GSC Adaptive Beamforming Using Fast NLMS Algorithm for Speech Enhancement

S. Siva Priyanka
Research Scholar, Department of ECE
National Institute of Technology Warangal
Warangal, India
e-mail: sivapriyanka703@gmail.com

T. Kishore Kumar
Professor, Department of ECE
National Institute of Technology Warangal
Warangal, India
e-mail: kishoret@nitw.ac.in

Abstract—**This paper presents speech enhancement using Generalized Side lobe Canceller (GSC) beamforming with Fast Normalized Least Mean Square (FNLMS) adaptive algorithm under various real-time noisy conditions. GSC beamforming provides spatial filtering to reduce the noise in the desired signal. In this paper, fast convergence and low complexity, FNLMS adaptive algorithm is implemented and is applied to GSC structure to have enhanced speech at the output. The performance of the proposed GSC-FNLMS is compared with existing GSC-Least Mean Square (LMS) and Normalized LMS (NLMS) algorithms under various noisy conditions. GSC-FNLMS is evaluated using the performance metrics like Signal to Noise Ratio (SNR), Perceptual Evaluation of Speech Quality (PESQ) and Log Likelihood Ratio (LLR). An enhanced speech with improved performance is achieved by GSC-FNLMS compared to GSC-LMS, and GSC-NLMS algorithms.**

Keywords—Speech Enhancement, GSC Beamforming, LMS, NLMS, and FNLMS

I. Introduction

In many speech communication applications like mobile phone, hearing aids, teleconference, etc., speech enhancement plays a crucial role. The main aim of speech enhancement is to improve the quality and intelligibility of degraded speech under various noisy conditions in real-time environments. Many speech enhancement techniques are introduced from the past few decades. Microphone array based speech enhancement methods improves the quality as well as the intelligibility of corrupted speech when compared to single channel speech enhancement techniques [1].

Single channel speech enhancement techniques fail to find the direction of arrival and are unsuccessful in improving desired signal strength. To overcome this drawback spatial information is required which is addressed by a technique called beamforming. Beamforming [2] gives the spatial information and improve the signal to noise ratio of the desired speech signal at the output. Due to the usage of the temporal and spatial character, the beamforming can also be called as spatial filtering. With the concept of beamforming and adaptive filter theory, speech enhancement techniques are developed for noisy environments, in this paper adaptive beamforming technique using fast convergence adaptive filter is implemented for a real-time noisy environment.

This paper is organized as follows. In Section II, beamforming, fixed and adaptive beamforming is described. In Section III proposed GSC beamforming using FNLMS algorithm is implemented. Simulation results are presented in Section IV and finally, the conclusion is given in Section V.

II. Beamforming

Beamforming is the classical method in the multichannel speech enhancement domain to suppress the noise. Beamformer forms a beam-like a pattern towards source direction and enhances the degraded speech from background noise. Using a microphone array, the output of each microphone is combined to have an enhanced signal. There are two types of beamforming techniques: fixed beamforming and adaptive beamforming.

A. Fixed Beamforming

Fixed beamforming is a fundamental technique to enhance the speech signal where the weight of each microphone is fixed. The types of fixed beamformer include Delay and Sum Beamformer (DSB) [3], weighted and sum beamformer, filter and sum beamformer. Among these DSB is the most commonly used fixed beamformer where the desired signal is enhanced based on the delay on the microphone and is summed finally to have enhanced signal at the output. In fixed beamformer the weights are fixed, due to this the desired speech signal is not enhanced completely using fixed beamformers. It fails in the reverberant environment to eliminate the diffuse noise. Due to this drawback, adaptive beamforming is addressed to enhance the quality of speech signal in noisy and reverberant environments [4].

