Speech Signal Enhancement Using Adaptive Noise Cancellation Techniques

Allam Mousa, Marwa Qados, Sherin Bader

Abstract *- Speech signal enhancement is an important topic in speech processing where signal changes its characteristics with time depending on various conditions. An important problem that affects the signal enhancement is the background noise which is a major source of quality degradation in speech and audio signals. Adaptive noise cancellation algorithms are used to reduce this noise with relatively fast convergence as desired. Minimization techniques like LMS, NLMS and RLS are widely used due to its simplicity in computation and implementation. These algorithms are evaluated under several conditions like sensitivity for language, text gender and noise power. Certain parameters were designed to obtain the best performance under various conditions where the RLS algorithm has outperformed the other two algorithms when noise power is fixed and that noise power has more influence on the RLS algorithm.*

Keywords***: adaptive filters, noise cancellation, LMS, NLMS, RLS, SNR, speech enhancement***

I. Introduction

Speech is a very basic way for humans to convey information, it has a bandwidth of only 4 kHz; it can convey information with the emotion of a human voice. The speech signal has certain properties such that it is a one-dimensional signal, with time as its independent variable, it is random in nature, it is non-stationary, and the frequency spectrum is not constant in time. Although human beings have an audible frequency range of 20Hz to 20 kHz, the human speech has significant frequency components only up to 4 kHz [1].

The most common problem in speech processing is the effect of interference noise in the signals. This noise masks the speech signal and reduces its intelligibility. It may be produced by acoustical sources such as ventilation equipment, traffic, crowds and commonly, reverberation and echoes. It can also arise electronically from thermal noise, tape hiss or distortion products. If the sound system has unusually large peaks in its frequency response, the speech signal can even end up masking itself [2].

It is important to cancel the noise which may combine the signal in order to obtain a good quality signal, this may be achieved using Active Noise Cancellation (i.e., reducing the noise by means of superposition of the same noise signal but in anti-phase) or by using Adaptive Noise Cancellation (i.e., improving the Signal-to-Noise Ration at the received noisy signal).

This paper is organized such that Section 2 describes the main characters of speech signal, Section 3 presents some main types of noise, Section 4 discusses the concept of adaptive algorithms and finally Section 5 provides the simulation results.

II. Speech Signal

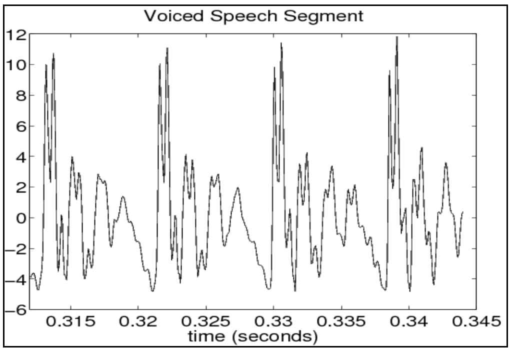
Speech is an acoustic waveform that conveys information from a speaker to a listener. Given the importance of this form of communication, it is no surprise that many applications of signal processing have been developed to manipulate speech signal [3]. At a linguistic level, speech can be viewed as a sequence of basic sound units called phonemes. The same phoneme may give rise to many different sounds or allophones at the acoustic level depending on the phonemes which surround it. Different speakers producing the same string of phonemes convey the same information, yet sound different as a result of differences in dialect and vocal tract length and shape.

Nearly all information in speech is in the range 200Hz to 8 kHz. Humans discriminate voices between males and females according to the frequency. Females speak with higher fundamental frequencies than males. The adult male is from has a fundamental frequency in the range of 50Hz to 250Hz, with an average value of about 120Hz. For an adult female, the upper limit of the range may be as high as 500Hz [1]. Different languages vary in its perceptibility due to differences in its phonetic contents and variations in distribution of different phonemes, stress level distribution among phonemes and of course intonation pattern, nasality usage, allophonic variants, contextual, phonotactic, or coarticulatory constraints.

Humans do have an intuitive understanding of spoken language quality, this may not be easy to quantify. In a number of studies, it has been shown that impact of noise on degradation of speech quality is non uniform. Since speech frequency content varies, across time, due to sequence of

Department of Electrical Engineering An Najah University, Nablus, P.O. Box 7, Palestine,
(e-mail: allam@najah.edu)

phonemes, needed to produce the sentence, impact of background distortion will also vary, causing some phone classes to get more effected than others, when produced in a noisy environment [4]. Speech can basically be identified as voiced and unvoiced speech where voiced speech is periodic with high amplitude and the unvoiced speech is random with lower amplitude as shown in Fig.1 [1].



