Implementation of Adaptive Filters for Enhancement of Speech Quality in Video Conference Systems

Santosh Kumar Akiri
ECE,GITAM UNIVERSITY
Visakhapatnam, INDIA

V Malleswara Rao
ECE,GITAM UNIVERSITY
Visakhapatnam, INDIA

***Abstract*— In Today's world communication plays key role in every aspect, people stays in their own place and want to communicate with the other persons, In this process the quality of speech which the people were delivering is gradually decreasing when reaching to other end due to many issues mainly echo, so we have to mainly concentrate on this echo to get a good quality of speech. In this we are considering a multi-party videoconferencing scenario, where there are different clients communicating with each other, In this videoconferencing we come across multi-echoes, there were many algorithms for eliminating these echoes, but rather eliminating echoes here we combine all the echo signals and make it as single signal with more strength, we can do this by estimating the delay between multi echo signals, so for this two algorithms were introduced namely MDF algorithm and Affine Projection algorithm, here these two algorithms were compared together for the same speech signal by speech parameters SNR ratio and ELRE. finally the results shows that the affine projection algorithm acquires a reasonable SNR ratio which is quite suitable for good communication.**

***Key-words*— Multi-echo, MDF Algorithm, AEC computation**

I. INTRODUCTION

In every sector communication is playing a prominent role, as now a days it's totally under one roof business so different country people participate in business dealing, for this they all need a perfect communication, for every small aspects business people can't gather at one place for discussing, anyway there were many issues for this the only answer is video conferencing systems, by these video conferencing system there is no need for people to move across countries for conferences, they can stay at their own offices and have conference with their clients. Actually in video conferencing we get both video and audio, these two must be maintained with good quality for effective communication, here we mainly concentrate on quality of audio, there were many issues which degrade the quality of audio like noise, reverberation, echo etc,. In this paper we concentrated on echo, which can be eliminated by using adaptive algorithms. here Multi delay block frequency adaptive algorithm (MDF algorithm) and Affine projection algorithm(AP algorithm) were used.

II. OCCURRENCE OF ECHO IN VIDEO CONFERENCE

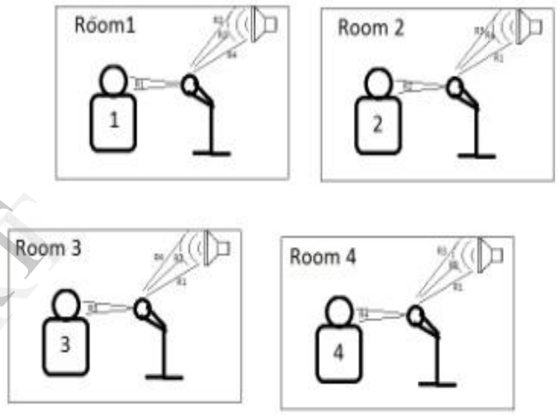

Figure 1: Occurrence of echo in video conferencing

In Figure1 we can see a diagrammatic form of video conferencing, here four clients present in different areas, talking with each other through video conference, for example consider here Room 1, person 1 who is delivering his speech enters into his microphone and it also receives the signals coming from the loudspeaker, these signals were carried to the other rooms, creating the echo over there, if it goes on increasing it creates like howling sound, which in-turn results poor communication. By analysis carried out till now the core problem is echo. Removal of this echo has many algorithms which complicate the structure of the system, so study of these algorithms provocateur combination of them to make an effective one. So here rather than cancelling echo, combining all the echo signals into a single signal can also increase the signal strength.

III. OVERVIEW OF ADAPTIVE ALGORITHM

Adaptive algorithms were frequently used in echo cancellation, as in speech the parameters are not stationary and varies with respect to the time, so we need a algorithm where the patch signal is mapped to the desired signal, this can be achieved only through adaptive algorithms.

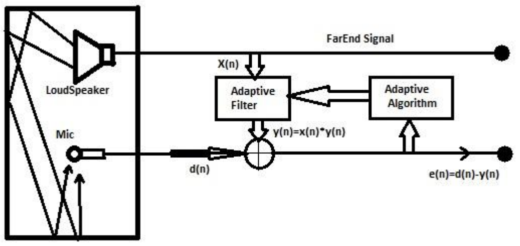

Figure 2: Block diagram of Adaptive filter

In the Figure2 it shows the basic block diagram of adaptive filter, where x(n) is a pure speech signal, d(n) is a desired signal which contains echo signal, and e(n) is a error signal. here the error signal should be minimum for efficient echo cancellation.

