ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20251217001846

Original scientific paper
Received: 2025-12-17, Accepted: 2026-01-16

Speech Signal Enhancement Using Linear Predictive Coefficients Adaptive Filter

Rehab I. Ajel*, Tariq A. Hassan

Abstract: The linear predictive coding (LPC) is a technique that is widely utilized in speech processing, especially in speech spectral envelope modelling, where the aim is to estimate the vocal tract resonances. Consequently, the LPC is used in a vast range of applications, like speech coding, transformation, compression, and speech and speaker recognition. However, with all advances in speech signal processing techniques, the existence of background noise poses a compelling challenge and has a huge impact on the speech quality. This paper proposes an enhancement technique that benefits from the LPC coefficients and exploits them as parameters for an adaptive filter used in speech noise reduction. Unlike the other methods that used the LPC technique for vocal tract spectral envelope estimation or source residual modelling, the proposed method adopts the LPC coefficients to build a frame-dependent, bandpass-like adaptive digital filter that is completely derived from the LPC parameters. This filter will be generated from each speech segment and used to suppress additive noise while preserving real-speech components. The proposed technique is evaluated and compared with two traditional fixed-parameter filters, Wiener and Kalman filters, using speech samples of the TIMIT dataset. The samples are corrupted with additive Gaussian white noise with an input signal-to-noise ratio (SNR) of 20 dB. Practical results show that the performance of the proposed LPC-based adaptive filter in enhancing the speech signal is remarkable with the voiced speech segments, while producing comparable results to Wiener and Kalman filters with unvoiced speech. These points out the effect of frame-dependent LPC-based adaptive filter to deal with nonstationary signals like speech.

Keywords: adaptive filter; Linear Predictive Coding (LPC); low-pass filter; Kalman filter; speech enhancement; Wiener filter

1 INTRODUCTION

One of the most reliable and essential forms of human verbal message is speech. As well as, important information can be conveyed by the speech signal; such as linguistic content (the words being said, the language, and the grammar), emotional state, gender, and speaker identity [1]. One of the most popular frameworks adopted in the speech processing field is the source–filter model. In which, the speech is represented as a convolution of two major components: the source signal and the filter envelop.

The source signal approximates the oscillation in the vocal cords and dictates the fundamental frequency along with its harmonics. The filter envelope, on the other hand, describes vocal tract which composing the throat, mouth cavity, lips, and the nasal cavities. The spectral envelope of the speech signal represent the vocal tract shape or the filter envelope [2, 3].

Naturally, the speech signals are often vulnerable to various types of noise, such as channel noise, babble noise, and additive background noise that enormously affect speech clarity and quality.

During the production of the speech, a background noise is presented due to various conditions and states such as health and emotion, as well as due to low-quality communication channels or interferences with other signals [4].

As a result, speech signal enhancement becomes one of the most fundamental tasks in contemporary speech signal processing techniques.

Enhancing the audio signals by reducing or removing the unwanted noise in order to improve their clarity and quality is the major role of the audio enhancement techniques. The goal is to produce a clean signal for humans and ready for analysis by machines.

Noise reduction is crucial for many applications, such as speaker recognition [5], automatic speech recognition [6], emotion classification [7], and related speech-based systems.

In the past few years, many proposed approaches have been presented to enhance and improve the signal quality. Some adopt classical signal processing techniques, and others rely on data-driven methods.

For instance, deep denoising autoencoders have been used in learning the mapping between the noisy and the clean signals [8]. A wavelet-based method along with deep-learning for signal enhancement has been proposed in [9]. Swarm optimization, integrated particularly with convolutional neural networks, has been explored in [10].

Classic filtering approaches such as adaptive filters, Kalman filters, and Wiener filters are still excessively used because of their ability in noise reduction and signal enhancement [11–14].

Linear predictive coding (or LPC) has presented as one of the main techniques for speech enhancement. The LPC coefficients, as stated by the traditional LPC-based methods, are mainly adopted in spectral envelope estimation of the speech signal. Or, in other words, to model the vocal tract of the signal during speaking. Noise reduction, on the other hand, is basically accomplished via residual or source signal processing or via LPC-derived parameters embedding with Wiener or MMSE-based filtering frameworks. However, the LPC approach, in these methods, functions as an analytical procedure rather than serving directly in the denoising mechanism.

This paper proposes a different use of the LPC technique. As opposed to the traditional methods, the proposed work suggest to directly adopting LPC coefficients as frame-dependent adaptive filter weights for noise suppression and speech enhancement. In other words, rather than applying LPC exclusively for signal envelope approximation or residual or source signal modelling, the proposed method builds an adaptive filter generated from the LPC coefficients, which will be continuously updated according to the noisy speech frame.

This technique will effectively contribute to signal noise removal without the need for residual reconstruction or the estimation of the noise power, as required in traditional Wiener or Kalman filtering methods.

The evaluation of our proposed system will be under the conditions of additive Gaussian white noise, and the practical testing results show that the LPC-based adaptive filter. The proposed system is evaluated under additive Gaussian white noise conditions, and experimental results demonstrate that the LPC filter-based approach exceeds (in some cases) the performance of the conventional Wiener and Kalman filters in terms of speech enhancement performance.

In this work, the main contribute can be outlined as follows:

1) A new adoption of linear predictive coding (LPC) coefficients as a frame-dependent weighted adaptive filter for explicit speech enhancement, instead of spectral envelope approximation or source signal modeling.
2) A time-domain LPC-based speech noise removal system signal transformation to reduce complexity and time consumption.
3) Evaluation of the proposed system performance demonstrates that the LPC–based adaptive filter can exceed the performance of other filters like Wiener and Kalman filters under different noises levels.