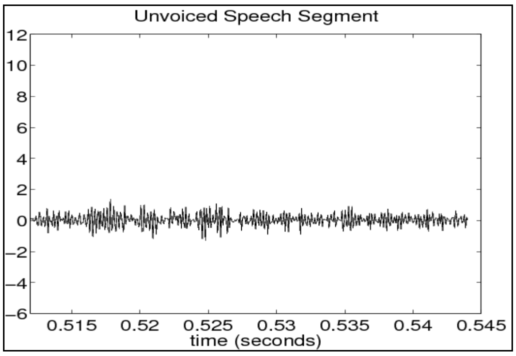


Fig.1 Voiced and unvoiced speech segment

III. Types of Noise

Noise is an unwanted electrical or electromagnetic energy that degrades the quality of signals and data. Noise occurs in digital and analog systems, and can affect files and communications of all types, including text, programs, images, audio, and telemetry [5]. The main effects of noise on a certain signal may include the following; (a) Loudness. (b) Frequency. (c) Continuity. (d) Variation with time. (e) Time of occurrence. (f) Information content. (g) Origin of the sound. (h) Recipient's state of mind and temperament. (i) Background noise level. In general, noise that affects the speech signals can be modeled as; White noise, Colored noise, and Impulsive noise [1], [2].

White noise is a sound or signal consisting of all audible frequencies with equal intensity. At each frequency, the phase of the noise spectrum is totally uncertain; it can be any value between 0 and 2π, and its value at any frequency is unrelated to the phase at any other frequency. When noise signals, arising from two different sources, are added up, the resultant noise signal has a power equal to the sum of the power components. Because of the broad-band spectrum, white noise has strong masking capabilities [5].

Any noise that is not white can be termed as colored noise. Unlike white noise colored noise has frequency spectrum that is limited within a range. There are different types of colored noise like brown, pink, orange noise and so on, depending upon the gradation in the Power Spectral Density (PSD) of the noise. This can be generated by passing white noise through a filter with required frequency response. Impulsive noise refers to sudden bursts of noise with relatively high amplitude. This type of noise causes click sounds in the signal of interest [5].

IV. Adaptive Noise Cancellation

Noise cancellation technology is a growing field that capitalizes on the combination of disparate technological advancements. This aims to cancel or at least minimize unwanted signal and so to remedy the excess noise that one may experience. There are already several solutions available [5], [6].

Adaptive noise cancellation is widely used to improve the Signal to Noise Ratio (SNR) of a signal by removing noise from the received signal. In this configuration the input x(n), a noise source, $N_1(n)$, is compared with a desired signal, d(n), which consists of a signal, s(n) corrupted by another noise, $N_0(n)$. The adaptive filter coefficients adapt to cause the error signal to be a noiseless version of the signal s(n) as shown in Fig. 2.

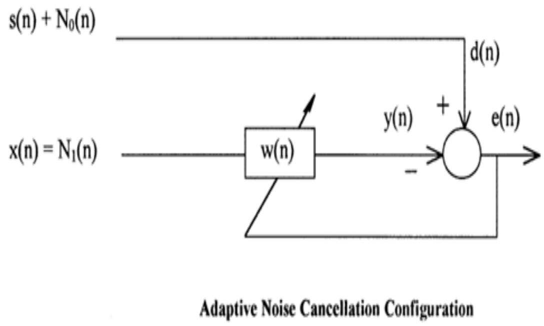


Fig.2: Adaptive noise cancellation configuration.

The noise signals for this configuration need to be uncorrelated to the signal s(n). In addition, the noise sources must be correlated to each other in some way, preferably equal, to get the best results. The error signal should converge to the signal s(n), but it will not converge to the exact signal. In other words, the difference between the signal s(n) and the error signal e(n) will always be greater than zero. The only option is to minimize the difference between those two signals using certain error minimization techniques [6].

1- Adaptive filters

An adaptive filter adapts itself to changes in its input signals automatically according to a given algorithm. The algorithm will change the coefficients according to a given criteria, typically an error signal to improve its performance. In essence an adaptive filter is a digital filter combined with an adaptive algorithm, which is used to modify the coefficients of the filter. Adaptive filters are used in many diverse applications like telephone echo canceling, radar signal processing, equalization of communication channels and biomedical signal enhancement [7] [8], [9].

