International Congress on Economics, Management and Business Studies
Hosted Online from New York, USA
Date: 23rd June , 2026
Website: https://econferencia.com

COMPARATIVE ANALYSIS OF ADAPTIVE DIGITAL FILTERING METHODS FOR SPEECH ENHANCEMENT IN NOISY ENVIRONMENTS

Isvandiyarova Madina Beknazar qizi
Doctoral Researcher (PhD) Department of Multimedia Technologies
Tashkent University of Information Technologies named after
Muhammad al-Khwarizmi

Supervisor: Saida Safibullayevna Beknazarova
E-mail: madimadi141094@gmail.com

Abstract

Speech enhancement in noisy environments remains one of the major challenges in digital signal processing and multimedia technologies. Environmental noise significantly degrades speech quality, reduces intelligibility, and negatively affects communication systems, automatic speech recognition, and intelligent audio applications. This paper presents a comprehensive comparative analysis of widely used adaptive digital filtering algorithms, including Least Mean Square (LMS), Normalized Least Mean Square (NLMS), Wiener filtering, Recursive Least Squares (RLS), and Kalman filtering. Their mathematical models, convergence characteristics, computational complexity, stability, and noise suppression capabilities are analyzed. Based on the identified limitations of existing approaches, a conceptual adaptive parameter optimization framework is proposed to improve speech enhancement performance under dynamically changing acoustic environments. The presented analysis provides a theoretical foundation for future research aimed at developing efficient adaptive filtering techniques suitable for real-time multimedia and communication systems.

Keywords: Speech Enhancement, Digital Signal Processing, Adaptive Filtering, LMS, NLMS, Wiener Filter, Kalman Filter, Noise Reduction, Audio Processing, Multimedia Technologies.

INTRODUCTION

Speech communication systems have become an essential component of modern information technologies and are widely used in telecommunications, intelligent transportation systems, virtual assistants, healthcare, multimedia applications, and human–computer interaction. However, the quality and intelligibility of speech signals are significantly degraded when they are transmitted or recorded in noisy acoustic environments. Environmental noise, industrial noise, traffic sounds, babble noise, and other interference sources reduce speech quality and negatively affect automatic speech recognition, speaker identification, hearing aids, and voice-controlled systems.

Speech enhancement has therefore become one of the most important research directions in digital signal processing (DSP). The primary objective of speech enhancement is to suppress unwanted noise while preserving the useful speech components with minimal distortion. Numerous signal processing techniques have been proposed during the past decades, including Wiener filtering, Least Mean Square (LMS), Normalized Least Mean Square (NLMS), Recursive Least Squares (RLS), Kalman filtering, spectral subtraction, and wavelet-based filtering.

Although these methods demonstrate satisfactory performance under stationary noise conditions, their effectiveness decreases considerably in non-stationary acoustic environments where the statistical characteristics of noise continuously change. This limitation creates a significant challenge for modern speech communication systems operating in real-world environments.

Consequently, improving adaptive digital filtering methods remains an актуal scientific problem. Instead of developing completely new filtering algorithms, a promising research direction is to improve the adaptability of existing algorithms by optimizing parameter selection according to varying acoustic conditions. Such an approach is expected to increase speech quality while maintaining computational efficiency suitable for real-time applications.

The objective of this study is to analyze existing adaptive digital filtering algorithms for noisy speech enhancement, identify their strengths and limitations, and determine promising directions for further improvement within speech signal processing systems.

2. Related Work

Speech enhancement has been one of the most actively studied areas of digital signal processing over the last four decades. Numerous algorithms have been proposed to suppress environmental noise while preserving the intelligibility and perceptual quality of speech signals. These methods can generally be categorized into classical digital filtering techniques, adaptive filtering algorithms, transform-domain methods, statistical estimation approaches, and recent data-driven techniques.

One of the earliest and most widely used approaches is the **Wiener filter**, which minimizes the mean square error (MSE) between the desired clean speech signal and the estimated output. The Wiener filter demonstrates high performance in stationary noise environments; however, its effectiveness decreases when noise characteristics change dynamically because accurate estimation of speech and noise power spectral densities becomes difficult.

The **Least Mean Square (LMS)** algorithm remains one of the most popular adaptive filtering techniques due to its computational simplicity and ease of implementation. LMS updates filter coefficients iteratively according to the instantaneous estimation error, allowing real-time adaptation to varying acoustic environments. Nevertheless, the convergence speed strongly depends on the selection of the step-size parameter. Improper parameter selection may lead to slow convergence or algorithm instability.

