Performance Comparison of STFT, WT, LMS and RLS Adaptive Algorithms in Denoising of Speech Signal

Mahbubul Alam, Md. Imdadul Islam, and M. R. Amin

***Abstract*—Different types of noise cancellation techniques are prevalent in recent literatures. The performance of a particular technique depends on mean, variance and maximum amplitude of error. At the same time the process time of signal and complexity of practical implementation of circuits is also a measuring tool for performance of a technique. The objective of this paper is to compare performance among the short time Fourier transform (STFT), wavelet transform (WT), least mean square (LMS) and recursive least square (RLS) methods in cancellation of noise from a speech signal. The analysis of the paper provides us the way of selection of the best denoising technique based on the statistical parameters of the above four mentioned techniques.**

***Index Terms*—Signal denoising, CPU time, statistical parameters of adaptive filter, wavelet transform, short-time Fourier transform, least mean square and recursive least square.**

I. Introduction

Denoising means removal of noise from a signal [1-4]. Different types of adaptive algorithms, e.g., least-measquare (LMS), recursive least square (RLS), Kalman filtering, etc. are widely used for denoising of low and medium frequency signals like biomedical signal, speech signals, passband fading signal of wireless communications, echo cancellation etc. Short-time Fourier transform (STFT) and wavelet transform (WT) are widely used for denoising of time dependent signals and images, image detection and synthesis, etc. Usual Fourier transformation (FT) is a mathematical technique for transforming a signal from time-domain to frequency domain. After transformation, time information of the signal is lost, i.e., Fourier transform cannot say which spectral component exits at what time. On the other hand, STFT provides both information on time-frequency plane [5, 6]. A small portion of the signal $f(t)$ is selected by multiplying it with a window function $W(t)$. Such technique is called windowing of signals. Next consecutive portion of the signal is picked introducing some delay τ to the window function $W(t)$, e.g.., $f(t,\tau)=f(t)\,W^*(t-\tau)$, where theasterisk indicates complex conjugation. The WT provides a time-frequency representation of the signal like spectrogram of STFT. In STFT, the time-frequency resolution is constant but WT overcomes the situation with provision of variable resolution [7-9].

The STFT can be used to denoise a speech signal in three steps: first compute the STFT of the noisy signal, make a threshold to the STFT and finally compute the inverse STFT. Noise on image or a time varying signal reveals itself as fine-grained structure results in small scale coefficient after wavelet transform. Image filtering is equivalent to discarding these coefficients. A threshold level is selected to discard the coefficients such that edge information remains unaltered. Because coefficients pertinent to edges are very close to that of noise hence discarding technique may leads huge distortion to the image. The whole job can be done by a single line code, “xd = wden(x, TPTR, SORH, SCAL, N, 'wname')” in MATLAB. In an adaptive algorithm, a desired signal is used which is correlated with one portion of the input signal. The adaptive algorithm updates the weighting factors of the filter to keep the difference between input and desired signal at a minimum level.

The paper is organized as follows: Section 2 deals with the mathematical analysis of denoising techniques on STFT, WT and adaptive algorithms, Section 3 provides the results making a comparison of four different techniques and finally, Section 4 concludes the entire analysis.

II. Theoretical Analysis

A. STFT

In STFT, a small portion of the signal $f(t)$ is selected by multiplying it with a window function $W(t)$:

$$f(t,\tau)=f(t)W^*(t-\tau)\,.$$

By taking Fourier transform on $f(t,\tau)$ [5, 6], we have

$$F(\tau,\omega)=\int_{-\infty}^{\infty} f(t,\tau)\,e^{-j\omega t}\,dt=\int_{-\infty}^{\infty} f(t)\,W^*(t-\tau)\,e^{-j\omega t}\,dt\,. \quad (1)$$

The transformed signal $F(\tau,\omega)$ depends on angular frequency ω, like ordinary Fourier transform; delay parameter, τ, and window function $W(t)$. The width $\Delta\tau$ of the window function $W(t)$ is called time resolution and the smallest difference $\Delta\omega$ in angular frequency of two sinusoidal waves for which two signals are resolvable is called frequency resolution. The squared magnitude of $F(\tau,\omega)$ is referred to as the spectrogram of the signal which gives the amplitude of the signal in time-frequency plane.