2 NOISE TYPES

Unwanted electrical or electromagnetic energy deteriorates data and signal quality. This is known as noise. Both digital and analog systems can experience noise, which can impact many file kinds and communications, such as text, programs, graphics, audio, and telemetry. The following are possible major impacts of noise on a given signal: information content, clarity, and loudness.

Generally speaking, three types of noise may be modelled that impact speech signals: impulsive, colored, and white noise. [15, 16]. A signal or sound that has all audible frequencies present at the same strength is known as white noise at every frequency. When two distinct noise signals are combined, the resulting noise signal has a power equal to the total of its power components. The broad-band spectrum of white noise gives it powerful masking powers.

Colored noise is any noise other than white noise. The frequency spectrum of colored noise is restricted within a range, in contrast to white noise. Various colored noises exist, such as brown, pink, orange, and so on, based on the noise's Power Spectral Density (PSD) gradation. Sudden noise bursts with a comparatively large amplitude are referred to as impulsive noise. In the signal of interest, this kind of noise results in click noises [17].

3 ADAPTIVE FILTER

In signal processing, the adaptive filter is a system in which the parameters are adjusted automatically in response to the input signal variation according to the surrounding environment. In contrast with the fixed-parameter filter (predefined filters), adaptive filters continuously upgrade their coefficients in accordance with an adaptation mechanism that is previously set in the system. The filter coefficients are updated to meet a selected performance criterion, which normally is an error signal.

Generally, the adaptive filtering technique is an algorithm that comprises the digital filter along with a refitting algorithm that manages the filter coefficient update [18].

Techniques like Least Mean Square (LMS), Normalized Least Mean Square (NLMS), and Recursive Least Squares (RLS) algorithms are normally used as adaptive algorithms. These algorithms are continuously update the filter coefficients until the optimal (or the best fit) solution is reached [19]. The main criterion of these algorithms is the errors between the desired version of the input signal (clean signal) and the filter output.

In this paper, a new strategy for adaptive filtering is proposed. In it, the LPC coefficients are adopted as parameters for the frame-dependent adaptive filter. This differs to the traditional error-driven algorithms normally used in this vicinity.

An adaptive all-pole filter with its parameters evaluated over time is constructed using the LPC coefficients of the short-term segment of the noisy speech signal. In this technique, the coefficients of the LPC-based filter are readjusted in accordance with the spectral characteristics of the speech signal that will help to estimate the enhanced version of the speech.

4 LPC PARAMETERS

Linear Predictive Coding (LPC) is a technique that is widely adopted in signal modelling and speech analysis, designed to efficiently approximate the spectral coefficients, or the envelope, of the speech signal.

One of the classical models of speech production is the source-filter model, where techniques like LPC, MFCC, and PLP are totally rely on this model. The speech, according to this model, is represented as a vocal tract all-pole filter output with a source signal excitation.

Because of its solid representation and approximate explanation, the LPC technique has been broadly adopted in areas like; speech analysis, coding, synthesis, and speech and speaker recognition [20].

Basically, any speech sample, according to the presumption of the linear prediction, can be represented as a mixture or combination of a pre-set number of the previous samples, by virtue of the contiguous speech samples high correlation. Mathematically, the speech sample $\hat{s}[n]$ can be approximated as:

$$\hat{s}[n] \approx \sum_{k=1}^{p} a_k s[n-k] \tag{1}$$

where p indicates the order of prediction and a_k represents the coefficients of the linear predictive.

The error between the predicted signal and the actual speech is called the residual or the excitation signal, mathematically is expressed as:

$$e[n] = s[n] - \hat{s}[n] = s[n] - \sum_{k=1}^{p} a_k s[n-k] \tag{2}$$

The aim is to determine the LPC coefficients $\{a_k\}$ by minimizing the mean squared error between the actual signal samples and the predicted ones. Normally, the autocorrelation-based technique, which employs the Levinson–Durbin algorithm, is used [21]. Modeling the resonances of the vocal tract requires such a stable all-pole filter established by this optimization.

The all-pole transfer function of the vocal tract can be expressed as:

$$H_{\mathrm{LPC}}(z) = \frac{1}{1 - \sum_{k=1}^{p} a_k z^{-1}} \tag{3}$$

According to this transfer function (or all-pole filter), the frequencies inherent with the speech signal will be concentrated while other spectral components (outside the speech envelope) will be attenuated. Therefore, the LPC-based adaptive filter will show a band-pass-like mode that keeps the most dominant speech frequencies.

In our proposed speech-enhancing system, the LPC-based filter coefficients will be determined based on a frame-by-frame scheme from the noisy signal and used as a parameters for a band-pass adaptive filter.

Each frame of the noisy signal $x[n]$ will pass through the LPC filter defined in (3) to estimate , to some extent, an enhanced (noise free) signal $\hat{s}[n]$.

For each frame, the LPC coefficients will continuously be modified, which will allow the filter to adapt in accordance with non-stationary spectral characteristics of speech to be able to suppress the additive noise without the need for residual or source signal reconstruction or stochastic noise modeling.

Assume that $x[n] = s[n] + v[n]$, denote a speech signal with additive white noise ($v[n]$). In our proposed framework, the speech signal should first be segmented into fixed unoverlabed frames. The LPC coefficients $\{a_k\}$ have to be estimated directly from each noisy frame. These coefficients are then utilized to generate an instantaneous adaptive band-pass filter. The LPC synthesis filter $H_{\mathrm{LPC}}(z)$ will be applied to the corresponding noisy speech frame in the time domain, all follow:

$$\hat{x}[n] = h_{\mathrm{LPC}} * x[n] \tag{4}$$

where h_{LPC} is the impulse response of the LPC filter and * is the convolution operation.

This filtering operation will selectively retain the speech formant- related frequency components while removing, to some extent, the noise-related components.

Fig. 1 shows one frame speech signal and its frequency envelope (formats) estimated using LPC coefficients.

