Speech Enhancement using Spectral Subtraction, Affine Projection Algorithms and Classical Adaptive Filters

Ravi Bolimera[1], Siva Prasad Nandyala[2]
[1, 2] Student Member IEEE, Department of ECE,
National Institute of Technology Warangal,
Warangal, Andhra Pradesh, India.
ravi_gamr@yahoo.com , nsprasad@nitw.ac.in

T. Kishore Kumar[3]
[3]Member IEEE, Department of ECE,
National Institute of Technology Warangal,
Warangal, Andhra Pradesh, India.
kishoret@nitw.ac.in

***Abstract*—In this paper, we analysed the performance analysis for speech enhancement using spectral subtraction, affine projection algorithms and classical adaptive filters. As the enhancement of speech signals is of very important in many applications like speech recognition, hearing aids, forensic applications and telephone conversations etc. The performance of the speech recognition also reduces if the speech signal is corrupted by noise. To remove the noise present in the speech signal, the adaptive filters shown the good improvement in increasing the Signal to Noise Ratio (SNR) values. The matlab simulations are performed using NOIZEUS speech corpus for different SNR values using Spectral Subtraction(SS), Least Mean Square (LMS), Normalized Least Mean Square (NLMS), Recursive Least Squares (RLS), Affine Projection Algorithms (APA). From the results it is observed that RLS algorithm has shown good improvement in speech enhancement when compared to the other methods.**

***Keywords*— Spectral Substraction, Adaptive Filters, LMS, NLMS, Affine Projection, RLS,Speech Enhancement**

I. INTRODUCTION

Adaptive filtering is an important subfield of digital signal processing having been actively researched for more than five decades and having important applications such as noise cancellation, system identification, etc.. In such noise removal applicable systems, the signal characteristics are quite faster rate. NLMS and RLS [2]-[5] [11] algorithms are the most frequently and widely applied adaptive algorithms for noise cancellation [1].

Out of LMS, NLMS and RLS, it is sure that the RLS algorithm provides fast adaptation rate and high computational complexity. Computation complexity is the weakest point of RLS. Whereas, low computational complexity is the advantage of NLMS algorithm [2]-[5][11]. Even the Affine Projection algorithm can be interpreted as a generalization of the NLMS algorithm, the main advantage of the APA over the NLMS algorithm consists of a superior convergence rate, especially in speech. Thus, the choice of adaptive algorithm to be applied is always a tradeoff in between computational complexity and fast convergence. The convergence property of RLS algorithm is superior to that of the usual LMS, NLMS and the APA algorithm.

Organization of paper as follows:

Section II presents the need for speech enhancement, section III explains the classical adaptive algorithms such as LMS, NLMS and RLS algorithms will be reviewed, Section IV briefly introduced the sparse adaptive filter(APA) [8] algorithm. Section V investigates the experimental results and the performance evaluation done for the above work in section VI. At long last in section VII, conclusion is summarized.

II. SPEECH ENHANCEMENT

The goal of Speech enhancement is to empower speech quality by using several algorithms. It is one of the significant topics to enhance the performance of the systems of noisy in speech signal processing. It has many applications like hearing aids, forensic applications, cellular environments, front-ends for speech recognition system, telecommunication signal enhancement, military, etc. Communication systems have noise and distortions are the main limiting factors. Hence to sweep over these, their modeling and removal have been at the core of the theory and practice of communications and signal processing. Various techniques are modeled for this purpose to improve the speech signal-to noise ratio and the performances depend on quality and intelligibility of the processed speech signal. The following figure 1 shows the basic idea of speech enhancement.

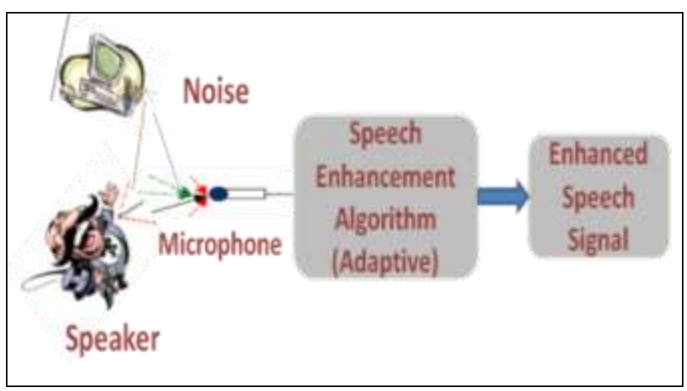

Figure 1. Basic idea of speech enhancement

Spectral Subtraction:

The spectral subtraction algorithm is historically one the first algorithms proposed for noise reduction. It is based on a simple principle. Assuming additive noise, one can obtain an estimate of the clean signal spectrum by subtracting an estimate of the noise spectrum from the noisy speech spectrum. The noise spectrum can be estimated, and updated, during the periods when the signal is absent or when only noise is present. The simple subtraction processing comes with a price. The subtraction process needs to be done carefully to avoid any speech distortion. If too much is subtracted, some speech information might be removed as well, if too little is subtracted, much of the interfering noise remains.

