INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & MANAGEMENT

ONLINE SPEECH ENHANCEMENT USING DIFFERENT ADAPTIVE FILTERS

Parag Acharya[1], Dr. Sandeep Agrawal[2]
M.E. Student, MIT, Ujjain[1], Prof. ECE Dept. MIT, Ujjain[2]
Department of Electronics and Communication Engineering
MIT Ujjain (Madhya Pradesh) India
paragachaya2011@gmail.com

ABSTRACT

Adaptive filtering has become a spacious area of researcher since last few decades in the field of communication. Adaptive noise cancellation is an approach used for noise reduction in speech signal. The speech signal easily gets contaminated with background noise. Channel noise addition makes this speech signal even poorer. Speech signal and noise signal both change continuously with time, then to separate them only adaptive filtering is desirable. This paper deals with cancellation of noise on speech signal using two old (LMS and NLMS) and one new UNANR algorithm. The UNANR (Unbiased and Normalized Adaptive Noise Rejection) model does not contain any bias unit, and the coefficients are adaptively updated by using the steepest-descent algorithm. Two modulation techniques, AM and FM are applied separately in combination with two communication channels i.e. AWGN and Rician. Signal quality parameter PSNR and RMSE measured and compared with respect to SNR. The results show that the performance of the UNANR based algorithm is superior to that of the LMS algorithm in noise reduction.

KEYWORDS: Adaptive filtering, LMS, NLMS, UNANR, PSNR, RMSE.

INTRODUCTION

Speech is a form of communication in daily life. Speech is most natural form of human communication. It existed since human civilizations began and even till now. The perception of speech signal is usually measured in terms of its quality and intelligibility. The quality is a subjective measure that indicates the pleasantness or naturalness of the perceived speech. Intelligibility is an objective measure which predicts the percentage of words that can be correctly identified by listeners. Enhancement means the improvement in the value or quality of something. When applied to speech, this simply means the improvement in intelligibility and/or quality of a degraded speech signal by using signal processing tools. By speech enhancement, it refers not only to noise reduction but also to de reverberation and separation of independent signals. This is a very difficult problem for two reasons. First, the nature and characteristics of the noise signals can change dramatically in time and between applications [1]. It is also difficult to find algorithms that really work in different practical environments. Second, the performance measure can also be defined differently for each application. Two criteria are often used to measure the performance quality and intelligibility.

A. Speech Enhancement

The main objective of speech enhancement technique is to improve the quality and minimize the loss in intelligibility of the signal and listener fatigue. The basic overview is shown in Figure 1. Continuous improvement of communication and multimedia systems has led to the widespread use of speech recording and processing devices, e.g., mobile phones, speech recognition tools. In most practical situations, these devices are being used in environments where undesirable background noise exists. Degraded speech can cause problems for both mobile communication and speech recognition systems. Nowadays, all the people use the communication devices almost as a primary good: telephones, mobiles, internet and the customers demand a high coverage and quality [2,3].

Fig. 1: Basic Speech Enhancement System

Speech enhancement in general has three major objectives

(a) To improve the perceptual aspects such as quality and intelligibility of the processed speech, i.e. to make it sound better or clearer to a human listener, this will in turn result in the reduction of listening fatigue.
(b) To improve the robustness of speech coders which tend to be severely affected by the presence of noise.
(c) To increase the accuracy of speech recognition systems operating in less than ideal locations. However, the background noise is an important handicap.

SIMULATION BLOCK DIAGRAM

The following block diagram gives the complete idea of the project. The major units are modulation, communication channel and adaptive filters. This project comprises of two types of input voice signal: stored voice signal and microphone voice signal.

