Volume 5, No. 2, March 2014 (Special Issue)

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Estimated Loud Speaker Impulse Response to Cancel Aural Reverberation

Ranga rao Orugu[1]
Assistant Professor
Electronics and Communication Engineering
Vishnu institute of Technology Bhimavaram, India
ranga.mouni@gmail.com

K.Neelima[2]
Electronics and Communication Engineering
Vishnu institute of Technology
Bhimavaram, India
kneelima1493@gmail.com

A.V.R.Rajeswari[2]
Electronics and Communication Engineering
Vishnu institute of Technology
Bhimavaram, India
rajeswari.av7@gmail.com

G.Bangar Raju[2]
Electronics and Communication Engineering
Vishnu institute of Technology
Bhimavaram, India
gadiraju.bangarraju@gmail.com

D.Harshitha[2]
Electronics and Communication Engineering
Vishnu institute of Technology Bhimavaram, India
harshi.minu@gmail.com

Abstract: Man machine interaction requires an acoustic interface for providing full duplex hands-free communication. Adaptive filtering techniques are used in a wide range of applications, including echo cancellation, adaptive equalization, adaptive noise cancellation, and adaptive beam forming. In this paper, we present a new approach to acoustic echo cancellation for a teleconferencing system including a loudspeaker for which an estimate of the loudspeaker impulse response is available. We show that the new approach reduces the computational complexity for echo cancellation algorithms.

Keywords: Acoustic echo cancellation, adaptive filtering, loudspeaker impulse response.

I. INTRODUCTION

In many speech communication applications, e.g., audio-conference and hands-free IP telephony, the received multi-microphone speech signals are corrupted by acoustic background noise as well as by echo signals. The noise and echo components significantly degrade the intelligibility of the desired signal, and restrict the performance of subsequent speech processing systems, e.g., speech coding and speech recognition systems. Therefore, efficient methods for joint noise reduction and echo cancellation are generally desirable. Acoustic Echo Cancellers are needed for removing the acoustic echoes resulting from the acoustic coupling between the loudspeaker(s) and the microphone(s) in communication systems. In Fig. 1, a typical setup for AEC is shown.

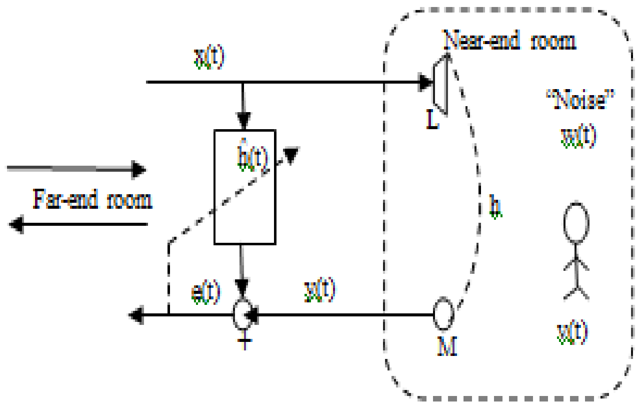

Figure.1. Typical AEC setup

The main purpose of the setup is that the near-end speech signal *v(t)* is to be picked up by the microphone is to be picked up by the microphone M and propagated to the far-end room while far-end speech is to be emitted by the loudspeaker L into the near-end room. During doubletalk, which is the case when both near-end and far-end speech is present, the near-end speech in the microphone signal *y(t)*is corrupted by the echo of the far-end speech signal *x(t)* that is propagated in the near-end room from the loudspeaker L to the microphone M. Therefore, during doubletalk, the resulting microphone signal *y(t)* consists of near-end speech mixed with far-end speech filtered by the near-end room impulse response **h** from the loudspeaker to the microphone.