III. Classical Adaptive Filters

Whenever there are either unknown fixed specifications or unsatisfied specifications by time-invariant filters, an adaptive filter is required. Since the characteristics are dependent on the input signal, an adaptive filter is a nonlinear filter and consequently the homogeneity and additivity conditions are not satisfied. Adaptive filters are time-varying since filter parameters are continually changing to meet performance requirement. The existence of a reference signal which is hidden in the fixed-filter approximation step, defines the performance criterion. The general adaptive-filter configuration is shown in figure 2.

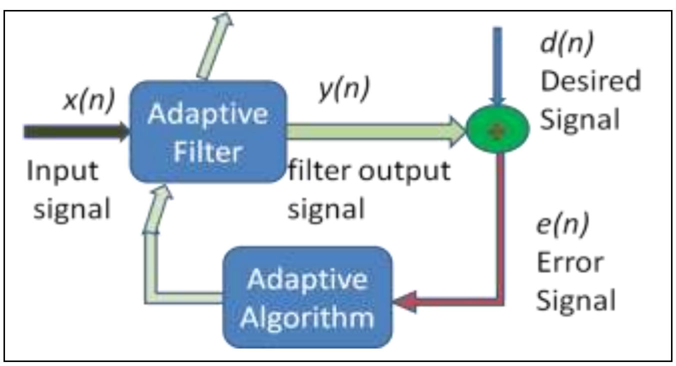

Figure 2. General adaptive-filter configuration.

Least Mean Square (LMS) Algorithm:

Widrow [2] and Hoff were developed the Least Mean Square (LMS) algorithm, is the first and most used adaptive algorithm. The LMS, itself established as the workhorse of adaptive signal processing for two primary reasons:

- Easy to implement and computational efficiency i.e., linear in the number of adjustable parameters.
- Robust performance

LMS is a gradient descent algorithm and it modifies adaptive filter taps by an amount proportional to the instantaneous estimate of the gradient of the error surface [7]. The following operations are performed in the standard LMS algorithm to update the filter coefficients [9]:

- Computes the output signal y(n) from the adaptive filter.
- Computes the error signal e(n) by the equation

$$e(n) = d(n) - y(n) \qquad (1)$$

- Modifies or updates the filter coefficients by the equation

$$\overline{w}(n+1) = \overline{w}(n) + \mu e(n)\bar{u}(n) \qquad (2)$$

Where $\overline{w}(n)$ is the filter coefficients vector, μ is the step size of the adaptive filter, and ū(n) is the filter input vector [6]. To minimize the cost function, LMS algorithm adjusts the filter coefficients. They do not demand any matrix operations and therefore less computational resources and memory requires.

Normalized Least Mean Square (NLMS) Algorithm:

The modified form of the standard LMS algorithm is the NLMS algorithm. By using the following equation, NLMS updates the coefficients of an adaptive filter

$$\overline{w}(n+1) = \overline{w}(n) + \mu e(n)\frac{\bar{u}(n)}{||\bar{u}(n)||^2 + \delta} \qquad (3)$$

Where μ(n) = μ / (||ū(n)||2 +δ) and δ is regularization constant. The value of δ is very small which is approximately 0.01.

The Main difference of NLMS algorithm to standard LMS algorithm is a time-varying step size μ(n). By this, it can establish that amplitude of the mean square error of the error signal is lesser than that of the standard LMS algorithm [9]. This needs N more multiplication operations. Also it is found that the impulse response has peaks of double the amplitude of the LMS algorithm after the same number of iterations, which implies the higher convergence rate of the NLMS than standard LMS.

Recursive Least Squares (RLS) Algorithm:

If the Wiener filter implementation is recursive which uses to find the filter coefficients that relates to producing the recursively least squares of the error signal i.e. the difference between the desired and actual signal is the RLS Filter.

