

# Nonlinear Acoustic Echo-Free System

Deshpande Tanavi A. and Dube R. R.

**Abstract**—Acoustic Echo Cancellation (AEC) is one of the most popular applications for adaptive filters. The goal of the adaptive algorithms is to identify the acoustic echo path between a loudspeaker and a microphone placed in an acoustic enclosure (i.e., the room acoustic impulse response) and generate the estimate of the acoustic echo that is removed from the output of the microphone. Most of the adaptive echo cancellation techniques proposed in literature are based on an assumption that the echo path is linear. However, recent research has shown that the acoustic echo path is better to be modeled as a nonlinear system due to the nonlinear characteristic of a loudspeaker[1].The nonlinearity of a power amplifier or loudspeaker in a large-signal situation gives rise to a nonlinear distortion of acoustic signal. A conventional acoustic echo canceller using linear adaptive filters is not able to eliminate the nonlinear echo component. Here, a novel nonlinear echo cancellation technique is presented by using a nonlinear transformation in conjunction with a conventional linear adaptive filter. The nonlinear transformation is derived from a raised-cosine function and is exploited to compensate for the nonlinearity of a loudspeaker. The algorithms which were derivates out of the Wiener Filter (like LMS, NLMS, steepest descent, RLS) in contrast to the Genetic algorithms (GAs) is of special interest[6]. With the comparison, NLMS algorithm is better solution considering trade-offs. The transformation parameters are updated using the normalized least mean square algorithm according to the unknown nonlinear characteristic of the loudspeaker.

**Index Terms**—Acoustic echo canceller, adaptive algorithm, NLMS, non linear transformation, raised cosine function.

## I. INTRODUCTION

Today's telecommunication systems system consists of mostly full duplex mode. In that, coupled acoustic input and output devices, both are active concurrently in full duplex mode. The system then acts as both a receiver and transmitter in full duplex mode. In this situation the received signal is output through the telephone loudspeaker (audio source) from Far End Speech, (FES), this audio signal is then reverberated through the physical environment and picked up by the systems microphone (audio sink) at near end speech (NES). This signal is reverberated within the environment and returned to the system via the microphone input. These reverberated signals contain time delayed

images of the original signal, which are then returned to the original sender (Fig. 1,  $a_k$  is the attenuation,  $t_k$  is time delay). [2] This makes the remote speaker hear her own voice distorted and delayed by the communication channel, which is known as echo. The longer the channel delay, the more annoying the echo becomes until it makes natural conversation impossible and decreases the perceived quality of the communication service.

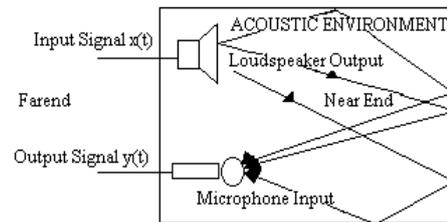


Fig. 1. Origins of acoustic echo.

Modern full-duplex communication systems make use of an acoustic echo canceller (AEC) to prevent the echo from being transmitted back to the channel. The AEC basically estimates the echo and subtracts the estimated echo from the microphone signal.

The resulting signal is transmitted to the far end speaker through the communication channel.

The method used to cancel the echo signal is known as adaptive filtering. Adaptive filters are dynamic filters which iteratively and algorithmically alter their characteristics in order to minimize a function of the difference between the desired output  $d(n)$  and its actual output  $y(n)$ .

This function is known as the cost function of the adaptive algorithm. Fig. 2 shows a block diagram of the adaptive echo cancellation system. Here the filter  $H(n)$  represents the impulse response of the acoustic environment,  $W(n)$  represents the adaptive filter used to cancel the echo signal. The adaptive filter aims to equate its output  $y(n)$  to the desired output  $d(n)$  (the signal reverberated within the acoustic environment). At each iteration the error signal,  $e(n) = d(n) - y(n)$ , is fed back into the filter, where the filter characteristics are altered accordingly.

Acoustic echo cancellers (AEC) greatly enhance the audio quality of a multipoint hands-free communications system. They allow conferences to progress more smoothly and naturally, keep the participants more comfortable, and prevent listener fatigue.

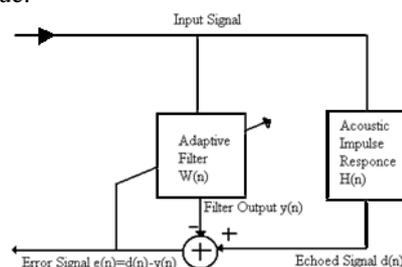


Fig. 2. Block diagram of an adaptive echo cancellation system.

