IPEC Journal of Science & Technology, Vol. 02 (02), Dec 2023
ISSN: 2583-3286(Online)

Performance Analysis of Adaptive Speech Enhancement Algorithms Using Real Time Simulation

Arpit Sharma*, Khushi Awasthi, Shafali Jagga****
**Delhi Technological University, New Delhi, India*
***Inderprastha Engineering College, Ghaziabad, India*

***Abstract*: In this study, real-time simulation of digital filters has been done using the concepts and methods of adaptive signal processing. An adaptive filter system based on the two important algorithms, namely least mean square (LMS) and recursive least square (RLS), has been simulated for speech enhancement. In real-time, one English speech signal has been taken as test case, which is mixed with Gaussian noise. The realization of both algorithms is quantitatively and qualitatively analyzed using MATLAB Simulink and its DSP system toolbox. The power spectral density (PSD) values of the output signals, along with their transient performances, have been compared. It has been found that at 20 kHz frequency, the simulated PSD with RLS is -79.5 dBm, which is better than the simulated PSD of -75.2 dBm measured with LMS.**

***Keywords*: LMS, RLS, PSD, real-time, speech Enhancement.**

I. INTRODUCTION

Speech enhancement is an important field of digital signal processing that aims to improve the intelligibility of a speech signal by suppressing the noise components and amplifying the speech components of the signal [1]. Owing to advancements in the telecommunications industry, distance is no longer a restriction for speech or voice communication. Today, the whole world is a user of telephonic communication. To enhance the experience of the end user in communication, it is necessary to create significant algorithmic techniques for optimizing speech signals. The goal of speech enhancement is to use audio signal processing techniques to make the deteriorated speech signal more understandable and/or substantially better. The most crucial area of speech processing is noise reduction, which is utilized in a variety of applications, including speech recognition [2] as in Alexa, Siri, and ok Google, VoIP, teleconferencing [3], mobile phones, speech processing for hearing aids, and improvement in the performance of digital communication system. Therefore, this study demonstrates the use of speech processing algorithms using advanced digital signal processing (DSP) [4] methods.

Date of Submission: 19 Nov 2023
Date of Acceptance: 22 Dec 2023
Corresponding Author: Ms. Shafali Jagga
(e-mail: shafali.jagga@ipec.org.in)

MATLAB Simulink with DSP toolbox is used to carry out the signal processing on sampled input speech signals.

II. THEORY AND DESIGN METHODOLOGY

In the transmission channel, noise is directly added to the speech signal and forms a noisy speech signal. Speech enhancement techniques are required for the removal of the noise and the extraction of clean speech from the contaminated signal. This is basically a filtering step, where all the filtering intelligence and signal processing techniques need to be employed, upon which the quality of the filtered signal depends.

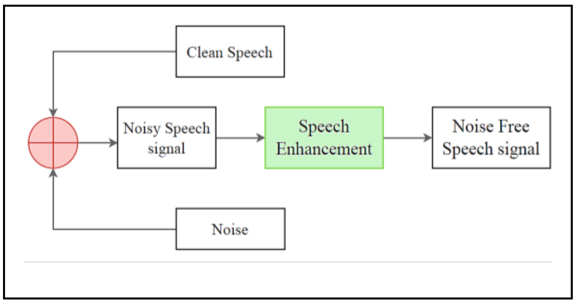

Figure. 1. Speech Enhancement Process [5,6]

ISSN: 2583-3286(Online)

Speech enhancement mainly includes the design of adaptive noise cancellation and enhancement filters using algorithms such as the spectral subtraction method [1,7], the LMS method [8,9], the RLS method [10], etc. Practical signals and systems are always easier to analyze in the frequency spectrum than in the time domain, especially in the case of signal processing, so the Fourier transform is the best technique used for the transformation from the time domain to the frequency domain. Digital filters, essentially linear and time-invariant (LTI) systems, act as frequency-selective systems; therefore, their magnitude and phase spectrum can always be plotted using the Fourier transform. Every LTI system can be articulated as a filter in the frequency domain, but digital filters that operate on signals to produce desired results suffer from the following limitations.