Denoising in STFT composed of three steps:

(a) Compute the STFT of the noisy signal:

$$F(\tau,\omega)=STFT\{f(t)\})$$

(b) Make a threshold on $F(\tau,\omega)$:

Manuscript received February 28, 2011; revised May 19, 2011.

Mahbubul Alam and Md. Imdadul Islam are with the Computer Science and Engineering Department, Jahangirnagar University, Savar, Dhaka 1342, Bngladesh. M. R. Amin is with the Department of Electronics and Communications Engineering, East West University, 43 Mohakhali, Dhaka 1212, Bangladesh corresponding author, phone: +880-1715296735; fax: +880-2-8812336; (e-mail: ramin@ ewubd.edu).

$F_d(\tau,\omega) = Th(F(\tau,\omega))$; where $Th(b)$ is the threshold function:

$$Th(b) = \begin{cases} 0, & |b| \le r \\ a, & |b| > r. \end{cases}$$

(c) Compute the inverse STFT:

$$y(t) = STFT^{-1}\left[F_d(\tau,\omega)\right].$$

B. Wavelet

In WT, any signal $f(t)$ (square integrable function) has the continuous-time wavelet transform (CWT) with respect to a wavelet $\psi(t)$ is defined as [7, 9],

$$W(a,b) = \int_{-\infty}^{\infty} f(t)\frac{1}{\sqrt{|a|}}\psi^*\left(\frac{t-b}{a}\right)dt, \quad (2)$$

where a and b are real and * denotes complex conjugation. Equation (2) can be written in a more compact form by defining,

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right). \quad (3)$$

Combining (2) and (3), we get

$$W(a,b) = \int_{-\infty}^{\infty} f(t)\psi_{a,b}(t)dt. \quad (4)$$

The WT provides a time-frequency representation of the signal like spectrogram of STFT. In STFT, the time-frequency resolution is constant but WT overcomes the situation with provision of variable resolution.

Inverse CWT operation can be expressed as

$$f(t) = \frac{1}{C}\int_{a=-\infty}^{\infty}\int_{b=-\infty}^{\infty}\frac{1}{a^2}W(a,b)\psi_{a,b}(t)\,da\,db, \quad (5)$$

where

$$C = \int_{-\infty}^{\infty}\frac{|\psi(\omega)|^2}{|\omega|}d\omega, \quad \psi(t) \leftrightarrow \psi(\omega),$$

and $0<C<\infty$.

One drawback of the CWT is that the representation of the signal is often redundant, since a and b are continuous over R (the real number). If scales and positions are chosen based on powers of two (called dyadic scales and positions) then the analysis will be much more efficient and accurate. Such analysis of WT is called the discrete wavelet transform (DWT). The original signal can be completely reconstructed by a sample version of $W(b,a)$ in the DWT. Typically, $W(b,a)$ is sampled in dyadic grid, i.e., $a = 2^m$ and $b = n2^m$. We have

$$\Psi_{a,b}(t) = \Psi\left(\frac{t-b}{a}\right) = \Psi\left(\frac{t-n2^m}{2^m}\right) = \Psi\left(2^{-m}t-n\right) = \Psi_{m,n}(t)$$

is the dilated and translated version of the mother wavelet.

The DWT or CWT of $f(t)$ with respect to a wavelet $\psi(t)$ is defined as ,

$$d(m,n) = \frac{1}{2^m}\int_{2^m n}^{2^m(n+1)} f(t)\psi(2^{-m}t-n)dt. \quad (6)$$

Inverse operation i.e., inverse DWT (IDWT) is expressed as

$$f(t) = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} d(m,n)2^{-m/2}\psi(2^{-m}t-n). \quad (7)$$

Similar threshold operation of STFT is applied in WT to denoise a signal or image.

