Amended Adaptive Algorithm for Corpus Based Improved Speech Enhancement

P. Shanmuga Priya, S. Selva Nidhyananthan and R. Shantha Selva Kumari

Abstract—Speech enhancement objective is to improve the noisy speech signals for human perception. The intention of speech enhancement algorithm is to improve the performance of the communication, when the signal is occluded by noise. The quality and intelligibility of speech is reduced because of the presence of background noise. There are various adaptive filtering algorithms for speech enhancement. The existing least mean square and normalised least mean square algorithms have the problem of choosing the step size that guarantees the stability of the algorithm. To overcome this problem, we focus on speech enhancement by amended adaptive filtering. The proposed algorithm follows blind source separation strategy using adaptive filtering. Comparison of existing adaptive filtering algorithms with proposed algorithm justifies the amendment incorporated in this paper. Taking the objective criteria into account the algorithms has been tested for segmental signal to noise ratio (SegSNR), segmental mean square error (SegMSE), signal to noise ratio and mean square error. The proposed algorithm can be used for hand-free cell phone, hearing aids and teleconferencing systems.

Keywords—Least mean square, Normalised least mean square, Amended normalised least mean square, Blind source separation

I. INTRODUCTION

SPEECH is used for communication and it is considered as one of the vital utilities of human beings. To improve the information from the outside world or to communicate with each other the human needs three most significant sources of information .i.e., speech, images and written text. The speech processing systems which are used for the purpose of communication is more often designed for an ideal noise-free environment but practically the presence of background interference in the form of additive background and channel noise greatly degrades the performance of these systems, causing erroneous information exchange and listener fatigue. Over the period, researchers have recognized a number of approaches to enhance speech from the degraded speech. Several algorithms like spectral subtraction, wiener filtering is proposed during the decades [1,2,3]. A simple algorithm in time domain [4] estimates the noisy speech spectrum frame by frame basis. Several authors have made an effort to reduce residual noise and unpleasant musical noise, for example [5], with less computational complexity. The moving car environment through a blind source separation system involves the problem of two closely spaced microphones [6]. The frequency-domain [7] approaches are used to calculate the filters situated at the output. Based on the usage of the recursive least square algorithm [8], forward blind structure is to update the cross-filters. Feed forward and feedback structures are the types of blind source separation [9] to

Authors are with Mepco schlenk engineering college, Sivakasi, Tamilnadu. (e-mail:panchamapriya@gmail.com, nidhyan@mepcoeng.ac.in and rshantha@mepcoeng.ac.in)

resolve speech enhancement issues. A DFNLMS algorithm [10] aimed at speech quality improvement. In this paper we discussed on Amended NLMS (ANLMS) algorithm. This Paper is framed as follows: in section 2 we discuss about adaptive filters. In section 3 we discuss about the convolution mixture model and in section 4 we discuss about the full mathematical equation for the proposed (ANLMS) algorithm and then in section 5 the comparison results of the various algorithms are given. Finally the conclusion is given in section 6.

II. ADAPTIVE FILTER

Adaptive filtering is used to tune the filter coefficients to the changing noise characteristics. Often the signal or noise types are non-stationary and the numerical constraints are varying with time. Based on the received signals, the filter is designed. The adaptive filter is a filter which can design itself. LMS algorithm [11] which is a type of adaptive filter and is used to act like a desired filter by finding the coefficients of the filter.

The weight update filter of the LMS is given by,

$$w_{n+1} = w_n + \mu\, e(n) x^*(n) \qquad (1)$$

By substituting the weight update equation in the filter output [11], we get the minimum error. Selecting the step size is tedious in LMS algorithm. This is because the stability of the system is guaranteed by the step size chosen. Hence, as a remedy it is good to use the NLMS filter which portrays the problem of the normalisation. This is advantageous comparatively because, the NLMS uses a varying step-size parameter. To further increase the convergence we go for the proposed method Amended Normalised Least Square.