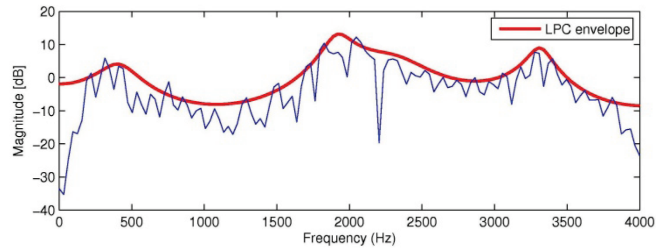


Figure 1 Speech signal frequency envelop and its formant

During the production of speech, the characteristics of the vocal tract resonances provoke a set of distinguished peaks in the envelope of the speech spectrum, which are known as the formant frequencies. The transferee function shows that the location of the resonance peaks is closely surrounded by the formant frequencies. The LPC with the all-pole filer framework can efficiently pattern this deed. The structure of the speech signal formant will be rounded up by the LPC filter poles.

This method takes advantage of the ability of the LPC to model the speech signal and acts as a band-pass adaptive filter that can modify its parameters according to time-varying speech characteristics in order to suppress the noise components that are located far from the modelled spectral structure.

5 THE PROPOSED MODEL

Our proposed system is to enhance and clarify the corrupted speech signal with additive white Gaussian noise. First, the input signal is segmented into short-time frames with a period of 20–30 ms. the speech characteristics tend to be quasi-stationary within this duration, so that the system can capture, almost, stable features of the speech.

To keep the frames in their original temporal structure, no windowing function is adopted.

Following segmentation, the next step is to classify the frames into speech-containing (or active) frames or silent frames. This classification is attained based on the calculation of frame energy. Practically, in noisy environments, the classification is highly challenging as the existence of background noise would make it hard to select the proper range of the energy threshold that adapts to separate the different frames. Yet, practically, an energy-based criterion was adopted to give an appropriate separation between these two types of frames. This would be appreciated for regulating the subsequent processing stages.

To eliminate the noise component within the high-frequency range, and prior to the LPC analysis step, each speech frame must pass through a low-pass filter. This is due to the fact that most of the speech energy is accumulated in a band within the low-frequency region of the speech, typically below 3.5 kHz. Additive noise, on the other hand, tends to concentrate around high frequencies [22]. Because of its maximum flat passband and smooth frequency response, a sixth-order Butterworth low-pass filter with a cut-off frequency of 3 kHz is adopted in our system.

Following low-pass filtering, the next step is the LPC-based filtering scheme. For each frame of the noisy signal, the LPC procedure has to estimate the prediction coefficient set (order p=15 in our experiment) and use them to construct the LPC-based band-pass filter. These frame-dependent LPC coefficients are the main ingredients for the adaptive time-domain filters that serve to remove noise while maintaining the most speech-relevant components.

In this regard, we would like to draw attention to two important points: First, in the process, there is no overlap between the adjacent frames. In other words, the frames are processed sequentially to prevent discontinuity in the enhanced speech signal. Second, the continuous updating of the LPC coefficients with each frame will enable the adaptive filter to adapt to the non-stationary features of the speech signal.

The general processing steps of the proposed LPC-based speech enhancement model is illustrated in Fig. 2.

6 EXPERIMENTS AND RESULTS

As explained in the earlier sections, the input speech (noisy signal) will first be segmented into fixed-length frames (in our experiment, between 20 and 30 ms duration) as a preprocessing stage of our proposed speech enhancement scheme. The short-time strategy setup will enable the system to take advantage of the semi-stationary nature of speech within a short period of time while allowing frame-dependent adaptation of the LPC-based filter.

Before starting with LPC analysis, the high frequencies of the input frames should be attenuated using a low-pass filter, which will assist in noise reduction as the additive noise normally concentrates within the high-frequency area of the speech. Butterworth low-pass filter is used due to its steady frequency response and maximally flat passband. Based on this step, the system will seek the superiority of speech-relevant parameters, including frequency envelope or formant structure, within the low-frequency part of the spectrum.

The LPC-based adaptive filtering approach is now ready to extract the clean and clear signal out of the noisy frames. The LPC- based filter will be examined with different levels of additive white noise. The main objective of this empirical evaluation is to investigate the extent to which the LPC parameters can effectively participate in constructing a highly potent adaptive filter. This would help to eliminate the undesirable additive noise while retaining the speech characteristics.

Important notes have to be mentioned about the speech signals used in our system: First, all signals have the same sampling frequency, to ensure that the system stays consistent for all inputs. Second, the frames are evaluated separately according to whether they are voiced or unvoiced since they have different spectral characteristics and need a bit different management for noise removal. Also, the practical analyses focus on examining the ability of the proposed system to secure the real speech component while suppressing the noise artifacts.