C. LMS Algorithm

Least mean square filter is a type of digital adaptive filter where a desired filter is obtained by finding the filter coefficients that relate to producing the least mean squares of the error signal. In LMS, error estimation is determined by comparing the output of the linear filter in response to the input signal to the desired response. LMS then involves an adaptive process for automatic adjustment of the parameters of the filter according to the error estimation. The filtering process of LMS algorithm is performed by a transversal filter and the adaptive process is carried out by a weight control mechanism [10-13].

LMS algorithm can be summarized as follows based upon wide-sense stationary stochastic signal [14, 15]:

Parameters:

M = number of taps,

μ = step size parameter,

where $0 < \mu < 2/MS_{\max}$ and S_{max} is the maximum value of the power spectral density of the tape input $\mathbf{u}(n)$ and filter length M are moderate to large.

Initialization:

If prior knowledge of the tap-weight vector $\hat{\mathbf{w}}(n)$ is available, use it to select an appropriate value of $\hat{\mathbf{w}}(0)$; otherwise set $\hat{\mathbf{w}}(0) = \mathbf{0}$.

Data:

Given

$\mathbf{u}(n) = M - \text{by} - 1$ tap input vector at time n

$= [u(n), u(n-1), \cdots, u(n-M+1]^T$,

$\boldsymbol{d}(n)$ = desired response at time $t = n$.

To be computed:

$\hat{\mathbf{w}}(n+1)$ = estimate of tap-weight vector at time n+1.

Computation:

$\boldsymbol{e}(n) = \boldsymbol{d}(n) - \hat{w}^H(n+1)\,\mathbf{u}(n)$,

$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \mu\,\mathbf{u}(n)e^*(n)$,

where the superscript H denote Hermitian conjugate.

D. RLS Algorithm

The RLS algorithm recursively finds the filter coefficients that minimize a weighted linear least squares cost function relating to the input signals, i.e., given the least squares estimate of the tap weight vector of the filter at iteration (n-1), we compute the updated estimate of the vector at iteration n upon the arrival of the new data. This, in contrast to LMS algorithm aims to reduce the mean square error. In the derivation of the RLS, the input signals are considered deterministic, while for the LMS, they are considered stochastic. Compared to most of its competitors, the RLS exhibits extremely fast convergence due to the fact that the RLS filter whitens the input data by using the inverse correlation matrix of the data, assumed to be of zero mean. However, this benefit comes at the cost of high computational complexity [14-16].

The RLS algorithm can be summarized as follows, Initialize the algorithm by setting

$$\hat{\mathbf{w}}(0) = 0\,,$$

$$\mathbf{P}(0) = \delta^{-1}\mathbf{I}\,,$$

and

$$\delta = \begin{cases} \text{small positive constant for high SNR} \\ \text{large positive constant for low SNR}. \end{cases}$$

For each instant of time, n = 1, 2, … … … compute

$$\boldsymbol{\pi}(n) = \mathbf{P}(n-1)\,\mathbf{u}(n)\,,$$

$$\mathbf{k}(n) = \frac{\boldsymbol{\pi}(n)}{\lambda + \mathbf{u}^H(n)\,\boldsymbol{\pi}(n)}\,,$$

$$\boldsymbol{\xi}(n) = \mathbf{d}(n) - \hat{\mathbf{w}}^H(n-1)\,\mathbf{u}(n)\,,$$

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}^H(n-1) + \mathbf{k}(n)\,\boldsymbol{\xi}^*(n)\,,$$