The **Normalized Least Mean Square (NLMS)** algorithm improves the LMS approach by normalizing the adaptation step with respect to the input signal power. As a result, NLMS generally exhibits faster convergence and greater stability under varying signal amplitudes. Despite these advantages, the algorithm may still experience performance degradation under rapidly changing non-stationary noise conditions.

The **Recursive Least Squares (RLS)** algorithm provides significantly faster convergence than LMS and NLMS by minimizing a weighted least squares cost function. Although RLS achieves excellent noise reduction performance, its computational complexity and memory requirements are considerably higher, limiting its applicability in embedded and real-time speech processing systems.

Another important technique is the **Kalman filter**, which estimates speech signals using state-space mathematical models. Kalman filtering has demonstrated remarkable performance for dynamic signal estimation and tracking applications. However, practical implementation requires accurate modeling of both speech and noise processes, which is often difficult under real acoustic conditions.

Transform-domain approaches such as **Spectral Subtraction**, **Short-Time Fourier Transform (STFT)**, and **Wavelet Transform** have also become widely adopted in speech enhancement systems. These techniques analyze speech in the frequency domain and suppress spectral components associated with noise. Although frequency-domain processing significantly improves speech quality, it often introduces undesirable artifacts such as musical noise, especially at low signal-to-noise ratios.

Recent advances in artificial intelligence have introduced deep neural networks, convolutional neural networks (CNNs), recurrent neural networks (RNNs), long short-term memory (LSTM) networks, Transformer architectures, and DeepFilterNet models for speech enhancement. These approaches demonstrate superior performance under complex acoustic conditions. Nevertheless, they require extensive training datasets, substantial computational resources, and high-performance hardware, making them less suitable for lightweight embedded audio processing systems.

Therefore, despite remarkable progress in speech enhancement research, there remains a significant need for improving adaptive digital filtering algorithms capable of maintaining high speech quality under dynamically changing acoustic environments while preserving low computational complexity.

Research Gap

A comprehensive review of existing adaptive speech enhancement techniques indicates that each filtering algorithm performs effectively only under specific acoustic conditions. Traditional adaptive filters such as LMS, NLMS, Wiener, Kalman, and RLS employ fixed adaptation mechanisms that cannot automatically adjust to rapidly varying environmental noise characteristics. Consequently, speech quality deteriorates when operating under highly non-stationary acoustic conditions.

Current research increasingly focuses on integrating intelligent parameter adaptation strategies into classical adaptive filtering algorithms rather than replacing them entirely. However, existing studies rarely investigate a unified adaptive framework capable of dynamically selecting or tuning filter parameters according to continuously changing acoustic environments while maintaining computational efficiency suitable for real-time implementation.

Therefore, the present research is directed toward improving adaptive digital filtering algorithms through dynamic parameter optimization for noisy speech enhancement, aiming to increase speech intelligibility, reduce residual noise, and preserve low computational complexity for practical multimedia and communication applications.

3. Materials and Methods

3.1 Speech Signal Model

A speech signal recorded in a noisy acoustic environment can generally be represented as the sum of the clean speech signal and environmental noise. The observed noisy speech signal is expressed as

$$y(n)=s(n)+v(n)$$

where

- $y(n)$– observed noisy speech signal;
- $s(n)$) – original clean speech signal;
- $v(n)$– additive environmental noise.

The primary objective of speech enhancement is to estimate the clean speech signal s^(n)\hat{s}(n)s^(n) from the observed noisy signal while minimizing speech distortion.

where

- d(n)d(n)d(n) is the desired signal.

The error signal determines how filter coefficients are updated during each iteration.

p(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-frac{(x\mu)^2}{2\sigma^2}\right)

where

- μ is the mean value,
- $\sigma 2$ is the variance of the noise.

Although real acoustic environments rarely contain purely white noise, this model is extensively used for algorithm evaluation and performance comparison.

Babble Noise

Babble noise is generated by the simultaneous speech of multiple speakers and is commonly encountered in public places such as conference halls, airports, classrooms, restaurants, and shopping centers. Unlike stationary noise, babble noise exhibits strong temporal and spectral variations, making speech enhancement considerably more challenging.

Adaptive filtering algorithms often experience slower convergence under babble noise because the statistical characteristics continuously change over time.

Pink Noise

Pink noise possesses a power spectral density inversely proportional to frequency. Its spectral characteristic is represented as S(f)\propto\frac{1}{f}

Compared with white noise, pink noise contains more energy in low-frequency regions and therefore better represents many natural acoustic environments.