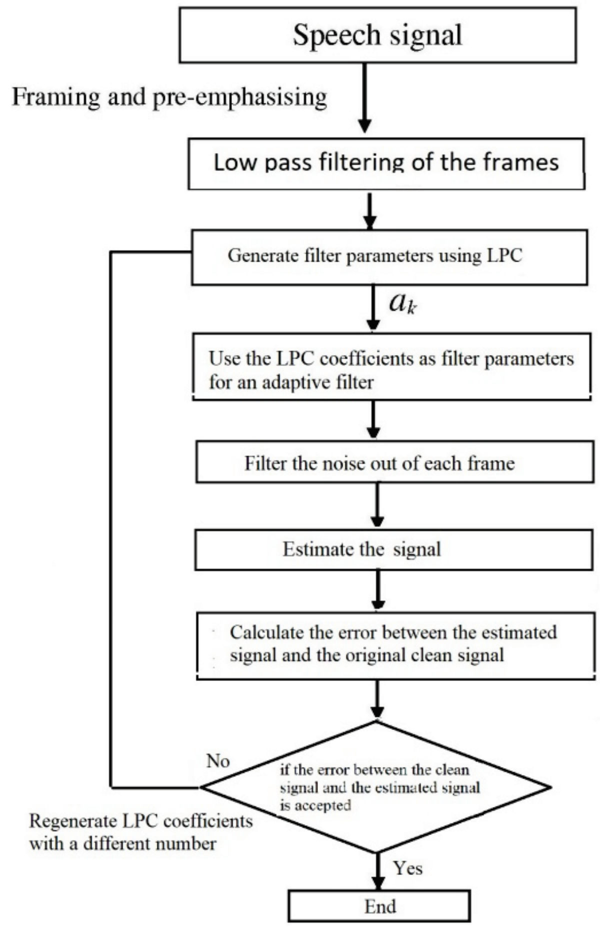


Figure 2 Block diagram of the suggested filtering model

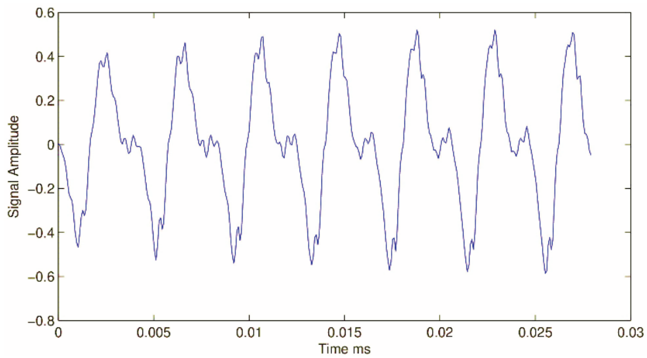
This section shows the performance evaluation of the proposed framework for voice and unvoiced speech segments with a different range of the additive Gaussian white noise, presenting the level of enhancement in speech quality under noise effect.

The setup started with Fig. 3, which shows depictive examples of both voiced and unvoiced speech frames in clean and noisy states in the time domain. The frames are taken from the BT Millar speech database of the spoken words "one" and "six". Both are corrupted by additive Gaussian white noise at a ratio of 50%. The figure also emphasizes the explicit temporal typical feature of voiced and unvoiced speech segments and the noise impact on each type.

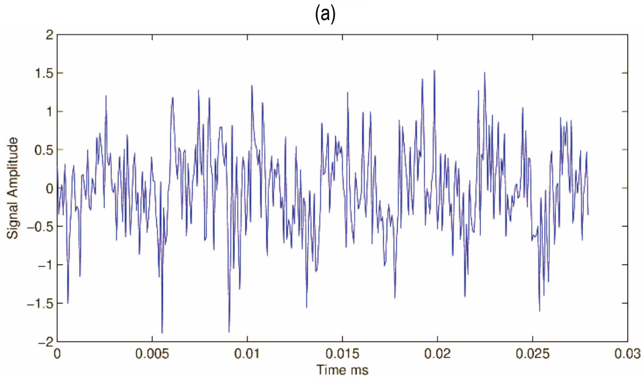
The corresponding frequency-domain depiction for these frames is presented in Fig. 4 illustrate the spectral distinction between voiced and unvoiced speech as well as the noise additive effect on the speech spectrum.

Fig. 5 illustrates the effect of low-pass filtering (at 3 kHz cut-off) on the noisy speech spectrum. It is clear how the low-pass filter significantly attenuates the unwanted noise.

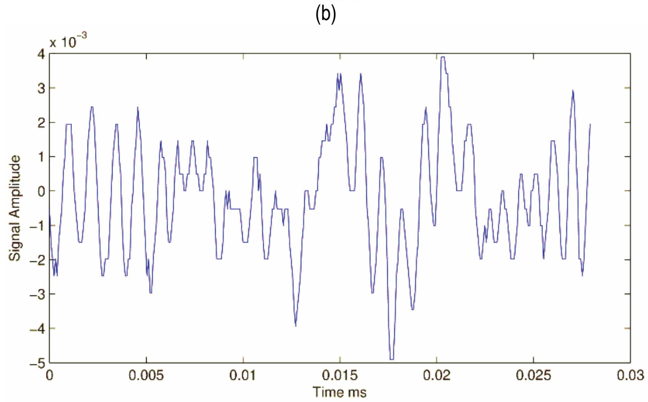
The LPC scheme is applied to each low-pass filtered frame to estimate the speech spectral envelope. The extracted envelope is consequently utilized to construct a band-pass-like filter. Fig. 6 shows the voiced and unvoiced spectra together with their corresponding LPC envelopes. It is very clear how the spectral envelope is concentrated around the resonance or the formant frequency of the speech.



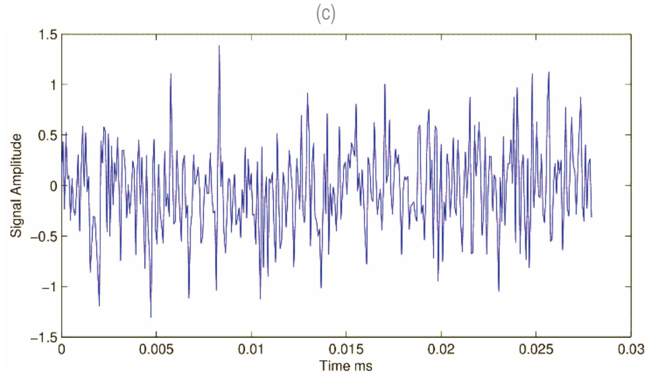
(a)



(b)