IV. MULTI DELAY BLOCK FREQUENCY DOMAIN ADAPTIVE (MDF) ALGORITHM

The Multidelay block frequency domain adaptive filter (MDF) algorithm[1] is a block-based frequency domain implementation of the (normalised) Least mean squares filter (LMS) algorithm

The advantages of MDF over the (N)LMS algorithm are:

1.Lower algorithmic complexity

2.Partial de-correlation of the input (which 'may' lead to faster convergence)

we already have NLMS algorithm equation i.e.

$$W(n+1) = W(n) + 2\mu \frac{e(n)x(n-k)}{N\sigma^2} \quad (1)$$

Where

n is the sample number

W k is the k th coefficient of the filter

µ is the learning rate

e(n) is the residual echo error signal

x(n) is the far end signal

N is the length of the filter

σ^2 is the reference signal power

A modified MDF algorithm[1][4] is a method of varying the optimal learning rate for the general case of complex NLMS. The derived learning rate is

$$\mu_{opt}(k,l) = \min\left(\frac{\sum_k R_{EY}(k,l)}{\sum_K R_{YY}(K,l)} \times \frac{|Y(k,l)|^2}{|E(k,l)|^2}\right) \quad (2)$$

V. AFFINE PROJECTION ALGORITHM

The concept of affine projection algorithm comes between the two basic algorithms i.e. NLMS and RLS algorithms, which takes features from these two algorithms in the areas of performance and complexity, this algorithm depends upon projection order , if the projection order is high then the rate of convergence is fast or else the rate of convergence will be slow, and also the affine projection algorithm does not make any delay between input and output signals. these all features together make affine projection algorithm, a well suited for echo cancellation

The Affine projection algorithm[2][3] updates it coefficients for every iteration so that the error signal is mapped to the desired signal and gives out a echo free signal to a maximum extent, here the weights updating is given by[5][6]

$$w(k+1) = w(k) + \mu X(k)T(k) \quad (3)$$

$$T(k) = (X^T(k)X(k) + \gamma I)^{-1} e(n) \quad (4)$$

where, γ *is a* small constant *and* µ value lies between 0 and 2, X(k) is a toeplitz matrix.

$$d(k) = [d(k), d(k-1), \ldots d(k-L)]^T \quad (5)$$

$$e(k) = [e(k), e(k-1), \ldots . e(k-L)]^T \quad (6)$$

where L is the projection order.

$$e(k) = d(k) - x^T(k)w(k) \quad (7)$$

here, *e(k)* is the error signal which is obtained from the affine projection algorithm.

VI. SPEECH QUALITY MEASURE

The speech quality measurements are required to calculate how far the adaptive filters are worked on echo cancellation, here we are considering the signal to noise ratio (SNR) and Echo return loss enhancement(ELRE) as two parameters.

A. *Signal To Noise Ratio:*

The signal to noise ratio level is one of the important parameter to know the performance of the particular algorithm, here it gives the amount of desired signal power to noise power. it can be calculated as follows

$$SNR = 10 \log_{10} \left| \frac{d(k)^2}{e(k)^2} \right| db \quad (8)$$

here the d(k) is reference signal, and e(k) is the recovered echo signal

B. *Echo Return Loss Enhancement*:

Echo return loss enhancement is another parameter to calculate the performance of the adaptive filter, for this we have to calculate the instantaneous power of the echo signal and instantaneous power of the residual error signal, the ratio of these two powers is considered as echo return loss enhancement.

In general form let us consider, y(n) is a echo signal and e(n) is a error signal, then the echo return loss enhancement is given by

$$ELRE = 10 \log \frac{E[y^2(n)]}{E[e^2(n)]} \text{ db} \quad (9)$$

VII. SIMULATIONS & RESULTS

In this paper, the input speech signal taken is "1 2 3 4 5 6 7 8 9 10" of sampling frequency is 8000hz, with time duration of 10sec, for this speech signal the echo is created using the matlab code, and the simulated waveforms are shown in Figure 3.and given as input to the MDF and AP adaptive algorithms and the respective output speech waveforms are shown in Figure 4 and Figure 5

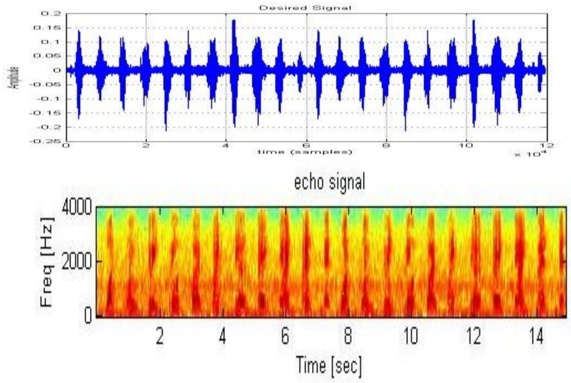

Figure 3: Echo signal and its corresponding spectrogram

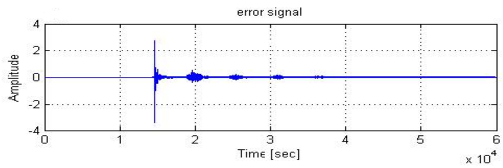

Figure 4:Recovered echo signal using mdf algorithm

In Figure 4 we can see the recovered echo signal using the MDF algorithm where the outcome speech signal is highly distorted, in general sense we can say signal is not much clear, which is not suitable for communication,

For this we have alternate algorithm i.e affine projection algorithm where we can overcome this problem and make much better for voice communication, in the below figure we can observe the occurrence of numbers where the echo is greatly reduced.