Manuscript received April 10, 2012; revised June 3, 2012. This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada. H. Dai is with SAP China, Shanghai 200003, China. W.-P. Zhu is with the Center for Signal Processing and Communications, Department of Electrical and Computer Engineering, Concordia University, Montreal, QC H3G 1M8, Canada (e-mail: weiping@ece.concordia.ca).

A. Deshpande Tanavi is with the Electronics Department, WIT, Solapur 413006, India (e-mail: tanvi.deshpande2003@gmail.com).

R. R. Dube is with the Electronics Department, WIT, Solapur 413006, India (e-mail: rajendradube67@rediffmail.com).

The primary beneficiaries of an echo canceller are the people at the far (or remote) end of the transmission path. The near (or local) echo canceller prevents the echo of the remote peoples' voices from being returned (i.e. echoed) to them through the audio system.

Focusing on heart of paper[1], we must say that, most of the AEC are designed considering that room impulse response is linear and tries to replicate it. But, recent research shows that loud speaker also processes far-end speech. This change in reverberated signal should be taken into account not to AEC converge badly. So, here raised cosine function is used to compensate the non linearity of loudspeaker. we will discussed it in detail in Section 3..

A. In Summary

AEC is Most popular applications for adaptive filters. The goal of the adaptive algorithms-

- To identify the AE path between a loudspeaker and uphone placed in an acoustic enclosure.
- To generate the estimate of the AE that is removed from the o/p of the microphone.

B. Drawback of conventional Linear AEC-

- Assumption that "The echo path is linear."
- However, better modeling with a nonlinear system gives rise to a nonlinear distortion of acoustic signal.
- A conventional AEC is not able to eliminate the nonlinear echo component.

C. Proposed Method-

- A novel nonlinear AEC technique - nonlinear transformation + a conventional linear adaptive filter.
- The nonlinear transformation - A raised-cosine function.
- Updation of the transformation parameters – NLMS

D. Algorithm

This thesis is divided into a number of sections: Section 2 deals with background signal processing theory as it pertains to the application of adaptive filters. Section 3 elaborate proposed method. Section 4 details summarizes the simulation work . Section 5 outlines the conclusion .Section 6 is of References.

II. IBACKGROUND THEORY [2],[3]

A. FIR Filter, Transversal FIR Filer and Adaptive Filter

FIR filters have the following primary advantages: They can have exactly linear phase .They are always stable. The design methods are generally linear. They can be realized efficiently in hardware. The filter startup transients have finite duration. Neglecting, the much higher length of filter and corresponding delay comparing to IIR filter, we have selected FIR filter following hardware implementation by transversal FIR filters[4]. The characteristics of a adaptive transversal FIR filter can be expressed as a vector consisting of values known as tap weights in column vector form as,  $w(n) = [w_0(n) w_1(n) w_2(n) .. w_{N-1}(n)]^T$  (impulse response of the FIR filter.). It is these tap weights which determine the performance of the filter. The number of elements on this impulse response is order of filter ,N. The aim of an adaptive

filter is to calculate the difference between the desired signal and the adaptive filter output,  $e(n)$ .  $e(n)$  (estimation error) is fed back into the adaptive filter and its coefficients are changed algorithmically in order to minimise a function, known as the cost function. Here, the echoed signal would be completely cancelled and the far user would not hear any of their original speech returned to them. Implementation of filter may be as Single filter, sub-band filter and dyadic synthesis sub band filter in time domain and also in frequency domain. But trade – off within complexity and performance, we are using time domain single filter Structure.

B. Summary and Selection of Algorithm

It can be seen from Table I [2] that when considering the attenuation values and the number of multiplication operations for each algorithm, the NLMS algorithm is the obvious choice for the real time acoustic echo cancellation system. The system is capable of cancelling echo with time delays of up to 75 ms, corresponding to reverberation off an object a maximum of 12 meters away.

III. PROPOSED METHOD [1]

A. UPDATE OF Adaptive Filter:

Here, we use the NLMS algorithm for the linear adaptive filter part. The cost function,  $\zeta(n)$ , is a function of mean of squares of the difference between a desired output and the actual output of the FIR filter. Let  $d(n)$  and  $y(n)$  denote the desired signal and the output of the adaptive filter, respectively. The NLMS algorithm can be briefly described as:

$$e(n) = d(n) - y(n) = d(n) - W_n^T x_n \tag{1}$$

$$W_{n+1} = W_n + \mu \frac{e(n)x_n}{x_n^T x_n + C} \tag{2}$$

where  $\mu$  is a step size, and c is a small constant. where  $W_n$  represents the coefficient vector of the transversal filter, and  $X_n$  is the signal vector input to the nonlinear transformer.

