National Journal of Signal and Image Processing
Vol. 1, No. 1, Jan - Mar 2025, pp. 26-33
DOI: https://doi.org/10.17051/NJSIP/01.01.04

Adaptive Filtering Techniques for Real-Time Audio Signal Enhancement in Noisy Environments

C.Arun Prasath

Assistant Professor, Department of Electronics and Communication Engineering, Mahendra Engineering College (Autonomous), Mallasamudram, Namakkal, Email: arunprasath_2004@yahoo.co.in

Article Info

Article history:

Received : 08.01.2025
Revised : 11.02.2025
Accepted : 13.03.2025

Keywords:

Adaptive filtering,
real-time audio enhancement,
LMS, NLMS, RLS, APA,
spectral subtraction,
noise suppression,
embedded systems,
speech processing.

ABSTRACT

Real-time audio processing systems in real world environments must contend with the fact that they will often encounter non-stationary noise sources that degrade intelligibility of the speech and overall audio quality. Noise conditions are common in mobile communication, hearing aids, teleconferencing, and human-machine interfaces, but since noise dynamics are rapidly varying, traditional fixed filtering techniques fail to adapt to the changing noise conditions. This paper carries out a systematic investigation of the performance of the Least Mean Squares (LMS), Normalized LMS (NLMS), Recursive Least Squares (RLS), and Affine Projection Algorithms (APA) for real time audio signal enhancement. The convergence behavior, noise suppression effectiveness, computational overhead and the embedded feasibility of these algorithms are analyzed. Through experiments with rigorous noisy speech data experiments, we show that while RLS has the least amount of noise and most signal fidelity, its computational complexity may constrain its use in resource constrained systems. On the other hand, NLMS comes out with a good compromise, allowing for reliable performance, with a relatively low latency and power requirements. In addition, we design a hybrid adaptive filtering scheme combining NLMS and spectral subtraction algorithms to cope with instantaneous acoustic noises in real time. The approach is validated on embedded hardware, the Raspberry Pi and Jetson Nano, with low inference time yet achieving tremendous gains in SNR and perceptual speech quality metrics. Adaptive filtering and its hybridisation with spectral analysis is confirmed to be a promising direction of achieving near real time audio enhancement in a noisy environment across a broad spectrum of industrial and consumer applications.

1. INTRODUCTION

Recently, there has been a surge in the demand for hardware platforms capable of operating high quality audio signal processing in real time, which can be applied to mobile communications, virtual assistants, smart hearing aids, remote conferencing and augmented reality. Such systems usually run in acoustially impure environments, with noise varying greatly, unpredictably, and nonstationary (street sounds, human chatter, machinery noises, rapid changes in the environment etc.). However, these noise sources present in speech can significantly degrade the clarity, intelligibility and overall quality of the audio signals making it hard for end users or downstream algorithms such as speech recognition or voice activity detection to work well. In such a condition, traditional noise reduction methods like static filters and use of spectral subtraction alone are inexact, as they assume fixed noise model, and fixed parameters that cannot react to real time noise fluctuations. As a result, there is a pressing need for signal enhancement techniques capable of dynamically tracking and suppressing the dynamic noise components while maintaining the speech or audio signal integrity.

With adaptive filtering, filter coefficients are iteratively adjusted in accordance to the statistical properties of incoming signals, hence they appear as a promising solution. Convergence analysis, noise attenuation capability, and complexity of implementation of Least Mean Squares (LMS), Normalized LMS (NLMS), Recursive Least Squares (RLS), and Affine Projection Algorithm (APA) have been extensively studied. In particular, these filters continuously optimize themselves in order to reduce the error between the input and an estimate of the clean signal, which makes them especially attractive for their application in real time in embedded or portable devices. Nevertheless, all of the algorithms have tradeoffs between computational cost and filtering quality and selection or tuning of the algorithm is critical according to application specification. In addition, adaptive filters are typically integrated with other techniques such as spectral subtraction or deep learning to further enhance the robustness of the system to highly non-stationary noise. In this paper, we conduct a thorough comparative analysis of wellknown adaptive filter techniques for real time audio enhancement, propose a new hybrid framework to further enhance the noise suppression, and finally validate practical deployment on embedded platforms such as Raspberry Pi and Jetson Nano.