1. Practical signals contain time-varying noise. *Fig. 3. General Structure of Adaptive Filter [1]*
2. As noise is a random process, the static function implemented by the fixed coefficient filter would be useless as it can only operate on deterministic frequency components. Also, when noise becomes time-varying, its probability distribution and statistical properties change with time.
3. Fixed coefficient filters would cause loss of information as they cannot be used to preserve or restore the spectral structure of aliasing, where overlapping of the bands of noise and signal takes place.

To overcome these limitations, adaptive filters are used. The adaptive filter is a type of digital filter which is capable of adapting to the changes in the inputs according to the requirements. These filters are versatile in nature and widely used in many digital signal processing techniques including speech processing and especially in the application of noise removal. If the adaptive filter is characterized as a system, then the inputs to the adaptive filter are the correlated noise, x(n) and the desired signal, d(n). The output is e(n) called as error signal. This system is depicted in Fig. 2.

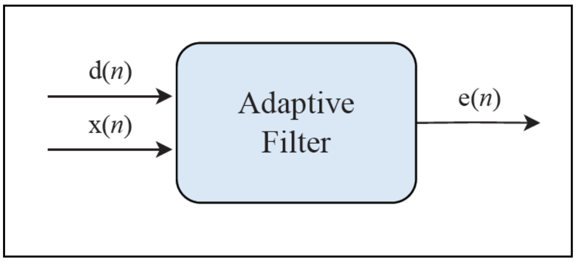

Figure. 2. Inputs and output of adaptive filter system [1]

Fig. 3 describes the general configuration of the adaptive filter. It mainly involves two components:

1) Digital filter (weight adjustable)

2) Adaptive algorithm [11]

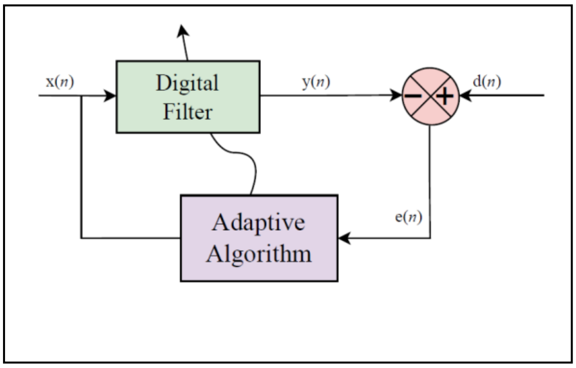

Figure. 3. General Structure of Adaptive Filter [1]

The adaptive algorithm defines how exactly the filter coefficients will be adjusted, for optimal filtering. Here, y(n) is the estimation of the noise present in d(n), e(n) is the error signal which is obtained by subtracting y(n) form d(n).

The aim of adaptive algorithm is to determine the optimal set of filter weights, also known as filter coefficients for the weight adjustable digital filter. By the virtue of optimal coefficients, the digital filter produces the nearest estimation or copy of the actual noise content present in d(n).

e(n) is called error signal not because it is an error signal, but because it is the signal whose power is to be minimized in order to achieve the noise removal. Now to understand the working principle of adaptive filters, let's consider a typical noise removal system [5] using adaptive filter as depicted in Fig. 4.

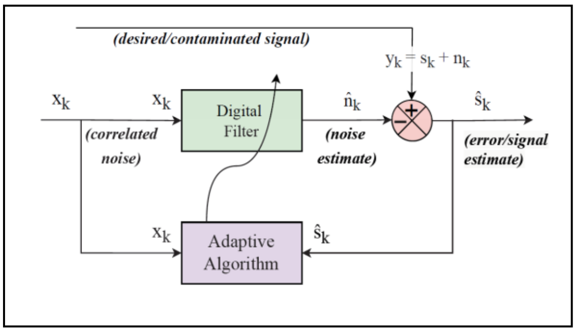

Figure. 4. Example of Denoising system using adaptive filter [11].