TABLE I: SUMMARY OF ADAPTIVE FILTER ALGORITHM.

Algorithm	Average Attenuation	Multiplication Operations	Comments
LMS	-18.2 dB	2N+1	Is the simplest to implement and is stable when the step size parameter is selected appropriately. This requires prior knowledge of the input signal which is not feasible for the echo cancellation system.
NLMS	-27.9dB	3N+1	Simple to implement and computationally efficient. Shows very good attenuation, and variable step size allows stable performance with non-stationary signals. This was the obvious choice for real time

			implementation.
VSLMS	-9.8 dB	4N+1	Displays very poor performance, possibly due to non-stationary nature of speech signals. Only half the attenuation of the standard LMS algorithm. Not considered for real time implementation.
LMS	-18.2 dB	2N+1	Is the simplest to implement and is stable when the step size parameter is selected appropriately. This requires prior knowledge of the input signal which is not feasible for the echo cancellation system.
RLS	-34.2 dB	4N <sup>2</sup>	Has the greatest attenuation of any algorithm studied, and converges much faster than the LMS algorithm. This performance comes at the cost of computational complexity and considering the large FIR order required for echo cancellation, this is not feasible for real time implementation.

**B. Method Proposed:**

Fig. 3 shows the proposed nonlinear echo canceller.

It consists of a nonlinear transformation and a conventional linear adaptive filter. The signal from a far-end user is assumed to be nonlinearly distorted in the power amplifier/loudspeaker[5], [6], [7], [8]. It then passes through a linear echo path characterized by the room impulse response to form the nonlinear acoustic echo. In the mean time, the undistorted far-end signal is input to a nonlinear transformation to compensate for the nonlinearity of the loudspeaker. The nonlinearly transformed signal is then processed by a conventional linear adaptive filter. In proposed method, the input to the linear adaptive filter is  $F(x_n)$ .

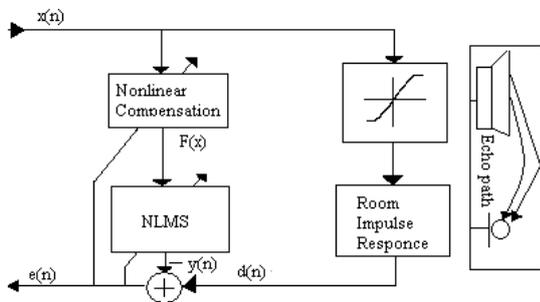


Fig. 3. Proposed nonlinear AEC.

When a transversal filter is used, the estimation error  $e(n)$  can be written as (3) given below:

$$e(n) = d(n) - y(n) = d(n) - W_n^T F(x_n) \quad (3)$$

Note that  $F(Xn)$  represents a vector that is the non-linearly distorted version of  $Xn$ . So, here NLMS algorithm is applied

$$W_{n+1} = W_n + \mu \frac{e(n)F(x_n)}{F^T(x_n)F(x_n) + C} \quad (4)$$

The nonlinear transformation ( $Xn$  to  $F(Xn)$ ) to be developed should be versatile and be able to cover a wide range of nonlinear distortion[5], [6], [7], [8]. It is expected to provide a good approximation for both linear mapping (when the amplitude of the input signal is small or the speaker system does not introduce nonlinear distortion) and nonlinear saturation characteristic (when the input has large amplitude or the nonlinear distortion occurs). Non linearity can be modeled by sigmoid function. We consider that this non-linearity can be modeled by raised-cosine function.