2- Adaptive algorithms

There are many algorithms used to adjust the coefficients of the digital filter in order to match the desired response as well as possible. This includes the following [7];

A. The LMS algorithm

The simplicity of the Least Mean Square (LMS) algorithm and ease of implementation makes it the best choice for many real-time systems [10]. The implementation steps for this algorithm can be stated as;

1. Define the desired response and set each coefficient weight to zero.

$$w(n) = 0, n = 1,2,3, \ldots, N \qquad (1)$$

For each sampling instant (n) carry out the following steps;

2. Move all samples in the input array one position to the right, now load the current data sample (n) into the first position in the array. Calculate the output of the adaptive filter by multiplying each element in the array of filter coefficients by the corresponding element in the input array and all the results are summed to give the output corresponding to that data that was earlier loaded into the input array, such that the output y(n) is;

$$y(n) = \sum_{n}^{N}{}_{0}^{1} w(n)x(n) \qquad (2)$$

3. Before the filter coefficients can be updated, the error must be calculated, simply find the difference between the desired response, d(n), and the output of the adaptive filter, y(n).

$$\mathrm{e(n) = y(n) - d(n)} \qquad (3)$$

4. To update the filter coefficients multiply the error by the learning rate parameter, μ, and then multiply the results by the filter input and add this result to the values of the previous filter coefficients.

$$\overline{\mathrm{w}}(\mathrm{n+1}) = \overline{\mathrm{w}}(\mathrm{n}) + \mu.\mathrm{e(n)}.\overline{\mathrm{x}}(\mathrm{n}) \qquad (4)$$

Where:

μ is the step size of the adaptive filter, $\overline{w}(n)$ is the filter coefficients vector, $\overline{x}(n)$ is the filter input vector

Then LMS algorithm calculates the cost function J (*n*) by using the following equation:

$$\mathrm{J(n) = e^2(n)} \qquad (5)$$

Where $\mathrm{e^2(n)}$ is the square of the error signal at time (n).

The resources required to implement the LMS algorithm for a transversal adaptive FIR filter of L coefficients in real time is given in Table I. The computations given are those required to process one sample.

B. Normalized LMS Algorithm

The normalized LMS (NLMS) algorithm is a modified form of the standard LMS algorithm. This algorithm updates the coefficients of an adaptive filter by using the following equation:

$$\overline{\mathrm{w}}(\mathrm{n+1}) = \overline{\mathrm{w}}(\mathrm{n}) + \mu.\mathrm{e(n)}.\frac{\overline{\mathrm{u}}(\mathrm{n})}{\left\|\overline{\mathrm{u}}(\mathrm{n})\right\|^2} \qquad (6)$$

This form can be rewritten as,

$$\overline{\mathrm{w}}(\mathrm{n+1}) = \overline{\mathrm{w}}(\mathrm{n}) + \mu(\mathrm{n}).\mathrm{e(n)}.\overline{\mathrm{u}}(\mathrm{n}) \qquad (7)$$

Where $\mu(\mathrm{n}) = \mu / \left\|\overline{u}(n)\right\|^2$

In this equation, the NLMS algorithm becomes the same as the standard LMS algorithm except that the NLMS algorithm has a time-varying step size $\mu(n)$. This step size can improve the convergence speed of the adaptive filter. The NLMS algorithm is a potentially faster converging algorithm compared to the LMS algorithm which may come at a price of greater residual error. The main drawback of the pure LMS algorithm is that it is sensitive to the scaling of its input x(n). This makes it very hard to choose a learning rate μ that guarantees stability of the algorithm. The NLMS algorithm is a variant of the LMS algorithm that solves this problem by normalizing with the power of the input.

C. Recursive Least Squares (RLS) algorithm

The standard RLS algorithm performs the following operations to update the coefficients of an adaptive filter.