B. Adaptive Beamforming

In an adaptive beamforming technique, the weights are updated using adaptive algorithms. This beamforming is also called as data dependent beamformer as the weights are dependent on previous iteration weight coefficients. GSC [5] is the most prominently used for noise reduction. Due to spatial filtering GSC beamformer suppresses the noise coming from different directions. GSC beamformer with

978-1-7281-3663-9/19/$31.00 ©2019 IEEE

LMS [6] has shown limited performance in removing the background noise, the performance of DSB and GSC beamformer with LMS is shown considering only with white and pink noise. In order to minimize the error using

GSC beamformer, the adaptive filter block of GSC structure must be maintained with fast convergence filters. The traditional LMS and NLMS adaptive filters give limited noise reduction, low convergence and high computational complexity [7]. To overcome this fast convergence adaptive filter is applied to GSC beamformer. In this paper, GSC beamforming using FNLMS[8] adaptive filter is proposed to enhance the corrupted speech under different noisy conditions. FNLMS algorithm gives fast convergence with low complexity for acoustic echo cancellation. The performance of the proposed algorithm is compared with applying traditional adaptive filters LMS and NLMS [9].

III. PROPOSED GSC ADAPTIVE BEAMFORMING

In this paper GSC with FNLMS adaptive filter is proposed under various noise conditions. Generalized Sidelobe Canceller structure is shown in Fig. 1. GSC structure has two parts. The upper part has a fixed beamformer. In this paper, DSB is considered as fixed beamformer. The lower part has a blocking matrix and FNLMS adaptive filter. Each block is explained in the below subsections.

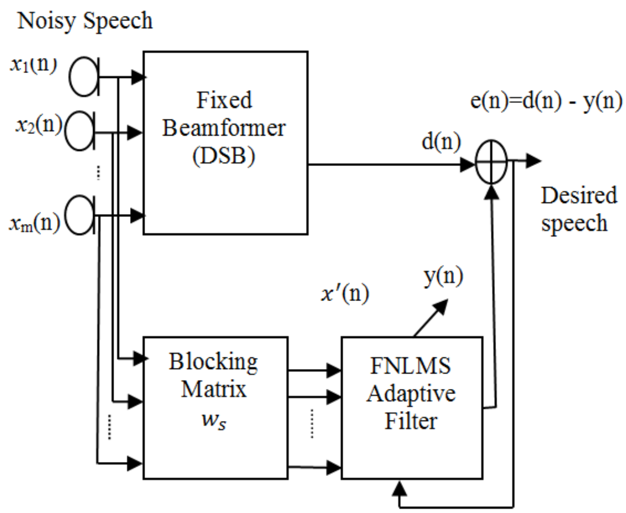

Figure. 1. Structure of GSC Beamformer

A. DSB Beamformer

DSB beamformer is a fixed beamformer used in improving SNR and finding the direction of arrival of the source signal. In the DSB structure, the microphones are placed in a linear manner by giving 'd' as the spacing between each microphone and angle 'θ' for receiving the input signal from a particular direction. $x_1(n), x_2(n) \ldots . x_m(n)$ are inputs to microphone which combines the desired speech signal with a different type of real-time noises. The delay between each microphone can be defined by angle θ of the incoming signal. The input of each microphone is delayed with angle θ and then summed to have an enhanced speech at the output of DSB which is shown in Fig.2.

$$d(n) = \frac{1}{M} \sum_{m=1}^{M} x_m(n - \tau_m) \qquad (1)$$

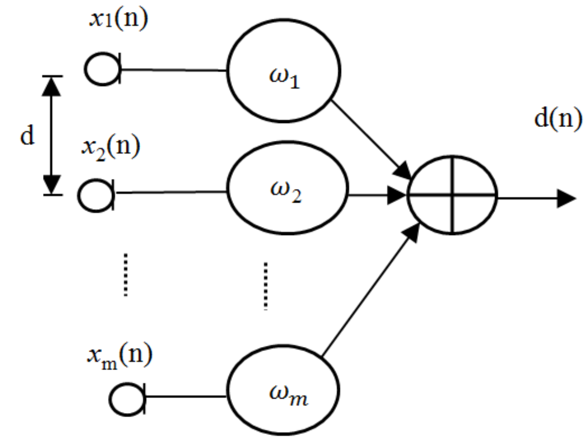

Figure. 2. Delay and Sum Beamformer (DSB)

where $d(n)$ is the DSB output, M is the number of microphones, $x_m(n)$ is an incoming signal at the mth microphone and $\tau_m$time is the delay taken from source to each microphone.