Raised Cosine function is given as:

$$f(x) = \begin{cases} -1, & x < -\frac{1+\beta}{2T} \\ Tx - \frac{1-\beta}{2} - \frac{\beta}{\pi} \cos\left(\frac{2Tx\pi + \pi}{2\beta}\right), & -\frac{1+\beta}{2T} \leq x \leq -\frac{1-\beta}{2T} \\ 2Tx, & -\frac{1-\beta}{2T} \leq x \leq \frac{1-\beta}{2T} \\ Tx + \frac{1-\beta}{2} + \frac{\beta}{\pi} \cos\left(\frac{2Tx\pi - \pi}{2\beta}\right), & \frac{1-\beta}{2T} \leq x \leq \frac{1+\beta}{2T} \\ 1, & x > \frac{1+\beta}{2T} \end{cases} \quad (5)$$

where  $T$  and  $\beta$  are the free parameters to determine the shape and the dynamic range of the function. A close observation of the curve given by Fig. 4 and equations yields a family of nonlinear transformation, which is very suitable for modeling the saturation behavior of a loudspeaker. In general, when the input signal has small amplitude, the amplifier/loudspeaker works in a linear region which produces a linear mapping  $2Tx$ . As the amplitude of the input signal increases, the nonlinear distortion or saturation occurs. When the amplitude of the input signal is larger than  $\frac{1+\beta}{2T}$ , the output is hard

limited to 1, (clipping). The soft-clipping case is similar to the sigmoid function, but it has two parameters jointly dictating the saturation curve while the sigmoid function contains only one argument. Moreover, hard clipping and linear mapping are not obtainable from the sigmoid function. In this way, the transformation given by above is able to cover a wide range of saturation curves depending on the choice of  $T$  and  $\beta$ .

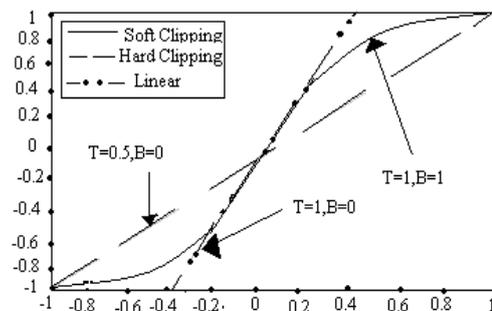


Fig. 4. Three types of saturation curves of a loudspeaker.

The squared error can be modified as here :

$$J(W_n, \beta, T) = [e(n)]^2 = [d(n) - W_n^T F(x_n)]^2 \quad (6)$$

Here, the cost function above contains two more unknowns, i.e.  $T$ , and  $\beta$ , in addition to the filter coefficient vector  $W_n$ . Applying the steepest descent criterion to minimize cost function and considering as constant in each iteration of  $\beta$  and  $T$ , we can obtain the following update formulas for  $\beta$  and  $T$ :

$$\beta_{n+1} = \beta_n - \frac{\mu\beta}{2}, \frac{\partial J(W_n, \beta, T)}{\partial \beta} \Big|_{\beta=\beta_n} \quad (7)$$

$$\beta_{n+1} = \beta_n + \mu\beta, e(n), W_n^T, \frac{\partial F(x_n)}{\partial \beta} \Big|_{\beta=\beta_n} \quad (8)$$

where the vector  $\frac{\partial F(x_n)}{\partial \beta}$  has the element  $\frac{\partial F(x)}{\partial \beta}$ , as shown below

$$\frac{\partial F(x)}{\partial \beta} = \begin{cases} 0, & \text{for } x < \frac{1-\beta}{2T} \text{ or } x > \frac{1+\beta}{2T} \\ \frac{1}{2} - \frac{1}{\pi} \cos\left(\frac{2Tx\pi + \pi}{2\beta}\right) - \frac{2Tx\pi + 1}{2\beta} \sin\left(\frac{2Tx\pi + \pi}{2\beta}\right), & \text{for } -\frac{1+\beta}{2T} \leq x \leq -\frac{1-\beta}{2T} \\ -\frac{1}{2} + \frac{1}{\pi} \cos\left(\frac{2Tx\pi - \pi}{2\beta}\right) + \frac{2Tx\pi - 1}{2\beta} \sin\left(\frac{2Tx\pi - \pi}{2\beta}\right), & \text{for } \frac{1-\beta}{2T} < x \leq \frac{1+\beta}{2T} \end{cases} \quad (9)$$

and

$$T_{n+1} = T_n + \mu_T, e(n), W_n^T, \frac{\partial F(x_n)}{\partial T} \Big|_{T=T_n} \quad (10)$$

$$T_{n+1} = T_n - \frac{\mu_T}{2}, \frac{\partial J(W_n, \beta, T)}{\partial T} \Big|_{T=T_n} \quad (11)$$

where each element of  $\partial F(x_n)/\partial T$  is given by

$$\frac{\partial F(x_n)}{\partial T} = \begin{cases} 2x, & x < \frac{1-\beta}{2T} \\ x + \sin\left(\frac{2Tx\pi + \pi}{2\beta}\right), & -\frac{1+\beta}{2T} \leq x \leq -\frac{1-\beta}{2T} \\ x - \sin\left(\frac{2Tx\pi - \pi}{2\beta}\right), & \frac{1-\beta}{2T} \leq x \leq \frac{1+\beta}{2T} \\ 0, & x > \frac{1+\beta}{2T} \end{cases} \quad (12)$$