III. METHODOLOGY

A. Convolution Mixture Model

The Convolution mixture model is given in Fig 1. Here we consider two independent sources, one is clean speech signal $d(n)$ and other is noise Signal $x(n)$. From the above figure, we assume that the direct acoustic paths are equal to unit impulse responses.
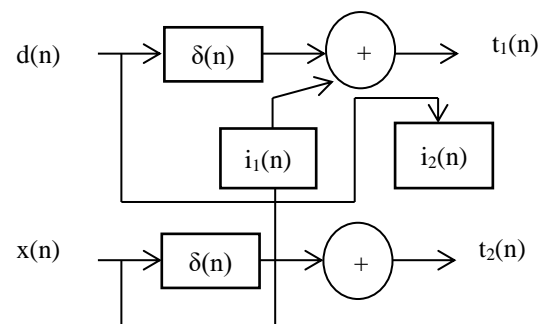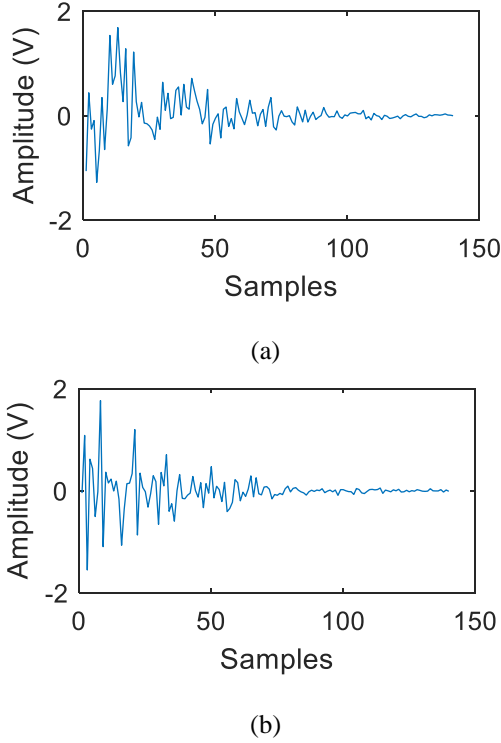

Fig.1. Convolution mixture model

(a)

(b)

Fig.2. Simulated impulse response (a) $i_{12}(n)$ and (b) $i_{21}(n)$

The two impulse responses $i_l(n)$ and $i_2(n)$ with two source signals are produced at the output of the convolution mixtures.

The output is given by the following relations:

$$t_1(n) = d(n) + i_1(n) * x(n) \tag{2}$$

$$t_2(n) = x(n) + i_2(n) * d(n) \tag{3}$$

Where, (*) represents the convolution operation.

IV. PROPOSED ALGORITHM

In this section we discuss about the mathematical formula for the amended normalised least mean square error (ANLMS). The existing algorithm has the problem of selecting the step size. To overcome this we go for the amended normalised least mean square algorithm and also to improve the performance of the speech.

The enhanced output $v_1(n)$ and $v_2(n)$ of the proposed algorithm is given as,

$$v_1(n) = d(n) - F_1^T(n)g_2(n) \tag{4}$$

$$v_2(n) = x(n) - F_2^T(n)g_1(n) \tag{5}$$

Where, $g_1(n)=[t_1(n),t_1(n-1)\ldots.,t_1(n-L+1)]^T$ and
$g_2(n)=[t_2(n),t_2(n-1)\ldots.,t_2(n-L+1)]^T$.

The update weights of the adaptive filter $F_1(n)$ and $F_2(n)$ is given by the relations,

$$F_1(n + 1) = F_1(n) - \mu_1(v_1(n)t_1(n)) \tag{6}$$

$$F_2(n + 1) = F_2(n) - \mu_2(v_2(n)t_2(n)) \tag{7}$$

Where $0<\mu_1$, $\mu_2<2$ are the two step sizes control of convergence behaviour of the two cross adaptive filter $F_1(n)$ and $F_2(n)$, respectively.

