Performance Analysis of Different DTD Methods for AEC

Harsh Kumar[1] Sanjeev Dhull[2]
[1]Student [2]Associate Professor
[1,2,3]Department of Electronics & Communication Engineering
[1,2]Guru Jambheshwar University of Science and Technology Hisar (Haryana), India

Abstract— In this paper, an acoustic echo canceller based on a system identification scheme with adaptive algorithms is presented. These papers compare the different DTD methods for acoustic echo cancellation. To remove the double talk from the signal using different DTD algorithms: Geigel, Benesty and NCC algorithms. Therefore compare the different DTD algorithms used to find the mean-square error and echo return loss enhancement based upon NLMS algorithm.

Key words: AEC, DTD, NCC, Benesty, Geigel, Adaptive Algorithm

I. INTRODUCTION

Echo is the phenomenon in that a postponed and distorted edition of an early sound is imitated back to the source. In echo cancellation a replica of the echo is endowed and subtracted from the consented signal [1]. The replica is generated by system arrangement identification. In applications such as desktop conferencing and hands free telephony provide telepresence to users by where multiple parties might be converting at the same time. When the near-end signal is active or when the speech comes from the far-end and near-end, the filter coefficients will diverge from the true echo path impulse response if adaptation is enabled. A double-talk detector is used to stop the AEC's filter adaptation during periods of near-end speech. Double-talk detection plays a very important part in acoustic echo cancellation [3]. A double-talk detection algorithm 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 double-talk from the echo-path variations. To solve this problem, this paper presents three different techniques for double-talk detection. An optimum decision variable ξ for double-talk detection should behave as follows [6]:

- If double-talk is not present i.e. $v = 0$, then $\xi \geq T$.
- If double-talk is present i.e. $v \neq 0$, then $\xi < T$.

The threshold T must be a constant independent of the data and the decision statistic ξ must be insensitive to echo-path variations when $v = 0$. Figure. 1 shows the basic structure of the adaptive acoustic echo canceller. The far-end signal x is filtered through the room impulse response h to get the echo signal [8].

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

where

$$h = [h_0\, h_1 \;\ldots\ldots ,\; h_{L-1}]^T , \tag{2}$$

$$x = [x(n)\;\; x(n-1)\;\; \ldots.. , x(n-L+1)]^T , \tag{3}$$

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

$$d(n) = y(n)+v(n) \tag{4}$$

The error signal at time n is defined as

$$e(n) = d(n)-h^T x \tag{5}$$

and is used to adapt the L taps of the AEC's adaptive filter h.

This paper is structured as follows. In section II, proposed basic AEC with DTD. In section III, the normalized cross-correlation double talk detection algorithm is formulated. In section IV, Benesty algorithm is formulated. In section V, Geigel algorithm is formulated. Next a comprehensive study on the proposed algorithms in section VI, which is followed by a summary and conclusion.

II. AEC WITH DTD

The configuration of an acoustic Echo canceller is given in figure 1. The echo canceller identifies the transfer function of the acoustic echo path i.e. the impulse response h(n) between the loudspeaker and the microphone [6]. The echo replica y(n) is then subtracted from the echo signal d(n) to give the error e(n)= d(n)-y(n).The adaptive FIR filter w(n) is adjusted to decrease the error power in every sampling interval. When near-end signal comes the signal then double talk presence found in the signal. To remove the double talk freezes the adaptation step in near-end signal by filter coefficients. The diagram of AEC with DTD is shown in figure 2 in which near-end signal is shown y implementing the DTD and Updating filter blocks in the figure above, the DTD will estimate the statistic decision may depend on far-end speech, near-end speech and error signal [7].