An exact minimization of sum of the squares of the desired signal estimation errors are performed by RLS [6] at each instant. The Exponential weighing factor should be in range 0 to 1 for proper memory organization, 1 specifies an infinite memory. The RLS approach offers faster convergence and smaller error with respect to the unknown system, at the expense of requiring more computations when compared to LMS and NLMS.

The Steps involved in RLS algorithm are:

Initialize the algorithm by setting

$$\hat{\mathbf{w}}(0) = \mathbf{0},$$
$$\mathbf{P}(0) = \delta^{-1}\mathbf{I},$$

and

$$\delta = \begin{cases} \text{small positive constant for high SNR} \\ \text{large positive constant for low SNR} \end{cases}.$$

For each instant of time, $n = 1, 2, \ldots,$ compute

$$\boldsymbol{\pi}(n) = \mathbf{P}(n-1)\mathbf{u}(n),$$
$$\mathbf{k}(n) = \frac{\boldsymbol{\pi}(n)}{\lambda + \mathbf{u}^H(n)\boldsymbol{\pi}(n)},$$
$$\xi(n) = d(n) - \hat{\mathbf{w}}^H(n-1)\mathbf{u}(n),$$
$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) + \mathbf{k}(n)\xi^*(n),$$

and

$$\mathbf{P}(n) = \lambda^{-1}\mathbf{P}(n-1) - \lambda^{-1}\mathbf{k}(n)\mathbf{u}^H(n)\mathbf{P}(n-1).$$

IV. SPARSE ADAPTIVE FILTER

The sparse system i.e., a small percentage of the impulse response components has a significant magnitude while the rest are zero or small. The sparseness of an acoustic impulse response is more problematic because it depends on many factors. However, acoustic echo paths are in general less sparse as compared to their network counterparts, but their sparseness can also be exploited.

Affine Projection Algorithm (APA):

Affine Projection algorithm (APA) [12] was derived as a generalization of the NLMS algorithm. In APA, the projections are made in multiple dimensions where as one dimensional in NLMS, in the sense that each tap weight vector update of the NLMS is viewed as a one dimensional affine projection, while in the APA the projections are made in multiple dimensions. As increasing the projection dimension, increases the convergence rate of the tap weight vector. However, it leads to an increased computational complexity. The equations that define the classical APA [8] [12] are

$$e(n) = d(n) - X^T(n)\ \hat{h}(n-1) \quad (4)$$

$$\hat{h}(n) = \hat{h}(n-1) + \alpha X(n)\,[\delta I_p + X^T(n)\,X(n)]^{-1} e(n), \quad (5)$$

Where $\hat{h}(n)$ is the coefficients vector and $d(n)=[d(n)\ d(n-1)\ldots\ d(n-P+1)]^T$ is a vector containing the most recent P samples of the reference signal, the matrix

$$X(n)=[x(n)\ x(n-1)\ldots x(n-P+1)] \quad (6)$$

Where ‘P’ is the Projection order

V. EXPERIMENTAL RESULTS

LMS and NLMS algorithms: The step size plays the key role which determines the amount of correction applied, as the filter adapts from one iteration to the next iteration.

- If the adaptive filter having the step size is too small that raises the filter time to converge on a set of coefficients, which affects the speed and accuracy of the filter.
- If it is too large that, it may campaign the adapting filter to diverge, and never reaching convergence, which results the filter might be unstable.

Keeping this in mind, experimentation was done and it was found from that the results are highly depending on the step size value. By testing various step values with the different datasets, it clears that to cope with the characteristics of the unknown to adapt; the smaller step sizes improve the accuracy of the convergence of the filter. It is also observed that a faster response attains for larger step size, but if it is too large, the result is not satisfactory [7]