x_k & n_k are correlated to each other. $\hat{n}_k$ is the estimation of n_k and $\hat{s}_k$ is the estimation of s_k. The output of the filter is

$$\hat{s}_k = y_k - \hat{n}_k \tag{1}$$

$$\hat{s}_k = (s_k + n_k) - \hat{n}_k \tag{2}$$

$$\hat{s}_k = s_k + (n_k - \hat{n}_k) \tag{3}$$

IPEC Journal of Science & Technology, Vol. 02 (02), Dec 2023
ISSN: 2583-3286(Online)

To perform SNR analysis, it is needed to find the power of each signal in (3) given by the following expressions.

$$E[(\hat{s}_k)^2] = E[(s_k)^2] + E[(n_k - \hat{n}_k)^2] + 2\,E[(s_k)(n_k - \hat{n}_k)] \quad (4)$$

Since s_k & n_k are uncorrelated $\therefore E[(s_k)(n_k - \hat{n}_k)] = 0$

$$\therefore E[(\hat{s}_k)^2] = E[(s_k)^2] + E[(n_k - \hat{n}_k)^2] \quad (5)$$

Where,

- $E[(\hat{s}_k)^2]$ is the power of signal estimate ($\hat{s}_k$).
- $E[(s_k)^2]$ is the power of desired signal (s_k).
- $E[(n_k - \hat{n}_k)^2]$ is the power of the remnant noise ($n_k - \hat{n}_k$).

Since $E[(s_k)^2]$ is constant, to maximize the SNR,

$$\min\{E[(\hat{s}_k)^2]\} = E[(s_k)^2] + \min E[(n_k - \hat{n}_k)^2]\} \quad (6)$$

There are no sources in the current document. It states that to minimize the noise content from the desired signal, minimize the power of the signal estimate. Many adaptive denoising algorithms [11] based on this approach are approximations of the discrete Wiener filter [12,13].

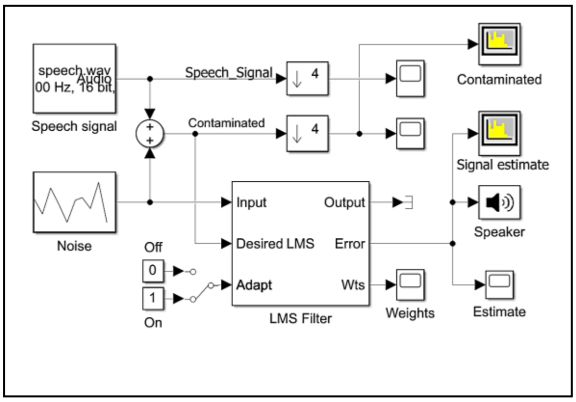

Figure. 2. LMS filter system in Simulink

Fig. 5 shows the developed adaptive filter which is a LMS [8,9] type filter which aims to reduce the error signal power to the minimum, to increase SNR. The LMS algorithm [14,15] can be derived as follows:

$$y(k)=w_1.x_1(k)+...+w_n.x_n(k) \quad (7)$$

In the vector form:

$$y(k)=x^T(k)w(k) \quad (8)$$

where y(k) is the filtered signal, w(k) is the vector of filter adaptive weights and $x^T(k)$ is the transposed input vector for the filter of size n, for the k^{th} sample.

The LMS weights adaptation is defined as:

$$w(k+1)=w(k)+\Delta w(k) \quad (9)$$

where $\Delta w(k)$ is given by

$$\Delta w(k)=\frac{1}{2}\mu\frac{\partial e^2(k)}{\partial \omega(k)}=\mu e(k)x(k) \quad (10)$$

where μ is the step size and e(k) is the error signal defined by

$$e(k)=d(k)-y(k). \quad (11)$$

The block parameters used for the LMS Filter in the system are given in the Table 1 as follows.

Table 1. LMS filter block parameters

Parameter	**Value**
Algorithm	LMS
Filter length	32
Step size (mu)	0.0008
Leakage factor	1.0
Initial value of filter weights	0

The noise specifications used are described below in Table 2.

The Speech signal used for simulation has the following specifications tabulated in Table 3.