The prediction errors $m_1(n)$ and $m_2(n)$ are

$$m_1(n) = t_2(n) - t_2(n - 1) \tag{8}$$

$$m_2(n) = t_1(n) - t_1(n - 1) \tag{9}$$

By subtracting the previous values from the current value, we get a significant improvement in the quality of the speech.

The forward prediction error variances can be calculated by using the following relation:

$$u_1(n) = \beta_a u_1(n - 1) + m_1^2(n) \tag{10}$$

$$u_2(n) = \beta_a u_2(n - 1) + m_2^2(n) \tag{11}$$

The variables $k_1(n)$ and $k_2(n)$ can be calculated as:

$$k_1(n) = \frac{x_1(n)}{x_2(n)+q_a} \tag{12}$$

$$k_2(n) = \frac{x_3(n)}{x_4(n)+q_a} \tag{13}$$

Where, $x_1(n)$ and $x_2(n)$ represents the first autocorrelation mixture of $t_1(n)$ and $x_3(n)$ and $x_4(n)$ represents the first autocorrelation mixture of $t_2(n)$.

The following relations $x_1(n)$, $x_2(n)$, $x_3(n)$ and $x_4(n)$ are estimated recursively by :

$$x_1(n) = \phi_a x_1(n - 1) + t_2(n)t_2(n - 1) \tag{14}$$

$$x_2(n) = \phi_a x_2(n - 1) + t_2^2(n) \tag{15}$$

$$x_3(n) = \phi_a x_3(n - 1) + t_1(n)t_2(n - 1) \tag{16}$$

$$x_4(n) = \phi_a x_4(n - 1) + t_1^2(n) \tag{17}$$

The mixture t_1 is obtained by convolving the impulse response with noise signal and then added with the speech signal and mixture t_2 is obtained by convolving the impulse response with speech signal and then added with noise signal.

