International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.8, No.9 (2015), pp.69-76
http://dx.doi.org/10.14257/ijsip.2015.8.9.07

Enhancement of Speech Signals in a Noisy Environment based on Wavelet based Adaptive Filtering

G. Renisha[1] and T. Jayasree[2]

[1,2]Assistant Professor,
Dept. of Electronics and Communication Engineering,
Govt. College of Engineering, Tirunelveli, India
renisha@gcetly.ac.in, jayasree@gcetly.ac.in

Abstract

This paper presents the enhancement of speech signals in a noisy environment based on Wavelet Based Adaptive filtering (WAF). In this technique, the speech signals contaminated with noise are processed through the WAF. The adaptive algorithm based on least squares, updates the weighting factors of the filter to keep the difference between input and desired signal at a minimum level to eliminate the unwanted noise and interferences. The performance of the proposed method is evaluated by computing the Signal-to-Noise Ratio (SNR), Peak Signal-to-Noise Ratio (PSNR), Root-Mean-Square-Error (RMSE), Percentage Root Mean Square Difference (PRD) after denoising. The investigation on speech signals contaminated by noise has demonstrated that the performance of the proposed method is stable and reliable in the noisy environment.

Keywords: *Speech, Discrete Wavelet Transform, Denoising, Adaptive filter*

1. Introduction

Speech signals carry a great deal of information that can be coded for transmission, and many other applications. However, under conditions with a great deal of background noise, for example, with many speakers in a room it may be difficult to get clear signal of a single user. Moreover, the audio signals are often contaminated by environmental noise and buzzing or humming noise from audio equipments [1-4,11].

The aim of the speech enhancement system is to gradually reduce the noise while retaining the underlying signals. Time frequency denoising techniques are adopted to eliminate the unwanted noise. Cedric Fevotte *et al.* [7] used Fourier Transform based techniques for audio signal denoising. After transformation, time information of the signal is lost, i.e., Fourier transform cannot say which spectral component exists at certain time. On the other hand, STFT provides information on time-frequency plane. Short-Time Fourier transform (STFT) can be used for denoising of audio signals [6]. But in STFT, the time frequency resolution is constant. But Wavelet Transform overcomes the situation with provision of variable resolution [6].

The effective noise reduction method using wavelet based on thresholding technique is used to remove the noise from the audio signals. Shima *et al.* [1] adopted wavelet thresholding method based on symmetric Kullback–Leibler divergence method for speech enhancement. Ting-Hua Yi *et al.* [2] adopted sigmoid- function based wavelet thresholding for smoothing vibration signals.

Especially, adaptive signal processing algorithms are under an intensive research due to their ability to handle spatially varying noise. The band-pass filtering assumes that random process representing the received signal consists of uncorrelated random process corresponding to the noise and signal, where their frequency spectra densities do not overlap [12]-[14].

ISSN: 2005-4254 IJSIP

V´ıtor *et al.* [8] showed that a combination of two filters from the same family (i.e., two LMS or two RLS filters) cannot improve the performance over that of a single filter of the same type with optimal selection of the step size or forgetting factor.

In this paper, speech signal enhancement based on WAF method is proposed which can detect the speech signals in the noisy environment with low SNR values.

2. Proposed Methodology

The proposed methodology for the enhancement of speech signals under noisy conditions is based on Wavelet based Adaptive Filtering (WAF) method. The proposed block diagram for the speech signal enhancement is shown in figure 1.

The speech signal corrupted with noise is the reference signal. It is decomposed in wavelet domain using Mallet's algorithm [3] producing approximation and detailed coefficients.

In wavelet decomposition, the signal x(n) is passed through a high pass filter g(n) and a low pass filter h(n). After that, half of the signal samples are eliminated by the process of down sampling. The scaling function is associated with a low pass filter combined with the filter coefficients h(n) and the wavelet function is associated with a high pass filter combined with the filter coefficients g(n).

The low pass filter coefficients are called the approximations and the high pass filter coefficients are called the details. The approximation coefficients and the detail coefficients are given by the following equations.

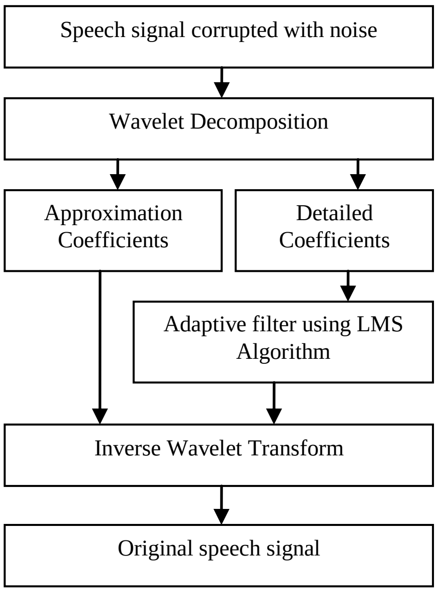

Figure 1. Block diagram for WAF based speech signal Enhancement system

$$A_j(k) = \sum_k x(n)h(-k+2n) \quad (2.1)$$

$$D_j(k) = \sum_k x(n)g(-k+2n) \quad (2.2)$$

Most of the noise is concentrated in the detailed (high frequency) coefficients. Adaptive filter is used to suppress the noise there by modifying the detailed coefficients [6].