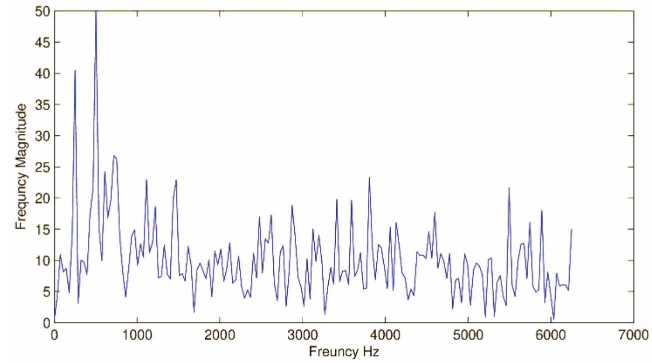


(c)

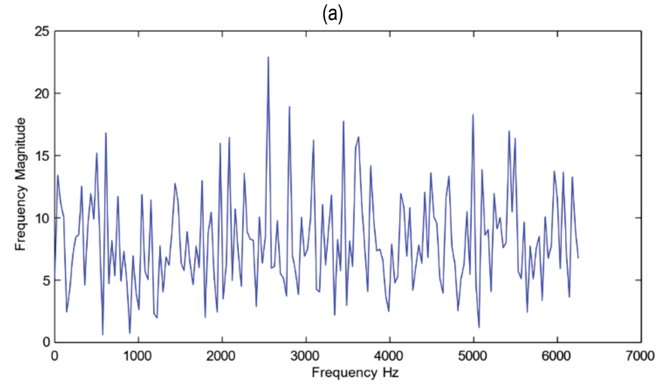


(d)

Figure 3 Examples of some speech signal frames: (a) clean Voiced frame, (b) Voiced frame with additive noise, (c) clean unvoiced frame, (b) Noisy unvoiced frame.

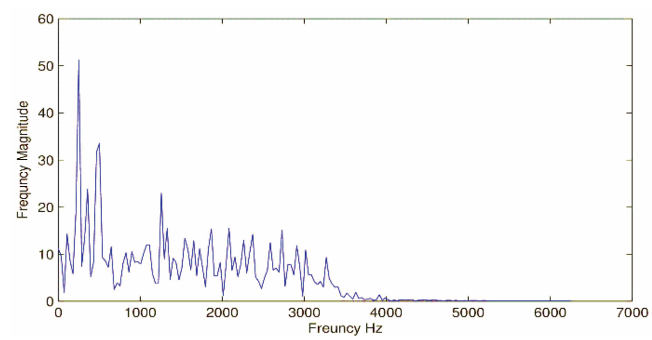


(a)

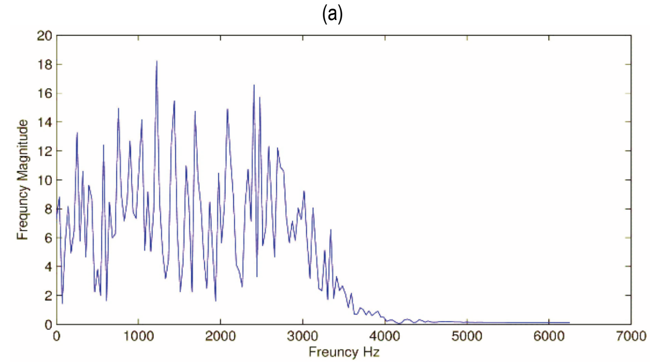


(b)

Figure 4 The frequency response of the voice and unvoiced frames: (a) voiced with additive noise frame, (b) unvoiced frame with noise.



(a)



(b)

Figure 5 The spectrum of the voice and unvoiced frames after low-pass filtering: (a) Low-pass frequency response of voiced frame, (b) Low-pass frequency response of an unvoiced frame.

As mentioned earlier, the key concept for the proposed framework is to take advantage of the LPC coefficients to construct a band-pass filter that is able to adapt its parameters according to the frame in hand. The idea is that the bands have to concentrate around the formant frequencies and to suppress unwanted spectral components related to the noise.

The potential strength of the proposed method lies in the way of processing each frame. It utilizes a frame-depending strategy in extracting the real-speech part of the corrupted signal. This proposed strategy enables the system to generate a new adaptive filter for each speech frame, which in turn permits the system to look distinctly at each corrupted frame in order to generate the proper filter.

This is particularly crucial for signals (like speech), whose spectral features vary rapidly over time.

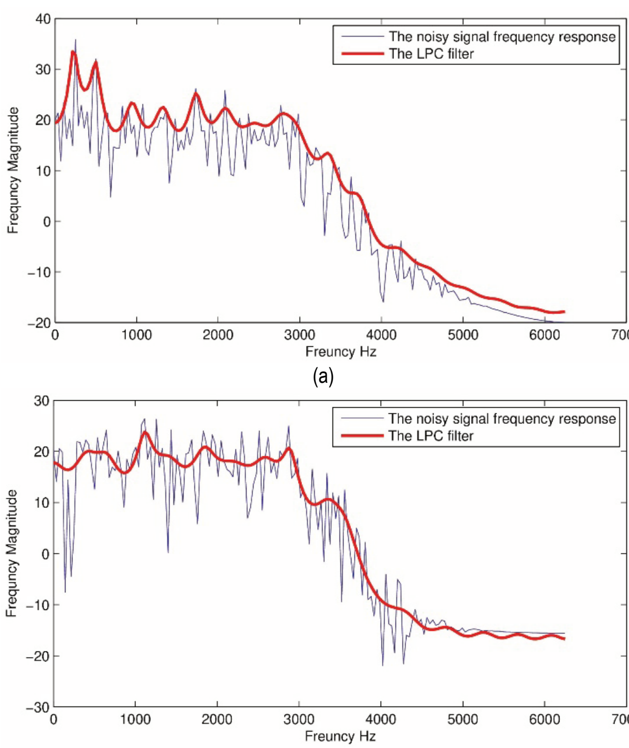
To evaluate the performance of the proposed system, standard objective and perceptual metrics, such as Signal-to-Noise Ratio (*SNR*) and Peak Signal-to-Noise Ratio (*PSNR*) (expressed in decibels (dB)), and Mean Squared Error (*MSE*), are used to qualify the system's capability to enhance the speech signal buried under additive noise.