1. Calculates the output signal y (n) of the adaptive filter.
2. Calculates the error signal *e* (*n*) as,

$$\mathrm{e(n) = y(n) - d(n)} \qquad (8)$$

3. Updates the filter coefficients as:

$$\overline{\mathrm{w}}(\mathrm{n+1}) = \overline{\mathrm{w}}(\mathrm{n}) + \mathrm{e(n)}.\overline{k}\,(\mathrm{n}) \qquad (9)$$

Where $\overline{w}\,(n)$ is the filter coefficients vector and $\overline{K}\,(n)$ is the gain vector defined as;

$$\overline{k}\,(n) = \frac{P(n)\overline{u}(n)}{\lambda + \overline{u}^{T}\,(n)P(n)\overline{u}(n)} \qquad (10)$$

Where λ is called the forgetting factor and *P*(*n*) is the inverse correlation matrix of the input signal.
The matrix *P* (*n*) has the following initial value *P* (0):

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

Where δ is the regularization factor. The standard RLS algorithm uses the following equation to update this inverse correlation matrix.

$$P(n+1) = \lambda^{-1}P(n) - \lambda^{-1}\overline{k}\,(n)\overline{u}^{T}\,(n).P(n) \qquad (11)$$

RLS algorithms calculate *J* (*n*) by using the following equation

$$\mathrm{J(n)} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i=0}}^{\mathrm{N\text{-}1}} \lambda^{\mathrm{i}}\,\mathrm{e}^{2}(\mathrm{n\text{-}i}) \qquad (12)$$

Where *N* is the filter length and *λ* is the forgetting factor.

This algorithm calculates not only the instantaneous value e^2 (*n*) but also the past values, such as e^2 (*n*–1), e^2 (*n*–2)... e^2 (*n*–*N*+1). The value range of the forgetting factor is (0, 1]. When the forgetting factor is less than 1, this factor specifies that this algorithm places a larger weight on the current value and a smaller weight on the past values.

The resources required to implement the LMS, NLMS and RLS algorithms for a transversal adaptive FIR filter of L coefficients in real time are given in Table I. The computations given are those required to process one sample. A typical comparison among the LMS, NLMS and RLS algorithms is given in Table II.

Table I: Algorithm resources

	LMS	NLMS	RLS
Memory	2L+1	2L+7	L^2 +2L+4
Multiply	2L	2L+7	2 L^2 +4L
Add	0	2L+2	1.5 L^2 +2.5L
Divide	0	1	L

Table II A comparison between the LMS, NLMS and RLS algorithms

LMS	• More iteration is needed for convergence. • Has mediate value of MSE.	• Less iteration is needed for convergence as mu is increased. • Has larger value of MSE than NLMS.	• As decrease number of weights is reduced, one needs more iteration to reach steady state. • More iteration is needed to converge than NLMS and RLS. • Decreasing mu requires more iteration to converge; steady state error will be decreased.
NLMS	• Needs less iteration than LMS and RLS to converge. • Has the lowest value of MSE.	• Large mu may results in unstable systems. • Has less value of MSE than LMS.	• Decreasing mu requires more iteration to converge but steady state error will be decreased. • Need less iteration to converge than LMS and RLS. • Decreasing number of weights needs more iteration to reach steady state.
RLS	• Need less iteration than LMS to converge and reach steady state. • Has the highest value of MSE.		• Need less iteration to converge than LMS and more iteration than NLMS. • As the number of weights is decreased, one needs more iteration to reach steady state.

V. Simulation results

In order to measure the performance of noise cancellation system, and to compare among certain adaptive algorithms, one may check certain performance measures like Convergence rate, Mean Square Error (MSE), Computational complexity, Stability, Robustness, Filter length, SNR and Segmental SNR. The original signal, the MSE and the output filtered signals were presented for LMS, NLMS and RLS algorithms as shown in Fig.3, Fig.4, Fig.5 and Fig.6 respectively [11]. These algorithms were then examined against certain parameters.