Note that  $\beta(n+1)$  and  $T(n+1)$  apply to each sample of input signal. The subscript has been omitted for notational simplicity. The above iterations for and are performed on sample-by-sample basis, namely, for each input sample, and are updated once. Since all the updates are given in an explicit form, the computation of and can easily be implemented. Due to the piecewise defined nonlinear function  $F(x)$  involved in the algorithm, a theoretical analysis of the convergence of  $T$ ,  $\beta$ , and  $W_n$  is rather complicated. Fig. 5 shows algorithm of proposed method.

#### IV. SIMULATION RESULTS

We are coding the proposed method in MATLAB® software. MATLAB® is a high-performance language for technical computing. MATLAB features a family of add-on application-specific solutions called toolboxes. We are using the Filter Design Toolbox which is a collection of tools that provides advanced techniques for designing, simulating, and analyzing digital filters and The Signal Processing Toolbox that supports a wide range of signal processing operations.

The echo return loss enhancement (ERLE) is used to measure the performance of the proposed nonlinear echo canceller. The ERLE is defined as A measure to express the effect of an echo cancellation filter is the echo-return loss enhancement (ERLE):

$$ERLE(db) = 10 \log_{10} \left( \frac{\sigma_d^2}{\sigma_e^2} \right) \quad (13)$$

where  $\sigma_d^2$  and  $\sigma_e^2$  are the variance of the desired signal  $d(n)$  and that of the error  $e(n)$ , respectively. So, more ERLE more better results are obtained. Here, the convergence behavior of parameters  $T$  and  $\beta$  and the ERLE performance of the proposed method along with that of canceller in conjunction with pure NLMS is examined.

Fig. 5 shows ERLE performance with proposed method. Here, for simulation of speech  $T = 1$ ,  $\beta = 1$  is chosen. And FIR of length 70 is taken for simulation of room impulse response. Hard limiting loudspeaker is thus simulated. To mimic exactly same and to eliminate undermodeling case, FIR of 70 is taken and initial value of  $T$  and  $\beta$  are 0.5 and 0.01 respectively chosen. Step size has formula

$$\mu = \frac{\text{alpha}}{(c + y \times y)} \quad (14)$$

where alpha is chosen 0.08 and c is small constant of value 0.01 to avoid divide by 0 case. Error is so minimized, it can't be audible to human ear. ERLE performance is within 40db to 110db mainly.

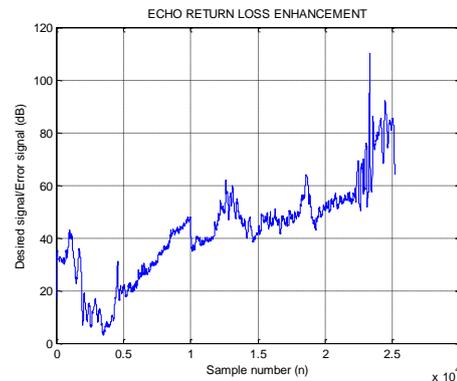


Fig. 5. ERLE performance with proposed method.

Fig. 6 gives ERLE performance with Pure NLMS without conjunction with nonlinear compensation. Here alpha is chosen same, that is 0.08. Here, ERLE performance is within 15 db to 40 db only. Close observation shows proposed method is much efficient.

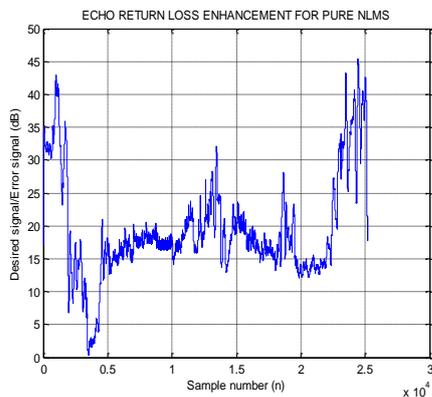


Fig. 6. ERLE performance using pure NLMS.