$$y(t) = \mathbf{h}^T \mathbf{x}(t) + v(t) + w(t). \qquad (1)$$

In (1), *w(t)* is noise and the input data vector, *x(t)* is defined as

$$\mathbf{x}(t) = [x(t) \quad x(t-1) \quad \cdots \quad x(t-n+1)]^T \qquad (2)$$

Where *n* is the order of the room impulse response modeled as a finite impulse response (FIR) filter (in this paper we will only consider FIR filters which is the most common filter type for AEC)

$$\mathbf{h} = [h_0 \quad h_1 \quad \cdots \quad h_{n-1}]^T. \qquad (3)$$

The room impulse response is varying with time since movements (e.g., people moving around) may occur in the room. Thus, usually in order to remove the undesired echo an adaptive filter estimate$\hat{h}$(t) of **h** is used to predict the far-end speech contribution $\mathbf{h}^T$*x(t)* and subtract it from the microphone signal ***y(t).*** Thereby, we get the error signal

$$e(t) = y(t) - \hat{\mathbf{h}}^T(t)\mathbf{x}(t) = v(t) + \mathbf{h}^T\mathbf{x}(t) - \hat{\mathbf{h}}^T(t)\mathbf{x}(t) + w(t) \quad (4)$$

That ideally should be equal to the near-end speech signal ***v(t).*** Note that in (4), for simplicity, we have assumed that $\hat{h}(t)$ and ***h*** are of the same length. If that is not the case, then (4) has to be modified accordingly.

When no near-end speech is present the error signal ***e(t)*** can be used to adapt the adaptive filter $\hat{h}(t)$ using some algorithm for filter adaptation. Several different algorithms for filter adaptation in AEC have been proposed [2]. The most common one is perhaps the normalized least-mean squares (NLMS) algorithm [3] which has been shown to perform well for the AEC problem while at the same time having a rather low computational complexity.

The AEC algorithms are to be run in real-time on a digital signal processor with limited memory and computational power. As the numerical complexities of these algorithms usually are proportional to a power of ***n*** (the length of the impulse response ***h***), and ***n*** usually is very large, ranging from several hundred to several thousand, it is important to minimize the computational complexity. The main purpose of this paper is to show how the knowledge of the impulse response for the loudspeaker *L* can be used to reduce the computational complexity of existing AEC algorithms while at the same time increasing the performance.

II. NORMALIZED LEAST MEAN SQUARE (NLMS) ALGORITHM

The normalized least mean square algorithm (NLMS) is an extension of the LMS algorithm which bypasses this issue by calculating maximum step size value. Step size value is calculated by using the following formula.

Step size=1/dot product (input vector, input vector)

This step size is proportional to the inverse of the total expected energy of the instantaneous values of the coefficients of the input vector x(n). This sum of the expected energies of the input samples is also equivalent to the dot product of the input vector with itself, and the trace of input vectors auto-correlation matrix, R [6-8].

$$tr[\mathbf{R}] = \sum_{i=0}^{N-1} E\left[x^2(n-i)\right] = E\left[\sum_{i=0}^{N-1} x^2(n-i)\right] \quad (5)$$

The recursion formula for the NLMS algorithm is stated in equation below

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \frac{1}{\mathbf{x}^T(n)\mathbf{x}(n)} e(n)\mathbf{x}(n) \quad (6)$$

Here ***x(n)*** is the input vector of time delayed input values, ***x(n) = [x(n) x(n-1) x(n-2) .. x(n-N+1)]***^{T}. The vector ***w(n) = [***$w_0(n)$ $w_1(n)$ $w_2(n)$ .. $w_{N-1}(n)$***]***^{T} represents the coefficients of the adaptive FIR filter tap weight vector at time ***n***. The parameter μ is known as the step size parameter and is a small positive constant.

A. *Implementation of the NLMS algorithm:*

The NLMS algorithm has been implemented in Matlab. As the step size parameter is chosen based on the current input values, the NLMS algorithm shows far greater stability with unknown signals. This combined with good convergence speed and relative computational simplicity makes the NLMS algorithm ideal for the real time adaptive echo cancellation system [9].