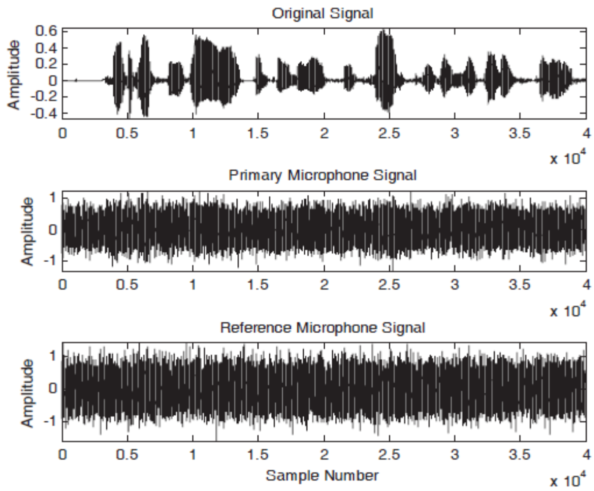


Fig. 3. Original, primary and reference signals

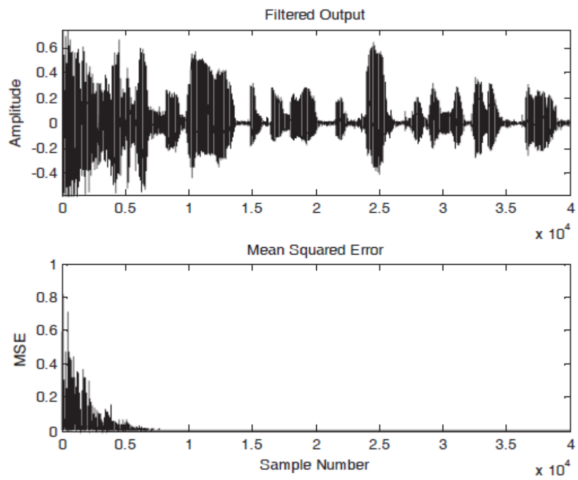


Fig.4 Filtered output signal and MSE curve of the LMS algorithm

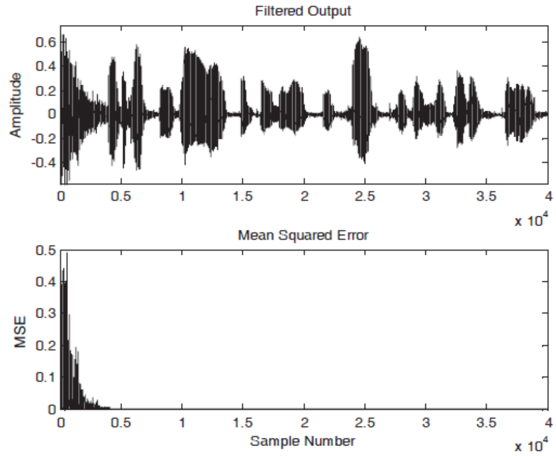


Fig.5 Filtered output signal and MSE curve of the NLMS algorithm

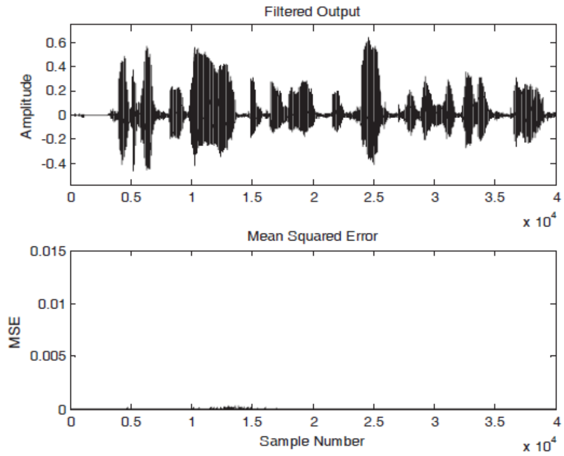
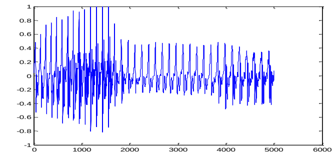
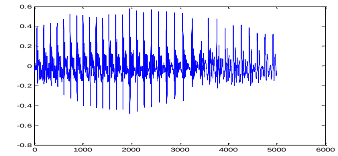
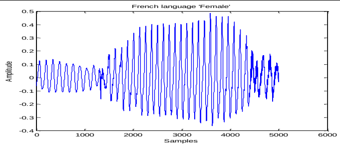
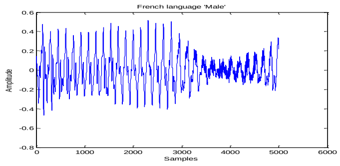
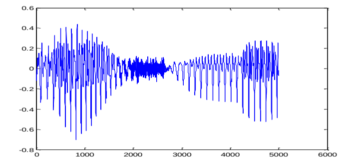
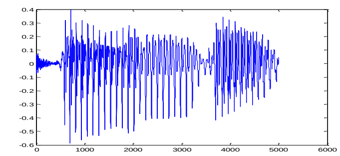
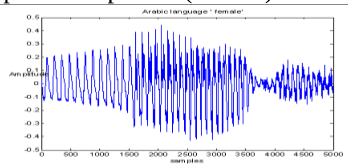
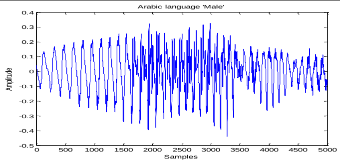


Fig.6 Filtered output signal and MSE curve of the RLS algorithm.