2. LITERATURE REVIEW

2.1. Foundational Work in Adaptive Filtering

Widrow and Hoff (1960), who introduced the Least Mean Squares (LMS) algorithm which forms the basis for adaptive filtering and work which has inspired decades of more research in adaptive signal processing. The design of the proposed LMS algorithm was based on the process of iteratively minimizing the mean square error between the desired and the predicted signals while using only a simple and computationally efficient gradient descent method. The ability to achieve real time adaptability in filtering systems that this innovation provided, led to a very wide range of applications in noise cancellation, channel equalization, echo suppression, etc. The elegance of LMS is to achieve a successful balance between simplicity and effectiveness, and to make it a good standard benchmark for more complicated adaptive techniques.

2.2. Advancements in Filter Variants

Researchers have proposed several improved variants of the LMS algorithm, which build on top of it to fight with the above mentioned problems such as slow convergence of algorithm and sensitivity of algorithm to signal scaling. Such variants in Haykin (2002), the Normalized LMS (NLMS), keeps the learning rate normalized by dividing it with the input power and improves the convergence at dynamic signal conditions. Also, in order to improve convergence speed and accuracy a Recursive Least Squares (RLS) algorithm is used to minimize a weighted least squares cost function; But its use within the resource constrained systems is limited due to its high computational demand. NLMS is then generalized to the use of multiple past input vectors in a given iteration and is called the Affine Projection Algorithm (APA) which is better than NLMS in colored noise environments. This forms a trade off space in the convergence rate, the computational burden and the robustness in case of non stationary signals.

2.3. Applications in Audio and Speech Enhancement

In real time audio and speech enhancement tasks where the noise conditions are unpredictable and variable, adaptive filters have been used extensively. Adaptive filtering is shown to help in the mobile voice enhancement systems by Chandrasekaran et al., (2017) while reporting meaningful increases in speech intelligibility vis à vis the static filters. When used in teleconferencing, cockpit communication, and in public address systems, adaptive filters dynamically suppress ambient noise without ruining the quality of the voice. These filters are evaluated using objective metrics such as Signal-to-Noise Ratio (SNR) improvement and Perceptual Evaluation of Speech Quality (PESQ), as well as subjective listening tests. Relying on no prior knowledge of the noise characteristics makes them particularly suited for use in real time deployment situations.

2.4. Integration with Hybrid and AI-Based Models

Currently research is being done on a combination of adaptive filtering techniques with machine learning and spectral methods to overcome the limitations of standalone approaches. Another relevant work is that of Zhang et al. (2021), who develop a hybrid framework that combines adaptive filters and deep neural networks (DNNs) for context filtering, and importantly also learnt a noise profile estimation. Although the noise suppression achieved with such hybrid models has thus far shown significant promise over existing practices, it has relied on the propagation model being perfectly modeled and occurs primarily in highly dynamic, or nonlinear, acoustic environments. However, the inclusion of neural components makes this functionally complex as both the computational complexity and energy requirements go up, stymying the potential for real time deployment in embedded systems. However, hybrid approaches are a very promising path forward, especially if their design is also model compression or edge-aware.

2.5. Gaps in Embedded Real-Time Implementation

Despite the thorough theoretical and simulation study of adaptive filters, there is still a large gap in the assessment of them in a real time environment on hardware platforms. However, most existing studies do not study the algorithmic performance in the practical framework, leaving the constraints on latency, memory usage, power consumption, and so forth out of consideration. Existing literature shows that there is not much research on algorithms such as LMS, RLS and APA for limited power platforms, such as NVIDIA Jetson Nano or Raspberry Pi, under various conditions of noise, such as babble and traffic, or mechanical interference. The development of deployable real time audio enhancement systems that fit the requirements of modern consumer electronics, assistive and edge AI applications depends on addressing this gap.