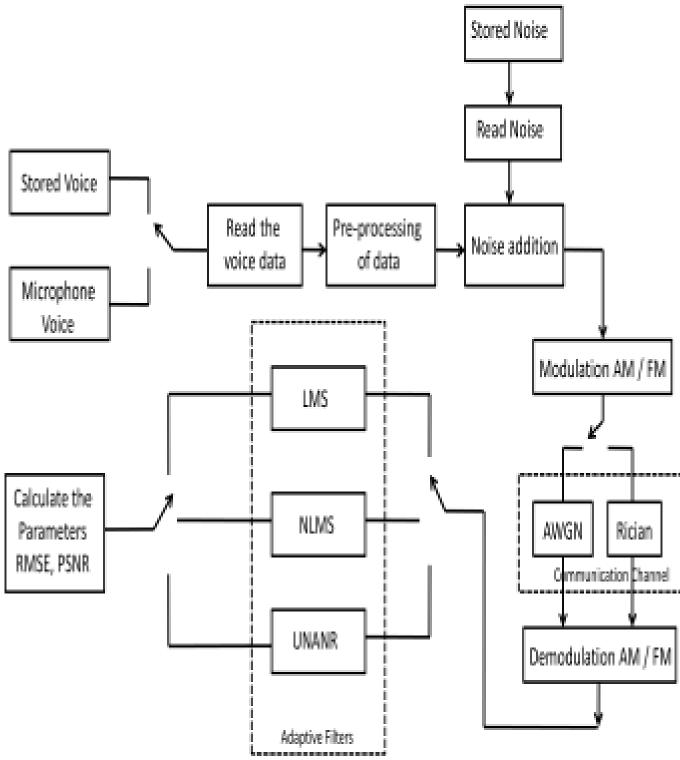

Fig. 2: Simulation Block Diagram

The stored voice signal is a wave file which is stored in the computer and microphone voice signal is a speech input from microphone. At a time only one type of input signal is selected. This input speech signal is processed and audible with signal graph in graph window of the project UI. A noise signal is mixed with this speech signal. The SNR value can be of user's choice as it is controlled by user while mixing noise signal. In other words noise signal level is set by user while mixing the noise into the speech signal. After addition of noise and pure speech signal modulation technique is applied. This modulated signal can be seen in graph window of the project UI. Then this modulated signal is sent through the one of the communication channel. At the receiver side received signal is first demodulated and the filtered with one of the three adaptive filters.

SIMULATION RESULTS AND COMPARISON

The performances of the adaptive filters are compared with respect to the variation in SNR (dB). The used modulation techniques are AM and FM and the considered channels are AWGN and Rician fading channel. Under speech enhancement techniques, for improving quality of adaptive filters a newly emerging filter is used i.e. UNANR. This filter's performance is compared with two traditionally used adaptive filters; LMS and NLMS. The above considered technologies have been combined using the MATLAB software (Version 7.10). Now for different cases for the performance evaluation, the selected range of SNR is 38 to -2.5. However there is no restriction of the SNR range. But, if SNR range increases then the simulation time will increase and the considered noise removal capability may decrease. It is necessary to evaluate the performance of the system, and PSNR and RMSE provide a base for comparing the performances of different filters.

Table 1: Simulation Parameters

S. No.	Parameter	Type/ Value
1.	Technology	Speech Enhancement
2.	Speech Input signal	Online
3.	Modulation Techniques	AM, FM
4.	Communication Channel	AWGN, Rician
5.	SNR Range	38 to -2.5
6.	Adaptive Filters	LMS, NLMS, UNANR
7.	Measuring Entity	PSNR, RMSE

A. Simulation results and analysis for speech enhancement using adaptive filters for online voice

In previous experiment speech enhancement is done with stored voice data. Now input had been taken online speech from microphone. For this experiment also there can be total four possible configurations possible, as we have two types of modulation techniques and two types of communication channels. So let's start analyzing the performances of different filters with all 4 possible combinations.

A.1 AM is used with AWGN channel

Consider the case of online input speech signal from microphone. In this case AM is selected to transmit the whole speech signal after addition of background noise at the transmitter side. AWGN channel is selected as a communication channel for transferring the speech signal. In AWGN channel, channel noise gets added to the speech signal. At the receiver side first AM demodulation is performed then speech signal is passed through one of the adaptive filter. First, LMS filter is selected and PSNR and RMSE signal parameters are recoded. Second, NLMS filter is selected for the same received demodulated speech signal. And at the last UNANR filter is selected for the same received demodulated speech signal. Graphs have been plotted to check the performance of the adaptive filters. Graphs are plotted between SNR v/s PSNR and SNR v/s RMSE. From the given graph it is seen that NLMS and UNANR perform much better than the LMS filter. Though performance of NLMS and UNANR are in same pattern but UNANR gives best results.