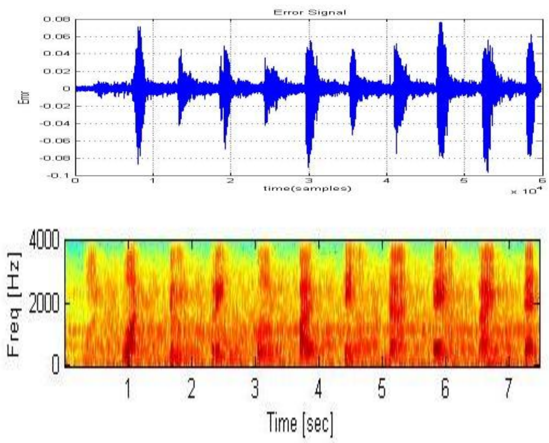

Figure 5:Recovered echo signal and its spectrogram

TABLE I. COMPARING SNR VALUES FOR MDF ALGORITHM & AP ALGORITHM

Filter order	SNR before Echo cancellation	SNR after Echo cancellation using MDF algorithm	SNR after Echo cancellation using AP algorithm
2500	-1.1784	-3.535e3	6.262
2700	-1.1784	-3.673e3	6.253
2900	-1.1784	-3.794e3	6.245
3100	-1.1784	-3.974e3	6.236

In TABLE I, the SNR values of both MDF Algorithm and Affine projection algorithm are shown of different filter order 2500 2700 2900 3100 respectively, the signal to noise ratio before echo cancellation is -1.1784 db and after echo cancellation it is increased highly i.e 6.262 db by affine projection algorithm compared to the MDF algorithm.

TABLE II. COMPARING ERLE VALUES FOR MDF ALGORITHM AND AP ALGORITHM

Filter order	ERLE for MDF algorithm	ELRE for Affine projection
2500	12.23 dB	44.20 dB
2700	13.74 dB	46.50 dB
2900	15.01 dB	48.50 dB
3100	17.65 dB	49.85 dB

In TABLE II. Echo return loss enhancement values are shown for both MDF algorithm and Affine projection algorithm of different filter order length, the ERLE value of Affine projection algorithm shows improvement than the MDF algorithm.

From TABLE I & II, we can say that the Affine projection algorithm shows decent results than the MDF algorithm, so it is suitable for echo cancellation for maintaining the efficient voice communication

VIII.CONCLUSION & FUTURE WORK

In Video conferencing, the echo is major problem, so for this elimination of echo, MDF Algorithm and Affine projection algorithm were implemented and from the results, it shows Affine projection algorithm works efficiently.

As video conferencing is subjected to closed rooms, a more efficient algorithm can be designed which is capable of eliminating echo in highly distorted speech signal.

REFRENCES

1. Xu Deng The Design and Implementation of Acoustic Echo Cancellation Subsystem for Multi- party videoconferencing systems, ISBN: 978-1-4673-5613-8©2013 IEEE.
2. Marcello L.R.de campos, Jose Antonio Apolinario, "The constrained affine projection algorithm-Development of convergence issues"
3. Paulo S.R.Diniz, "Adaptive filtering Algorithms and practical implementation" kluwer academic publishers, second edition ISBN: 978-0-387-31274-3
4. Jean-Marc Valin, "On adjusting the learning rate in frequency domain echo cancellation with double-talk", IEEE Transactions on Audio, Speech, and Language Processing, volume 15, Issue 3, pages:1030-1034, March 2007.
5. K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electronics and Communications in Japan*, vol. 67-A, no. 5, pp. 19–27, 1984.
6. S. L. Gay and S. Tavathia, "The fast affine projection algorithm," in *Proc. International Conference of Acoustics, Speech, and Signal Processing*, Detroit, USA, pp. 3023–3026, 1995.

A Santosh Kumar is presently pursuing M.Tech in the specialization of Digital System and Signal Processing. He received his B.E degree from Andhra University. .His areas of interest are Digital Systems and Signal Processing.

Dr. V. MALLESWARA RAO is presently working as Professor & Head of Dept. in the dept. of ECE, GITAM University. He submitted his PhD in speech Microwave Engineering. He received his M.E from Andhra University. He has 28 years of teaching experience. His areas of interest are Microwave Engineering.