Mathematically, these equations can be expressed as:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}\left(s[n]-\hat{s}[n]\right)^2 \tag{5}$$

$$SNR(\mathrm{dB}) = 10\log_{10}\left(\frac{\sum s^2[n]}{\sum\left(s[n]-\hat{s}[n]\right)^2}\right) \tag{6}$$

$$PSNR(\mathrm{dB}) = 10\log_{10}\left(\frac{(\max s[n])^2}{MSE}\right) \tag{7}$$

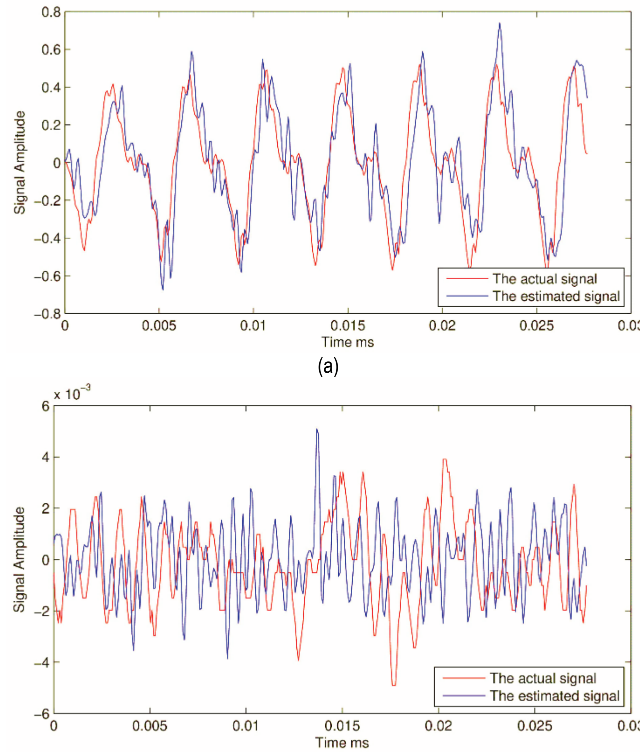


(a)

(b)

Figure 6 The frequency response and LPC envelope for voiced and unvoiced frames in db scale: (a) The voiced frame, (b) The unvoiced frame.

Fig. 7 presents a voiced and unvoiced corrupted frames with additive Gaussian white noise along with their corresponding enhanced frames achieved using the LPC-based adaptive filtering procedure with prediction order (p = 25).



(a)

(b)

Figure 7 The actual and the estimated signals for voiced and unvoiced frames utilizing the LPC-based adaptive filter: (a) The voiced frame, (b) The unvoiced frame.

The presented results are for frames with 25 ms duration and Gaussian white noise with an input *SNR* of 20 dB.

As shown in the figure, using the proposed system, the obtained *PSNR* is approximately 85% for the voiced speech and about 50% for the unvoiced speech. This indicates that the frames with voiced tones can be handled well by the LPC-based filter compared to the unvoiced speech.

Fig. 8(a)–(d) summarizes the results of speech estimation and enhancing the voiced and unvoiced frames using Wiener and Kalman filtering approaches. The assessment is done under the same noise conditions (*SNR* = 20 dB).

Using the Wiener filter, the obtained PSNR values are approximately 75% for the voiced speech and 40% for the unvoiced speech. Equivalently, the Kalman filter achieved *PSNR* values of approximately 77% for voiced speech and 42% for unvoiced speech.

According to the obtained results, we can note that all filters can enhance the corrupted signal, but with different levels of accuracy.

The problem is that the Wiener and Kalman filters are relying on fixed filter coefficients (or pre-set parameters) that make them unable to cope with the signal's unpredictability and unstable changes with time. So the need for an instantaneous-adaptation filter that is able to modify its parameters according to the signal at hand is very essential for unstable and highly fluctuated signal like speech.

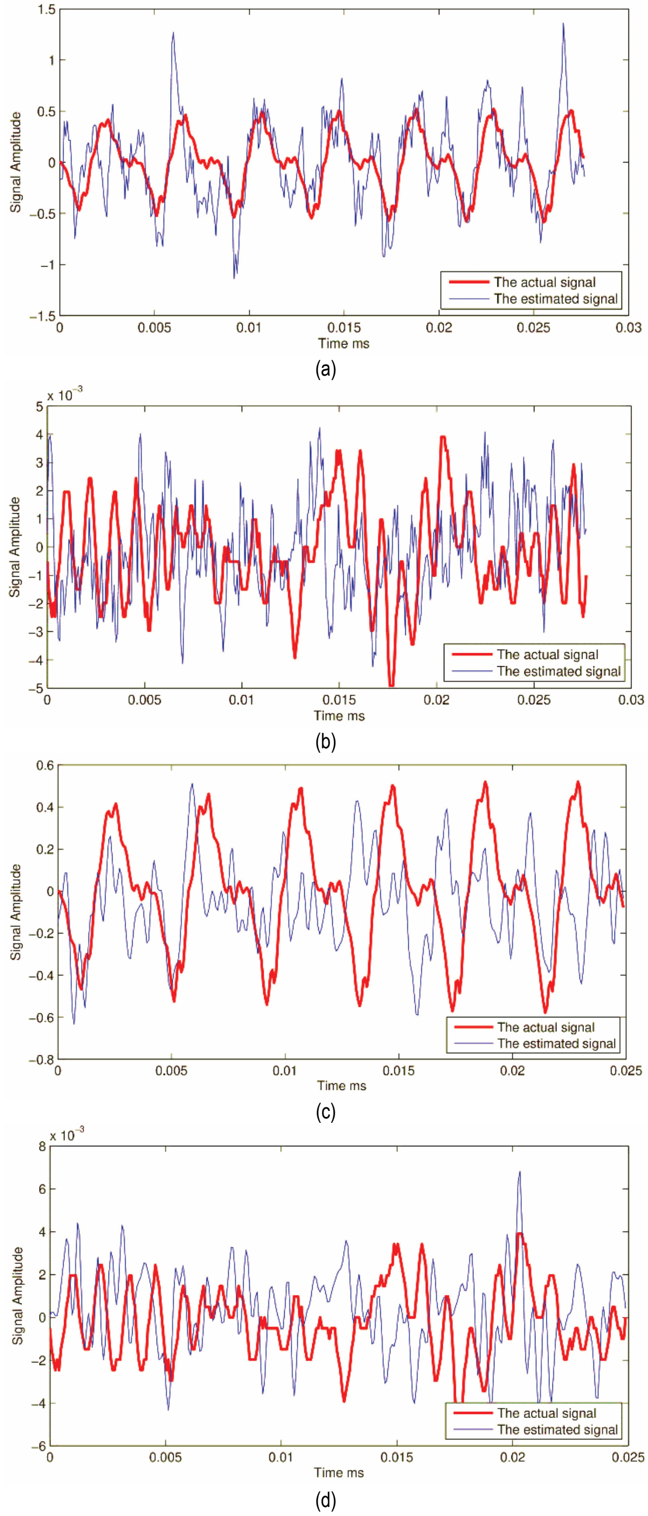


Figure 8 The real signals with their estimated counterparts using Kalman and Wiener filters: (a) real and estimated signal for voiced speech using Kalman filter, (b) real and estimated signal for unvoiced speech with Kalman filter, (c) real and estimated signal for voiced speech using Wiener filter, (d) real and estimated signal for unvoiced speech using Wiener filter.

Quantitatively, in the case of voiced speech, the mean squared error (*MSE*) values are approximately 0.0092, 0.018, and 0.024 for the LPC-based adaptive filter, Kalman filter, and Wiener filter, respectively. This demonstrates the exceptional performance of the proposed LPC-based approach in retaining voiced speech characteristics compared with other fixed-parameter filters.

As indicated in Figs. 7 and 8, the three filtering techniques, LPC-based, Kalman, and Wiener, have the ability to suppress the noise and enhance the speech signal with different levels of accuracy. The status is that nearly all methods exhibit excellent performance in terms of voiced speech. However, with the unvoiced speech frames, they all show limited performance.

Actually, we can note that the performance decline of the LPC-based technique to enhance the unvoiced speech attribute to the fact that this system is almost totally relying on the formant frequency envelope or the structure of the vocal tract. Since, it's well-know, that the unvoiced speech lacks a well-defined formant structure it's nearly noise, like speech.

Consequently, the LPC-based filter, particularly in this case, will be unable to distinguish between the real speech and the additive noise. Thus, the performance will decline dramatically.

We can attribute the remarkable performance of the LPC-based adaptive filter, especially with the voiced speech, to many key factors. First, it is different from the traditional Wiener and Kalman filters that totally utilize a previously set, unchangeable, optimized parameters.

The proposed method, on the other hand, builds up a dedicated filter that adapts its parameters depending on the frame it processes. This frame-wise adaptation allows the system to better track and identify the local spectral characteristics of the speech signal.

Second, the coefficients of the adaptive filter are estimated directly from the signal under analysis. This will allow the system to take advantage of the intrinsic acoustic information carried by the speech segment. Thus, the produced filter would be highly correlated with the instantaneous spectral envelope, or vocal tract resonant, of the speech signal, particularly in the case of voiced frames.

Nevertheless, the progress in the system performance is less noticeable in the case of unvoiced speech, which is because of the absence of the formant structures that adversely affect the LPC-based formants spectral modelling.

7 CONCLUSION

The paper presents three strategies for speech enhancement and background noise suppression: the LPC-based adaptive filter, the Wiener filter, and the Kalman filter. The latter filters are famous and widely used in the signal processing field, especially in signal enhancement and noise reduction. However, these filters are performed well under stationary noise conditions since they utilized a fixed or pre-set filter parameters that will never be changed or updated during the signal process. This will reduce their ability to tune with the nonstationary nature of speech signals.

The experimental results show that a frame-dependent adaptive filter with parameters that are directly derived from the speech signal provides excellent enhancement performance particularly for the voiced segments of speech. In other words, the proposed LPC-based adaptive filter technique exceeds the traditional fixed-parameter filtering approaches in noise suppression while retaining real-speech structure.

These practical results emphasize the importance of a signal-dependent adaptive filter for noise reduction and speech enhancement. Also, they point out the significant characteristics of the speech that can be carried by voiced segments. Consequently, LPC-based adaptive filtering could be a promising substitute to conventional fixed-parameters filters (Wiener and Kalman filters) for enhancing nonstationary signals.

8 SYSTEM LIMITATIONS AND PLANS FOR FUTURE

Although the proposed system produces imperfect performance with unvoiced segments of speech, practically, due to the absence of formant structure that is normally connected with the voiced speech. This will restrict the efficacy of the LPC-based filtering. Future work needs to focus on producing hybrid approaches that take advantage of the LPC-based filtering and prominent features of the unvoiced speech.

Acknowledgment

My colleagues and I in this article would like to thank the University of Mustansiriyah for supporting this work and the reviewers for providing useful suggestions, allowing for the improved presentation of this work.

9 REFERENCES

[1] Haque, M., & Bhattacharyya, K. (2018). Speech background noise removal using different linear filtering techniques. In S. Bhattacharyya, T. Gandhi, K. Sharma & P. Dutta (ed.), *Advanced computational and communication paradigms* (pp. 297–307). Springer Singapore. https://doi.org/10.1007/978-981-10-8240-5_33

[2] Dutoit, T. (1997). Linear prediction synthesis. In *Speech synthesis* (pp. 201–228). Springer Netherlands. https://doi.org/10.1007/978-94-011-5730-8_8

[3] Migimatsu, K., & Tokuda, I. (2019). Experimental study on nonlinear source–filter interaction using synthetic vocal fold models. *The Journal of the Acoustical Society of America, 146*(2), 983–997. https://doi.org/10.1121/1.5120618

[4] Proakis, J. G., & Manolakis, D. G. (1996). *Digital signal processing: Principles, algorithms, and applications* (3rd Ed.). Prentice Hall.

[5] Taherian, H., Wang, Z.-Q., Chang, J., & Wang, D. (2020). Robust speaker recognition based on single-channel and multi-channel speech enhancement. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 28*, 1293–1302. https://doi.org/10.1109/TASLP.2020.2986896

[6] Ochiai, T., Iwamoto, K., Delcroix, M., Ikeshita, R., Sato, H., Araki, S., & Katagiri, S. (2024). Rethinking processing distortions: Disentangling the impact of speech enhancement errors on speech recognition performance. *arXiv*. https://arxiv.org/abs/2404.14860 https://doi.org/10.1109/TASLP.2024.3426924

[7] Wijayasingha, L., & Stankovic, J. A. (2021). Robustness to noise for speech emotion classification using CNNs and attention mechanisms. *Smart Health, 19*, 100165. https://doi.org/10.1016/j.smhl.2020.100165

[8] Lu, X., Tsao, Y., Matsuda, S., & Hori, C. (2013). Speech enhancement based on deep denoising autoencoder. In *Proceedings of Interspeech 2013* (pp. 436–440). https://doi.org/10.21437/Interspeech.2013-130

[9] Gutiérrez-Muñoz, M., & Coto-Jiménez, M. (2022). An experimental study on speech enhancement based on a combination of wavelets and deep learning. *Computation, 10*(6). https://doi.org/10.3390/computation10060102

[10] Karthik, A. M. I. J. (2023). An optimized convolutional neural network for speech enhancement. *International Journal of Speech Technology, 26*(4). https://doi.org/10.1007/s10772-023-10073-6

[11] Haque, M., & Bhattacharyya, K. (2018). Speech background noise removal using different linear filtering techniques. In S. Bhattacharyya, T. Gandhi, K. Sharma & P. Dutta (Ed.), *Advanced computational and communication paradigms* (pp. 297–307). Springer Singapore. https://doi.org/10.1007/978-981-10-8240-5_33

[12] Das, O., Goswami, B., & Ghosh, R. (2016). Application of the tuned Kalman filter in speech enhancement. In *IEEE First International Conference on Control, Measurement and Instrumentation (CMI2016)* (str. 62–66). https://doi.org/10.1109/CMI.2016.7413711

[13] El-Fattah, M., Dessouky, M., Abbas, A., Diab, S., El-Rabaie, E.-S., Al-Nuaimy, W., Alshebeili, S., & Abd El-Samie, F. (2014). Speech enhancement with an adaptive Wiener filter. *International Journal of Speech Technology, 17*, 53–64. https://doi.org/10.1007/s10772-013-9205-5

[14] Sailaja, B. L. V., & Sunitha, P. (2019). Wiener filtering in frequency domain to enhance speech corrupted by colored noise. *International Journal of Recent Technology and Engineering*.

[15] Kondoz, A. M. (2004). Efficient LPC quantization methods. U *Speech coding and synthesis* (Chapter 5, pp. 87–147). John Wiley & Sons. https://doi.org/10.1002/0470870109.ch5

[16] Gold, B., & Morgan, N. (2011). *Speech and audio signal processing: Processing and perception of speech and music*. John Wiley & Sons. https://doi.org/10.1002/9781118142882

[17] Oudre, L. (2015). Automatic detection and removal of impulsive noise in audio signals. *Image Processing On Line, 5*, 267–281. https://doi.org/10.5201/ipol.2015.64

[18] Haykin, S. S. (2014). *Adaptive filter theory*. Pearson.

[19] Kalamani, M., & Krishnamoorthi, M. (2021). Modified least mean square adaptive filter for speech enhancement. In N. Dey (Ed.). *Applied speech processing* (pp. 47–73). Academic Press. https://doi.org/10.1016/B978-0-12-823898-1.00004-7

[20] Makhoul, J. (1975). Linear prediction: A tutorial review. *Proceedings of the IEEE, 63*(4), 561–580. https://doi.org/10.1109/PROC.1975.9792

[21] Huang, X., Acero, A., & Hon, H.-W. (2001). *Spoken language processing: A guide to theory, algorithm, and system development*. Prentice Hall PTR.

[22] Das, D. (2022). A systematic review of the effects of noise characteristics on human mental performance. In D. Chakrabarti, S. Karmakar & U. R. Salve (Ed.), *Ergonomics for design and innovation* (pp. 1705–1717). Springer International Publishing. https://doi.org/10.1007/978-3-030-94277-9_145

Authors' contacts:

Rehab I. Ajel
(Corresponding author)
Department of Computer Science, College of Science,
Mustansiriyah University,
Palastine St, 10064 Baghdad, Iraq
E-mail: rihabajel@uomustansiriyah.edu.iq

Tariq A. Hassan
Department of Computer Science, College of Education,
Mustansiriyah University,
Palastine St, 10064 Baghdad, Iraq
E-mail: tariq-mono2016@uomustansiriyah.edu.iq