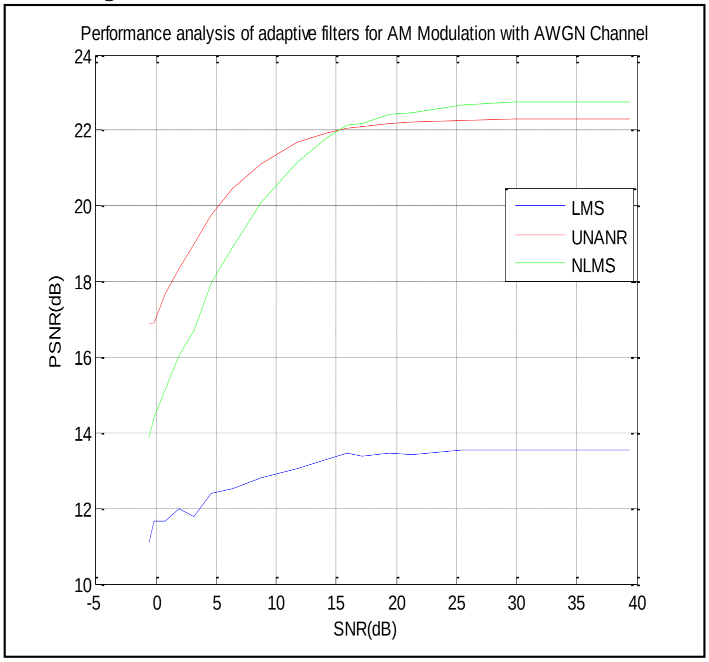

Fig. 3: Adaptive filtering for AM with AWGN channel (SNR v/s PSNR) for online voice

A.2 FM is used with AWGN channel

In this case FM is selected to transmit the whole online speech signal after addition of background noise at the transmitter side. AWGN channel is selected as a communication channel for transferring the speech signal. In AWGN channel, channel noise gets added to the speech signal.

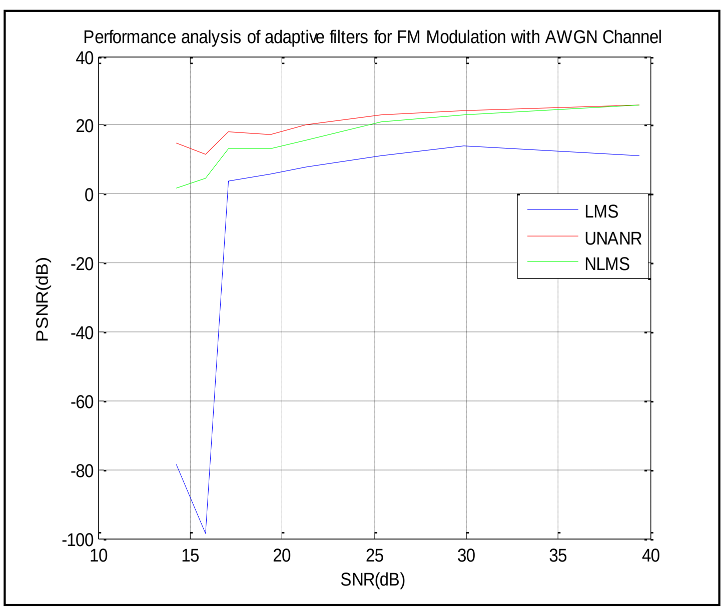

Fig. 4: Adaptive filtering for FM with AWGN channel (SNR v/s PSNR) for online voice

At the receiver side first FM demodulation is performed then speech signal is passed through one of the adaptive filter. Firstly LMS filter is selected and PSNR and RMSE signal parameters are recoded. Secondly NLMS filter is selected for the same received demodulated speech signal. And at the last UNANR filter is selected for the same received demodulated speech signal. Graphs have been plotted to check the performance of the adaptive filters. Graphs are plotted between SNR v/s PSNR and SNR v/s RMSE. for FM modulation technique with AWGN channel, performance of LMS, NLMS and UNANR are shown and the graph is plotted between SNR and PSNR.

A.3 AM is used with Rician channel

In this case AM is selected to transmit the microphone speech signal after addition of background noise at the transmitter side. Rician fading channel is selected as a communication channel for transferring the speech signal. In Rician fading channel, channel noise gets added to the speech signal. At the receiver side first AM demodulation is performed then speech signal is passed through one of the adaptive filter. Firstly LMS filter is selected and PSNR and RMSE signal parameters are recoded. Secondly NLMS filter is selected for the same received demodulated speech signal. And at the last UNANR filter is selected for the same received demodulated speech signal. Graphs have been plotted to check the performance of the adaptive filters. Graphs are plotted between SNR v/s PSNR and SNR v/s RMSE. for AM modulation technique with Rician fading channel, performance of LMS, NLMS and UNANR are shown and graph is plotted between SNR and PSNR.