V.   ANALYSIS OF SIMULATION RESULTS

The clean speech signal and noise signal is taken from Noizeus database. The mixing signals $t_1(n)$ and $t_2(n)$ are generated by two cross coupling impulse responses $i_1(n)$ and $i_2(n)$. These impulse responses in Fig 2 are generated according to the exponential functions [12], by generating the random sequences. The clean speech is of about 4 sec, with the sampling frequency 8 KHz. Fig 3 shows the clean speech, noisy speech signal and its spectrogram, mixing samples of mixture1 and enhanced signal.
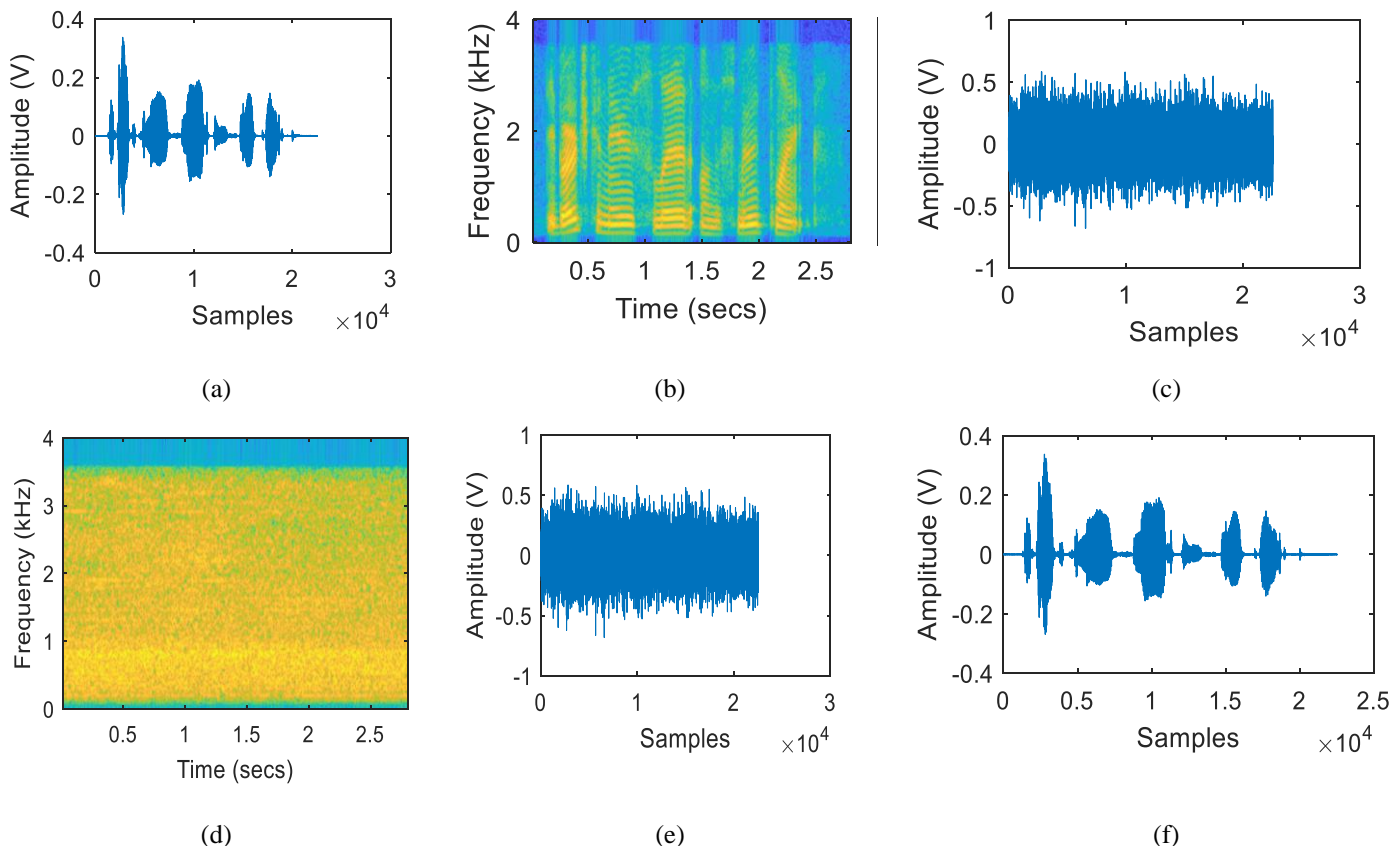
Fig. 3. (a) Clean speech signal (b) Spectrogram of speech signal (c) Noise signal (d) Spectrogram of noise signal (e) Mixing samples of mixture 1 and (f) Enhanced speech signal.

The quality of the enhanced speech signal is present in the next subsection. The comparative results of the proposed algorithm and existing algorithms are tested in terms of the performance measures of (i) segmental signal to noise ratio (ii) segmental mean square error (iii) signal to noise ratio and (iv) mean square error.

B.   A. Segmental SNR evaluation

The output segmental SegSNR of the proposed algorithm, and existing algorithms are tested under different noisy observation at 0 dB and 5 dB. The segmental SNR estimation is based on the following relation:

$$segSNR_{db} = \frac{10}{M}\sum_{m=0}^{M-1} log10\left(\frac{\sum_{n=0}^{N-1}|d(n)|^2}{\sum_{n=0}^{N-1}|d(n)-v_1|^2}\right)$$

(18)

Where, $d(n)$ and $v_1(n)$ are the clean speech and the enhanced speech signals. The parameters M and N are the number of segments and the segment length. The Segmental SNR is calculated for only in the presence of speech region. From the Fig 4, we can observe that good behaviour of ANLMS in comparison with the LMS and NLMS algorithms. The proposed algorithm performs better when compared to the existing algorithms.
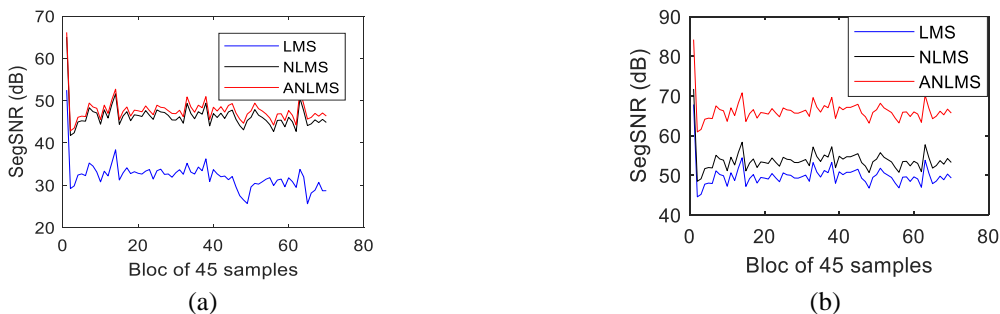
Fig..4. Shows the simulated results of segmental SNR of the proposed algorithm, LMS and NLMS for the adaptive filter of $L=$ 64 at (a) 0 dB and (b) 5 dB.