Table 2. Noise parameters

Parameter	**Value**
Source type	Gaussian
Method	Ziggurat
Mean	0
Variance	0.1
Repeatability	Not Repeatable
Sample Mode	Discrete
Sample Time	1/44100
Samples per frame	1024

ISSN: 2583-3286(Online)

Table 3. Speech parameters

Parameter	Value
Sample rate	44100 Hz
Bit Depth	16 bits per sample
Stereo/Mono	Mono
Variance	0.1

The contaminated signal waveform, both in time domain and frequency domain can be seen in Fig. 6 and Fig. 7.

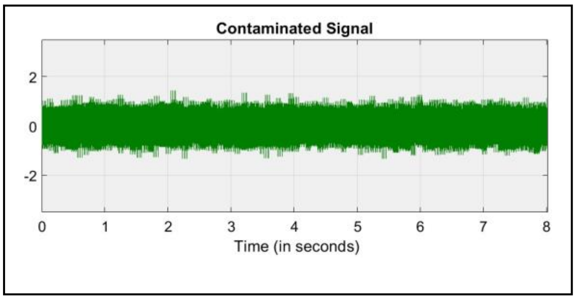

Figure. 3. Contaminated signal in the time domain.

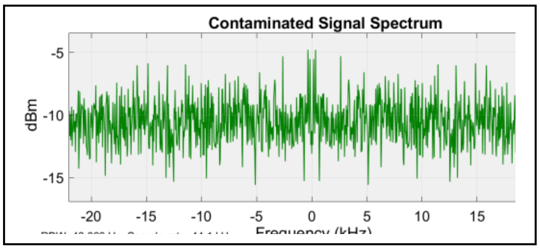

Figure. 4. PSD of the contaminated signal in the frequency domain.

The PSD of the gaussian noise, is depicted in the Fig. 8.

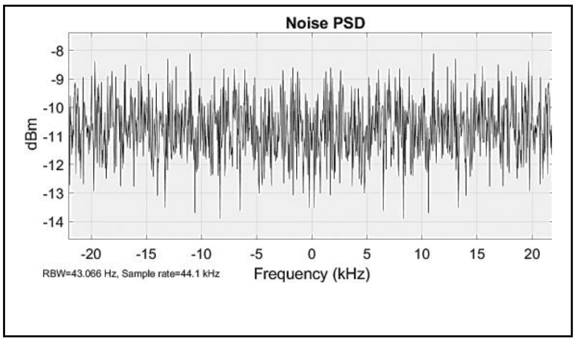

Figure. 5 PSD of the input correlated noise

It is observed that the weights of the filter start rising from zero to a specific optimal value. This change in the weights of the filter is the adaptation process of the coefficients of the filter, and the time taken by the filter coefficients to rise from zero and to reach the optimal value for filtering is called the convergence time. And the filter is said to be converged. Ideally, it should be as small as possible for better performance. This is the main disadvantages of the LMS filter, that it takes a significant time to converge, which is not good because by the time it is being converged, the information is noisy and hence information is lost in the real time. But the advantage of using this filter is that is not very mathematically complex and hardware intensive because of its low computational complexity. This makes the filter simple to implement and provides robustness and stability for different application conditions.

To take care of the slow convergence rate of the LMS filter, RLS [16,17] adaptive filter algorithm is used. The convergence rate of RLS filter is very high in comparison to LMS. But this convenience comes with high computational complexity and correspondingly difficult hardware implementation of the RLS filter.

For RLS weights adaptation $\Delta w(k)$ in (9) is expressed as [18]

$$\Delta w(k) = R(k)x(k)e(k) \qquad (12)$$

Where R(k) is the inverse of the autocorrelation matrix given by

$$R(k) = \frac{1}{\mu}\left[R(k-1) - \frac{R(k-1)x(k)x(k)^T R(k-1)}{\mu + x(k)^T R(k-1)x(k)}\right] \qquad (13)$$

The initial value of R(k) which is R(0) should be kept at

$$R(0) = \frac{1}{\delta}\mathbf{I} \qquad (14)$$

where **I** is the identity matrix and δ is a positive integer.