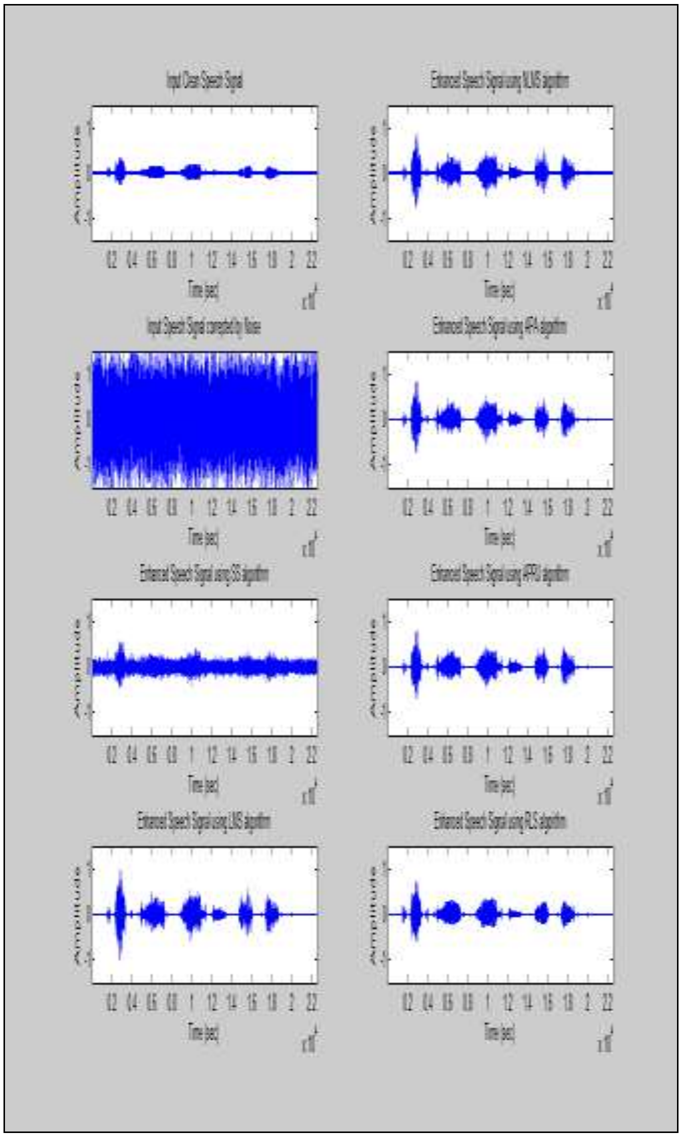

Figure 3. Enhanced speech signal for white noise at 5dB SNR using SS, LMS, NLMS, APA, RLS algorithms respectively.

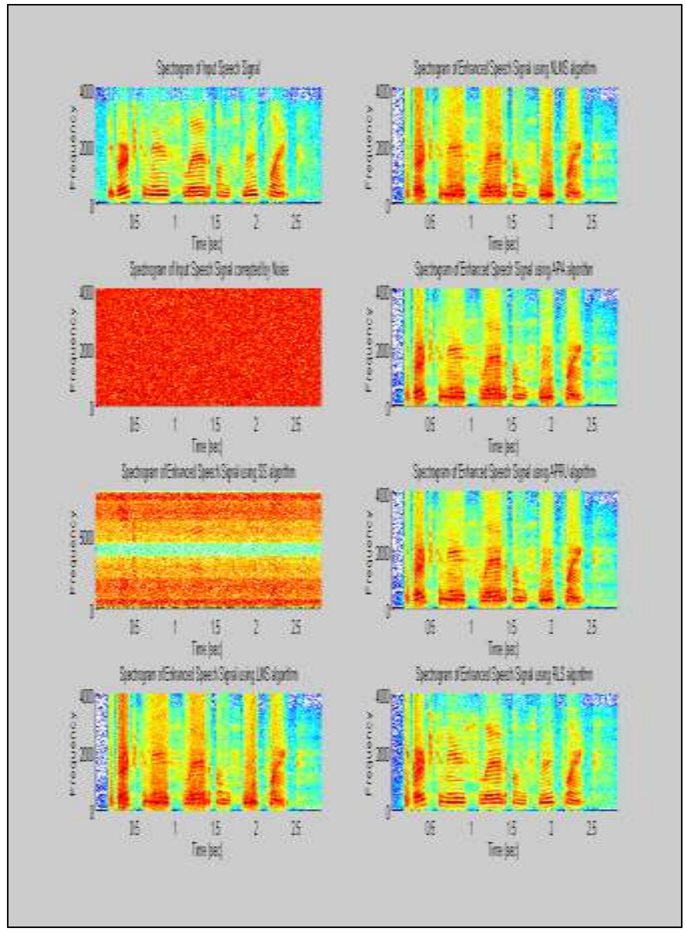

Figure 4. Spectrograms for white noise at 5dB SNR using LMS, NLMS, APA, RLS algorithms respectively.