By modifying the phase weight, $\varphi m(f)$ the main lobe position in the directivity pattern will be changed which is given by

$$\varphi m(f) = 2\pi\alpha(m-1)d \qquad (2)$$

where

$$\alpha = \frac{\sin\theta}{\lambda}$$

θ is the direction of arrival of incoming signal and λ is used to determine the wavelength of frequency. The phase shift in the frequency domain in equation (2) can be analyzed as a time delay in the time domain, which is given as

$$\tau_m = \frac{\varphi m(f)}{2\pi f} \qquad (3)$$

$$\tau_m = \frac{2\pi\alpha(m-1)d}{2\pi f} \qquad (4)$$

$$\tau_m = \frac{2\pi\alpha(m-1)d \sin\theta}{2\pi f} \qquad (5)$$

$$c = f\lambda$$

$$\tau_m = \frac{2\pi\alpha(m-1)d \sin\theta}{c} \qquad (6)$$

The fixed beamformer DSB with fixed amplitude weights denoted by $w_1, w_2, ..w_M$. From each microphone the, fixed weights are summed to have enhanced speech reference which is defined by $d(n)$. The output of the DSB beamformer is given by

$$d(n) = \bar{w}^T \bar{x}(n) \tag{7}$$

where

$\bar{w}^T = [w_1, w_2, \ldots\ldots w_M]$ is fixed weights vector of DSB and $\bar{x}(n)$ is the microphone input vector.

B. Blocking Matrix

The lower part of the GSC structure has two blocks, the first block is Blocking Matrix (BM). $\overline{w_s}$ is used to block the speech signal and give output as noise reference. The blocking matrix is represented by $\begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & -1 & -1 \end{bmatrix}$. This is used to know the spatial information on the adjacent microphones. The efficiency of the BM is decided by the number of microphones. In this paper number of microphones is taken as four, so the efficiency of the blocking matrix is three. The blocking matrix output is given by

$$\overline{x'}(n) = \overline{w_s}\bar{x}(n) \tag{8}$$

$\overline{x'}(n)$, is an array of order $M-1$, the incoming microphone signal is placed in the rows of the BM, the spatial information is completely used by the GSC beamformer because of BM in its structure.

C. Adaptive Filters

An adaptive filter with robust convergence rate is essential in speech enhancement. In the lower part of GSC, the second block is an adaptive filter. In this paper, an adaptive filter plays a prominent role in reducing the error between the desired and noisy reference of a GSC. This can be achieved by using FNLMS adaptive algorithm in the adaptive filter block. The blocking matrix noisy reference is given as input to the adaptive filter, where the weights are updated to enhance the degraded speech at the GSC output. Here, the FNLMS algorithm is applied to the adaptive filter block of GSC. FNLMS algorithm gives fast convergence when compared to traditional LMS[10] and NLMS[11] adaptive algorithms

1) Fast NLMS algorithm

A fast convergence and low complexity adaptive algorithm named FNLMS [12] is presented in this section, where updating the filter coefficients depend on adaption gain and likelihood variable of the fast transversal filter. In FNLMS, the forward prediction error e(n) of fast transversal filter[13] is calculated by applying a de-correlated technique to the input signal. This is used in analyzing the dual Kalman gain. The step by step procedure of FNLMS is as follows:

Initialization:

Initialize $C_N(0)$ adaptation gain vector, $h_N(0)$ estimated filter coefficient vector, and also $\gamma_N(0)$ and $\gamma_1(0)$ likelihood variables for N samples.

$$\begin{gathered} C_N(0) = h_N(0) = 0 \\ \gamma_N(0) = 0 \\ \gamma_1(0) = 0 \end{gathered} \tag{15}$$

$$\alpha(0) = \gamma_1(0) = E_0 \tag{16}$$

where E_0 is an initialization constant and $\alpha(0)$ is the forward prediction errors variance.