Fig. 7 shows results for  $T = 1$ ,  $\beta = 1$ . alpha is chosen 0.32. Fig. shows clearly improved ERLE performance for proposed method than only pure NLMS algorithm.

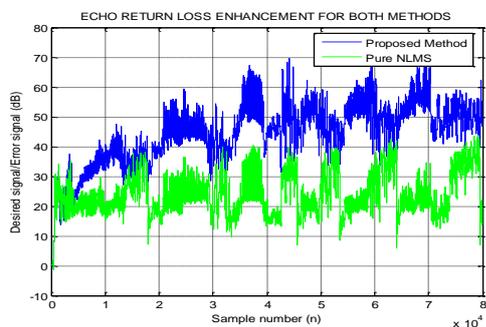


Fig. 7. ERLE performance.

Fig. 8 shows convergence of B and T parameter with simulated loudspeaker has  $T = 1$ ,  $\beta = 1$ . B has some misalignment but T is approximately converges to 1. Error is non perceptible to human ears.

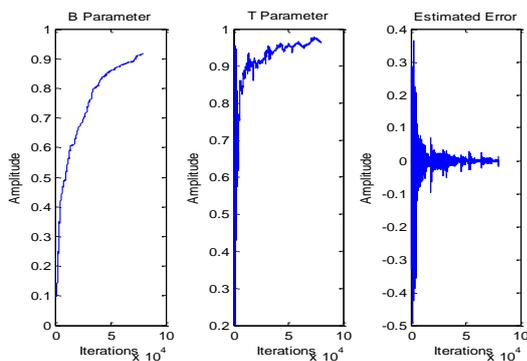


Fig. 8. B and T parameter learning curves and Error graph.

V. CONCLUSION

Due to acoustic echo in speech transmission, quality of communication is degraded. Most of the AEC techniques proposed are based on an assumption that the echo path is linear. Recent research has shown that the acoustic echo path is better to be modeled as a nonlinear system due to the nonlinear characteristic of a loudspeaker. The saturation characteristic of a loudspeaker is modeled by a sigmoid function. However, the sigmoid function fails to approximate the linear mapping due to its capability of modeling various saturation phenomena of loudspeaker is heavily limited. On other hand, Raised cosine function has two parameters jointly dictating the saturation curve and able to cover a wide range of saturation curves depending on the choice of  $T$  and  $\beta$ . Discussed update mechanism for the parameters is to best compensate for the unknown loudspeaker distortion. Simulation has shown that, proposed method With compensation of nonlinearity of loudspeaker is much more efficient than method with only conjunction with pure NLMS for same speech signal and same set-up.

REFERENCES

- [1] H. Y. Dai and W.-P. Zhu, *Senior Member, IEEE*, "Compensation of Loudspeaker Nonlinearity in Acoustic Echo Cancellation Using Raised-Cosine Function," *Published in IEEE Transaction ON Circuits and System —II: Express Briefs*, vol. 53, no. 11, November 2006.
- [2] M. Hutson, "Acoustic echo cancellation using digital signal processing a thesis from the school of information technology and electrical engg," The University of Queensland. *Submitted for the degree of Bachelor of Engineering (Honours) in the division of Electrical and Electronic Engineering*. November 2003.
- [3] Book Adaptive Filter Theory - Simon Haykin, 3<sup>rd</sup> edition
- [4] A. P. Liavas and P. A. Regalia, *Senior Member, IEEE*, "Acoustic Echo Cancellation: Do IIR Models Offer Better Modeling Capabilities than Their FIR Counterparts," *IEEE Transactions On Signal Processing*, vol. 46, no. 9, September
- [5] C. A. Lagrange and A. Arliaud, "Acoustic Echo cancellation with cascade filters," *IEEE paper presented in proceedings of ICASSP 2003*, pp. 389-392
- [6] Stenger and W. Kellermann, "Nonlinear Acoustic echo cancellation with fast converging memoryless preprocessor," *A Paper Published in IEEE Transaction* pp. 805
- [7] A. Guarin, G. Faucon, and R. L. Bouquin-Jeannes, "Nonlinear Acoustic Echo Cancellation Based on Volterra Filters," *IEEE Transaction on Speech and Audio processing*, vol. 11, no. 6, Nov-2003 pp. 672-683
- [8] M. I. Mossi and W. D. Nicholas, "Evans An Assessment of Linear Adaptive Filter performance With Nonlinear Distortion," *Paper Presented in Proceedings of IEEE, ICASSP 2010*, pp. 313-31.