As the NLMS is an extension of the standard LMS algorithm, the NLMS algorithms practical implementation is very similar to that of the LMS algorithm. Each iteration of the NLMS algorithm requires these steps in the following order.

a. The output of the adaptive filter is calculated.

$$y(n) = \sum_{i=0}^{N-1} w(n)x(n-i) = \mathbf{w}^T(n)\mathbf{x}(n) \quad (7)$$

b. An error signal is calculated as the difference between the desired signal and the filter output.

$$e(n) = d(n) - y(n) \quad (8)$$

c. The step size value for the input vector is calculated.

$$\mu(n) = \frac{1}{\mathbf{x}^T(n)\mathbf{x}(n)} \quad (9)$$

d. The filter tap weights are updated in preparation for the next iteration.

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu(n)e(n)\mathbf{x}(n) \quad (10)$$

Each iteration of the NLMS algorithm requires 3N+1 multiplications, this is only N more than the standard LMS algorithm. This is an acceptable increase considering the gains in stability and echo attenuation achieved.

III. AEC USING ESTIMATED LOUD SPEAKER IMPUILSE RESPONSE

A. *AEC-LIME approach:*

The loudspeaker impulse response ***h*** in (1) includes both the unknown time-varying impulse response $\boldsymbol{h_E}$ of the echo path in the near-end room, and the time-invariant impulse response $\boldsymbol{h_L}$ of the loudspeaker of which an estimate $\boldsymbol{\hat{h}_L}$ is assumed to be available. Assuming these impulse responses can be approximated as linear (which is a common basic assumption in AEC), we can write ***h*** as

$$\mathbf{h} = \mathbf{h}_L * \mathbf{h}_E \quad (11)$$

Where denotes convolution, the loudspeaker impulse response $\boldsymbol{h_L}$ of length is defined as

$$\mathbf{h}_L = [h_{L,0} \quad h_{L,1} \quad \cdots \quad h_{L,p-1}]^T, \quad (12)$$

And the echo path impulse response $\boldsymbol{h_E}$ is defined similarly. If ***m*** denotes the length of $\boldsymbol{h_E}$, we have from (5) that

$$n = p + m - 1. \quad (13)$$

Most AEC filter adaptation algorithmswork with the data model in (1). Since we have assumed that we know an estimate $\boldsymbol{\hat{h}_L}$ of $\boldsymbol{h_L}$, we can rewrite this equation as

$$y(t) = \mathbf{h}_E^T\bar{\mathbf{x}}(t) + v(t) + w(t), \quad (14)$$

Where,

$$\bar{\mathbf{x}}(t) = [\bar{x}(t) \quad \bar{x}(t-1) \quad \cdots \quad \bar{x}(t-n+1)]^T \quad (15)$$

$$\bar{x}(t) = \hat{\mathbf{h}}_L^T \check{\mathbf{x}}(t) \quad (16)$$

$$\check{\mathbf{x}}(t) = [x(t) \quad x(t-1) \quad \cdots \quad x(t-p+1)]^T. \quad (17)$$

Since (14) is almost identical to (1), the AEC filter adaptation algorithm can be applied to (14) instead of (1). As $\boldsymbol{h}_E$is shorter than, and the computational complexities of the AEC filter adaptation algorithms usually are proportional to the order of the filter to estimate, the transition from (1) to (14) results in a reduction of the computational complexity for the AEC filter adaptation algorithm. Note, however, that this reduction is only substantial if we have a good estimate of $\boldsymbol{h}_L$. If the estimate $\hat{\boldsymbol{h}}_L$is very poor, we still have to estimate a filter of similar length as $\boldsymbol{h}$ (using an input signal prefiltered by $\hat{\boldsymbol{h}}_L$)

B. Estimation of loud speaker impulse response:

The impulse response of a loudspeaker may be obtained in different ways. The best, and perhaps most direct way, is to compute it from measurements taken in an anechoic chamber. There are, however, also methods for computing the impulse response from measurements taken in an ordinary echoic room [4]. If the loudspeaker impulse responses were time-varying the LIME-approach would not be feasible. Fortunately, it seems that the loudspeaker impulse responses are relative time-invariant, at least for more sophisticated loudspeakers. However, no scientific results have been published about this, instead this property has simply been assumed by the industry and the assumptions seem to be correct. Indeed, this time-invariance is a property used by music products such as the *Dirac Research Corrector* that can compensate for the acoustic properties of loudspeakers [5].