Prediction error: $e(n)$

The prediction coefficient estimation can be calculated as

$$a(n) = \frac{r_1(n)}{r_0(n) + c_a} \tag{17}$$

where $r_1(n)$ and $r_0(n)$ can be estimated recursively according to

$$r_1(n) = \lambda_a r_1(n-1) + \bar{x}'(n)\bar{x}'(n-1) \tag{18}$$

$$r_0(n) = \lambda_a r_0(n-1) + \bar{x}'^2(n) \tag{19}$$

where $\bar{x}'(n)$ is input vector at time 'n', 'λ_a' is exponential forgetting factor and is 'c_a' a small positive constant.

To compute the prediction error using a first-order prediction model:

$$e(n) = \bar{x}'(n) - a(n)\bar{x}'(n-1) \tag{20}$$

The forward prediction error variance is defined as

$$\alpha(n) = \lambda\alpha(n-1) + e^2(n) \tag{21}$$

Adaption Gain is given by

$$\begin{bmatrix} \widetilde{C_N}(n) \\ c(n) \end{bmatrix} = \begin{bmatrix} -\frac{e(n)}{\lambda\alpha_N(n-1)+c_0} \\ \widetilde{C_N}(n-1) \end{bmatrix} \tag{22}$$

$\widetilde{C_N}(n)$ dual Kalman gain,

$$\delta(n) = c(n)\bar{x}'(n-N) + \frac{\bar{x}'(n)e(n)}{\lambda\alpha_N(n-1)+c_0} \tag{23}$$

$$\gamma_N(n) = \frac{\gamma_N(n-1)}{1+\gamma_N(n-1)\delta(n)} \tag{24}$$

Error for an adaptive filter is given as

$$\varepsilon_N(n) = d(n) - h_N^T(n-1)\bar{x}_N'(n) \tag{25}$$

Finally, FNLMS updating equation is defined as

$$h_N(n) = h_N(n-1) - \mu\varepsilon_N(n)\,\gamma_N(n)\widetilde{C_N}(n) \tag{26}$$

FNLMS algorithm converges faster compared to NLMS because of adaption gain. FNLMS exhibits faster convergence with low complexity compared to LMS and NLMS. The computational complexity of FNLMS is 3N multiplications whereas LMS with $2N+1$ and NLMS with $2N\hat{}2+3N$. The computational complexity of FNLMS is low compared to LMS and NLMS algorithms.

FNLMS along with existing LMS and NLMS are applied to the adaptive filter block of GSC and the error is minimized under various real-time noisy environments. The existing GSC-LMS and NLMS algorithm performance is less under real-time noisy conditions. The proposed GSC-FNLMS algorithm gives enhanced speech with a minimal error when compared to GSC-LMS and NLMS. GSC-FNLMS achieves faster convergence when compared with existing GSC-NLMS. The performance evaluation of the proposed algorithms is shown in the next section.

IV. Simulation Results

The performance of the proposed GSC-FNLMS compared with existing adaptive filters is described in this section. Considering a uniform linear array [14] of four microphones with a spacing of 0.04m between each microphone and a sound with a speed of 344m/s is generated in MATLAB. To estimate the GSC-FNLMS beamformer performance, the clean speech is degraded with Car, Restaurant, Babble, Station, Airport, Street, White noises under different SNR level like -10, -5, 0, 5, 10 dB using NOISEX-92 database [15].

The input degraded speech is given to the fixed beamformer i.e., a delay and sum beamformer where the degraded speech delay is calculated with weight coefficients and is summed to have speech reference at the output of DSB. In the next stage, BM gives a noise reference which is given as input to FNLMS adaptive algorithm in the first stage and the filter is updated to have enhanced speech at the GSC output. In the next stages, the LMS and NLMS adaptive algorithms are placed in the adaptive filter block of GSC to have performance comparison for the proposed algorithm. GSC-FNLMS gives a high quality enhanced speech at the GSC output shown in Fig. 1. To show the performance improvement of the proposed GSC beamformer over other techniques, the objective measures like Perceptual evaluation of speech quality PESQ [16] for calculating intelligibility of the enhanced speech, Signal to Noise Ratio SNR, Log Likelihood Ratio (LLR), lower the LLR higher will be the speech quality, these parameters are calculated using the following equations (27) and (28)