Several speech samples were applied to the noise cancellation system where the samples were changed such that the effects of various parameters were examined; this includes changing the language, gender and text. The noise power was set to 6.15 dB and the same set of samples was processed by different algorithms (LMS, NLMS, RLS), and the SNR and SNR_{seg} values were obtained as illustrated in Table III.

Table III Algorithm dependence on speech text

Algorithm	LMS		NLMS		RLS	
Speech text	SNR	SNR_{seg}	SNR	SNR_{seg}	SNR	SNR_{seg}
 Speech1: English (female)	24.4	23.9	24.0	19.1	38.1	27.8
 Speech 2: English (male)	19.1	18.7	22.0	18.0	25.6	27.8
 Speech3: French (female)	23.0	21.8	23.2	22.0	25.5	27.0
 Speech4: French (male)	21.6	20.2	19.0	18.8	26.1	30.3
 Speech5: Spanish (male)	19.6	17.5	24.1	16.7	27.1	21.0
 Speech6: Spanish (female)	20.0	18.3	22.2	14.3	25.4	20.2
 Speech7: Arabic (female)	20.1	18.0	19.2	17.3	25.9	18.6
 Speech7: Arabic (female)	19.1	18.2	19.6	17.9	23.9	18.9
Average	20.8	19.6	21.7	18.0	27.2	23.9

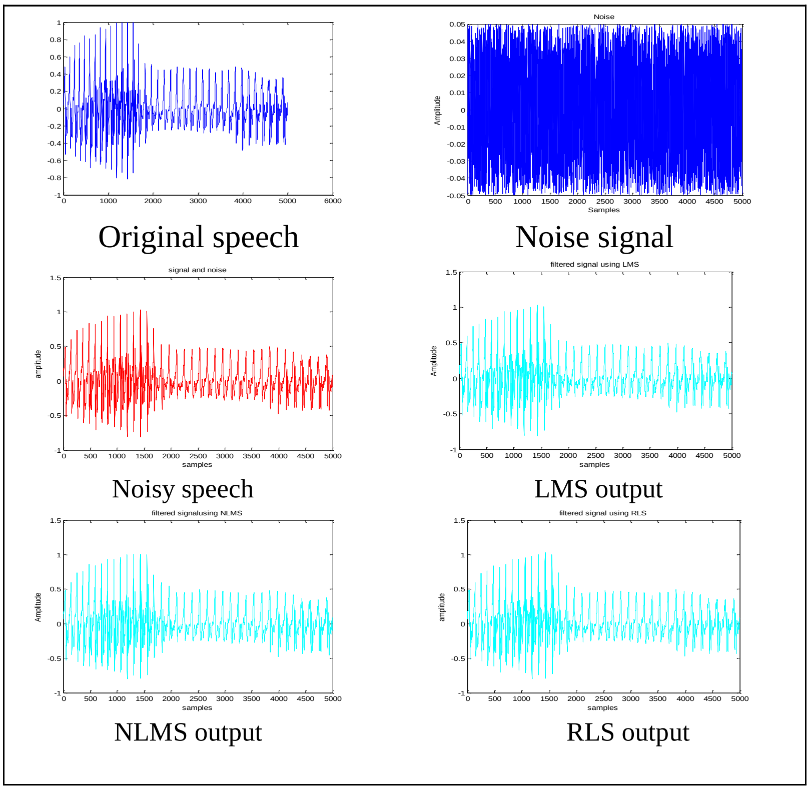


Fig .7 Different algorithms applied to English text spoken by female

As an average, RLS algorithm has achieved the best SNR and SNR_{seg} with an average of 27.2 and 22.9 respectively, which is 6.4 dB and 4.3 dB greater than the average of LMS algorithm values, and 5.5 dB and 5.9 dB greater than those of the NLMS algorithm values.

These algorithms are not affected by changing language, gender or text. A typical performance for the noise cancellation system using different algorithms applied for an English text by female speaker is shown in Fig.7. Moreover, the error signal, defined as the difference between the original speech signal and the output of the noise cancellation system is illustrated for various algorithms as shown in Fig.8. The female speaker texts (English, French, Spanish, and Arabic) were investigated per frame of speech as illustrated in Fig.9, Fig.10 and Fig.11 for LMS, NLMS and RLS respectively.