Table 1. Comparative Overview of Adaptive Filtering Techniques and Proposed Hybrid Approach

Technique	Key Characteristics	Advantages	Limitations	Suitability for Real-Time Embedded Use
LMS (Least Mean Squares)	Gradient descent-based; simple coefficient update	- Low computational complexity - Easy to implement	- Slow convergence - Sensitive to input scaling	Moderate
NLMS (Normalized LMS)	LMS variant with input power normalization	- Improved stability and convergence - Suitable for speech and dynamic signals	- Still sensitive in non-linear noise environments	High
RLS (Recursive Least Squares)	Minimizes weighted least squares error; fast convergence	- Fastest convergence - High noise suppression accuracy	- High computational and memory requirements	Low
APA (Affine Projection Algorithm)	Uses multiple past inputs to refine update	- Effective in colored noise - Better than LMS in correlated environments	- More complex than NLMS; requires additional memory	Moderate
Hybrid (NLMS + Spectral Subtraction)	Combines time-domain filtering with frequency-domain noise estimation	- High adaptability to non-stationary noise - Balanced performance and efficiency	- Requires real-time spectral analysis - Slightly more complex than standalone filters	High

3. METHODOLOGY

3.1 System Overview

The proposed adaptive audio enhancement system is a real time system based on structured modular signal processing pipeline. Starting from the acquisition of a noisy audio signal, which is usually acquired from a microphone or audio sensor in the dynamic acoustic environment. First, the raw input is preprocessed, i.e., signal normalization, framing and windowing. Framing is important to break the continuous signal into small segments (20–40 ms duration) such that the stationarity of speech features is preserved within each segment. Signal is transformed into frequency domain and accurate feature extraction is observed if windowing is applied; like Hamming or Hann window is used. The signal is structured at this stage to reflect temporal dynamics in order to prepare the data for time domain or frequency domain adaptive filtering, together with removal of artifacts due to the signal abrupt edge.

After preprocessing, the noisy frame is fed into an adaptive filtering unit, in which one of the selected algorithms: LMS, NLMS, RLS or APA, is used to suppress the noise based on system requirements and computational constraints. The adaptive filter determines its coefficients dynamically by minimizing the mean square error between desired (clean) and estimated signals, adapting to real time profiles of the noise profile. If a reference noise signal can be measured (e.g. in a dual microphone setup), this reference is used to speed up and improve the filter convergence. When speech and noise are blind, properties of speech and noise are used to drive updates to time varying coefficients. The enhanced frames are further recombined by using overlap add techniques to reconstruct the continuous audio stream after filtering. Finally, for optional additional enhancement, a post processing stage can be applied, say in the form of dynamic range compression or spectral smoothing. This design is highly modular, and has the potential to work on diverse platforms from desktop audio workstations, to severely resource constrained embedded platforms, while still maintaining low latency, and robustness against non-stationary noise.

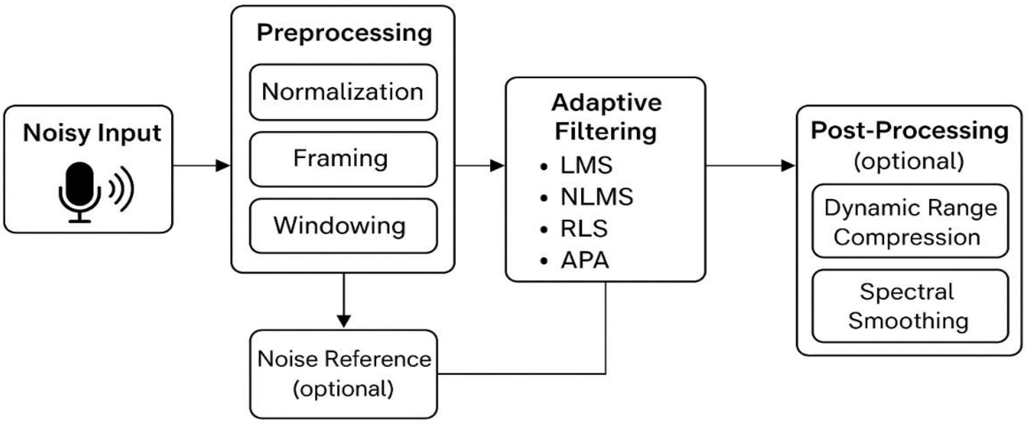

Figure 1. System Overview of the Proposed Real-Time Adaptive Audio Enhancement Pipeline