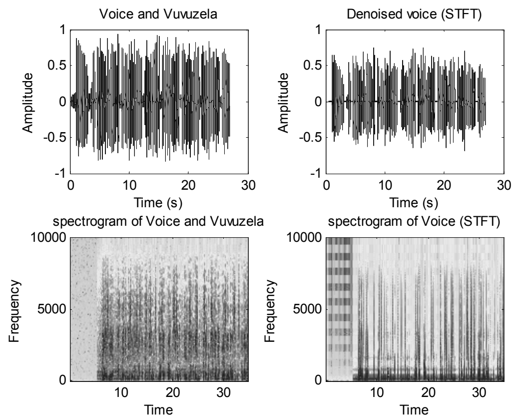
and

$$\mathbf{P}(n) = \lambda^{-1}\,\mathbf{p}(n-1) - \lambda^{-1}\mathbf{k}(n)\,\mathbf{u}^H(n)\,\mathbf{P}(n-1)\,.$$

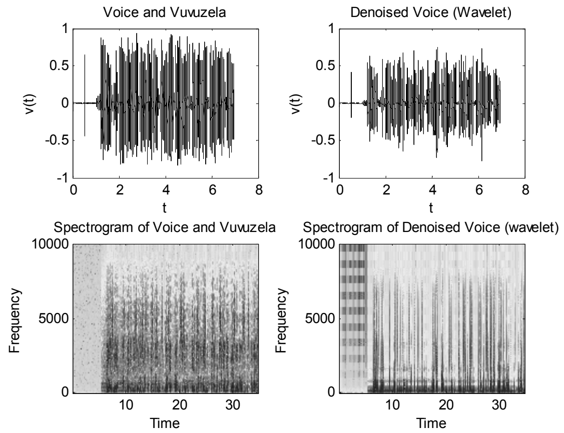
III. Results and Discussions

In this paper a voice signal contaminated by Vuvuzela is selected as a noisy signal. Our objective is to denoise the signal, i.e., to get the voice signal without the tone of Vuvuzela. Fig.1 (a)-(d) shows the contaminated signal, denoised signal and their spectrogram for STFT, WT, LMS adaptive algorithm and RLS adaptive algorithm respectively. In Fig. 1, it is really difficult to make a comparison of the four different techniques. Five parameters: mean error, maximum error, variance, CPU time and practical implementation are considered here to make a comparison of four techniques used in this paper. All the five parameters are shown in Table 1. In the context of mean error, adaptive RLS algorithm is best and STFT is the worst. For the case of maximum error, the adaptive algorithms LMS and RLS are the worst. In an adaptive algorithm, initial error is very high because of assumed values of desired signal and takes time for the convergence. Once the adaptive filter catches the profile of the change of the signal, the men square error is very low; therefore, average error is smaller. Because of initial values of large error, the variance of the error is also higher in adaptive algorithm cases compared to STFT and WT, which is also visualized from the Table 1.

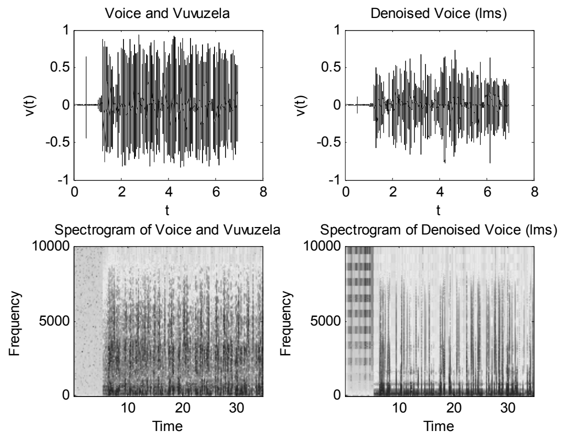
In context of CPU time, STFT and WT provide very close results but much better than both the adaptive algorithms. Adaptive algorithm is a recurrence algorithm where each sample of output, weighting factor matrix of the filter has to be upgraded which incurs huge time for entire output. On the other hand, in STFT and WT, the entire noisy signal is transformed and a thresholding is applied on the transformed signal, and finally inverse operation is done on the threshold signal requires much smaller process time. In context of practical implementation, adaptive algorithm is simple because of computation of adaptive algorithm is in favor of VLSI technology.