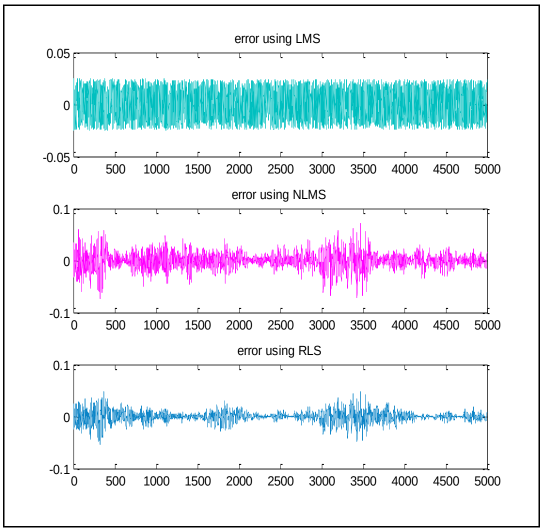


Fig .8 Error signal using different algorithms

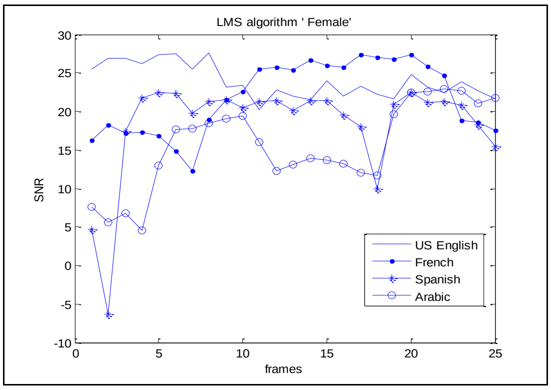


Fig.9 Segmental SNR for different languages (female) using LMS algorithm

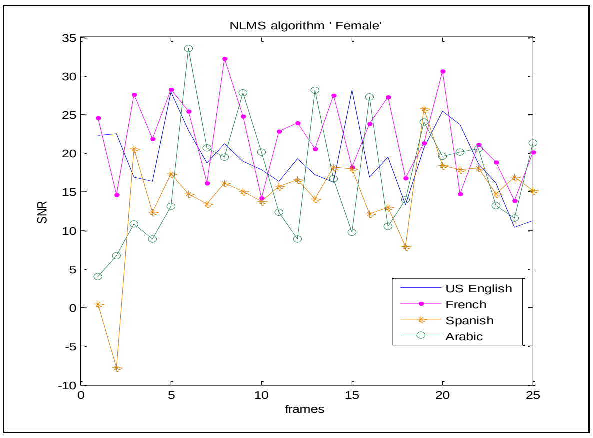


Fig.10 Segmental SNR for different languages (female) using NLMS algorithm

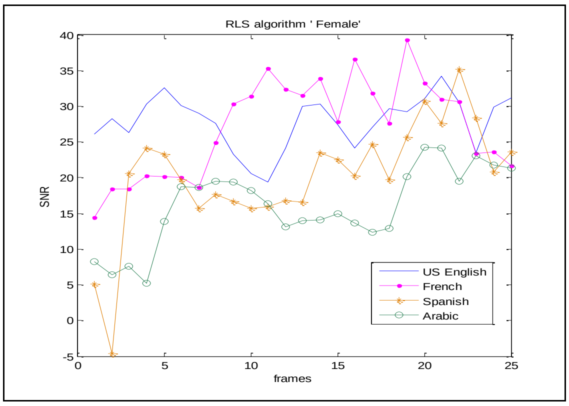


Fig.11 Segmental SNR for different languages (female) using RLS algorithm

The behavior of noise cancellation system using LMS, NLMS and RLS algorithms for different speakers, languages and genders does not change the trend of the SNR response, and it varies according to the power of the input signal. The signal characteristic may change the response of the algorithm at particular energy levels of the input signal.

Noise power has been another parameter to consider in this analysis, this power been fixed while other parameters were changed for the various algorithms. As these parameters are fixed and the noise power is changed then the sensitivity of the respected algorithms to the increment in noise power is illustrated in Fig. 12.