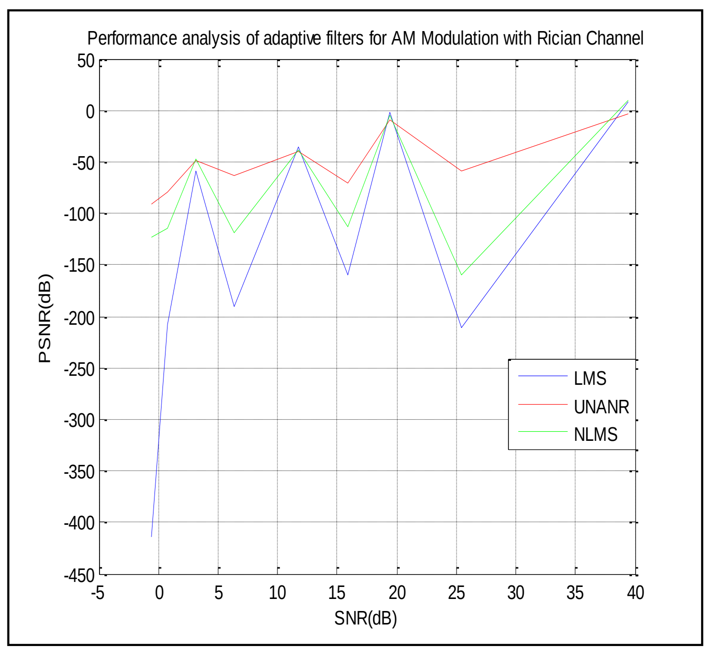

Fig. 5: Adaptive filtering for AM with Rician channel (SNR v/s PSNR) for online voice

A.4 FM is used with Rician channel

In this case FM is selected to transmit the whole speech signal after addition of background noise at the transmitter side. Rician fading channel is selected as a communication channel for transferring the speech signal. In Rician fading channel, channel noise gets added to the speech signal. At the receiver side first FM demodulation is performed then speech signal is passed through one of the adaptive filter. Firstly LMS filter is selected and PSNR and RMSE signal parameters are recoded. Secondly NLMS filter is selected for the same received demodulated speech signal. And at the last UNANR filter is selected for the same received demodulated speech signal. Graphs have been plotted to check the performance of the adaptive filters. Graphs are plotted between SNR v/s PSNR and SNR v/s RMSE. for FM modulation technique with Rician fading channel, performance of LMS, NLMS and UNANR are shown and graph is plotted between SNR and PSNR.

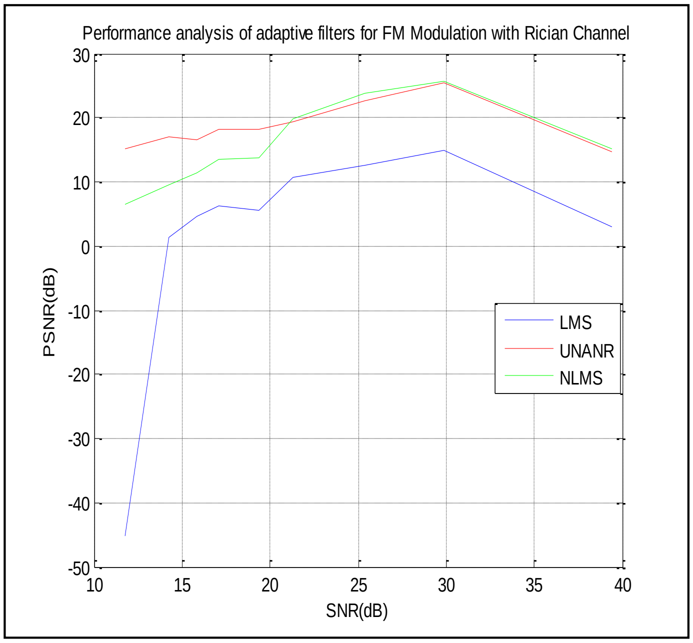

Fig. 6: Adaptive filtering for FM with Rician channel (SNR v/s PSNR) for online voice