It should also be noted that what we mean by the loudspeaker impulse response is the part of the impulse response that corresponds to the electronics in the loudspeaker and the amplifier.

It is clear that the loudspeaker impulse response is highly dependent on what direction to the loudspeaker it is measured for. What we are interested in is, however, the part that is directional independent (the case is the same for the *Corrector* product mentioned above).

IV. SIMULATION RESULTS

The NLMS algorithm was simulated using MATLAB. Fig.2. shows the input signal. Fig.3. shows the desired signal. Fig.4. shows the adaptive filter output. The adaptive filter is a 1025th order FIR filter. The step size was set to 0.1.

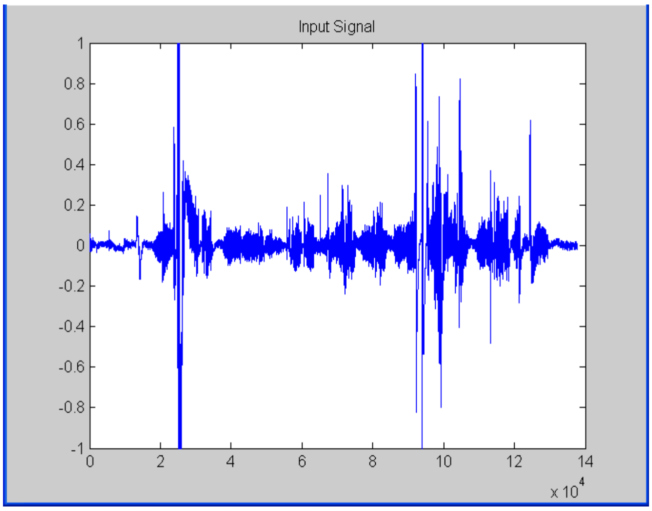

Figure.2. Input signal

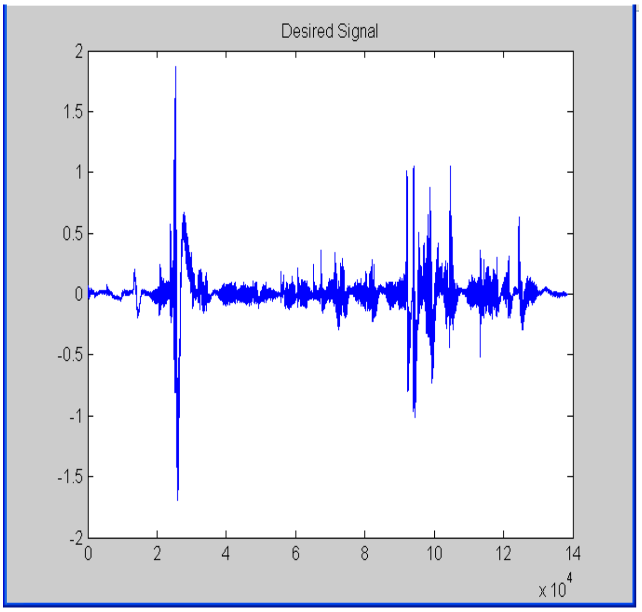

Figure.3. Desired signal

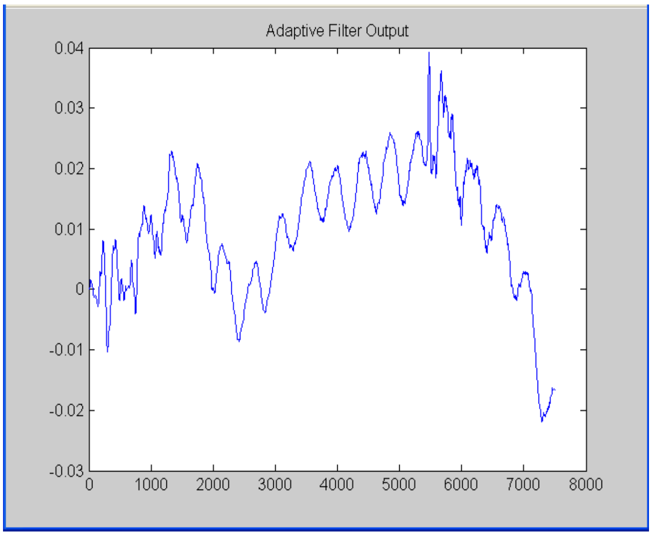

Figure.4. Adaptive filter output

The NLMS with AEC-LIME approach is simulated using MATLAB. Fig.5. shows the input signal. Fig.6. shows the desired signal. Fig.7. shows the adaptive filter output.

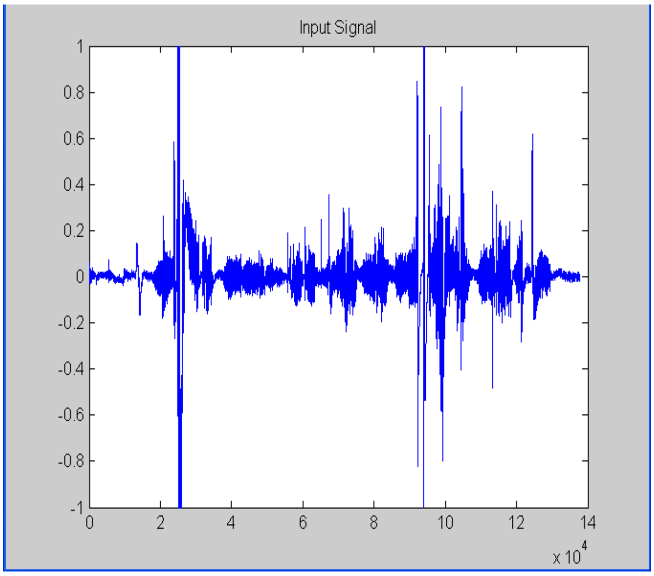

Figure.5. Input signal

CONFERENCE PAPER
Two day National Conference on Advanced Trends and Challenges in Computer Science and Applications
Organized by: Shree Vishnu Engineering College for Women, Bhimavaram A.P.
Schedule: 18-19 March 2014