Fig. 9 shows the RLS filter system developed in Simulink.

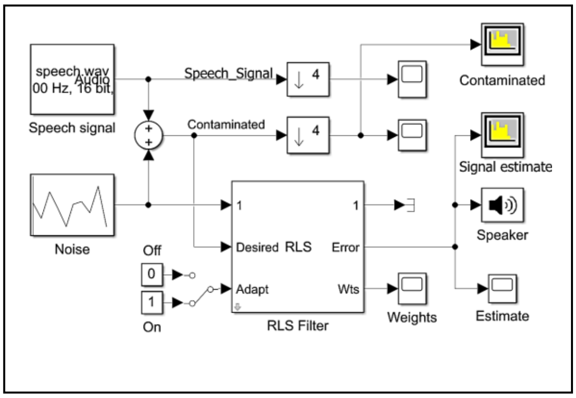

Figure. 6. RLS filter system in Simulink

The block parameters used for the RLS Filter in the system can be seen in the Table 4 as follows.

Table 4. RLS filter block parameters

Parameter	Value
Algorithm	RLS
Filter length	16
Forgetting Factor	1.0
Initial value of filter weights	0
Initial input variance estimate	0.1

The speech signal and the correlated noise signal specifications are same as mentioned in Table 2 and Table 3 respectively. Since the same inputs are given to this filter also, therefore the input waveforms are also same as shown in the Fig. 6, Fig. 7, and Fig. 8.

III. RESULT AND DISCUSSION

The LMS and RLS outputs of the signal estimate using MATLAB Simulink with DSP toolbox is shown in Fig. 10 and Fig. 11.

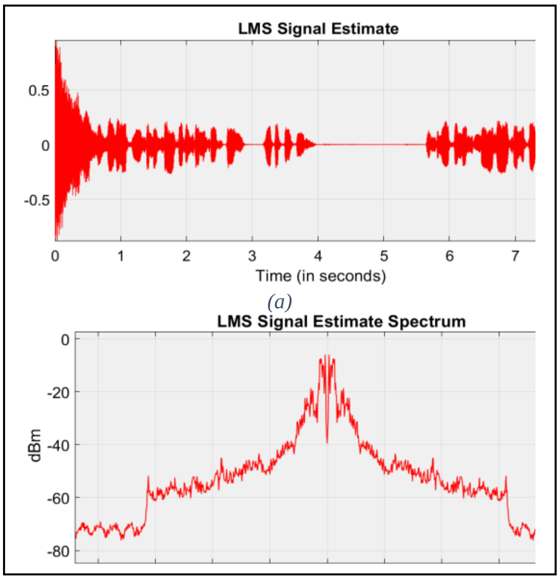

Figure. 7. (a) Signal estimate as observed in the time domain.
(b) PSD of the signal estimate produced by LMS filter.

When the filter is turned 'ON' while the simulation is running, the filter weights start adapting and adjusting in the following manner as depicted below in Fig. 11.

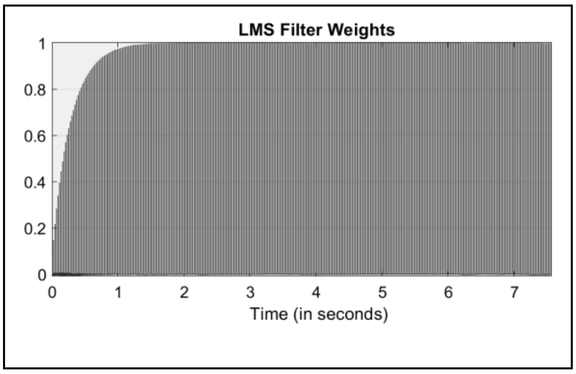

Figure. 8. LMS Filter coefficients changing as soon as the filter is turned 'ON'

It is observed that the signal is initially noisy for about 1.5 seconds because of the late convergence's influence, but after these 1.5 seconds the filter output has become clean and the signal obtained at the output is essentially a cleaned-out speech signal, which is better understandable now. The RLS output of the signal estimate is shown in Fig. 12.