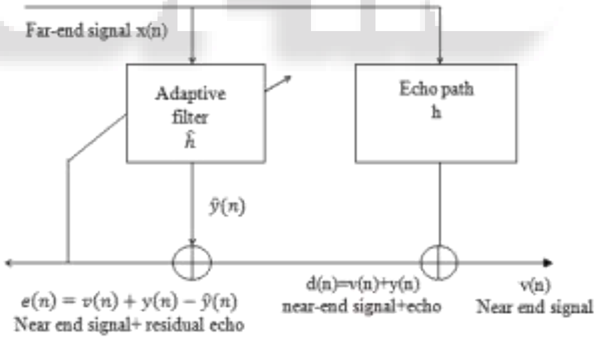

Fig. 1: Basic Model of AEC

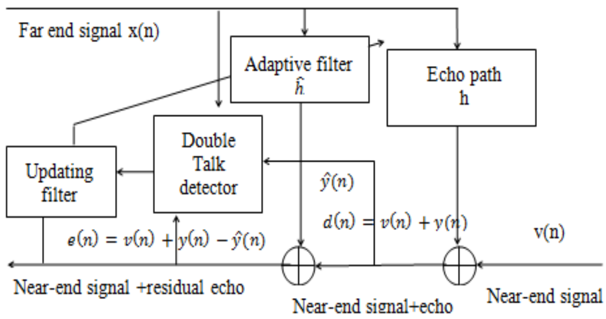

Fig. 2: Double Talk Detector with AEC

After that, it will compare to the threshold to make the DTD decision to control the Updating filter (freeze the adaptation or not). The "Updating filter" block here has the meaning as a switch (on or off) which permits to update weight vector or not. There are several methods of DTD, one can use the basic algorithm as Geigel, one bases on the cross correlation calculations (Benesty and Normalized Cross-Correlation algorithms).

III. NORMALIZED CROSS- CORRELATION ALGORITHM

The NCC algorithm computes the decision statistic depending on the relations of microphone signal and error signal [6]. It can be approached by considering the values of variance of near-end signal and cross-correlation between error signal and microphone signal. The cross-correlation r_{ed} between the error signal e(n) and microphone signal d(n) which is given as,

$$r_{ed} = E\{e(n)d(n)\} \quad (6)$$
$$=E[(y(n)+v(n)-h^T x(n))(y(n)+v(n))^T] \quad (7)$$
$$=E[(h^T x(n)-\hat{h}^T x(n)+v(n))(h^T x(n)+v(n))^T] \quad (8)$$
$$=E[(h^T x(n)-\hat{h}^T x(n))x(n)^T h+v(n)v(n))^T] \quad (9)$$
$$=(h^T - \hat{h}^T)R_x h+\sigma_v^2 \quad (10)$$

Now one introduces the normalized decision statistic as,

$$\xi_{NCC} = 1 - \frac{r_{ed}}{\sigma_d^2} \quad (11)$$

By substituting the values of r_{ed} and σ_d^2 from above relations into Equation, decision statistic is,

$$\xi_{NCC} = 1 - \frac{\hat{h}^T R_x h}{h^T R_x \sigma_v^2} \quad (12)$$

When the adaptive filter works well to converge to an estimate echo path that approximately equal to the true echo path [8]. Therefore, easily to obtain bellow conclusion,

- If near-end speech is present (v(n) = 0), then $\xi_{NCC} \approx 1$
- If near-end speech is not present (v(n) ≠ 0), then $\xi_{NCC} < 1$

Thus, finally we get the double-talk decisions as,

$$\text{Decision} = \begin{cases} \xi_{NCC}(t) < T & \text{double talk} \\ \xi_{NCC}(t) > T & \text{no}-\text{double talk} \end{cases}$$

Where, T is a threshold with the chosen value approximately is 1.

The values of r_{ed} and σ_v^2 are not available in practice, so we define the new estimated decision statistic as,

$$\xi_{NCC} = 1 - \frac{\hat{r}_{ed}}{\hat{\sigma}_d^2} \quad (13)$$

Where,

$\hat{r}_{ed}$ is the estimate of r_{ed}

$\hat{\sigma}_d^2$ is the estimate of σ_d^2

We can found these estimates by using the exponential recursive weighting algorithm.

$$\hat{r}_{ed}(n) = \lambda\hat{r}_{ed}(n-1) + (1-\lambda)e(n)d^T(n) \quad (14)$$
$$\hat{\sigma}_d^2(n) = \lambda\hat{\sigma}_d^2(n-1) + (1-\lambda)d(n)d^T(n) \quad (15)$$

Where,

- e(n) is the captured cancellation error sample at time n
- d(n) is the captured microphone signal sample at time n
- λ is the exponential weighting factor (λ<1 and λ≈ 1)

In NCC algorithm, threshold value is taken as 0.98. Mean square error is also shown at a threshold value [6]. ERLE result is also calculated. The waveform is shown at the figure 3 for DTD, figure 4 for MSE and figure 5 for ERLE. NCC detect the doubletalk with error signal and microphone signal. NCC freezes the adaptation step with near-end signal to remove the error in the signal.