Impulse Noise

Impulse noise consists of short-duration, high-amplitude disturbances caused by electrical switching, communication errors, or hardware malfunctions.

It can be modeled as x(n)=\begin{cases}A,& n=n_0\\ 0,& \text{otherwise} \end{cases}

where

- A denotes impulse amplitude,
- n0 is the occurrence time of the impulse.

Impulse noise is particularly difficult to suppress because of its sudden appearance and high energy.

Colored Noise

Unlike white noise, colored noise exhibits frequency-dependent spectral characteristics. Examples include brown noise, blue noise, violet noise, and grey noise. These noise models are frequently encountered in practical multimedia systems and acoustic signal processing applications.

Because colored noise is non-uniformly distributed across frequencies, adaptive filtering algorithms require more sophisticated parameter adjustment strategies for effective suppression.

Summary of Noise Characteristics

The diversity of environmental noise demonstrates that no single speech enhancement algorithm can provide optimal performance under all acoustic conditions. Consequently, adaptive digital filtering methods capable of adjusting their parameters according to changing noise characteristics remain an important research direction in modern speech signal processing.

3.3 Digital Filtering Theory

Digital filtering is one of the fundamental techniques in digital signal processing used to improve the quality of speech signals by suppressing unwanted noise while preserving useful speech information. Unlike analog filters, digital filters provide high accuracy, flexibility, stability, and easy implementation in modern multimedia and communication systems.

A digital filter processes a discrete-time input sequence and generates an output signal according to its transfer function. In general, the relationship between the input and output signals can be represented as

$$y(n)=\sum_{k=0}^{M}b_kx(n-k)-\sum_{k=1}^{N}a_ky(n-k)$$

where

- $x(n)$ is the input speech signal;
- $y(n)$ is the filtered output signal;
- bkdenotes the feedforward coefficients;
- ak denotes the feedback coefficients.

This mathematical representation forms the basis of most digital filtering techniques used in modern speech enhancement systems.

Finite Impulse Response (FIR) Filters

Finite Impulse Response (FIR) filters are characterized by the absence of feedback components. Consequently, they are always stable and possess linear phase characteristics, making them highly suitable for speech processing applications where waveform preservation is important.

The FIR filter output is expressed as $y(n)=\sum_{k=0}^{M}b_kx(n-k)$

The advantages of FIR filters include:

- unconditional stability;
- linear phase response;
- simple implementation;
- high numerical robustness.

However, achieving sharp frequency selectivity often requires a large number of filter coefficients, increasing computational complexity.

Infinite Impulse Response (IIR) Filters

Infinite Impulse Response (IIR) filters employ both feedforward and feedback coefficients, enabling them to achieve desired frequency responses with fewer coefficients than FIR filters.

The mathematical model of an IIR filter is $y(n)=\sum_{k=0}^{M}b_kx(n-k)-\sum_{k=1}^{N}a_ky(n-k)$

Compared with FIR filters, IIR filters require fewer computations; however, inappropriate coefficient selection may lead to instability.

Therefore, careful parameter design is necessary when applying IIR filters to speech enhancement systems.

Frequency Response of Digital Filters

The frequency response of a digital filter is obtained by evaluating its transfer function on the unit circle in the complex plane.

The transfer function is defined as $H(z)=\frac{Y(z)}{X(z)}$

Substituting $z=e^{j\omega}$ gives the frequency response $H(e^{j\omega})$

where

- ω represents the normalized angular frequency.

The frequency response determines how different frequency components of the speech signal are amplified or attenuated during filtering.

Adaptive Digital Filtering

Unlike conventional digital filters, adaptive filters automatically adjust their coefficients according to changes in the surrounding acoustic environment. This property makes adaptive filtering especially suitable for speech enhancement under time-varying noise conditions.

The coefficient vector is represented as $\mathbf{w}(n)=[w_0,w_1,\ldots,w_M]^T$

The adaptive filtering process continuously minimizes the estimation error between the desired speech signal and the filter output.

Adaptive Filter Structure

A typical adaptive speech enhancement system consists of four main components:

1. Noisy speech input;
2. Adaptive filter;

3. Error estimation block;
4. Coefficient update algorithm.

The filter coefficients are iteratively adjusted until the estimation error reaches its minimum value.

This adaptive mechanism enables the filter to follow continuously changing acoustic environments and significantly improves speech intelligibility compared with fixed digital filtering approaches.