Signal to Noise Ratio (SNR):

$$\mathrm{SNR(dB)} = 10\log_{10}\frac{\sum_{K=0}^{N-1}x^2(k)}{\sum_{K=0}^{N-1}[\hat{x}(k)-x(k)]^2} \tag{27}$$

Log Likelihood Ratio (LLR)[16]:

$$d_{LLR}(\alpha_p,\alpha_c) = \log\left(\frac{\alpha_p R_{c\alpha_p}{}^T}{\alpha_c R_{c\alpha_c}{}^T}\right) \tag{28}$$

GSC with FNLMS, NLMS, and LMS beamformer performances are estimated by combining the speech signal with real-time noise like Car, Restaurant, Babble, Station, Airport, Street, White noises with different SNR level and their output SNR and PESQ is shown in Table 1.

TABLE I. Performance comparison of GSC using LMS, NLMS, and FNLMs

SNR in dB	**Noise Types**	**GSC with LMS**		**GSC with NLMS**		**GSC with FNLMS**	
		SNR	***PESQ***	***SNR***	***PESQ***	***SNR***	***PESQ***
-10	Car	7.0	2.198	7.8	2.369	8.4	2.444
	Restaurant	9.3	2.338	11.5	2.516	12.7	2.687
	Babble	8.5	2.239	9.3	2.412	10.1	2.574
	Station	6.3	2.494	6.8	2.753	6.9	2.76
	Airport	6.3	2.273	7.5	2.465	8.8	2.612
	Street	10.3	2.378	11.0	2.601	12.0	2.768
	White	7.8	2.037	9.0	2.135	10.8	2.35
-5	Car	11.0	2.474	12.3	2.651	12.9	2.698
	Restaurant	13.3	2.58	17.3	2.820	18.8	2.88
	Babble	12.3	2.49	13.5	2.705	14.9	2.961
	Station	11.0	2.764	11.8	3.039	12.0	3.05
	Airport	9.8	2.576	11.8	2.818	12.8	2.894
	Street	14.5	2.62	17.0	2.886	17.6	2.91
	White	12.3	2.195	13.5	2.331	15.1	2.442
0	Car	18.3	2.707	18.5	2.935	18.9	2.941
	Restaurant	20.0	2.785	21.5	3.089	22.2	3.099
	Babble	18.3	2.729	19.8	3.000	22.8	3.100
	Station	17.0	3.012	17.9	3.344	18.7	3.410
	Airport	14.0	2.872	16.5	3.132	18.6	3.436
	Street	20.8	2.822	21.5	3.160	23.2	3.360
	White	18.8	2.403	20.0	2.566	22.3	2.770
5	Car	22.5	2.972	24.0	3.255	26.6	3.350
	Restaurant	24.0	3.039	26.8	3.451	27.5	3.578
	Babble	22.8	2.972	24.5	3.301	25.0	3.421
	Station	21.0	3.221	21.8	3.691	22.4	3.700
	Airport	21.8	3.094	22.0	3.487	23.7	3.513
	Street	25.5	3.043	27.0	3.464	28.1	3.686
	White	24.0	2.659	24.8	2.835	25.0	2.886
10	Car	27.5	3.158	28.8	3.500	31.3	3.535
	Restaurant	28.8	3.26	30.8	3.711	31.7	3.740
	Babble	27.3	3.174	29.5	3.576	29.9	3.598
	Station	25.0	3.407	26.0	3.960	31.4	4.010
	Airport	26.0	3.315	6.8	3.774	27.0	3.824
	Street	29.8	3.256	1.0	3.740	32.5	3.785
	White	28.5	2.894	9.5	3.127	30.0	3.410

At 10dB input SNR with street noise, the GSC-LMS gives an output SNR of 29.8dB, whereas for GSC-NLMS and GSC-FNLMS gives an output SNR of 31.0dB and 32.5dB. Intelligibility measure PESQ at 10 dB for GSC-FNLMS is 4.010dB whereas for GSC-NLMS and GSC-LMS is 3.960dB and 3.407dB under station noise, similarly for same station noise at -5dB the PESQ for GSC-LMS, GSC-