TABLE I
SIMULATION PARAMETER OF THE PROPOSED ALGORITHM

Serial No.	Proposed Algorithm Parameters
1	Step-sizes of the adaptive filters $\mu_1 = \mu_2 = 0.04$
2	Exponential forgetting factor: $\Phi_a = 0.9$
3	Positive constant: $q_a = 1$.
4	Initialization constant: $E_0 = 5$.

TABLE II
PERFORMANCE MEASURE OF SIGNAL TO NOISE RATIO

Noise type	Input SNR (dB)	Signal to noise ratio (dB)		
		Existing algorithm		Proposed algorithm
		LMS	NLMS	ANLMS
Car	0	23.49	38.52	**38.78**
	5	39.47	44.34	**51.11**
	10	47.50	49.53	**62.84**
	15	58.49	58.97	**68.07**
Street	0	22.13	36.93	**45.07**
	5	45.65	48.97	**57.35**
	10	53.12	55.61	**67.18**
	15	58.50	58.55	**73.47**
Station	0	22.60	39.80	**40.08**
	5	37.13	42.09	**52.90**
	10	48.40	50.62	**63.32**
	15	58.46	58.62	**69.76**

B. Segmental MSE (SegMSE) evalu9ation

The Segmental mean square error is given by the following relation:

$$SegMSE_{dB} = \frac{10}{M}\sum_{m=0}^{M-1} log10\left(\frac{1}{N}\sum_{n=0}^{N-1}|d(n) - v_1(n)|^2\right) \quad (19)$$

Where, N is the average frame length and M is the number of segments. From the Fig 6, we see the faster performance of the ANLMS with the LMS and NLMS respectively. The simulation parameters are given in Table I. Segmental MSE is calculated only in the absence of speech region.

TABLE III
PERFORMANCE MEASURE OF MEAN SQUARE ERROR

| Noise type	Input SNR (dB)	Mean square error (dB)		
		Existing algorithm		Proposed algorithm
		LMS	NLMS	ANLMS
Car	0	-69.08	-76.39	**-84.11**
	5	-85.06	-89.92	**-96.69**
	10	-93.09	-95.11	**-108.43**
	15	-104.08	-104.56	**-113.66**
Street	0	-67.71	-82.52	**-90.66**
	5	-91.24	-94.56	**-102.94**
	10	-98.71	-101.20	**-112.77**
	15	-104.09	-104.14	**-119.06**
Station	0	-68.19	-85.67	**-85.39**
	5	-82.72	-87.67	**-98.49**
	10	-93.98	-96.21	**-108.91**
	15	-104.05	-104.21	**-115.35**
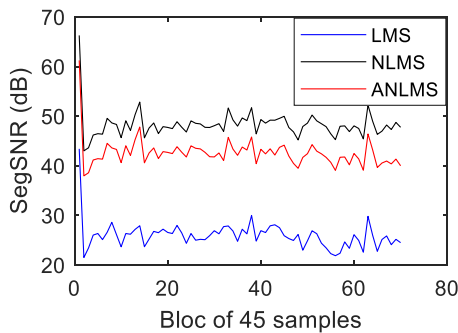
(a)
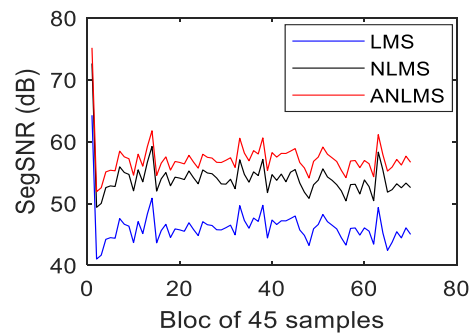
(b)