Discussion

The theoretical analysis presented above demonstrates that conventional FIR and IIR filters provide efficient frequency-selective processing but lack the ability to respond dynamically to changing acoustic environments. Adaptive digital filters overcome this limitation by continuously updating filter coefficients according to the characteristics of the input signal. For this reason, adaptive filtering has become the dominant approach in modern speech enhancement systems and serves as the theoretical foundation for the algorithms analyzed in the following sections.

3.4 Least Mean Square (LMS) Algorithm

The Least Mean Square (LMS) algorithm is one of the most widely used adaptive filtering techniques for speech enhancement due to its simplicity, robustness, and low computational complexity. The algorithm updates filter coefficients iteratively by minimizing the instantaneous mean square error between the desired signal and the estimated output. Because of its computational efficiency, the LMS algorithm has been extensively applied in acoustic echo cancellation, active noise control, speech enhancement, biomedical signal processing, and wireless communication systems. The operation of the LMS algorithm is based on the steepest descent optimization method. Instead of calculating the exact gradient of the cost function, the algorithm estimates the gradient using the instantaneous error signal, thereby significantly reducing computational requirements.

The cost function is defined as $J(n)=E[e^2(n)]$

where

- $J(n)$ is the objective function,
- E denotes the expectation operator,
- $e(n)$ is the estimation error.

The estimation error is calculated as $e(n)=d(n)-\mathbf{w}^T(n)\mathbf{x}(n)$

where

- $J(n)$ is the objective function,
- Edenotes the expectation operator,
- $e(n)$ is the estimation error.

The estimation error is calculated as $e(n)=d(n)-\mathbf{w}^T(n)\mathbf{x}(n)$

where

- $d(n)$ is the desired speech signal;
- $x(n)$ is the input vector;
- $w(n)$ is the adaptive coefficient vector.

The coefficient update rule is expressed as $\mathbf{w}(n+1)=\mathbf{w}(n)+\mu e(n)\mathbf{x}(n)$

where

- μ represents the adaptation step size.

The convergence behavior of the LMS algorithm strongly depends on the selection of the adaptation parameter. Stable convergence is guaranteed when $0<\mu<\frac{2}{\lambda_{max}}$

where

- λmax\lambda_{max}λmax denotes the largest eigenvalue of the input signal autocorrelation matrix.

A larger step size accelerates convergence but increases the risk of instability, whereas a smaller step size provides more stable convergence at the expense of slower adaptation.

The computational complexity of the LMS algorithm is O(N)
where N represents the filter length.

This low computational complexity makes LMS particularly attractive for real-time speech enhancement applications.

Advantages of LMS Algorithm

- Simple mathematical formulation.
- Low computational complexity.
- Fast implementation.
- Suitable for real-time systems.
- Easy hardware realization.

Limitations of LMS Algorithm

Despite its popularity, the LMS algorithm suffers from several limitations:

- Slow convergence in highly non-stationary environments.
- Strong dependence on the adaptation step size.
- Reduced performance under low Signal-to-Noise Ratio conditions.
- Sensitivity to correlated input signals.

These limitations motivate the development of improved adaptive filtering strategies.

3.5 Normalized Least Mean Square (NLMS) Algorithm

The Normalized Least Mean Square (NLMS) algorithm is an improved version of the conventional LMS algorithm. The primary objective of NLMS is to eliminate the sensitivity of LMS to variations in input signal power by normalizing the adaptation coefficient during each iteration.

The adaptive coefficient update equation is expressed as $\mathbf{w}(n+1)=\mathbf{w}(n)+\frac{\mu}{\delta+\|\mathbf{x}(n)\|^2} e(n)\mathbf{x}(n)$

where

- δ is a small positive constant preventing numerical instability.

The normalization process allows the algorithm to maintain stable convergence even when the input speech signal exhibits large amplitude variations.

The adaptive learning rate becomes $\mu_n=\frac{\mu}{\delta+\|\mathbf{x}(n)\|^2}$

Consequently, the algorithm automatically adjusts its learning behavior according to the instantaneous signal energy.

Compared with LMS, NLMS generally provides

- faster convergence;
- improved numerical stability;
- better tracking capability;
- higher robustness under varying acoustic conditions.

However, the algorithm still experiences reduced performance in environments where noise characteristics change abruptly.

Its computational complexity remains $O(N)$

which makes NLMS highly suitable for practical multimedia applications.