NLMS and GSC-FNLMS are 2.764dB, 3.039dB, 3.05dB respectively. These measures show that GSC-FNLMS beamformer gives an improved performance (high output SNR and maximum PESQ) in terms of quality and intelligibility compared to GSC-LMS and GSC-NLMS beamformers. Another parameter LLR measure for the GSC-FNLMS gives lower LLR of 0.582dB at 10dB and 1.035dB at -10dB under station noise, it means the proposed GSC-FNLMS. The LLR performance of GSC- FNLMS, GSC-LMS, and GSC-NLMS is shown in Fig.3. The time domain and frequency domain plots of GSC beamformer using LMS, NLMS, and FNLMS for 5dB input SNR, corrupted with station noise are shown in Fig. 4 and Fig. 5. The proposed GSC beamformer using FNLMS adaptive filter gives better performance compared to GSC-NLMS and LMS. As the FNLMS algorithm has high convergence capacity the GSC beamformer gains high noise reduction whereas LMS and NLMS have slow convergence rate, In GSC-LMS and GSC-NLMS the performance is partially reduced. Due to GSC beamformer spatial information, the speech is enhanced under various noise conditions.

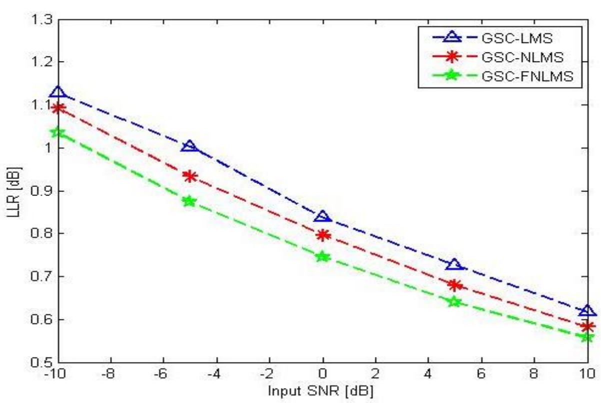

Figure. 3. Performance of Log Likelihood Ratio (LLR)

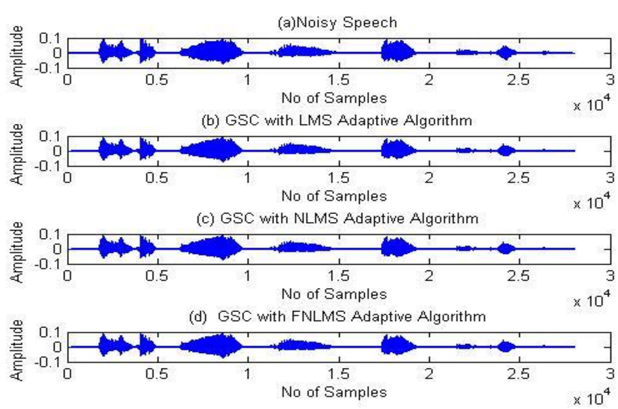

Figure. 4. Time domain representation of GSC Beamforming. (a) Received Noisy Signal at 5dB SNR, (b) Enhanced Signal of GSC-LMS, (c) Enhanced Signal of GSC-NLMS, (d) Enhanced Signal of GSC-FNLMS

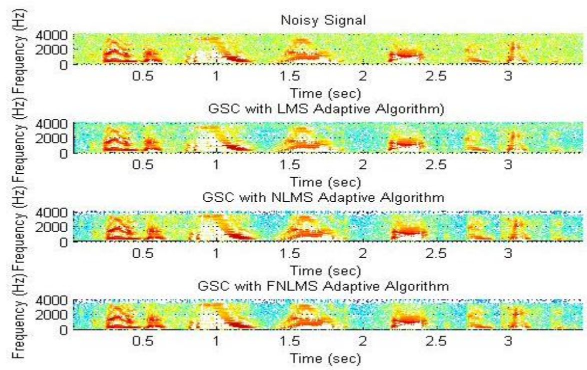

Figure. 5. Frequency domain representation of the noisy signal, GSC-LMS, GSC-NLMS, GSC-FNLMS.