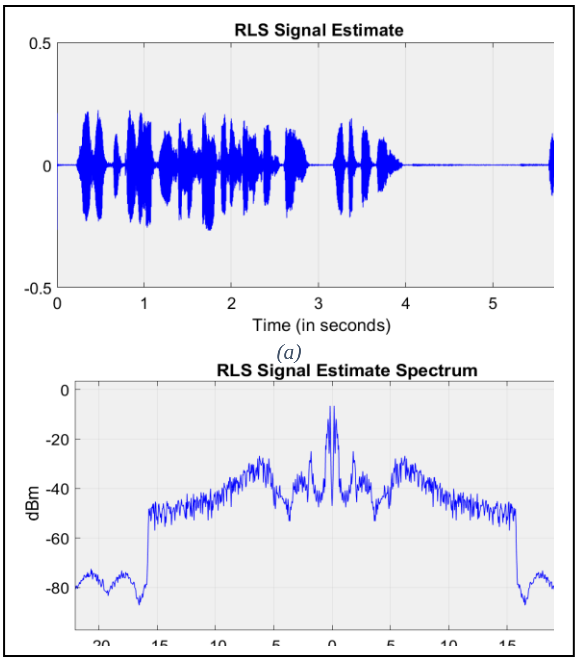

Figure. 9. (a) Signal estimate as observed in the time domain. (b) PSD of the signal estimate produced by RLS filter.

The filter weights adaptation process as soon as the simulation starts is depicted for the RLS filter in Fig. 13.

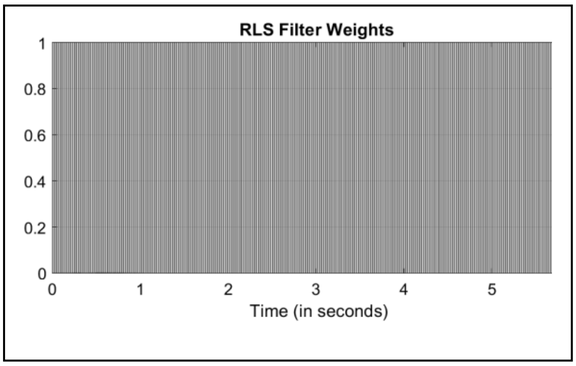

Figure. 10 RLS Filter Coefficients changing as soon as the filter is turned 'ON'

Table 5 shows the quantitative comparison of the LMS and RLS filters based on their output PSD of the signal estimate at different frequencies. This proves the higher efficiency and superiority of the RLS algorithm for the same task of filtering the same input signals.

Table 5. Comparison of PSD of output estimate signal

Frequency (kHz)	PSD of contaminated signal (dBm)	PSD of LMS signal estimate output (dBm)	PSD of RLS signal estimate output (dBm)	PSD of actual speech signal (dBm)
0	-7.4	-35.0	-45.6	-23.85
5	-10.7	-53.8	-54.8	-48.15
10	-14.0	-55.7	-56.0	-55.03
15	-8.7	-63.43	-64.0	-56.84
20	-11.4	-75.2	-79.5	-59.46

I. CONCLUSION

Speech enhancement is successfully carried out using LMS and RLS algorithms. The results for speech enhancement for English speech sample affected with Gaussian noise have been demonstrated successfully. The efficiency of the RLS algorithm is found better than the LMS algorithm when it comes to the convergence time taken by both of their corresponding filters. The RLS filter almost converged immediately as soon as the filter was turned 'ON'. Whereas the LMS filter has taken a significant amount of time to completely converge to the optimal filter weights. It can be observed that the RLS filter coefficients take little or no time to adapt to the optimal set of filter coefficients as opposed to the 1.5 seconds convergence time taken by the LMS algorithm. The main gist of the conclusion is that for the real-life cases of noises, RLS filter is a much better option to implement rather than the LMS filter due to its slow performance could result in loss of information. The flip side of RLS is its higher computationally complexity, therefore its hardware implementation needs to be optimal correspondingly to satisfy its computational needs.