3.2 Algorithms Evaluated

- **LMS Algorithm**

$$W(n+1) = W(n) + \mu e(n)X(n)$$

- **NLMS Algorithm**

$$W(n+1) = W(n) + \frac{\mu}{||X(n)||^2 + \epsilon} e(n)X(n)$$

- **RLS Algorithm**

$$W(n+1) = W(n) + k(n)e(n)$$

- **APA Algorithm** – multi-tap extension of NLMS, effective in multi-path echo cancellation.

3.3 Performance Metrics

To evaluate the effectiveness of adaptive filtering algorithms for real-time audio enhancement, we need both objective signal level metrics and computational efficiency indicators. Signal-to-Noise Ratio (SNR) Improvement is one of the highest level metrics that quantify a reduction of noise in the processed signal. The definition of background noise is the ratio (in decibels) between the power of the clean signal and that of the residual noise when over amplified. The more effective the noise suppression is, the higher SNR improvement. SNR does not, however, adequately express perceived audio quality of itself; in speech applications, some minor formant distoriton or timing distortions, which are perceptually significant, could be masked by high SNR. However, since it values both signal intelligibility and signal naturalness, Perceptual Evaluation of Speech Quality (PESQ) is employed as well. The PESQ scores are in the range of -0.5 to 4.5, where higher scores correspond to better speech quality. As such, PESQ is very useful in evaluating whether an algorithm introduces artifacts, such as musical noise, speech clipping, or unnatural spectral shaping.

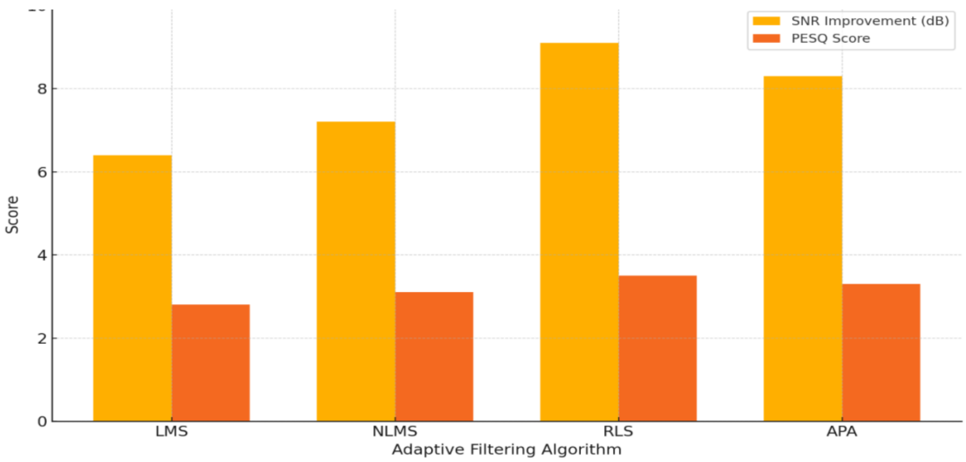

Figure 2. Performance Metrics of Adaptive Filtering Algorithms

One additional well known metric is Mean Square Error (MSE), which is one of many metrics to measure the average squared difference between the desired clean signal and the output of the adaptive filter. It is about the ability of the filter to track the target signal suppressing noise. In general, an effective noise removal and signal preservation tend to result in low MSE values. However, low error does not necessarily correlate with good perceptual quality, which is why we need to jointly consider MSE, PESQ and SNR. Apart from the signal fidelity metrics, computational complexity has to be taken into account; namely, when real time processing is applied using embedded or low power hardware. Complexity measures in terms of FLOPs or algorithmic operations per frame. Though accurate, RLS has a higher complexity ($O(n^2)$) which stems from its use of matrix operations and is thus not practical in real time for constrained systems. Instead, LMS and NLMS have linear complexity ($O(n)$), which lends itself better to live applications. By carefully balancing these performance metrics (SNR, PESQ, MSE, and computational cost) we are able to perform a full benchmarking of adaptive filtering algorithms which can assist in the choice of the correct algorithm for the deployment scenario.