Fig.5. Shows the simulated results of segmental SNR of the proposed algorithm, LMS and NLMS for the adaptive filter of L= 128 at (a) 0 dB and (b) 5 dB.
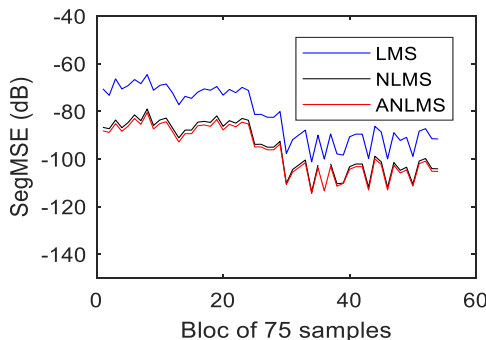
(a)
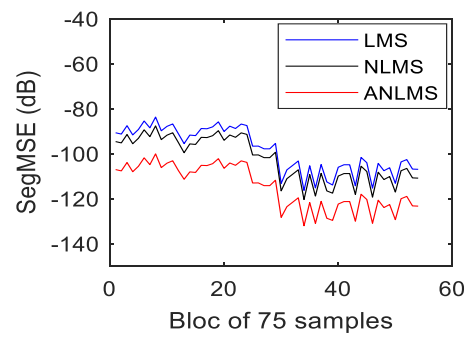
(b)

Fig.6. Shows the simulated results of segmental MSE of the proposed algorithm, LMS and NLMS for the adaptive filter of L= 64 at (a) 0 dB and (b) 5 dB.
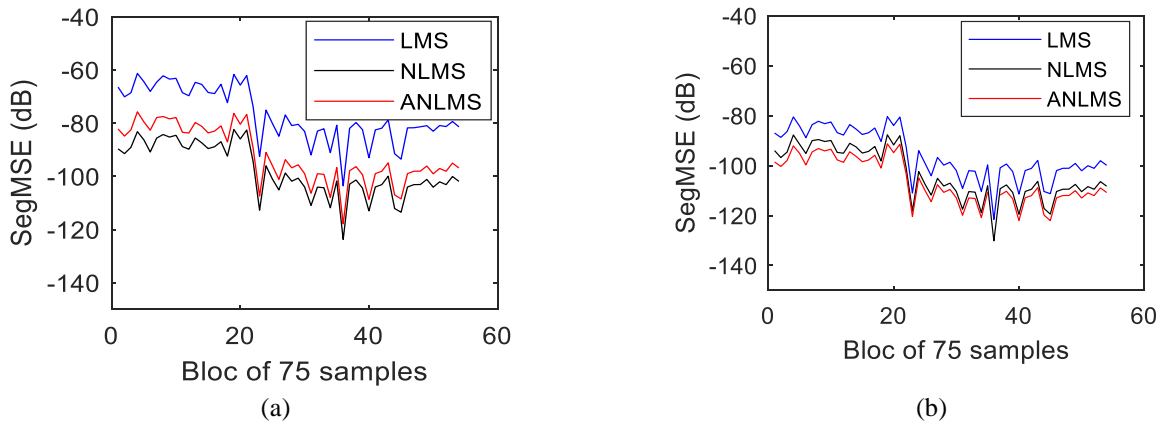
Fig.7. Shows the simulated results of segmental MSE of the proposed algorithm, LMS and NLMS for the adaptive filter of *L*= 128 at (a) 0 dB and (b) 5 dB.

