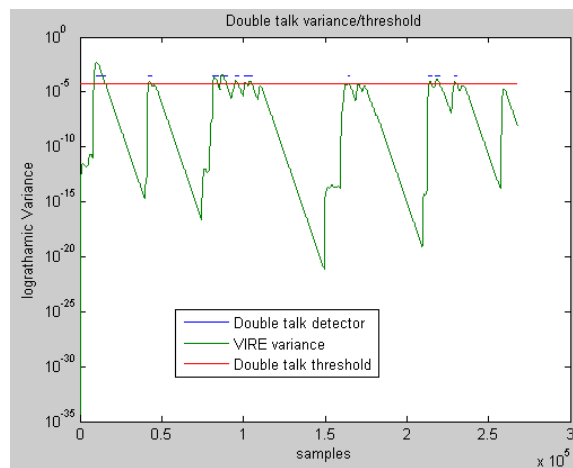
Now ERLE (Echo return loss enhancement) can be formulated as:

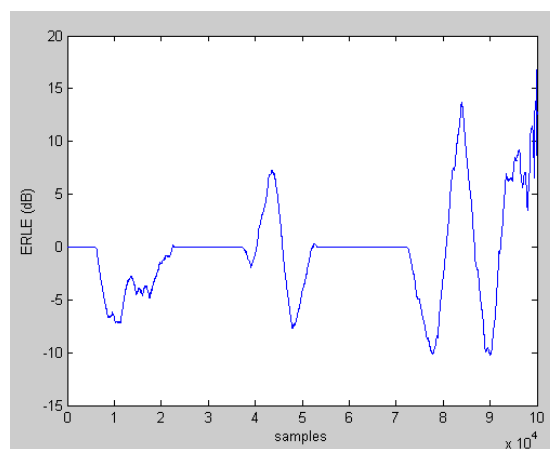$$\text{ERLE} = -10 \, log_{10} \frac{E[d^2\,(n)]}{E[e^2\,(n)]}$$

Here $d^2\,(n)$ is power of the microphone signal and $e^2\,(n)$ is the power of residual echo.
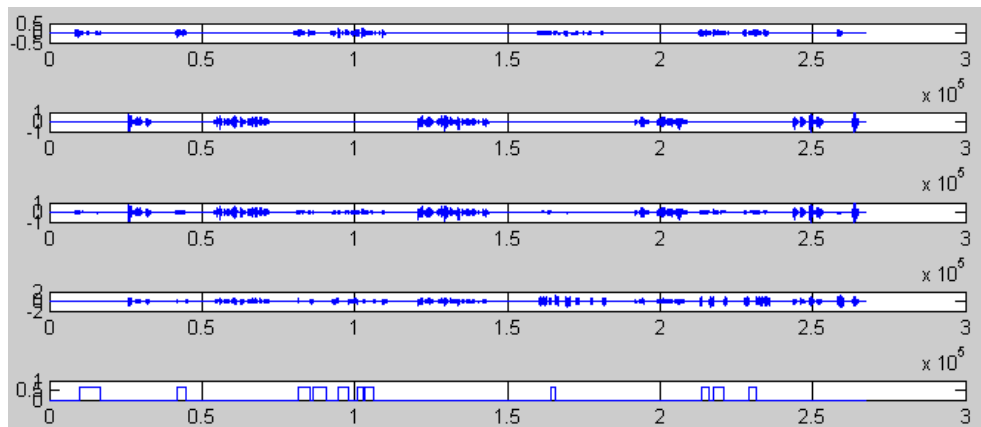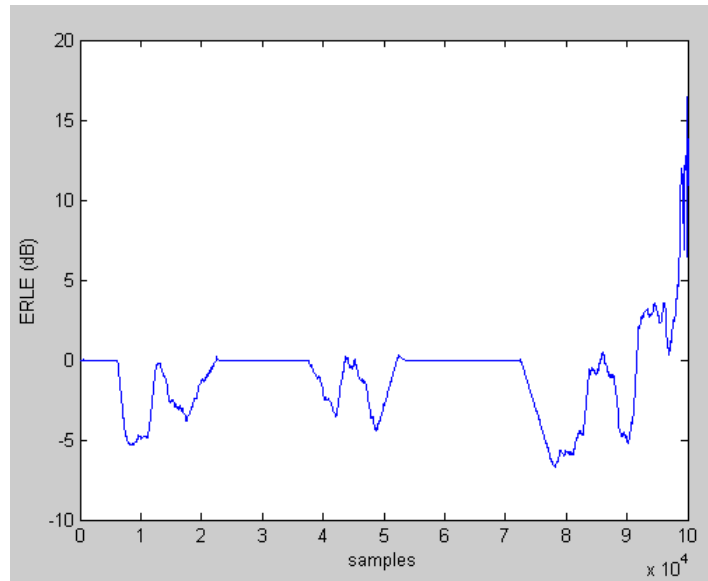
Figure 3: Waveform from top to bottom far end signal, near end signal, microphone pickup signal, filtered signal and double talk detection.

Figure 4: The VIRE Variance

Figure 5: ERLE of AEC without using DTD.

Figure 6: ERLE of AEC with DTD.

Conclusion

In this paper, we propose a double talk detection using the VIRE DTD in Acoustic Echo Canceller. The experimental results have shown the robustness of proposed DTD as well as the performance improvement of acoustic echo canceller.

References

[1]   Hirotsugu Urakami and Yoshinobu Kajikawa, 2010, "A Double-Talk-Detector Using Sound and Image Information" IEEE Proceeding.

[2]   Jae Ha Yoo*, Soo Chan Kim, 2010, "A new echo canceller using double talk detector based on the accumulated number of selection of the scale factor" ICSP Proceeding.

[3]   Artur Ferreira Paulo Marques, 2008, "An Efficient Long Distance Echo Canceller" ICSES.

[4]   Wonchul Heo, Taehwan Kim, and Keunsung Bae, 2007, "Robust Double-talk Detection in the Acoustic Echo Canceller using Normalized Error Signal Power" IEEE.

[5]   H. Haghshenas and M. MalmirChegini, *Member, IEEE,* 2007 "A Novel Double-Talk Detector Based on Adaptive Estimation of Near-End Power" IEEE Proc.

[6]   Mohammad Asif Iqbal and Steven L. Grant, 2007, "A Novel Normalized Cross-Correlation Based Echo-Path Change Detector" IEEE Proc.