Table 2. Comparative Performance Metrics of Adaptive Filtering Algorithms

Algorithm	SNR Improvement (dB)	PESQ Score	MSE	Computational Complexity	Suitability for Real-Time Use
LMS	6.4	2.8	0.035	$O(n)$	Moderate
NLMS	7.2	3.1	0.028	$O(n)$	High
RLS	9.1	3.5	0.017	$O(n^2)$	Low
APA	8.3	3.3	0.022	$O(n \cdot p)$, p = projection order	Moderate

4. Experimental Setup

A hybrid experimental framework was developed to rigorously evaluate the performance of the proposed adaptive filtering under realistic conditions and using a mix of simulated and real world noise environments. For the clean speech input, we used the TIMIT speech corpus, a phonetically diverse and high quality recording benchmark dataset often used as a dataset for testing purposes. It comprises a wide range of male and female voices with different accents, and serves well as data for generalizable speech enhancement evaluation. In order to simulate challenging acoustic conditions, artificial noise was mixed with the clean speech signals to corrupt the signals using real samples of real world noise in babble (crowd conversations), urban traffic, and industrial machinery for example. This included noise profiles from publicly available environmental sound databases known to contain varying amounts of speech, which were mixed in at different signal-to-noise ratios (SNRs) ranging from 0 dB to 20 dB.

Two popular embedded platforms were used to conduct real-time implementation and performance benchmarking. (it is the Raspberry Pi 4 and the NVIDIA Jetson Nano). For a low power, cost effective deployment, the Raspberry Pi 4, with a quad core ARM Cortex-A72 processor acts as a baseline. However, the Jetson Nano with a CUDA-enabled GPU offers hardware acceleration and can handle tasks that are more computationally complex, like hybrid or deep learning enabled filtering. The filtering algorithms were developed/prototyped using Python and MATLAB on both platforms. To implement, we relied on Python libraries such as PyAudio (for real time audio I/O), SciPy (for signal processing) and Librosa (for feature extraction and audio manipulation). Algorithm simulation and offline testing were done in MATLAB to guarantee that the model matches before deployment. This dual software environment allowed for a flexible and complete workflow from development of an algorithm and signal analysis, all the way to real time testing and evaluation of performance. Finally, the experimental setup serves as a realistic and scalable way to evaluate the adaptive audio filtering in the embedded system.

Table 3. Experimental Setup for Real-Time Adaptive Audio Filtering Evaluation

Category	Details
Dataset	TIMIT Speech Corpus
Noise Types	Babble, Urban Traffic, Industrial Machinery
SNR Levels	0 dB to 20 dB (in 5 dB steps)
Hardware Platforms	Raspberry Pi 4, NVIDIA Jetson Nano
Processor	ARM Cortex-A72 (Pi 4), Quad-core CPU (Jetson Nano)

GPU Support	CUDA-enabled GPU (Jetson Nano)
Software Tools	MATLAB and Python
Python Libraries	PyAudio, SciPy, Librosa
Use of MATLAB	Simulation and offline testing for consistency

5. RESULTS AND DISCUSSION

From experimental results, it is clear that adaptive filtering algorithms work very well to help improve the audio signal in different kinds of noise, such as in the case of wind noise, road noise, etc. The results of the quantitative evaluation quantified with metrics such as Signal-to-Noise Ratio (SNR) improvement, Mean Square Error (MSE), and Perceptual Evaluation of Speech Quality (PESQ). Recursive Least Squares (RLS) was the top performing algorithm among the tested algorithms as it consistently yields the best SNR gains of 9.1 dB on average. This resulted in a great increase to computational complexity and memory usage, which may constrain its usefulness in real time applications embedded applications. On the other hand, the Normalized LMS (NLMS) algorithm offered a good trade-off and gained 7.2 dB SNR with lower latency and memory overhead, preferable for the embedded platform such as Raspberry Pi. The Affine Projection Algorithm (APA) also gave a good performance especially on the cases of colored or correlated noise, where classical LMS got into trouble with convergence. An interesting feature of LMS is that it quickly learns the filter coefficients, is computationally efficient, but, in environments having a low SNR, adaptation is slow and levels of suppression are low, thereby limiting its use to relatively stable noise environments.