3.1 LMS Algorithm

In this algorithm, the desired filter is obtained by finding the filter coefficients that relate to produce the least mean squares of the error signal. The error estimation is determined by comparing the output of the linear filter in response to the input signal to the desired response. Least Mean Square (LMS) then involves an adaptive process for automatic adjustment of the parameters of the filter according to the error estimation. The adaptive process is carried out by the weight control mechanism [14].

4. Performance Analysis

The performance of the improved denoising technique based on MDWT for audio signal enhancement is evaluated using the following quantitative measures.

4.1 Root Mean Square Error (RMSE)

RMSE is defined as square root of ‘mean of error squares’ and is calculated using the formula

$$RMSE = \sqrt{\sum_{k=1}^{N}[x(n) - x'(n)]^2} \tag{4.1}$$

where x´ (n) is the denoised signal, and x(n) is the original signal. The constant N is the number of samples composing the signal.

4.2 Signal to Noise Ratio (SNR)

Signal to Noise Ratio (SNR) is defined as the ratio of power between the signal and the unwanted noise. SNR is calculated using the formula.

$$SNR(db) = 10\ln \frac{\sum_{k=1}^{N} x^2(k)}{\sum_{k=1}^{N}[x(k) - x'(k)]} \tag{4.2}$$

where x´ (k) is the denoised signal, and x(k) is the original signal. The constant N is the number of samples composing the signal.

4.3 Peak Signal to Noise Ratio (PSNR)

PSNR is the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation.

$$PSNR = 10\log \frac{255}{1/N\sum\left|x(n) - x'(n)\right|} \tag{4.3}$$

4.4 Percentage Root Mean Square Difference (PDR)

The most prominently used distortion measure is the Percent Root Mean Square Difference (PRD) that is given by

$$PDR = \sqrt{\frac{\sum_{n=0}^{N-1}\left[x(n) - x'(n)\right]^2}{\sum_{n=0}^{N-1}\left[x(n)\right]}} \quad \text{x } 100 \tag{4.4}$$

5. Simulation Results and Evaluation

This section deals with the performance evaluation of WAF based denoising technique for speech signals in different noisy environments

The details of the simulation carried out for the noisy database of 30 IEEE sentences (produced by three male and three female speakers) corrupted by eight different real-world noises at different SNRs [10] are also discussed. The noise was taken from the AURORA database and includes suburban train noise, babble, car noise, exhibition hall noise, restaurant noise, street noise, airport noise and train-station noise. The signals are sampled at 8 kHz. The 30 clean speech signals are normalized and linked together, one after the other, obtaining a speech segment of 80 seconds length.

WAF based denoising technique is applied to the speech database embedded with noise to extract the Approximation coefficients and the Detailed coefficients. The speech signal corrupted with train noise (SNR = 10 db), their Detailed Coefficients and their Modified Detailed Coefficients after applying Adaptive filter are shown in Figure 2.

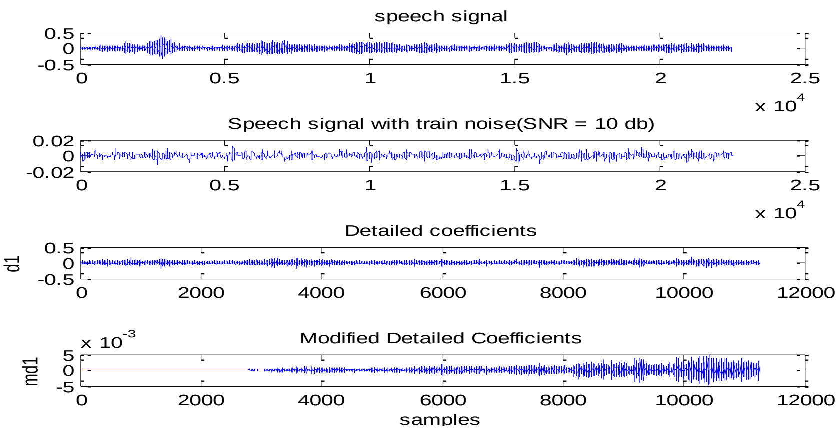

Figure 2. Speech Signal with Train Noise and its Detailed Coefficients

The denoised signal can be extracted by finding the IWT using the approximation coefficients and modified Detailed Coefficients of the noise corrupted speech signal. The effectiveness of this method is tested considering the speech database in the noisy environment for high and low SNR.

Figure 3 shows speech signal corrupted with noise (SNR=10 db, 5 db and 1 db) and the denoised signal using WAF.