The figures 3 and 4 shows the enhanced speech signals and their spectrograms for white noise at 5dB using SS, LMS, NLMS, APA, APRU and RLS respectively. The separate noise corpus from NOIZEUS [10] were collected and added to the clean Speech signals for the experimentation. At different noisy levels, performances of these evaluated for speech enhancement. Restaurant noise, Car noise and White noise at 0, 5, 10, and 15 dB SNR were experimented. A total of 12 datasets were generated for this research work.

VI. Performance Evaluation

The main objective of the adaptive filters is the error signal e(n) minimization. Its success will clearly depends on the length of the adaptive filter, the nature of the input signals, and the adaptive algorithm used. The signal is perceived by listeners reflects the subjective measure of quality of speech signals. At 0 dB the two signals are of equal strength and positive values are usually connected with better intelligibility where as negative values are connected with loss of intelligibility due to masking. Positive and higher SNR values are found in all the algorithms. The performances are measured based on the metrics namely MSE and SNR for all the algorithms.

Mean Squared Error (MSE):

MSE is defined as 'mean of error squares' and is calculated using the formula

$$MSE = \frac{\Sigma\left(y_i - \hat{y}_i\right)}{n-p} \qquad (7)$$

In order to quantify the difference between values implied and the true being estimated, the MSE of an estimator is used.

Signal-to-Noise Ratio (SNR):

SNR is defined as the ratio of power between the signal and the unwanted noise. SNR is calculated using the formula

$$\frac{S}{N} = \frac{n_{signal}}{n_{noise}} \qquad (8)$$

One of the most important goals of any speech enhancement technique is to achieve highest possible SNR. Higher the SNR ratios, better the performance of speech signal enhancement.

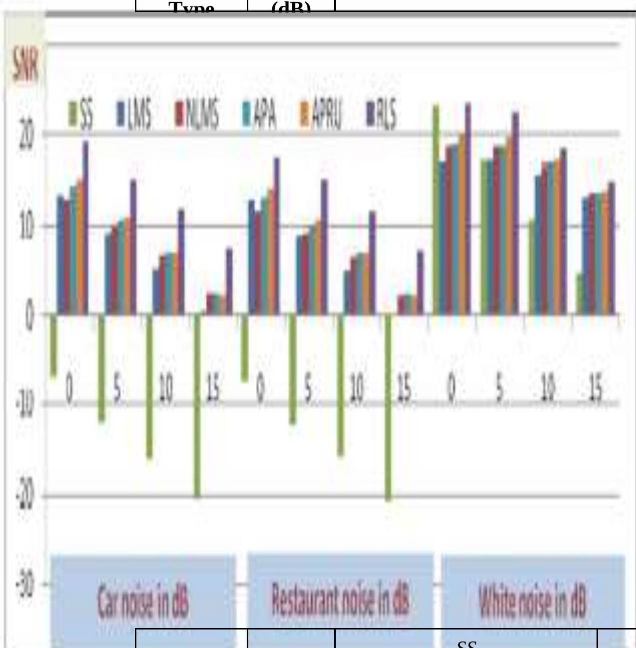

Figure 5. SNR comparison for SS, LMS, NLMS, APA, APRU and RLS algorithms

The figure 5 shows the SNR comparison of different values for SS,LMS, NLMS,APA,APRU and RLS algorithms.

The table 1 gives the MSE and SNR values for the algorithms for 0,5, 10,15 dB noise values of Car, Restaurant and White noise.

TABLE 1: MSE AND SNR COMPARISON FOR SS, LMS, NLMS, APA, APRU AND RLS ALGORITHMS

Noise Type	**SNR (dB)**	**Enhancement Method**					
		SS		*LMS*		*NLMS*	
		MSE	*SNR*	*MSE*	*SNR*	*MSE*	*SNR*
Car	0	0.0051	-6.7604	-0.0013	13.319	-0.001	12.9452
	5	0.0068	-12.2282	-0.0004	9.216	-4E-04	9.97055
	10	0.0054	-16.0588	-9E-05	5.0226	-1E-04	6.54745
	15	0.0047	-20.3911	-3E-06	0.2962	-2E-05	2.22815
Restaurant	0	0.0059	-7.26449	-0.0013	12.651	-0.001 11	0.0059
	5	0.007	-12.3414	-0.0004	8.9063	-4E-04	9.07674