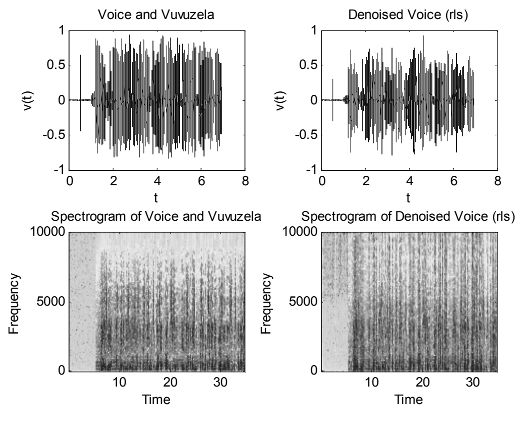
(a) STFT



(b) Wavelet



(c) LMS



(d) RLS

Fig. 1 Noisy and denoised signal with spectrogram

TABLE I: COMPARISON OF PERFORMANCE PARAMETERS

Parameters	**Wavelet**	**STFT**	**LMS**	**RLS**
Mean Error	0.0339	0.0342	0.0339	0.0247
Maximum error	0.5608	0.3883	0.8612	0.8661
Variance	0.0043	0.0018	0.0082	0.0083
CPU time	0.9844	1.6094	13.2656	28.1719
Practical Implementation	Complex	Complex	Simple	simple

IV. CONCLUSION

This paper makes a comparison of LMS and RLS adaptive algorithms with STFT and WT for the case of denoising a speech signal. Form the analysis of the paper we can conclude that for real time denoising operation (for example, live broadcasting of speech) of voice signal DWT is preferable (from CPU time) but for non-real time case (for example, denoising of an old song from a gramophone) RLS is preferable for precise denoising (from mean square error). STFT and WT transforms are suitable for denoising of a signal but in case of recovery of fading signals of wireless communication, adaptive algorithms are the best. The phenomenon can also be proved with same theoretical analysis used in the paper. Still there is a scope of the paper to make the same comparison on very low frequency signal like biomedical signals (EEG and ECG). The work of comparison can also be extended incorporating Kalman filter theory.

REFERENCES

[1] Swathi Nibhanupudi, Signal Denoising Using Wavelets, Thesis of Master of Science, K.L.College of Engineering, 2001

[2] Md. Anamul Haque, A.K.M Kamrul Islam and Md. Imdadul Islam, 'Demystifying the digital adaptive filters conducts in acoustic echo cancellation,' Journal of Multimedia, vol.5, no.6, pp.568-579, Academy Publisher, Dec'2010.

[3] Yi Hu, Philipos C. Loizou, "A Perceptually Motivated Approach for Speech Enhancement", IEEE Transactions on Speech and Audio Processing, VOL. 11, NO. 5, pp.-457-465, September 2003.

[4] Ing Yann Soon, Soo Ngee Koh, "Speech Enhancement Using 2-D Fourier Transform", IEEE Transactions on Speech and Audio Processing, VOL. 11, NO. 6, pp.-717-724, November, 2003

[5] Md. Imdadul Islam and M. Ruhul Amin Khandaker, Laboratory Works on Electronics and Communications, CBO Publications (pvt.) Ltd., ISBN: 984-70062-0008-6, July' 2007

[6] Simon Haykin and Barry Van Veen, 'Signals and Systems", Second Edition, John Wiley & Sons, Inc., New York, 2003.

[7] Md. Imdadul Islam, Nasima Begum, Mahbubul Alam and M. R. Amin, "Fingerprint detection using Canny filter and DWT, a new approach," Journal of Information Procession Systems, vo.6, no.4, pp.511-520, Dec' 2010

[8] Burrus, C. S.; Gopinath, R. A.; and Guo, H., 'Introduction to Wavelets and Wavelet Transforms': A Primer. Prentice Hall Inc., Upper Saddle River, New Jersey, 1998

[9] M. Tico, P. Kuosmanen, J. Saarinen, "Wavelet domain features for fingerprint recognition", IEE Electronics Letters, Vol.37, No. 1, pp.21-22, 4th January, 2001.