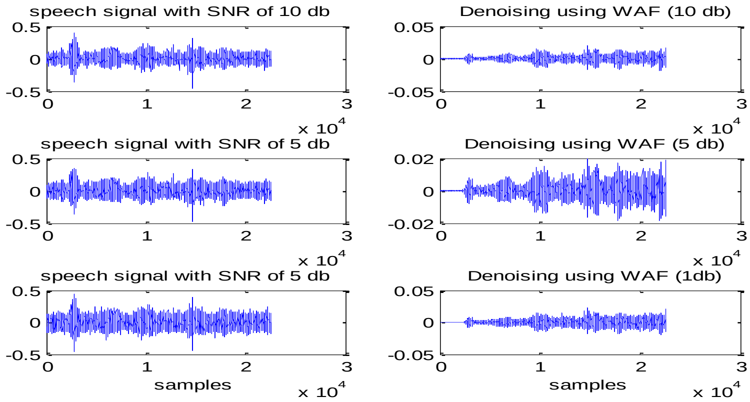

Figure 3. Speech Signal Corrupted with Noise (SNR=10 db, 5 db and 1 db) and the Denoised Signal using WAF

Figure 4 shows the comparison of RMSE for different SNRs using WAF based on LMS algorithm with other denoising techniques.

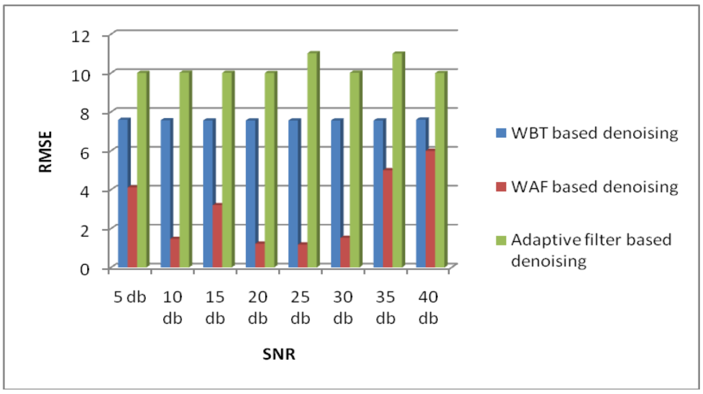

Figure 4. Comparison of RMSE for Different SNRs Using WAF Method with Other Denoising Techniques

From Figure 4, it is clear that, RMSE using WAF denoising is very less compared to the other denoising techniques. The performance of the advanced denoising technique is evaluated for the speech signals recorded in different noisy environments and the performance measures are depicted in Table 1.

Table 1. Performance Measures for speech signal under different noise Environments using MDWT

Type of noise added	RMSE	SNR	PSNR	PDR
Train noise	0.35	235.20	36.49	3.10
Car noise	1.18	-43.9	36.70	3.26
Airport noise	0.05	-657.43	36.29	3.11
Babble noise	0.66	66.04	36.11	3.27
Exhibition noise	0.62	100.6	36.52	2.94

The filter order also affects the performance of a noise cancellation system. Figure 5 illustrates how the RMSE changes as we change filter order. When filter order is increased (>15) the performance of WAF becomes good. It proves that the selection of right filter order is necessary to achieve the best performance.

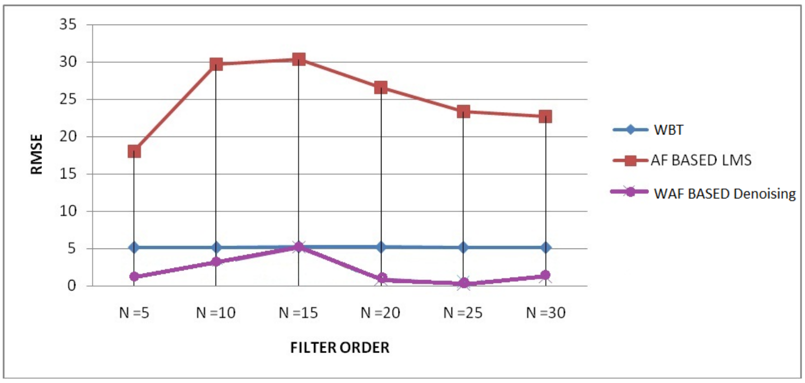

Figure 5. Comparison of RMSE for Different Values of N (Filter Order) MDWT based LMS and RLS Algorithm with Other Denoising Techniques

6. Conclusion and Future Work

This paper presented the enhancement of speech signals by Wavelet based Adaptive Filtering (WAF). The results of the improved denoising method are compared with other traditional methods. The implementation results have revealed that the process of WAF method offers better denoising effect and good error performance in the presence of any noise.

7. References

7.1. Journal Article

[1] S. Tabibian, A. Akbari and B. Nasersharif, "Speech enhancement using a wavelet thresholding method based on symmetric Kullback–Leibler divergence", International journal of signal processing, (**2015**), pp. 184–197.
[2] T.-H. