Performance Analysis of Adaptive Algorithms for Speech Enhancement Applications

Rashmirekha Ram* and Mihir Narayan Mohanty

Department of Electronics and Communication, ITER, Siksha 'O' Anusandhan University, Bhubaneswar - 751030, Odisha, India; ram.rashmirekha14@gmail.com, mihirmohanty@soauniversity.ac.in

Abstract

Background/Objectives: The improvement of the speech quality is an essential component in this modern world. Though different methods have been used by many researchers since a long period, still the research field is open to work further. In this paper, we have analyzed the same using adaptive algorithms. Methods/Statistical Analysis: The algorithms like Least Mean Square (LMS) and Recursive Least Square (RLS) are used to improve the quality. Further, State-Space Recursive Least Square (SSRLS) algorithm is developed for enhancement of the noisy speech. The existing Spectral Subtraction (SS) algorithm is verified for two types of noisy speech. Also the standard adaptive algorithms have been experienced for these signals. Findings: Spectral Subtraction method is implemented initially. Then LMS and RLS algorithms are tested for two different types of noisy speech signals. In case of LMS, the improvement of SNR is 8.2822 dB for random noise and 6.5245 dB for babble noise. Whereas for RLS algorithm, the SNR improvement is 11.7152 dB for random noise and 14.3883dB for babble noise. The SNR after enhancement is 43.5409 dB and 47.2456 dB for random noise and for babble noise respectively in case of SSRLS algorithm which is much higher than other algorithms as results suggest. Also the error minimization curve converges more rapidly than the others. So SSRLS algorithm provides better enhanced signal than the other three methods. **Application/Improvements:** The applications are in different field of communications like mobile telephony, video conferences as audio and video applications. Also the applications are of speech recognition, voice communications, information forensics and many more. The improvement of SNR in this proposed algorithm found better than SS, LMS and RLS algorithms as 31.2377 dB and 35.6381 dB.

Keywords: Adaptive Filter, Least Mean Squares, Noise, Recursive Least Squares, Spectral Subtraction, Speech Enhancement, State-Space Recursive Least-Squares

Introduction

Reduction of background noise is still a challenging task for speech enhancement which affects the performances of communication systems like mobile phones, voice communications etc. Speech enhancement is pertained with improving the perceptual part of speech that has been dissipated by various noises. In many applications, such as mobile telephony, speech communication, information forensics, speech recognition, video conferences etc., the speech enhancement algorithms improve the quality of dissipated speech. To obtain clean speech, different enhancement algorithms have been used since several decades¹. But still it is a challenging task. Among all the algorithms Spectral Subtraction (SS) method is the first method used for speech enhancement. But it is not suitable for nonstationary noise^{2,3}. As the characteristics of speech signal changes over time. So, adaptive algorithms are well suited for speech enhancement. Due to simplicity in computation and implementation, Least Mean Square (LMS) and Normalized Mean Square (NLMS) algorithms have been generally utilized in signal processing application. As the Mean Square Error is increased linearly with the signal power, so Recursive Least Square (RLS) algorithm is used. This method provides the better convergence rate but computational complexity is more. Different comparisons are also made among different adaptive algorithms^{4.5}.

Adaptive filtering methods are commonly used for spectral domain processing. Several filtering techniques have already been used for enhancement process. The Least Mean Square (LMS) and the Recursive Least Squares (RLS) algorithms are the most prevalent algorithms in the field of adaptive signal processing. RLS algorithm gives better performance in comparison with LMS algorithm in terms of convergence rate. But computational complexity is more than LMS algorithm. So the decision of choosing the adaptive algorithms always creates difficulty^{6.7}.

Spectral subtraction is the first and simple algorithms used for speech enhancement. To get the clean speech, the noise spectrum is estimated from the speech silence regions. The remaining musical noise is the main drawback of SS algorithms. A number of modifications have been made of this method since several years. Navneet Upadhyay and Abhijit Karmakar have presented different spectral subtractive algorithms. Different objective and subjective tests were performed for both stationary and non-stationary noises⁸.

To implement Minimum Variance Distortionless Response (MVDR) beam former Jwu Sheng Hu et.al. has developed a second order extended H_{∞} filter. The second order extended (SOE) H_m filter is designed to minimize the signal disturbance effects without a prior knowledge of the signal⁹. For real time signals, the statespace observer was derived from the beam former and from this the SOE Kalman filter was derived. This filter was helpful in assuming the statistics and dynamics of the noise signals. However, the denoising performance is not up to a remarkable level, when the signal affected by multiple noises. To solve such types of problems, adaptive self-switching noise removal techniques were used in¹⁰. The performances of different adaptive filters were compared with the optimization technique. Different types of noise are added to the speech signal and the denoising performance was evaluated in terms of SNR. So a speech enhancement method is developed using adaptive algorithms and optimization techniques.

A new perceptual speech enhancement method based on an improved spectral subtraction filter and a spectral attenuation filter is proposed in¹¹. Both musical noise and speech distortions are less in the enhanced speech signal. Speech signal is non-stationary in nature. So the speech signal is divided into frames based on Empirical Mode Decomposition (EMD) for speech filtering. The filtering method was developed by using the combined effects of EMD and Adaptive Centre Weighted Average (ACWA) filter. The Intrinsic Mode Functions (IMFs) was generated by EMD for each frame. The voiced and unvoiced frames are distinguished by filtered IMFs. The voiced frames are identified by energy whereas an unchanging index differentiates between unvoiced and transient sequences. In¹² shows that the frame class is amazingly convincing in expelling the noise constitutive from noisy speech signal.

The filter weights are updated using conventional LMS which helps to reduce the norm. SSRLS algorithm is an extension of state space representation of RLS algorithm. But the assumption is that the corrupted output signal is available for testing. To assuage matrix inversion, Ricatti equation using matrix inversion lemma has used which provides better convergence than RLS algorithm¹³.

The SSRLS algorithm is a new kind of adaptive filter. It provides better tracking operation than RLS algorithm. The authors have designed a state-space model for RLS which was capable of tracking linear time varying systems¹⁴. In order to palliate the matrix inversion associated in SSRLS, a Ricatti equation using matrix inversion lemma was derived and the observer gain for SSRLS based on Kalman gain was also developed^{15,16}.

For noise cancellation and enhancement of speech subband filter, spatial subtraction and voice activity detector have used. To test the performance of these algorithms two different real time speech signals are tested which contains different types of noise signals. Results show that the surrounding noise decreases by the proposed method. A LSB technique is used by the researchers embed text data in audio file by using the steganography techniques. Only in specific bit positions LSB technique is used. The information security is enhanced by this process. A very little distortion is there in the original audio signal and stegno audio signal^{17.18}.

Adaptive filtering technique is an important method of spectral domain processing. Several filtering techniques have been already used for enhancement process. The Recursive Least Squares (RLS) algorithm is the most prominent algorithms among all the adaptive algorithms in the field of adaptive signal processing. Among all the adaptive algorithms the RLS algorithm provides fast convergence rate than the other adaptive algorithms. The decision of selecting the adaptive algorithms always creates a trade-off^{19,20}.

The rest paper is organized as follows. The next section describes the concept of speech enhancement, spectral subtraction and RLS algorithm; Section 3 describes the SSRLS algorithm for speech enhancement. The results are summarized in Section 4 and Section 5 concludes the work.

2. Speech Enhancement Methods

To enhance the quality of dissipated speech signal by using several algorithms is the main objective of the speech enhancement methods. So the speech must be clear and understandable. Speech enhancement has many practical applications like hearing aids design, forensic labs, speech recognition, voice communication, etc. To get the clean speech signal different algorithms are designed for this purpose. Figure 1 shows the basic idea of speech enhancement.



Figure 1. Proposed speech enhancement method.

In this work the basic spectral subtraction method is implemented and verified. Next LMS, RLS, SSRLS adaptive algorithms are applied for enhancement of noisy speech. Then the comparison is made among them in the result section.

1.1 Spectral Subtraction Method

The speech and noise signals are uncorrelated in spectral subtraction method. Estimation of noise is attained from the pauses in the speech signal.

1.1.1 Power Spectral Subtraction

Assuming the phase difference between the clean speech and noise signal is zero. The power spectrum can be given by:

$$\left|S(e^{j\omega})\right|^{2} = \left|Y(e^{j\omega})\right|^{2} - E\left\{\left|N(e^{j\omega})\right|^{2}\right\}$$
(1)

1.1.1 Magnitude Spectral Subtraction

Assuming the phase difference between the clean speech and noise signal is one. The magnitude spectrum can be given by:

$$E\left\{\left|Y(e^{j\omega})\right|^{2}\right\} = E\left\{\left|S(e^{j\omega})\right|^{2}\right\} + E\left\{\left|N(e^{j\omega})\right|^{2}\right\}$$
(2)

Where $|S(e^{j\omega})| =$ power spectrum of clean signal, $|Y(e^{j\omega})| =$ power spectrum of the noisy signal and $N(e^{j\omega})| =$ power spectrum of the noise.

The general form of the spectral subtraction algorithm is:

$$\left|S(e^{i\omega})\right|^{a} = \left|Y(e^{i\omega})\right|^{a} - E\left\{\left|N(e^{i\omega})\right|^{a}\right\} \text{ for } a > 0$$
(3)

Where a is the power exponent,

When a = 1 results the magnitude spectral subtraction and a = 2 gives the power spectral subtraction algorithm.



Figure 2. Block diagram of spectral subtraction algorithm.

Following steps are to be followed for spectral sub-traction algorithm:

- The noisy speech signal is splited up into frames with an overlap of signal.
- The hamming window is multiplied with the frame based signal.
- The magnitude and phase spectrum of the windowed signal is calculated with the help of Fourier transform.
- Then the noise spectrum is estimated from the initial region of the speech (i.e. from speech pauses)
- The estimated noise spectrum is subtracted from the input noisy signal to get the clean speech.
- The signal is reconstructed by overlap add method and enhanced signal is achieved by using synthesis window.

The block diagram of the spectral subtraction algorithm is shown in Figure 2.

The spectral subtraction algorithm incorporates only forward and inverse fourier transform. So it is computationally simple and easy.

2.2 LMS Adaptive Filter

For noise reduction, the LMS adaptive filter is widely used. It is a gradient descent algorithm proposed by Widrow¹⁸. The filter coefficients are adjusted depending on the gradient of the error surface. The following steps are involved in LMS algorithm.

The output signal y(m) of the adaptive filter is calculated.

• The error signal *a*(*m*) is calculated by using the following Equation:

$$a(m) = y(m) - d(m) \tag{4}$$

Where d(m) is the desired signal.

• The filter coefficients are updated by utilizing the following equation:

$$\vec{w}(m+1) = \vec{w}(m) + \mu . a(m) i(m)$$
 (5)

Where $\vec{w}(m)$ is the filter coefficient vector, μ is the step size of the adaptive filter and $\vec{i}(m)$ is the filter input vector.

This algorithm is computational efficient with adjustable parameters. As there is no matrix operation is involved, so it is easy to implement.

2.3 RLS Adaptive Filter

In general the adaptive filters do not have constant filter coefficients. So the a priori information of the signal or noise is not needed. This filter has adaptable parameters so it is capable of adjusting the filter coefficients. In general a filter is designed to extract or enhance the desired information from a signal. Most adaptive filters are digital filters. It can adjust the transfer function by itself. In case of adaptive filters the error signal is used as feedback and cost function is used for optimum performance of the filter.

The FIR and IIR filters are used for adaptive filtering. But the FIR filter is generally more utilized. Because FIR filters have adjustable zeros, so there are no constancy problems. However the adaptive FIR filters are not always stable. The stability of the filter devolves on the algorithm used.

The general adaptive filter configuration is shown in Figure 3.



Figure 3. Adaptive filter configuration.

Recursive Least Squares (RLS) algorithm uses all the information present in the input signal. It is recursive because the present coefficients update by using the past coefficients. It provides faster convergence rate than the LMS algorithm. The RLS adaptive algorithm is mostly used for determining the coefficients of an adaptive filter. A weighting factor is used to reduce the influence of present sample from the past samples^{19,20}.

To update the coefficients of the filter the RLS algorithm has to follow the following steps.

The output signal y(m) of the adaptive filter is calculated.

• The error signal *a*(*m*) is calculated by using the Equation:

$$a(m) = y(m) - d(m) \tag{6}$$

Where d(m) is the desired signal.

• To update the filter coefficients following equation is applied:

$$\vec{w}(m+1) = \vec{w}(m) + a(m) \vec{M}(m)$$

$$(7)$$

Where the filter coefficients vector is w(m) and $\vec{M}(m)$ is determined by the Equation:

$$\vec{M}(m) = \frac{I(m)\vec{u}(m)}{\lambda + \vec{u}(m)I(m)\vec{u}(m)}$$
(8)

Where λ is the forgetting factor and I(m) is the inverse correlation matrix of the input signal.

I(m) as the following initial value I(0) used.

$$I(0) = \begin{bmatrix} \delta^{-1} & 0 \\ & \delta^{-1} \\ 0 & \delta^{-1} \end{bmatrix}$$

Where δ is the regularization factor. To update the inversion matrix the following Equation is used.

$$I(m+1) = \lambda^{-1} I(m) - \lambda^{-1} \vec{M}(m) \vec{u}(m) I(m)$$
(9)

2.4 SSRLS Adaptive Filter

The SSRLS algorithm is another expansion to the group of adaptive filters. It is a time varying filter which is computationally acute and based on least squares method. This filter is also recursive in nature. The present coefficients depend on the past coefficients. But difference is that a less weightage is given to the past coefficients when the observations become old. To compute the process state estimate s[p] from the output vector y[p], the prior state estimate $\overline{s}[p]$ is given by:

$$\overline{s}[p] = A \, s[p-1] \tag{10}$$

Where A is the state transition matrix and the assumption is that the matrix is uniquely determined and invertible. Similarly for output observation the *a prior* estimate is calculated by:

$$\overline{y}[p] = U\overline{s}[p] = UAs[p-1] \tag{11}$$

The pair (A, U) is considered to be *p*-step observable and the prediction error is calculated by the Equation:

$$e[p] = y[p] - \overline{y}[p] \tag{12}$$

Then the input vector estimation is calculated by using

$$s[p] = \overline{s}[p] + R[p]e[p] \tag{13}$$

Where R[p] is the observer gain and is decided by the least squares method.

After getting the input estimation vector and error vector t is required to calculate the recursive update of correlation matrices of $\Psi[p]$ and $\Psi^{-1}[p]$ as in (11) and the Equation is:

$$\Psi[p] = \beta \left(A^{-T} \Psi[p-1] A^{-1} + U^{T} U \right)$$
(14)

To calculate the observer gain the following Equation is used.

$$P[p] = \Psi^{-1}[p]U^{T}$$
⁽¹⁵⁾

A less weightage is given to the past values and a forgetting factor (λ) less than 1 is used¹³. The weighting matrix is given as:

$$\lambda^{k} I_{m} \quad 0 \quad \cdots \quad 0 \quad 0 \\ 0 \quad \lambda^{k-1} I_{m} \quad 0 \quad 0 \\ W[p] = \vdots \quad \ddots \quad \vdots \\ 0 \quad 0 \quad \lambda I_{m} \quad 0 \\ 0 \quad 0 \quad \cdots \quad 0 \quad I_{m}$$
(16)

Where I_m is the identity matrix of dimension $m \times m$.

3. Results and Discussions

Different speech signals are recorded by different students of ITER in noisy environment. To enhance the noisy signal by different adaptive algorithms the speech signals 'Hello everything is fine' and 'Can I know your name please' are tested in this work. Figure 4 and Figure 5 show the noisy speech signal. As the spectral characteristics of speech signal changes over time so recorded voices are of duration 3-4 seconds are taken. The random noise is added with the first and the second speech contains babble noise. Then the signal is splited up into frames. The framed speech signal is shown in Figure 6. The 'for loops' are processed through each frame individually. For getting a smooth and continuous signal, the frames are overlapped with each other. In SS algorithm an overlap of 40% is taken. Hamming window is multiplied to obtain the FFT of the resultant speech signal. The enhanced signal by SS algorithm is shown in Figure 7.



Figure 4. Original signal for sentence 'hello everything is fine'.

With an objective to enhance the original speech signal by LMS algorithm the order of the filter is 8 and the value of μ is 0.02 in LMS filter is taken. Figure 8 and Figure 9 shows the filtered output signal and the learning curve respectively. Figure 10 shows the comparison of desired signal and the LMS adaptive output signal. The SNR of the filtered signal is calculated and shown in Table 1 for random noise and Table 2 for babble noise.



Figure 5. Babble noise is added with speech 'can I know your name please'.



Figure 6. Noisy signal (signal+random noise) after framing.



Figure 7. Enhanced signal by SS algorithm.



Figure 8. Desired signal for adaptive algorithms.

The SNR improvement is 8.2822 dB for random noise and 6.5245 dB for babble noise are obtained.

In RLS algorithm the adaptive filter coefficients are to be updated. RLS algorithm is tested for the same speech signals. Figure 11 shows the desired signal which is input to the LMS, RLS and SSRLS algorithms. For RLS algorithm a forgetting factor of 1 and a regularization factor 0.01 are used to provide the desired accuracy and the filter is of order 10 are used. The output obtained using RLS algorithm is shown in Figure 12. After 15-20 iterations the MSE curve converges in RLS algorithm which is better than LMS algorithm. Figure 13 shows the MSE curve of RLS algorithm. Also the SNR improvement is better than the LMS algorithm as mentioned in Tables. The comparison of desired signal and the RLS output signal is shown in Figure 14.



Figure 9. MSE curve using LMS algorithm.



Figure 10. Comparison of desired signal and LMS output signal.



Figure 11. Enhanced signal by LMS algorithm.

A forgetting factor of 0.99 and the filter order of 15 is used in case of SSRLS. The enhanced signal is obtained using SSRLS algorithm is shown in Figure 15. Figure 16 shows the comparison of desired signal and SSRLS output signal of a single frame. The error minimization curve of random noisy signal using SSRLS algorithm is shown in Figure 17. The SNR after enhancement is 43.5409 dB for random noise and 45.2456 dB for babble noise. The improvement of SNR values is better than the other algorithms. A better SNR also we get from the other algorithms. The babble noise is removed more than the random noise. The error is also less in SSRLS algorithm as compared to LMS, RLS and spectral subtraction. So SSRLS is better than LMS, RLS and spectral subtraction. So SSRLS algorithm is better than LMS and RLS algorithms.



Figure 12. Enhanced signal by RLS algorithm.



Figure 13. MSE curve using RLS algorithm.



Figure 14. Comparison of desired signal and RLS output signal.

The error minimization curve of Figure 6 is shown in Figure 18 for LMS algorithm. After 30-40 iterations are taken place. But after 18-20 iterations the MSE curve converges. In SSRLS, the MSE curve converges better than the other two algorithms.



Figure 15. Adaptive output using SSRLS algorithm.



Figure 16. Comparison of desired signal and SSRLS output signal.



Figure 17. MSE curve using SSRLS algorithm.



Figure 18. MSE curve using LMS algorithm.



Figure 19. MSE curve using RLS algorithm.



Figure 20. MSE curve using SSRLS algorithm.

Table 1.	Comparison of different adaptive algorithms
in terms	of SNR for random noise

Adaptive Algorithms	SNR_ before (dB)	SNR_ after (dB)	SNR Improvement
			(ab)
SS	12.6252	16.7659	4.1407
LMS	12.3608	20.6430	8.2822
RLS	12.8974	24.6126	11.7152
SSRLS	12.3032	43.5409	31.2377

 Table 2. Comparison of different adaptive algorithms

 in terms of SNR for babble noise

Adaptive Algorithms	SNR_before (dB)	SNR_after (dB)	SNR Improvement (dB)
SS	11.0365	15.3400	4.3035
LMS	11.7074	18.2319	6.5245
RLS	11.3629	25.7512	14.3883
SSRLS	11.6075	47.2456	35.6381

4. Conclusion

In this work different adaptive algorithms are tested for two different types of noisy signals. First speech contains the random noise and the second speech contains the babble noise. Spectral subtraction method is first implemented to enhance the noisy speech signal. Then LMS, RLS and SSRLS adaptive algorithms are applied for enhancement. The MSE of both the signals are shown in different figures. The listening test has also performed for both signals. Table 1 and Table 2 shows the SNR value for different algorithms. By comparing the result, SSRLS algorithm provides better enhanced signal than the other algorithms. Also the improvement of SNR is found better in case of SSRLS algorithm.

5. References

- 1. Haykin S. Adaptive filter theory. 3rd ed. Upper Saddle River: Prentice Hall; 1996.
- Boll SF. Suppression of acoustic noise in speech using spectral subtraction. IEEE Transaction ASSP; 1979 Apr. p. 113–20.,
- 3. Li J, Sakamoto S, Hongo S, Akagi M, Suzuki Y. Adaptive β -order generalized spectral subtraction for speech enhancement. Signal Processing. Elsevier Journal No. 88; 2008. p. 2764–76.
- Hadei SA, Lotfizad M. A family of adaptive filter algorithms in noise cancellation for speech enhancement. International Journal of Computer and Electrical Engineering. 2010 Apr; 2(2):307–15.
- 5. Caiyun W. Study of speech enhancement system using combinational adaptive filtering. International Conference on Computer Distributed Control and Intelligent Environmental Monitoring; 2012.
- Santosh DHH, Pendyala VS, Kumar VNL, Rao NS. Statistical approach for noise removal in speech signals using LMS, NLMS, Block LMS and RLS Adaptive filters. International Journal of Modeling and Optimization. 2012 Jun; 2(3):351–5.
- 7. Ram R, Mohnaty MN. Speech enhancement using RLS algorithm based adaptive filter. IEEE Conf, ICECS; 2016.
- Upadhyay N, Karmakar A. Speech enhancement using spectral subtraction-type algorithms: A comparison and simulation study. Elsevier, Procedia Computer Science. 2015; 54:574–84.
- 9. Hu JS, Lee MT, Yang CH. Robust adaptive beamformer for speech enhancement using the second-order extended H_{∞} filter. IEEE Transactions on Audio, Speech and Language Processing. 2013; 21(1).

- Anoop V, Rao PV, Nidhin M. Performance analysis of speech enhancement methods using adaptive algorithms and optimization techniques. IEEE Conference, ICCSP; 2015.
- 11. Alaya S, Zoghlami N Lachiri Z. Adaptive filter for perceptual speech enhancement. IEEE International Multi-Conference on Systems, Signals and Devices; 2015.
- Khaldi K, Boudraa AQ, Turki M. Voiced/unvoiced speech classification-based adaptive filtering of decomposed empirical modes for speech enhancement. IET Signal Processing. 2016; 10(1):69–80.
- Malik MB. State-space recursive least-squares: Part I. Signal Processing. 2004; 84:1709–18.
- 14. Malik MB. State-space recursive least squares. IEEE Conf on Acoust Speech Signal Process; 2000. p. 645–8.
- Haykin S, Syed AH, Zeidler J, Yee P, Wei P. Adaptive tracking of linear time-variant systems by extended RLS algorithms. IEEE Trans Signal Process. 1997 May; 45(6):1118–28.

- 16. Sayed AH, Kailath T. A state-space approach to adaptive RLS filtering. IEEE Signal Process Mag. 1994; 11(3):18–60.
- Dhiman J, Ahmad S, Gulia K. Comparison between adaptive filter algorithms (LMS, NLMS and RLS). IJSETR. 2013 May; 2(5):1100–3.
- Soujanya BM, Rajasekhara Rao Ch, Sastry DVLN. Speech enhancement using combinational adaptive LMS algorithms. International Journal of Advanced Computer Research. 2015 Mar; 5(18):100–4.
- Ganesan V, Manoharan S. Surround noise cancellation and speech enhancement using sub band filtering and spectral subtraction. Indian Journal of Science and Technology. 2015 Dec; 8(33).
- Rameshkumar P, Monisha M, Santhi B. Enhancement of information hiding in audio signals with efficient LSB based methods. Indian Journal of Science and Technology. 2014 Apr; 7(S4):80–5.