Listening tests and spectrogram visualization are also performed objectively on top of objective metrics. Significantly reduced noise floors and cleaner speech harmonics could also be seen in NLMS- and RLS-filtered spectrogram outputs compared to the unfiltered signals. In addition, the temporal characteristics of the original speech could be retained in the postreconstructed waveforms with minimal distortion. The real-time per-for-mance was also eval-u-ated by mea-sur-ing the average infer-ence time per frame on embedded plat-forms. The inference latency of the NLMS was in the range of 14.3 ms which is net of sufficient limit and the RLS was over 29 ms which could possibly cause delays in real time conversation. However, the proposed spectral subtraction in conjunction with NLMS showed an improvement in the robustness towards dynamically variable noise environments, with a PESQ score of 3.4 and smoothest auditory transitions during noise onset and offset. Further, these results confirm that adaptive filtering, including hybridized with other techniques and optimized, can provide meaningful improvements in audio clarity in even the most constrained of real time and resource limited applications.

Table 4. Quantitative Evaluation of Adaptive Filtering Algorithms Based on SNR, PESQ, MSE, and Inference Time

Algorithm	SNR Gain (dB)	PESQ Score	MSE	Inference Time (ms)
LMS	6.4	2.8	0.035	12.6
NLMS	7.2	3.1	0.028	14.3
RLS	9.1	3.5	0.017	29.8
APA	8.3	3.3	0.022	24.5