V. CONCLUSION

Adaptive beamforming using FNLMS adaptive filters for speech enhancement is proposed in this paper. GSC-FNLMS beamformer gives fast convergence and low complexity when compared with existing GSC-LMS and GSC-NLMS algorithms under various noisy conditions. The quality of the speech signal for the proposed GSC-FNLMS gives superior performance compared to existing algorithms. At 10dB the output SNR for GSC beamformer using FNLMS is 33.5dB whereas GSC beamformer using LMS is 29.8dB under street noise condition. Both quality and intelligibility of speech are improved for GSC with FNLMS compared to LMS and NLMS under various noise types.

ACKNOWLEDGMENT

This work is financially supported by the Science and Engineering Research Board (SERB) New Delhi, India under EMEQ.

REFERENCES

[1] Michael Brandstein, Darren Ward, Microphone Arrays: Signal Processing Techniques and Applications. Springer 2013.

[2] Barry D.Van Veen and Kevin M. Buckley, “Beamforming: A versatile approach to spatial filtering,” IEEE Acoust. Speech, Signal Processing Magazine, vol. 5, pp. 4–24, April 1988.

[3] Israel Cohen, Jacob Benesty, Sharon Gannot, Speech Processing in Modern Communication: Challenges and Perspectives, Springer, Heidelberg (2009).

[4] S. Gannot D. Burshtein ; E. Weinstein, “Signal Enhancement Using Beamforming and Non-stationarity with applications to speech”, IEEE Transactions on Signal Processing, Volume: 49, Issue: 8, Aug 2001.

[5] John H.L. Hansen, Xianxian Zhang “Analysis of CFA-BF: Novel combined fixed/adaptive beamforming for robust speech recognition in real car environments”, ELSEVIER Speech Communication 2010.

[6] Pogula Rakesh, S Siva Priyanka, and Dr. T Kishore Kumar, Performance evaluation of beamforming techniques for speech Enhancement, Fourth International Conference on Signal Processing, Communication and Networking (ICSCN), Chennai, India 2017.

[7] Priyanka Gupta, Mukesh Patidar, Pragya Nema, Performance analysis of speech enhancement using LMS, NLMS and UNANR algorithms, International Conference on Computer, Communication and Control (IC4), Indore, India 2015.

[8] Akila Sayoud, Mohamed Djendi, Soumia Medahi, Abderrezak Guessoum, “A dual fast NLMS adaptive filtering algorithm for blind speech quality Enhancement”, Elsevier Applied Acoustics 135 2018 101–110
[9] Simon Haykin, Adaptive Filter Theory, Prentice- Hall, 2002
[10] M. Turki-Hadj Alouane and M. Jaídane-Saídane, “A new nonstationary LMS algorithm for tracking Markovian time-varying systems,” Signal Processing, Volume 86, No. 1, January 2006 pp. 50-70.
[11] A.H. Sayed, Fundamentals of Adaptive Filtering, Wiley, 2003.
[12] A. Benallal, and M. Arezki, “A fast convergence normalized least-mean-square type algorithm for adaptive filtering” Int. J. Adapt. Control Signal Process. 2013
[13] Cioffi J, Kailath, “T. Fast recursive least squares transversal filters for adaptive filtering”, IEEE Transactions on Acoustics, Speech and Signal Processing 1984; ASSP-32:304–337.
[14] Allen and D. Berkley, “Image method for efficiently simulating small-room acoustics,” *J. Acoust. Soc. Amer.*, vol. 65, pp. 943–950, Apr. 1979.
[15] http://ecs.utdallas.edu/loizou/speech/noizeus/.
[16] Y. Hu and P. Loizou, “Evaluation of objective quality measures for speech enhancement,” IEEE Trans. Audio, Speech, Lang. Process., vol. 16, no. 1, pp. 229–238, Jan. 2008.