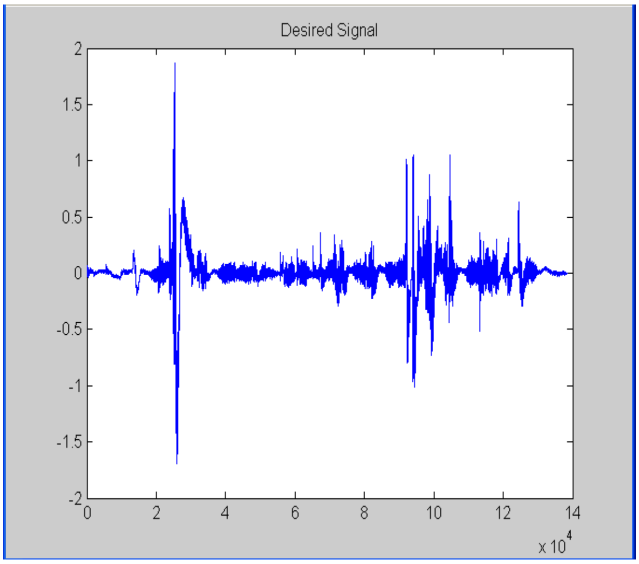

Figure.6. Desired signal

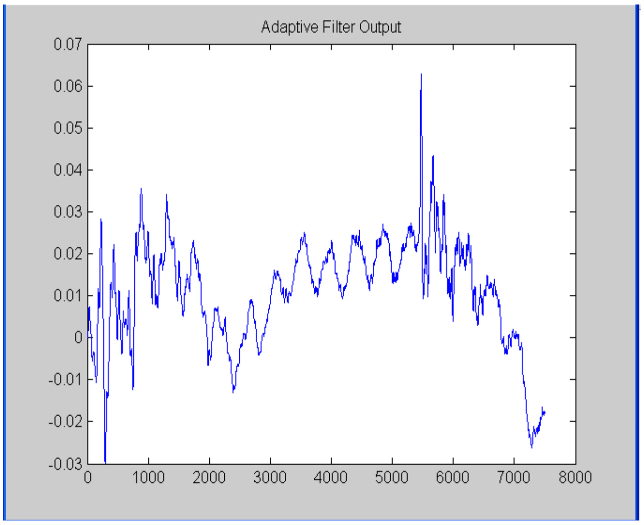

Figure.7. Adaptive filter output

The performance of the AEC-LIME approach applied to an AEC-setup where NLMS is the adaptive algorithm. As far-end speech signal a 10-s speech sample is used, and the total impulse responses are 550 long. The loudspeaker impulse response (that is common to all total impulse responses used in the simulation) is of length 100. The lengths of the filters estimated by NLMS with and without the LIME-approach are set to 450 and 500, respectively.

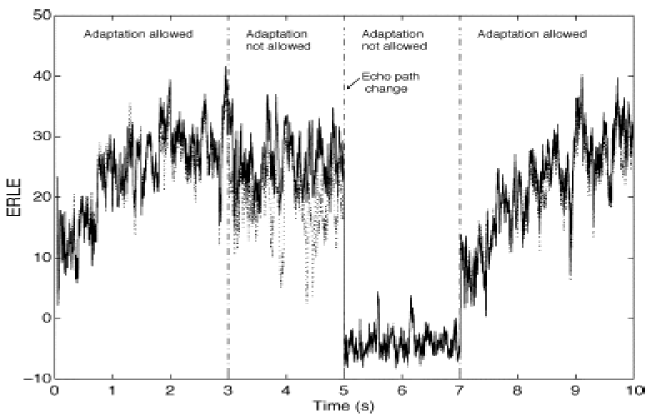

Figure.8. Echo cancellation performance in terms of ERLE as a function of time for NLMS with AEC-LIME (solid) and NLMS without AEC-LIME (dotted).

The SNR is set to 35 dB. In order to simulate a reasonably realistic AEC-setup, we introduced changes in $\boldsymbol{h}_{\boldsymbol{E}}$. During the first 5 s, $\boldsymbol{h}_{\boldsymbol{E}}$ is kept constant. After 5 s, $\boldsymbol{h}_{\boldsymbol{E}}$is changed abruptly (corresponding to somebody suddenly blocking or moving the loudspeaker or microphone) and then again kept constant for the rest of the simulation.Furthermore, filter adaptation is not allowed from 3 to 7 s, corresponding to a doubletalk situation. Note, however, that we did not add any near-end speech as the ERLE measure is only valid when there is no near-end speech present. This does, however, not modify the interpretation of the simulation results. The simulation results are shown in Fig. 6, where the ERLE is plotted as a function of time.

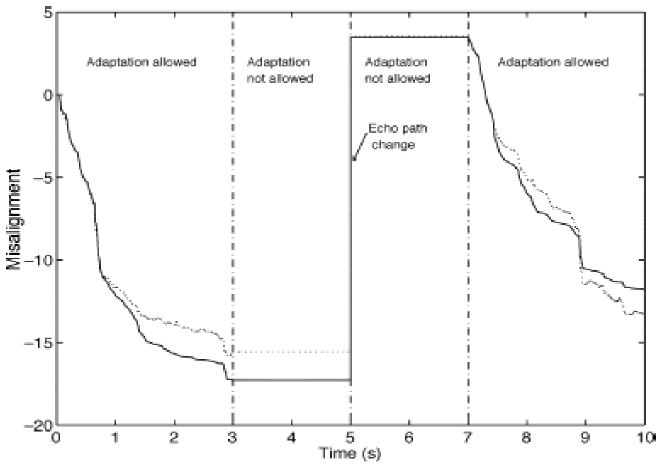

Figure.9. Echo cancellation performance in terms of misalignment as a function of time for NLMS with AEC-LIME (solid) and NLMS without AEC-LIME (dotted).