CONCLUSION

This dissertation work presented the speech enhancement through LMS, NLMS and UNANR. Performances of these filters are measured with respect to PSNR and RMSE when speech signal is affected by both background and channel noise. The comparison and analysis of performances of these adaptive filters have been done through plotted graphs. From all the performed experiments it is apparent that NLMS and UNANR filters have better performance than LMS. There is a fight between in NLMS and UNANR filter's performances in most of the graphs, but when online voice signal is taken as input to the speech enhancement system then UNANR performed slightly better than NLMS filter. Though UNANR takes little more time to filter speech signal from noise in comparison with LMS and NLMS but has better convergence rate than other two. As soon as the SNR improves performances of the filters get improved and when background noise level gets increased then performances of LMS, UNANR and NLMS filters gets degraded. For high SNR scenarios, the output quality is quite good but on the other hand for very low SNR scenarios, the output quality improves but it still suffers from either residual noise or distortion to speech or both. For these situations, intelligibility definitely improves as the speech is otherwise overwhelmed by noise and nothing could be understood without enhancement.

REFERENCES

[1] B. Widrow, J. Glover, J. M. McCool, J. Kaunitz, C. S. Williams, R. H.Hearn, J. R. Zeidler, E. Dong, and R. Goodlin, *"Adaptive noise cancelling: Principles and applications"* , Proc. IEEE, vol. 63, pp.1692-1716, Dec. 1975.
[2] Suleyman S. Kozat, Andrew C. Singer, Alper Tunga Erdogan, and Ali H. Sayed *"Unbiased Model Combinations for Adaptive Filtering",* IEEE Trans. on Signal Processing, Vol. 58, No. 8, August 2010.
[3] B. L. Sim, Y. C. Tong, J. S. Chang and C. T. Tan, *"A parametric formulation of the generalized spectral subtraction method"*, IEEE Trans. on Speech and Audio Processing, vol. 6, pp. 328-337, 1998.
[4] I. Y. Soon, S. N. Koh, and C. K. Yeo, *"Noisy speech enhancement using discrete cosine transform",* Speech Communication, vol. 24, pp. 249-257, 1998.
[5] Md Zia Ur Rahman , Sk. Khaja Mohedden, Dr. B V Rama Mohana Rao , Y. Jaipal Reddy and G.V.S. Karthik, *"Filtering Non-Stationary Noise in Speech Signals using Computationally Efficient Unbiased and Normalized Algorithm"* , International Journal on Computer Science and Engineering, Vol. 3 No. 3 Mar 2011.
[6] Marvin R. Sambur, *"Adaptive Noise Canceling for Speech Signals",* IEEE Transactions on Acoustics, Speech, And Signal Processing, Vol. Assp-26, No. 5, October 1978.
[7] Zhang Jun-chang, Zhang Yi, Ye Zhen, *"Adaptive Speech Enhancement Based on Classification of Voiced/Unvoiced Signal",* in International Conference on Multimedia and Signal Processing, IEEE Computer Society 2011
[8] H. Sheikhzadeh, and H. R. Abutalebi, *"An improved wavelet-based speech enhancement system"*, Proc. of the Eurospeech, 2001.
[5] S. Salahuddin, S. Z. Al Islam, M. K. Hasan, M. R. Khan, "Soft thresholding for DCT speech enhancement", Electron. Letters, vol. 38, no.24, pp.1605-1607, 2002.
[9] J. Homer, *"Quantifying the convergence speed of LMS adaptive filter with autoregressive inputs",* Electronics Letters, vol. 36, no. 6, pp. 585-586, March 2000.
[10] H. C. Y. Gu, K. Tang and W. Du, *"Modifier formula on mean square convergence of LMS algorithm",* Electronics Letters, vol. 38, no. 19, pp. 1147-1148, Sep 2002.
[11] M. Chakraborty and H. Sakai, *"Convergence analysis of a complex LMS algorithm with tonal reference signals",* IEEE Trans. on Speech and Audio Processing, vol. 13, no. 2, pp. 286 – 292, March 2005.
[12] M. A. Abd El-Fattah, M. I. Dessouky , S. M. Diab and F. E. Abd El-samie, *"Adaptivewiener Filtering Approach For Speech Enhancement",* Ubiquitous Computing and Communication Journal, Egypt, Volume 3 Number 2.
[13] S. Olmos, L. Sornmo and P. Laguna, *"Block adaptive filter with deterministic reference inputs for event-related signals: BLMS and*

BRLS", IEEE Trans. Signal Processing, vol. 50, pp. 1102-1112, May.2002.
[14] Jamal Ghasemi and Mohammad Reza Karami Mollaei, *"A New Approach for Speech Enhancement Based On Eigenvalue Spectral Subtraction",* Signal Processing: An International Journal, vol. 3, Issue. 4, pp. 34-41.