	10	0.0049	-15.6892	-9E-05	4.87	-1E-04	6.24954
	15	0.0049	-20.604	-2E-06	0.2401	-2E-05	2.01625
White	0	-0.999	23.21583	-0.9782	17.079	-0.984	18.7137
	5	-0.318	17.235	-0.3142	17.22	-0.316	18.7234
	10	-0.095	10.76188	-0.0975	15.491	-0.098	16.5933
	15	-0.024	4.554904	-0.0308	13.029	-0.031	13.6224

Noise Type	**SNR (dB)**	**Enhancement Method**					
		APA		*APRU*		*RLS*	
		MSE	*SNR*	*MSE*	*SNR*	*MSE*	*SNR*
Car	0	-0.00132	14.399	-0.001	15.047	-0.0013	19.106
	5	-0.00039	10.644	-4E-04	10.898	-0.0004	15.021
	10	-0.00011	6.7968	-1E-04	6.79	-0.0001	11.789
	15	-1.8E-05	2.3275	-2E-05	2.1748	-3E-05	7.1073
Restaurant	0	-0.0013	13.025	-0.001	14.122	-0.0013	17.535
	5	-0.00039	10.062	-4E-04	10.596	-0.0004	15.138
	10	-0.00011	6.7331	-1E-04	6.8106	-0.0001	11.735
	15	-1.7E-05	2.2125	-2E-05	2.1144	-3E-05	7.0419
White	0	-0.98483	18.893	-0.988	20.146	-0.9932	23.481
	5	-0.31605	18.81	-0.317	19.827	-0.3184	22.344
	10	-0.0982	16.728	-0.098	17.357	-0.0989	18.451
	15	-0.03096	13.573	-0.031	13.9	-0.0313	14.752

VII. CONCLUSION:

The speech communication system performs greatly when the input signal has no limits or no noise effects and is degraded when there is a fairly large level of noise input signal. In such cases system cannot meet speech intelligibility, speech quality, or recognition rate requirements. In this paper, among the adaptive filters such as SS, LMS, NLMS, APA, APRU and RLS algorithms which are used for speech enhancement, simple and effective to implement is LMS algorithm but it is slower one.

Despite the fact that, with increased step size, the rate of convergence obtained in NLMS and affine Projection is not at acceptable level and even in. But the Affine Projection. The experimental results had shown that when compared to the other algorithms the RLS algorithm provides better noise reduction at faster converging speed, improved speech quality and intelligibility respectively. As a result, RLS adaptive algorithm has more SNR as well efficient noise reduction than the other classical adaptive filters and sparse adaptive filters.

REFERENCES

[1] W. Harrison, J. Lim, E. Singer, "A new application of adaptive noise cancellation," IEEE Trans. Acoustic Speech Signal Processing, vol.34, Jan 1986.

[2] B. Widrow, S. Steam, Adaptive Signal Processing. Englewood Cliffs, NJ: Prentice-Hall, 1985.

[3] J. R. Treichler, C. R. Johnson, M. G. Larimore, Theory and Design of Adaptive Filters, Wiley, 1987.

[4] S. I. A. Sugiyama, "An adaptive noise canceller with low signal distortion for speech codes" IEEE Trans. Signal Processing, vol. 47, pp. 665-674, Mar 1999.

[5] P. S. R. Diniz, Adaptive Filtering Algorithms and Practical Implementation, 2 Editions, Kluwer, 2002.

[6] Komal R. Borisagar and Dr. G. R. Kulkarni, "Simulation and Comparative Analysis of LMS and RLS Algorithms Using Real Time Speech Input Signal", Global Journal of Researches in Engineering, Page 44, Vol.10, 2010.

[7] Siva Prasad Nandyala and T. Kishore Kumar, "Speech Enhancement Using Kernel Adaptive Filtering Method", IEEE International Conference on Microwaves, Communications, Antennas and Electronics Systems, COMCAS, November 2011.

[8] K.Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," Electron. Commun. Jpn., vol. 67-A, pp. 19–27, May 1984.

[9] A. Srinivasan, "Adaptive Echo Noise Elimination for Speech Enhancement of Tamil letter 'Zha'", International Journal of Engineering and Technology Vol.1 (3), 91-97, 2009.

[10] http://www.utdallas.edu/~loizou/speech/noizeus/index.html

[11] Vimala.C and Radha.V, "Optimal Adaptive Filtering Technique for Tamil Speech Enhancement", International Journal of Computer Applications,Volume 41– No.17, March 2012.

[12] Constantin Paleologu, Jacob Benesty, Silviu Ciochin˘a, "Sparse Adaptive Filters for Echo Cancellation", 2010, Morgan and Claypool publishers.