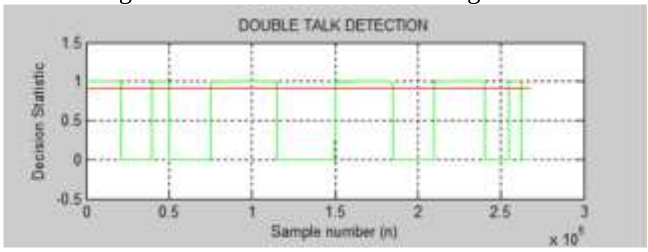

Fig. 3: Double Talk Detection with Normalized Cross-Correlation DTD Algorithm

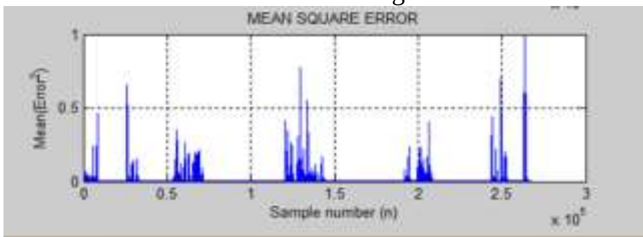

Fig. 4: Mean Square Error of AEC with NCC DTD Algorithm

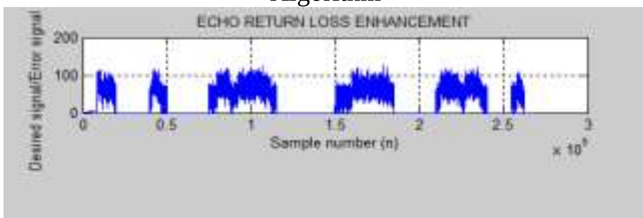

Fig. 5: ERLE of AEC with NCC DTD Algorithm

IV. THE CROSS-CORRELATION (BENESTY) ALGORITHM

The cross-correlation vector between the far-end signal x(n) and the error signal e(n) for doubletalk detection which is given as,

$$r_{ex} = E\{e(n)x(n)^T\} \quad (16)$$

Where, r_{ex} is the cross-correlation vector between far-end and error signal.

But Benesty worked on this with different approach and he claimed that the above approach does not work well for doubletalk detection [7]. He mentioned that both near-end speech v(n) and the far-end speech signal x(n) are independent and assume that all the signals are zero mean. According to him, the cross-correlation r_{xd} between far-end signal and microphone signal will be used to calculate the decision statistic.

$$r_{xd}=E\{x(n)d(n)^T\} \quad (17)$$
$$=E\{x(n)(y(n)+v(n))^T\} \quad (18)$$
$$=E\{x(n)(h^T x(n))^T\} \quad (19)$$
$$=R_x h \quad (20)$$

Where,

$R_x = E\{x(n)x(n)^T\}$ is the autocorrelation vector of far-end signal.

Benesty's decision statistic for double-talk detection is,

$$\xi_{Benesty} = r_{xd}^T(\sigma_d^2 R_x)^{-1} x_d \quad (21)$$

In this equation, the variance of the microphone signal σ_d^2 is,

$$\sigma_d^2 = E\{d(n)d(n)^T\} \quad (22)$$
$$=E\{(y(n)+v(n))(y(n)+v(n))^T\} \quad (23)$$
$$=E\{y(n)y(n)^T\}+E\{v(n)v(n)^T\} \quad (24)$$
$$=E\{ h^T x(n)(h^T x(n))^T\}+\sigma_v^2 \quad (25)$$
$$= h^T R_x h+\sigma_v^2 \quad (26)$$

Where,

σ_v^2 is variance of the near-end speech

Finally, the Equation of the decision statistic becomes,

$$= \frac{h^T R_x R_x h}{(h^T R_x h + \sigma_v^2) R_x} \qquad (27)$$

$$= \frac{h^T R_x h}{h^T R_x h + \sigma_v^2} \qquad (28)$$

Therefore, observe the above equation, easily to see that,

- If near-end speech is present (v(n)= 0), then $\xi_{Benesty} \approx 1$
- If near-end speech is not present (v(n) ≠ 0), then $\xi_{Benesty} < 1$

Thus, finally we get the double-talk decisions as,

$$\text{Decision} = \begin{cases} \xi_{Benesty}(t) < T & \text{double talk} \\ \xi_{Benesty}(t) > T & \text{no} - \text{double talk} \end{cases}$$

Where, T is a threshold with the chosen value approximately is 0.98.