[10] K. Banovic, A. R. Esam and M. A. S. Khalid,, "A Novel Radius-Adjusted Approach for Blind Adaptive Equalization*", IEEE Signal Processing Letters*, Vol. 13, No. 1, Jan. 2006.

[11] R. A. Valenzuela, "Performance of Adaptive Equalization for Indoor Radio Communications", *IEEE Tran. On Commun*., Vol. 31, No. 3, pp. 291-293, March 1989.

[12] X. Tang, M. S. Alounini and A. Goldsmith, " Effect of channel estimation error on M-QAM BER performance in Rayleigh fading", *IEEE Trans. Commun*., PP. 1856-64, December 1999.

[13] O. P. Sharma, V. Janyani and S. Sancheti, "Recursive Least Squares Adaptive Filter a better ISI Compensator", International Journal of Electronics, Circuits and Systems Vol.3, No. 1, pp 40-45, 2009

[14] Simon Haykin, "*Adaptive Filter Theory*", Fourth Edition, Pearson Education, ISBN 81-7808-565-8, pp.-231-237, pp.-345-48, pp. 466-485

[15] R. W. Lucky, "Techniques for adaptive equalization of digital communication system", Bell System Tech. Journal, Vol. 45, pp. 255-286, 1966.

[16] S. U. H. Qureshi, "Adaptive Equalization", *IEEE Procs*., Vol. 73. No. 9, pp. 1349-1387, Sept. 1985.

Mahbubul Alam completed his B.Sc. (Hons.) in Computer Science and Engineering from Jahangirnagar University, Savar, Dhaka, Bangladesh in 2006. He received the "Asadul Kabir Gold Medal" for securing the highest grade in the B.Sc. (Hons.) final examination 2006 at Jahangirnagar University. He is now working as a Lecturer, at the department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka, Bangladesh from 13th July 2009. His research interests are wavelet transform, wireless communications, adaptive filter theory, artificial intelligence, image processing and machine learning.

Md. Imdadul Islam has completed his B.Sc. and M.Sc Engineering in Electrical and Electronic Engineering from Bangladesh University of Engineering and Technology, Dhaka, Bangladesh in 1993 and 1998 respectively and has completed his Ph.D degree from the Department of Computer Science and Engineering, Jahangirnagar University, Dhaka, Bangladesh in the field of network traffic engineering. He is now working as a Professor at the Department of Computer Science and Engineering, Jahangirnagar University, Savar, Dhaka, Bangladesh. Previously, he worked as an Assistant Engineer in Sheba Telecom (Pvt.) LTD (A joint venture company between Bangladesh and Malaysia, for Mobile cellular and WLL), from Sept'94 to July'96. He has a very good field experience in installation of Radio Base Stations and Switching Centers for WLL. His research field is network traffic, wireless communications, wavelet transform, OFDMA, WCDMA, adaptive filter theory, ANFIS and array antenna systems. He has more than hundred research papers in national and international journals and conference proceedings.

M. R. Amin received his B.S. and M.S. degrees in Physics from Jahangirnagar University, Dhaka, Bangladesh in 1984 and 1986 respectively and his Ph.D. degree in Plasma Physics from the University of St. Andrews, U. K. in 1990. He is a Professor of Electronics and Communications Engineering at East West University, Dhaka, Bangladesh. He served as a Post-Doctoral Research Associate in Electrical Engineering at the University of Alberta, Canada, during 1991-1993. He was an Alexander von Humboldt Research Fellow at the Max-Planck Institute for Extraterrestrial Physics at Garching/Munich, Germany during 1997-1999.

Dr. Amin awarded the Commonwealth Postdoctoral Fellowship in 1997. Besides these, he has also received several awards for his research, including the Bangladesh Academy of Science Young Scientist Award for the year 1996 and the University Grants Commission Young Scientist Award for 1996. He is a member of the IEEE.