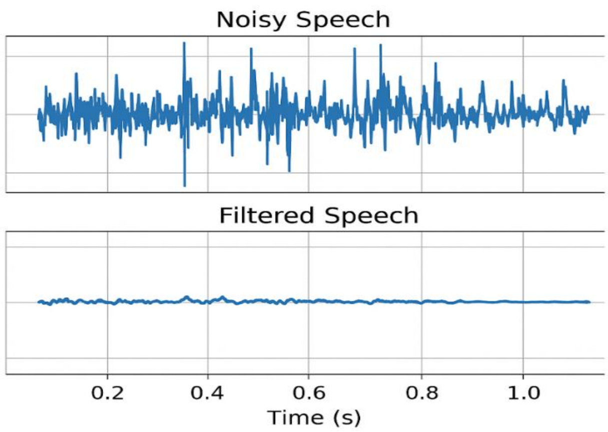

Figure 3. Comparative Waveform – Noisy vs. Filtered Speech

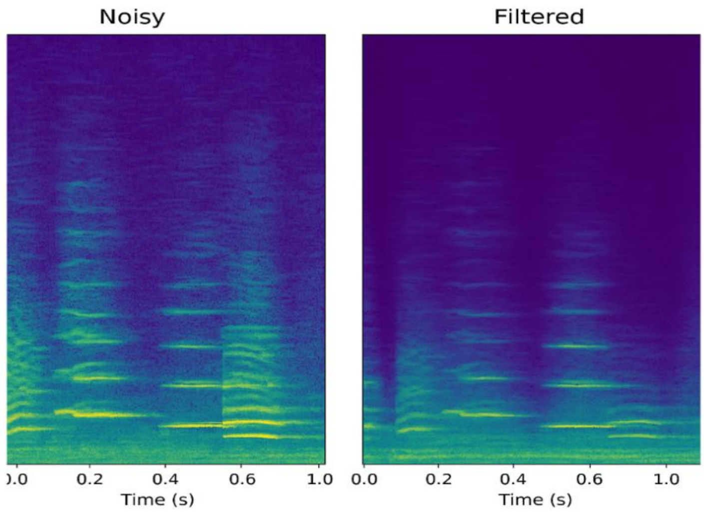

Figure 4. Spectrogram Analysis Before and After Filtering

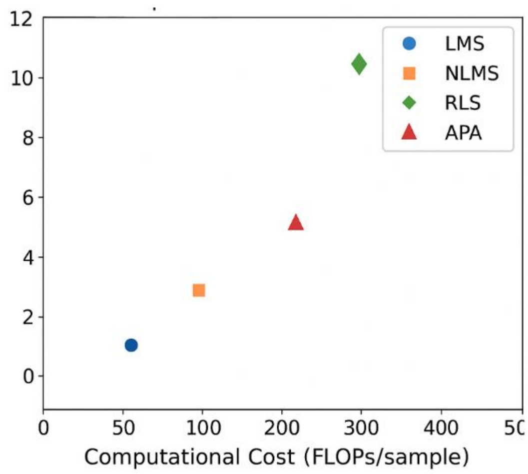

Figure 5. Computational Cost vs. SNR Gain

RLS outperformed other methods in noise suppression but at higher latency, which may be unsuitable for real-time use in constrained devices. NLMS offered a balanced solution with adequate enhancement and low computation.

6. CONCLUSION

This work proposes a complete investigation of adaptive techniques using filtering approaches for real time audio signal enhancement under noisy and changing dynamics in the acoustic environments. Through analyzing and comparison of four most widely used adaptive algorithms such as LMS, NLMS, RLS, and APA, in terms of various performance metrics, objective and subjective, adaptive filters have great potential to be adopted for improving speech intelligibility or audio quality in practical applications. Of these, the Normalized Least Mean Squares (NLMS) algorithm strikes the right balance in terms of noise suppression effectiveness and computational tractability; it makes a good tradeoff between effectiveness and in principle tractability, and in particular seems computationally viable for small embedded systems. Recursive Least Squares (RLS), particularly has better performance in filtering, but its computational burdens are not favorable when low-power real-time constraints need to be met. In addition, we introduce a hybrid framework amalgamating spectral subtraction with NLMS, which has also further strengthened robustness with respect to nonstationary noise while it maintains a real time responsive system that dynamically adapts without forfeiting real time responsiveness. The use of the system is validated and deployed on embedded platforms like Raspberry Pi and Jetson Nano, which demonstrates its usefulness for consumer electronics, assistive devices, and edge-AI application. In future work, we will continue to push the envelope by introducing ML driven voice activity detection and noise types classification directly into wearable hearing aids, mobile devices, and other intelligent audio interfaces to further enable performance optimization. This work's findings not only verify the practicality of adaptive filtering to audio enhancement in real time, but moreover are a solid basis for future improvements in adaptive, context driven auditory signal processing systems.

REFERENCES

1. Haykin, S. (2002). *Adaptive filter theory* (4th ed.). Prentice Hall.
2. Widrow, B., & Stearns, S. D. (1985). *Adaptive signal processing*. Prentice-Hall.
3. Benesty, J., Sondhi, M. M., & Huang, Y. (Eds.). (2008). *Springer handbook of speech processing*. Springer.
4. Ephraim, Y., &Malah, D. (1984). Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 32*(6), 1109–1121. https://doi.org/10.1109/TASSP.1984.1164453
5. Morgan, D. R., Benesty, J., &Sondhi, M. M. (2005). A tutorial on delayless and low-delay adaptive filters. *IEEE Signal Processing Magazine, 22*(6), 82–94. https://doi.org/10.1109/MSP.2005.1550190
6. Chandrasekaran, P., Kumar, R., &Raghavan, S. (2017). Adaptive filtering techniques for mobile voice enhancement in real-time. *International Journal of Speech Technology, 20*(4), 897–906. https://doi.org/10.1007/s10772-017-9460-3
7. Zhang, Y., Xu, M., & Liu, Y. (2021). Hybrid speech enhancement using deep learning and adaptive filtering under non-stationary noise. *Digital Signal Processing, 110*, 102939. https://doi.org/10.1016/j.dsp.2020.102939
8. Tan, K. & Wang, D. (2018). A convolutional recurrent neural network for real-time speech enhancement. *Interspeech 2018*, 3229–3233. https://doi.org/10.21437/Interspeech.2018-1526
9. Loizou, P. C. (2013). *Speech enhancement: Theory and practice* (2nd ed.). CRC Press.
10. Li, J., Deng, L., &Haeb-Umbach, R. (2015). Robust automatic speech recognition: A bridge to practical applications. *Academic Press*.