In order to perform comparison evaluation, the speech enhancement algorithm is subjected to all noise types and the performance is measured in terms of SNR, MSE. Signal to noise ratio is calculated by using (18) and mean square error is calculated by using (19). But the difference is SNR and MSE is calculated for total number of samples. Table II and III, represents the performance measure of signal to noise ratio and mean square error. The SNR achieved is better in ANLMS algorithm when compared to LMS and NLMS which is shown in the Table II. The MSE achieved is better in ANLMS algorithm when compared to LMS and NLMS which is shown in the Table III. We observe that the ANLMS performs better when compared with the other algorithms in terms of both signal to noise ratio and mean square error. Fig.5-7 portray the performance measure for the technique segmental SNR and segmental MSE for dispersive impulse responses at length $L = 64$, and 128. The impulse responses at 0 dB, and 5 dB are analysed for the bloc of 45 samples for segmental SNR and the bloc of 75 samples are analysed for segmental MSE. Among the existing and proposed algorithms, the ANLMS perform better when compared with the existing algorithm.

VI. CONCLUSION

In this paper, we have shown that our proposed ANLMS algorithm performs better than the existing algorithms. The existing algorithms have the problem of choosing the step size exhaustively. To overcome this problem ANLMS algorithm is designed, to select the step size within the range of 0 to 2 for updating the weight. By comparing the existing algorithm, the proposed ANLMS algorithm has shown superiority under different noisy conditions in terms of the performance measures.

VII. REFERENCES

[1] Loizou PC. Speech enhancement theory and practice.1ˢᵗ edition. CRC Press; 2013.

[2] Cohen I, Gannot S. Spectral enhancement methods, Springer handbook of speech processing.1ˢᵗ edition. Springer; 2008.

[3] Djendi M, Gilloire A, Scalart P. Noise cancellation using two closely spaced microphones experimental study with a specific model and two adaptive algorithms. In IEEE International conference. ICASSP, 2006 May; vol. 3; pp. 744–8.

[4] Steven F. Boll. Suppression of Acoustic Noise in Speech using Spectral Subtraction. IEEE Transaction on Accoustics, Speech and Signal Processing 1979 April; vol. 27; pp.873-902.

[5] Chin Tuan Tan et.al. A parametric Formulation of the Generalised Spectral Subtraction Method. IEEE Transaction on Speech and Audio processing, 1998 July; vol. 6; pp. 328-337.

[6] Suhadi, Carsten Last, and Tim Fingscheidt. A Data-Driven Approach to A priori SNR estimation. IEEE Transactions on Audio, Speech, and Language Process 2011 January; vol. 19; pp. 186-195.

[7] Djendi M, Gilloire A, Scalart P. New frequency domain post-filters for noise cancellation using two closely spaced microphones. In Proc EUSIPCO, Poznan, 2007 Sep; vol. 1; pp. 218–21.

[8] Djendi M, Henni R, Sayoud A. A new dual forward BSS based RLS algorithm for speech enhancement. In International conference on engineering and MIS, ICEMIS , 2016. May; vol.5; pp.302-309.

[9] Henni Rahima, Djendi Mohamed, Mustapha Djebari. A Dual Backward Adaptive Algorithm for Speech Enhancement and Acoustic Noise Reduction. Association for Computing Machinery, 2017 June; vol. 978; pp.4503- 6392.

[10] Akila Sayoud, Mohamed Djendi, Soumia Medahi, Abderrezak Guessoum. A dual fast NLMS adaptive filtering algorithm for blind speech quality Enhancement. Applied Acoustics, 2018 June; vol.135, pp.101–110.

[11] Jyoti Dhiman, Shadab Ahmad, Kuldeep Gulia. Comparison between Adaptive filter Algorithms. International Journal of Science, Engineering and Technology 2013 May; Research vol. 2; pp.503-515.

[12] Djendi M, Gilloire A, Scalart P. Noise cancellation using two closely space Microphones experimental study with a specific model and two adaptive algorithms. In IEEE International Conference. ICASSP, 2006 May; vol. 3; pp. 744–8.