3.6 Wiener Filtering

The Wiener filter is one of the classical optimal filtering techniques developed to minimize the mean square estimation error between the desired speech signal and the estimated output. Unlike adaptive filtering methods, Wiener filtering assumes that the statistical properties of speech and noise are either known or can be accurately estimated.

The optimal transfer function is expressed as $H(f)= \frac{P_s(f)}{P_s(f)+P_v(f)}$ where

- Ps(f) represents the speech power spectrum;
- Pv(f) represents the noise power spectrum.

The estimated speech spectrum is obtained as $\hat{S}(f)=H(f)Y(f)$
where

- Y(f) denotes the noisy speech spectrum.

The Wiener filter performs exceptionally well under stationary Gaussian noise conditions because it minimizes the global mean square error. However, accurate estimation of spectral densities remains a major challenge in real acoustic environments. The computational complexity of Wiener filtering is approximately$O(N\log N)$ due to the use of Fourier-domain processing.

Advantages of Wiener Filtering

- Excellent stationary noise suppression.
- Optimal MSE performance.
- High speech quality preservation.
- Strong theoretical foundation.

Limitations

- Requires accurate noise estimation.
- Less effective under rapidly changing environments.
- Performance depends on spectral estimation accuracy.

3.7 Kalman Filtering

The Kalman filter is one of the most effective recursive estimation techniques employed in speech enhancement and digital signal processing. Unlike conventional adaptive filters, the Kalman filter estimates the clean speech signal by combining prior knowledge of the system dynamics with noisy observations. Owing to its recursive nature, it continuously updates the estimated speech signal whenever new measurements become available, making it particularly suitable for non-stationary acoustic environments. The Kalman filter models the speech signal using a state-space representation consisting of a state equation and an observation equation.

The state equation is defined as $x(k)=Ax(k-1)+Bu(k)+w(k)$

where

- $x(k)$ represents the state vector;
- Ais the state transition matrix;
- B is the control matrix;
- $u(k)$ denotes the control input;
- $w(k)$ represents process noise.

The observation equation is$z(k)=Hx(k)+v(k)$

where

- $z(k)$ is the observed noisy speech signal;
- H is the observation matrix;
- $v(k)$ denotes measurement noise.

The prediction stage estimates the current state as

where

- $z(k)$ is the observed noisy speech signal;
- H is the observation matrix;
- $v(k)v(k)v(k)$ denotes measurement noise.

The prediction stage estimates the current state as $\hat{x}_{k}^{-}=A\hat{x}_{k-1}+Bu_k$ The prediction covariance matrix is $P_k^{-}=AP_{k-1}A^T+Q$

where

- Q denotes the covariance matrix of process noise. The Kalman gain is calculated as $K_k=P_k^{-}H^T (H P_k^{-}H^T+R)^{-1}$

where

- R represents the covariance matrix of measurement noise.

Finally, the updated state estimate is obtained by $\hat{x}_k= \hat{x}_k^{-}+ K_k(z_k-H\hat{x}_k^{-})$

The Kalman filter provides excellent tracking capability and high estimation accuracy under dynamic acoustic conditions. Nevertheless, its implementation requires accurate statistical models of both speech and noise, which may not always be available in practical applications.The computational complexity of Kalman filtering is significantly higher than LMS and NLMS because matrix inversion operations must be performed during each iteration. Consequently, the algorithm is mainly employed in applications where estimation accuracy is more important than computational efficiency.

3.8 Comparative Mathematical Analysis of Adaptive Filtering Algorithms

To identify the most appropriate speech enhancement technique, adaptive filtering algorithms must be compared according to objective mathematical criteria. These include convergence speed, computational complexity, stability, robustness against noise variation, and implementation cost.

The convergence rate of LMS depends directly on the adaptation step size

μ

hereas NLMS automatically normalizes the adaptation coefficient according to the signal energy.

The convergence condition of LMS is expressed as $0<\mu< \frac{2}{\lambda_{max}}$

The normalized adaptation factor of NLMS is $\mu_n= \frac{\mu}{\delta+\|x(n)\|^2}$

Consequently, NLMS exhibits faster convergence under varying signal amplitudes.

The computational complexity of the algorithms can be summarized as follows.

$C_{LMS}=O(N)$

$C_{NLMS}=O(N)$

$C_{Wiener}=O(N\log N)$

$C_{RLS}=O(N^2)$

$C_{Kalman}=O(N^3)$

These expressions indicate that LMS and NLMS require significantly fewer computational resources than Kalman and RLS algorithms.