In Benesty algorithm, mean square error and ERLE is also calculated. The waveform is shown at the figure 6 for DTD, figure 7 for MSE and figure 8 for ERLE. Benesty detect the doubletalk with far-end signal and microphone signal [8]. Benesty freezes the adaptation step with near-end signal to remove the error in the signal.

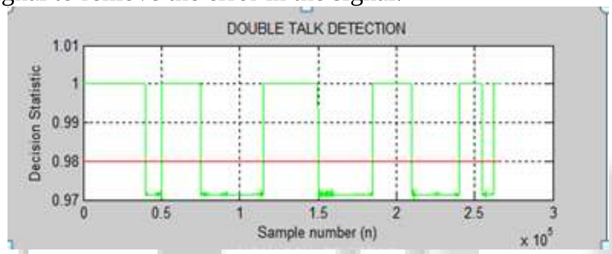

Fig. 6: Double Talk Detection with Benesty DTD Algorithm

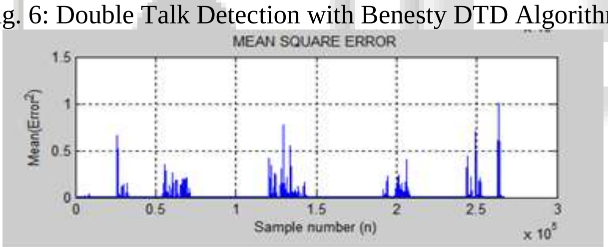

Fig. 7: Mean Square Error of AEC with Benesty DTD Algorithm

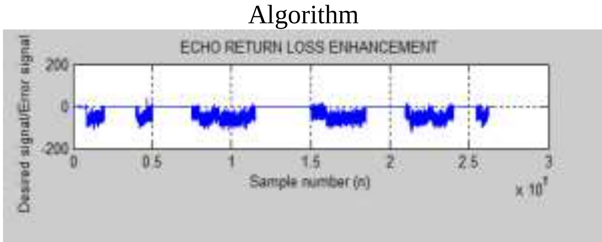

Fig. 8: ERLE of AEC with Benesty DTD Algorithm

V. THE GEIGEL ALGORITHM

The Geigel algorithm is introduced by A.A.Geigel. This algorithm, first measure the power of the received signal and compares this power to the power of the far-end signal [8]. Due to damping, the signal by acoustic filter, the power of received signal containing of echo and near-end speaker. The decision statistic for this algorithm is,

$$\xi_G(t) = \frac{\max\{|x(t)|,\ldots\ldots,|x(t-L+1)|\}}{|d(t)|} \qquad (29)$$

where, L is the length of adaptive filter

Compare the value of decision statistic with threshold T_G. If decision statistic is greater than threshold, then doubletalk is present and otherwise is not.

$$\text{Decision} = \begin{cases} \xi_G(t) < T_G & \text{double talk} \\ \xi_G(t) > T_G & \text{no} - \text{double talk} \end{cases}$$

The selection of T_G requires to be chosen carefully because it strongly affect the performance of the detector [6]. The Geigel detector has the benefit of being computationally simple and requiring very little memory. This detection approach is based on a waveform level comparison between microphone signal d(n) and the far-end signal x(n). And also assume that the near-end speech signal v(n) in the microphone signal will be stronger than the echo y(n) = $h^T x(n)$. For AEC, it is difficult to set threshold which works in any situation because the loss through the acoustic echo path depends on different factors. In general, this detector has quite poor performance.