REFERENCES

[1] Bahoura, Mohammed. "FPGA implementation of multi-band spectral subtraction method for speech enhancement." Midwest Symposium on Circuits and Systems (MWSCAS), pp. 1442-1445, IEEE, 2017.

[2] S. A. R . Al-Haddad, S. A. Samad, A. Hussain, K. A. Ishak and A. O. A. Noor, "Robust speech recognition using fusion techniques and adaptive filtering." American Journal of Applied Sciences 6.2, p. 290, 2009.

[3] M. Bekrani and M. Lotfizad, "A modified wavelet-domain adaptive filtering algorithm for stereophonic acoustic echo cancellation in the teleconferencing application." International Symposium on Telecommunications, pp. 548-554, IEEE, 2008.

[4] S. Carlo, T. Ogunfunmi, W. J. Kozacky and B. K. Mohanty. "FPGA implementation of LMS-based FIR adaptive filter for real time digital signal processing applications." IEEE International Conference on Digital Signal Processing (DSP), pp. 1251-1255, IEEE, 2015.

[5] V. R. Vijaykumar, P. T. Vanathi and P. Kanagasapabathy, "Modified adaptive filtering algorithm for noise cancellation in speech signals." Elektronika ir elektrotechnika 74, no. 2, pp: 17-20, 2007

[6] R. Ram, and M. N. Mohanty, "Performance analysis of adaptive algorithms for speech enhancement applications." Indian Journal of Science and Technology 9, no. 44, pp: 1-9, 2016.

[7] J. Whittington, K. Deo, T. Kleinschmidt, and M. Mason, "FPGA implementation of spectral subtraction for in-car speech enhancement and recognition." In 2008 2nd International Conference on Signal Processing and Communication Systems, pp. 1-8, IEEE, 2008.

[8] S. Dixit and D. Nagaria, "LMS adaptive filters for noise cancellation: A review." International Journal of Electrical and Computer Engineering (IJECE) 7, no. 5, pp: 2520-2529, 2017.

[9] H. Touati and K. Khaldi "Speech denoising by adaptive filter LMS in the EMD framework." In 2018 15th International Multi-Conference on Systems, Signals & Devices (SSD), pp. 1-4. IEEE, 2018.

[10] P. Rakesh and T. K. Kumar "A novel RLS based adaptive filtering method for speech enhancement." World Academy of Science, Engineering and

ISSN: 2583-3286(Online)

Technology, International Journal of Electrical, Computer, Electronics and Communication Engineering 9, no. 2, pp: 624-628, 2015.

[11] S. A. Hadei and M. lotfizad, "A family of adaptive filter algorithms in noise cancellation for speech enhancement." International Journal of Computer and Electrical Engineering, Vol. 2, No. 2, April 1106.0846, 2010.

[12] I. Almajai and B. Milner, "Visually derived wiener filters for speech enhancement." IEEE Transactions on Audio, Speech, and Language Processing, 19(6), pp.1642-1651, 2010.

[13] B. Widrow. "Thinking about thinking: the discovery of the LMS algorithm." IEEE Signal Processing Magazine, 22(1), pp.100-106, 2005

[14] J. Benesty and T. Gänsler. "New insights into the RLS algorithm." EURASIP Journal on Advances in Signal Processing, pp.1-9, 2004(3).

[15] Y. Hu and P.Loizou. "Subjective evaluation and comparison of speech enhancement algorithms." Speech Communication, 49, pp.588-601, 2007.

[16] Sudhakar, M. V., Charan, M. P., Pranai, G. N., Harika, L., & Yamini, P., "Audio signal noise cancellation with adaptive filter techniques." Materials Today: Proceedings, 80, pp: 2956-2963, 2023.

[17] Jadda, A. and Prabha, I.S., "Speech enhancement via adaptive Wiener filtering and optimized deep learning framework." International Journal of Wavelets, Multiresolution and Information Processing, 21(01), pp:2250032, 2023.