For real-time multimedia systems, computational complexity is one of the most important design parameters because speech enhancement must be performed within a few milliseconds.

Table 2. Comparative characteristics of adaptive filtering algorithms

Algorithm	**Convergence**	**Complexity**	**Stability**	**Noise Suppression**	**Real-Time Suitability**
LMS	Moderate	Low	High	Moderate	Excellent
NLMS	Fast	Low	Very High	Good	Excellent
Wiener	Fast	Medium	High	Very Good	Good
RLS	Very Fast	High	High	Excellent	Moderate
Kalman	Excellent	Very High	Excellent	Excellent	Limited

The comparison demonstrates that there is no universal filtering algorithm capable of providing optimal performance under all acoustic conditions. Each method performs effectively only within a specific operating environment.

Consequently, recent research has shifted toward adaptive parameter optimization rather than developing completely new filtering algorithms.

3.9 Performance Evaluation Metrics

The effectiveness of speech enhancement algorithms is evaluated using both objective and perceptual quality assessment metrics. The Signal-to-Noise Ratio (SNR) is defined as$SNR=10\log_{10}\left(\frac{\sum s^2(n)}{\sum(s(n)-\hat{s}(n))^2}\right)$

A higher SNR value indicates better noise suppression. The Mean Square Error (MSE) is$MSE=\frac1N\sum_{n=1}^{N} (s(n)-\hat{s}(n))^2$

Lower MSE values correspond to higher reconstruction accuracy.The Root Mean Square Error (RMSE) is $RMSE=\sqrt{MSE}$ Speech quality can also be assessed using the Peak Signal-to-Noise Ratio (PSNR). $PSNR=10\log_{10}\left(\frac{MAX^2}{MSE}\right)$

where MAXMAXMAX represents the maximum signal amplitude.

Perceptual quality is commonly evaluated using:

- PESQ (Perceptual Evaluation of Speech Quality);
- STOI (Short-Time Objective Intelligibility);
- SDR (Signal-to-Distortion Ratio);
- SI-SDR (Scale-Invariant Signal-to-Distortion Ratio).

Unlike SNR and MSE, these metrics correlate more closely with human auditory perception and therefore provide a more reliable assessment of speech intelligibility. Considering both objective and perceptual evaluation metrics enables a comprehensive comparison of speech enhancement algorithms and facilitates the identification of the most suitable filtering strategy for practical multimedia applications.

Proposed Improvement Framework

4.1 Motivation for Algorithm Improvement

The comparative analysis presented in the previous sections demonstrates that existing adaptive filtering algorithms exhibit satisfactory performance only under specific acoustic conditions. The LMS algorithm is computationally efficient but converges slowly under rapidly changing environments. Although NLMS improves convergence speed through input signal normalization, its performance still deteriorates under highly non-stationary noise. Wiener filtering provides excellent suppression for stationary noise; however, its effectiveness decreases significantly when accurate noise statistics cannot be estimated. Similarly, Kalman filtering achieves high estimation accuracy but suffers from high computational complexity, limiting its applicability in real-time embedded systems. These limitations indicate that the principal challenge in modern speech enhancement is not the lack of filtering algorithms but rather the absence of adaptive parameter selection mechanisms capable of responding dynamically to varying acoustic environments.

Therefore, instead of developing a completely new filtering algorithm, this research proposes an improved adaptive digital filtering framework that dynamically adjusts filter parameters according to environmental noise characteristics.

4.2 Proposed Research Concept

The proposed framework consists of five sequential processing stages:

1. Acquisition of noisy speech signals.
2. Acoustic environment analysis.
3. Noise characteristic estimation.
4. Adaptive parameter optimization.

5. Speech enhancement using an optimized adaptive filter.

Unlike conventional adaptive filtering approaches, the proposed framework continuously analyzes the acoustic environment and updates filtering parameters whenever significant changes in noise characteristics are detected.

The conceptual processing chain can be represented as

4.3 Adaptive Parameter Optimization

One of the most important factors influencing adaptive filtering performance is the adaptation coefficient. Conventional LMS algorithms employ a fixed learning rate, which cannot provide optimal convergence under all operating conditions. To overcome this limitation, the proposed framework introduces a dynamic adaptation strategy in which the learning coefficient varies according to the estimated acoustic conditions.

The adaptive learning coefficient is defined as \mu(n)=\mu_{0}\cdot \alpha(n) where

- μ0\mu_{0}μ0 is the initial learning coefficient;
- α(n)\alpha(n)α(n) is the adaptive correction factor determined from the estimated noise characteristics.