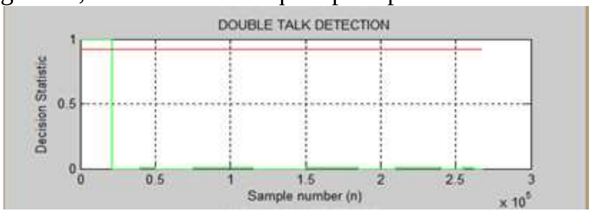

Fig. 9: Double Talk Detection with GEIGEL DTD Algorithm

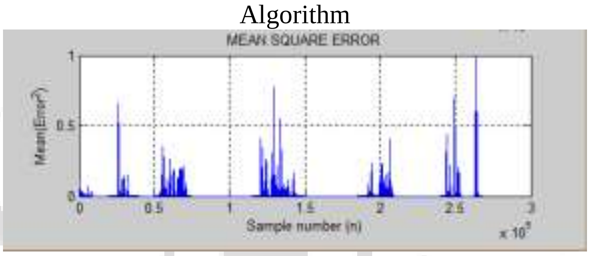

Fig. 10: Mean Square Error of AEC with GEIGEL DTD Algorithm

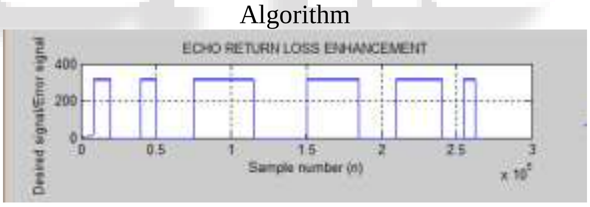

Fig. 11: ERLE of AEC with GEIGEL DTD Algorithm

VI. SUMMARY AND CONCLUSION

Acoustic echo cancellers typically employ control mechanisms to stop adaptive filter updates during echo events. Near-end echo creates a problem of double talk detection with error signal and microphone signal. To remove the double talk detection, freezes the adaptation step of near-end signal with adaptive filter coefficients. Adaptive echo cancellers are therefore used to nullify this effect using adaptive digital filter for which the amount of coefficients must be quite large in order to cancel the and near-end echo. Double Talk is detected with different methods: Geigel, Benesty and NCC algorithms. According to the techniques using in this paper NCC algorithm gives the best result for MSE and ERLE.

REFERENCES

[1] Fuyun Ling, "Achievable performance and limiting factors of echo cancellation in wirelesscommunications",DOI:10.1109/ITA.2014.

6804210, pp:1-8, IEEE , Information Theory and Applications Workshop (ITA), 2014,2014

[2] Sudo, T.; Tanaka, H.; Kohno, R.,"Spectrum selective nonlinear echo suppression based on the amount of linear echo cancellation", IEEE, Intelligent Signal Processing and Communications Systems (ISPACS), 2013Print ISBN:978-1-4673-6360-0,pp.315-319, International Symposium on,2013

[3] Mahbub, U.; Fattah, S.A.,"Gradient based adaptive filter algorithm for single channel acoustic echo cancellation in noise", ISBN:978-1-4673-1434-3,pp:880-883, IEEE, Electrical & Computer Engineering (ICECE), 2012 7th International Conference on,2012

[4] Sonika, Sanjeev Dhull., "Double Talk Detection in Acoustic Echo Cancellation based on Variance Impulse Response", ISSN 0974-2166 Volume 4, pp. 537-542, 2011

[5] Aimaduddin M.Y, M.; Zain Ismail, M.; Ali, A.H.; Mohammad, S., "Elimination of echoes from voice communication",ISBN:978-1-4244-7454-7,pp:1023-1026,IEEE,Circuits and Systems (APCCAS), 2010 IEEE Asia Pacific Conference on,2010

[6] Hung Ngoc Nguyen, Majid Dowlatnia, Azhar Sarfraz, "Implementation of the LMS and NLMS Algorithms for Acoustic Echo Cancellation in Teleconference system using MATLAB", ISSN1650-2647, 2009

[7] Enzner, Gerald, "Signal Models, Filter Structures, and Adaptive Algorithms for Acoustic Echo Control",ISBN:978-3-8007-3120-6,pp:1-4,VDE,Voice Communication (SprachKommunikation), 2008 ITG Conference on,2008

[8] Meng Wang; Li-Rong Dai; Yan Song; Ren-Hua Wang, "A Double-talk Robust and Fast Converging Echo canceller", IEEE, Information, Communications and Signal Processing, 2005 Fifth International Conference on,2005