As we can see NLMS without the LIME approach performs similarly to NLMS with the LIME approach when there is no doubletalk. However, when there is doubletalk (and filter adaptation is not allowed), NLMS with the LIME approach performs better than NLMS without the LIME approach. After the change in $\boldsymbol{h}_{\boldsymbol{E}}$ at 5 s, both algorithms performs poorly, but that is to be expected as the previous estimates for $\boldsymbol{h}_{\boldsymbol{E}}$ same simulation are displayed in terms of misalignment. Again we see that NLMS with the LIME approach performs similarly to NLMS without the LIME approach. It is clear that using the LIME-approach for AEC it is possible to reduce the length of the adaptive filter, and still get a comparable or even better, AEC performance.

V. CONCLUSION

The NLMS algorithm, an equally simple, but more robust variant of the LMS algorithm, exhibits a good balance between simplicity and performance. Due to its good properties the NLMS has been largely used in real-time applications.

We have proposed a new approach to acoustic echo cancellation that can be used for most echo cancellation algorithms. When applied to echo cancellation algorithms, the approach offers a minor improvement in computational complexity. However, as the simulations show, it may improve the echo cancellation performance. .

VI. ACKNOWLEDGMENT

The authors express their deep sense of gratitude to Sri K.V. Vishnu Raju, Chairman of Sri Vishnu Educational Society, Andhra Pradesh, India, Prof. K.Srinivas, head of the department and Prof. O.Ranga Rao, assistance professor for giving the chance to work in the labs of ECE, and also, for the constant encouragement by providing funds to complete this project.

VII. REFERENCES

[1]. M. M. Sondhi, “An adaptive echo canceler,” *Bell Syst. Tech. J.*, vol. XLVI, no. 3, pp. 497–510, 1967.

[2]. C. Breining, P. Dreiseitel, E. Hänsler, A. Mader, B. Nitsch, H. Puder, T. Schertler,G. Schmidt, and J. Tilp, “Acoustic echo control–an application of very-high-order adaptive filters,” *IEEE Signal Process. Mag.*, vol. 16, no. 4, pp. 42–69, Jul. 1999.

[3]. S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.

[4]. P. Åhgren and P. Stoica, “A simple method for estimating the impulse responses of loudspeakers,” *IEEE Trans. Consumer Electron.*, vol. 49, no. 4, pp. 889–893, Nov. 2003.

[5]. (2003, Jul.) Corrector. [Online] Available: http://www.diracresearch.se

[6]. J. Shynk, “Frequency-domain and multirate adaptive filtering,” IEEE Signal Processing Mag., vol. 9, pp. 15– 37, Jan. 1992.

[7]. Ahmed I. Sulyman and AzzedineZerguine, "Echo Cancellation Using a Variable Step-Size NLMS Algorithm", Electrical and Computer Engineering Department Queen's University.

[8]. D. L. Duttweiler, “A twelve-channel digital echo canceller,” IEEE Trans. Commun., vol. 26, no. 5, pp. 647–653, May 1978.

[9]. J. Benesty, H. Rey, L. Rey Vega, and S. Tressens, “A nonparametric VSS NLMS algorithm,” IEEE Signal Process. Lett., vol. 13, pp. 581–584, Oct. 2006.

[10]. Radhika Chinaboina, D.S.Ramkiran, Habibulla Khan, M.Usha, B.T.P.Madhav, K.PhaniSrinivas&G.V.Ganesh, “adaptive algorithms for acoustic echo cancellation in speech processing”, IJRRAS 7 (1), April 2011