This adaptive strategy allows the filter to increase convergence speed during rapid environmental changes while maintaining stability during stationary conditions.

4.4 Noise Classification Stage

Before filtering, the acoustic environment should be analyzed to estimate the dominant noise characteristics.

Typical noise categories include:

- White Gaussian Noise
- Babble Noise

- Traffic Noise
- Industrial Noise
- Office Noise
- Vehicle Noise
- Wind Noise
- Crowd Noise

Different noise categories require different adaptation strategies because their statistical characteristics differ considerably.

For example,

- stationary noise favors Wiener filtering;
- rapidly changing noise requires adaptive LMS/NLMS optimization;
- mixed environments require dynamic parameter adjustment.

4.5 Expected Advantages

Compared with conventional adaptive filtering techniques, the proposed framework is expected to provide:

- higher speech intelligibility;
- improved noise suppression;
- faster convergence;
- reduced residual noise;
- better adaptation to non-stationary environments;
- lower computational complexity compared with advanced deep learning models;
- suitability for real-time multimedia applications.

Consequently, the proposed approach represents a promising direction for future speech enhancement research.

5. Results and Discussion

Since the primary objective of this study is to investigate existing adaptive filtering techniques and identify promising improvement directions, the discussion focuses on the comparative evaluation of the analyzed algorithms. The literature analysis demonstrates that adaptive filtering continues to be one of the most effective approaches for speech enhancement in noisy environments. LMS remains attractive because of its computational simplicity, whereas NLMS provides improved convergence through input normalization. Wiener filtering achieves excellent performance under stationary acoustic conditions but requires reliable estimation of speech and noise statistics. RLS and Kalman filtering produce superior estimation accuracy; however, their computational complexity significantly limits real-time implementation.

Table 3 summarizes the overall comparison.

Table 3. Overall comparison of adaptive filtering techniques

Criterion	LMS	NLMS	Wiener	RLS	Kalman
Noise Reduction	Medium	Good	Very Good	Excellent	Excellent
Adaptability	Good	Very Good	Moderate	Excellent	Excellent
Computational Cost	Low	Low	Medium	High	Very High
Stability	Good	Excellent	Good	Excellent	Excellent
Real-Time Suitability	Excellent	Excellent	Good	Moderate	Limited

The comparison indicates that no single algorithm satisfies all practical requirements simultaneously. Consequently, future speech enhancement systems should focus on adaptive parameter optimization rather than replacing existing adaptive filtering techniques. The proposed framework provides an effective basis for such improvements and represents the next stage of the present doctoral research.

6. Conclusion

This paper presented a comprehensive comparative analysis of adaptive digital filtering algorithms for speech enhancement in noisy environments. Classical filtering techniques, including LMS, NLMS, Wiener filtering, RLS, and Kalman filtering, were analyzed from both theoretical and practical perspectives. The analysis demonstrates that each algorithm exhibits unique advantages and limitations depending on the acoustic environment and computational requirements. While LMS and NLMS remain suitable for real-time applications because of their low computational complexity, Wiener, RLS, and Kalman filters provide higher estimation accuracy under specific operating conditions.The study concludes that future research should concentrate on improving adaptive parameter optimization mechanisms capable of dynamically responding to continuously changing acoustic environments. The proposed adaptive filtering framework establishes the theoretical foundation for subsequent research devoted to developing an improved speech enhancement algorithm for modern multimedia communication systems.

REFERENCES

[1] S. Haykin, Adaptive Filter Theory, 5th ed. Pearson Education, 2014.

[2] B. Widrow and S. D. Stearns, Adaptive Signal Processing. Prentice Hall, 1985.

[3] P. C. Loizou, Speech Enhancement: Theory and Practice, 2nd ed. CRC Press, 2013.

[4] A. V. Oppenheim and R. W. Schafer, Discrete-Time Signal Processing, 3rd ed. Pearson, 2010.

[5] J. G. Proakis and D. G. Manolakis, Digital Signal Processing: Principles, Algorithms and Applications, 4th ed. Pearson, 2007.