I. Conclusion

This paper investigates the main parameters which may affect the performance of an adaptive noise cancellation system in speech processing applications. Particular attention is given to the input speech signal under various conditions such as changing the text, gender, speaker, and language. It also considers the effect of noise power on the related algorithms. The RLS algorithm has outperformed both LMS and NLMS algorithms in terms of SNR under certain conditions. The NLMS has better performance when noise power is considered. It has been shown that changing the input of the system does not change the performance of the system in general and that RLS outperforms LMS and NLMS algorithms at the cost of complexity and with a moderate noise power.

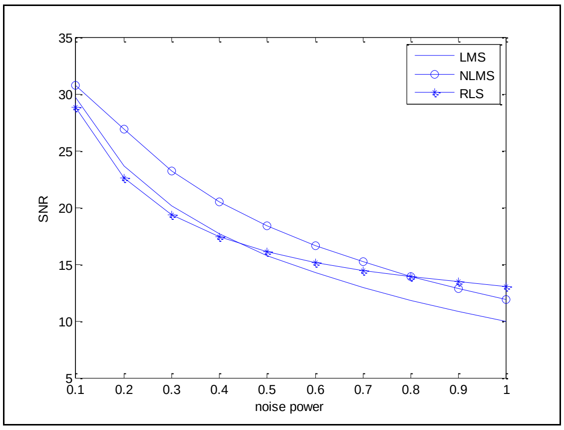


Fig. 12 Relation between noise power and SNR

II. References

1 A. M. Kondoz, "Digital Speech: Coding for Low Bit Rate Communication Systems" John Wiley, September 2004

2 Ben G., Nelson M., "Speech and Audio Signal Processing" John Wiley and Sons, Inc. 2000

3 Lawrence R Rabiner; Ronald W Schafer, "Theory and applications of digital speech processing", 1st ed, Upper Saddle River: Pearson, 2011

4 Sunita Maithani, J. S. Rawat, "Noise as a Tool for Spoken Language Identification", Metcalfe House, Delhi, Proceedings of the Third International Joint Conference on Natural Language Processing, pp. 811-816, Hyderabad, India, January 7-12, 2008

5 Colin H Hansen, "Understanding Active Noise Cancellation", Taylor and Francis Inc. Print ISBN: 0415231922, 2002

6 V. R. Vijaykumar, P. T. Vanathi, P. Kanagasapabathy, "Modified Adaptive Filtering Algorithm for Noise Cancellation in Speech

Signals", Electronics and Electrical Engineering, 2007 No 2(74) pp. 17-20, 2007.

7 Simon Haykin, "Adaptive Filter theory", 4th ed., Prentice Hall, September 2001.

8 J. M. Górriz, Javier Ramírez, S. Cruces-Alvarez, Carlos G. Puntonet, Elmar W. Lang, and Deniz Erdogmus, "A Novel LMS Algorithm Applied to Adaptive Noise Cancellation", IEEE Signal Processing Letters, Vol. 16, No. 1, pp. 34-37, January 2009

9 Allam Mousa, Marwa Qadus, Sherin Bader, "Adaptive Noise Cancellation Algorithms Sensitivity to Parameters", The 2nd International Conference on Multimedia Computing and Systems (ICMCS'11), IEEE Catalog Number CFP731-3-CDR, ISBN: 978-1-61284-731-3, pp. 1-5, Ouarzazate, Morocco, April 7-9 2011

10 Jan Vanus, "The Use of the Adaptive Noise Cancellation for Voice Communication with the Control System", International Journal of Computer Science and Applications, Vol.8, No. 1, pp. 54-70, 2011

11 N. Sonbolestan, S. A. Hadei, "A Fast Affine Projection Algorithm Based on Matching Pursuit in Adaptive Noise Cancellation for Speech Enhancement", International Conference on Intelligent Systems, Modeling and Simulation, pp. 193-198, Liverpool, England, 27 – 29 January 2010

Allam Mousa is an Associate Professor at An Najah University, Departments of Electrical Engineering and Telecommunication Engineering. He received his BSc, MSc and PhD in Electrical and Electronics Engineering from Eastern Mediterranean University in 1990, 1992 and 1996 respectively. His current research interests include Speech and Image Processing and Telecommunications. He has dozens of research papers in international journals and conferences. Dr. Mousa is also interested in higher education quality assurance and Planning. He is currently Deputy Presidents for Planning Development and Quality. He is also SMIEEE and member of Engineering Association.

Marwa Qados, Received her BSc in Electrical Engineering at An Najah University in 2010

Sherin Bader Received her BSc in Electrical Engineering at An Najah University in 2010