[6] S. F. Boll, "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 27, no. 2, pp. 113–120, 1979.
[7] Y. Ephraim and D. Malah, "Speech Enhancement Using a Minimum Mean Square Error Short-Time Spectral Amplitude Estimator," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 32, no. 6, pp. 1109–1121, 1984.
[8] R. Martin, "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics," IEEE Transactions on Speech and Audio Processing, vol. 9, no. 5, pp. 504–512, 2001.
[9] I. Cohen, "Noise Spectrum Estimation in Adverse Environments," IEEE Transactions on Speech and Audio Processing, vol. 11, no. 5, pp. 466–475, 2003.
[10] S. Vaseghi, Advanced Signal Processing and Digital Noise Reduction, 4th ed. Wiley, 2008.
[11] J. Benesty, Y. Huang, and J. Chen, Adaptive Signal Processing: Applications to Real-World Problems. Springer, 2003.
[12] R. G. Lyons, Understanding Digital Signal Processing, 3rd ed. Pearson, 2011.
[13] S. K. Mitra, Digital Signal Processing: A Computer-Based Approach, 4th ed. McGraw-Hill, 2011.
[14] L. R. Rabiner and R. W. Schafer, Digital Processing of Speech Signals. Prentice Hall, 1978.
[15] P. Vary and R. Martin, Digital Speech Transmission. Wiley, 2006.
[16] D. Wang and J. Chen, "Supervised Speech Separation Based on Deep Learning: An Overview," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 26, no. 10, pp. 1702–1726, 2018.
[17] E. Vincent, T. Virtanen, and S. Gannot, Audio Source Separation and Speech Enhancement. Wiley, 2018.

[18] D. Jurafsky and J. H. Martin, Speech and Language Processing, 3rd ed. Draft, 2024.
[19] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.
[20] S. Makino, Audio Source Separation. Springer, 2018.
[21] H. L. Van Trees, Detection, Estimation and Time Series Analysis. Wiley, 2001.
[22] ITU-T Recommendation P.862, Perceptual Evaluation of Speech Quality (PESQ).
[23] ITU-T Recommendation P.863, Perceptual Objective Listening Quality Assessment (POLQA).
[24] ITU-T Recommendation P.800, Methods for Subjective Determination of Transmission Quality.
[25] IEEE Signal Processing Society, "Speech Enhancement and Noise Reduction Techniques," 2023.
[26] X. Zhang, Y. Wang, and L. Chen, "Adaptive Noise Reduction for Real-Time Speech Enhancement Using Improved LMS Algorithm," IEEE Access, 2022.
[27] J. Li, S. Zhao, and H. Wang, "Performance Analysis of Adaptive Speech Enhancement Algorithms Under Non-Stationary Noise," IEEE Access, 2023.
[28] M. Tanveer and A. Kumar, "Comparative Analysis of Wiener and Adaptive Filters for Speech Enhancement," Applied Acoustics, 2021.
[29] Y. Xu et al., "A Regression Approach to Speech Enhancement Based on Deep Neural Networks," IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2015.
[30] K. Tan and D. Wang, "A Convolutional Recurrent Neural Network for Real-Time Speech Enhancement," INTERSPEECH, 2018.

[31] A. Pandey and D. Wang, “Dense CNN with Self-Attention for Time-Domain Speech Enhancement,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2021.

[32] C. Valentini-Botinhao et al., “Noisy Speech Database for Training Speech Enhancement Algorithms,” University of Edinburgh, 2017.

[33] J. Lim and A. Oppenheim, “Enhancement and Bandwidth Compression of Noisy Speech,” Proceedings of the IEEE, vol. 67, no. 12, pp. 1586–1604, 1979.

[34] S. Gannot, E. Vincent, S. Markovich-Golan, and A. Ozerov, “A Consolidated Perspective on Multimicrophone Speech Enhancement,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2017.

[35] T. Gerkmann and R. Hendriks, “Unbiased MMSE-Based Noise Power Estimation,” IEEE Transactions on Audio, Speech, and Language Processing, 2012.

[36] K. Kinoshita et al., “A Summary of the REVERB Challenge,” EURASIP Journal on Advances in Signal Processing, 2016.

[37] S. Doclo et al., “Acoustic Beamforming for Hearing Aid Applications,” Handbook on Array Processing and Sensor Networks, Springer.

[38] J. Benesty, M. M. Sondhi, and Y. Huang, Springer Handbook of Speech Processing. Springer, 2008.

[39] T. Virtanen, R. Singh, and B. Raj, Techniques for Noise Robustness in Automatic Speech Recognition. Wiley, 2012.

[40] D. L. Donoho, “De-Noising by Soft Thresholding,” IEEE Transactions on Information Theory, vol. 41, no. 3, pp. 613–627, 1995.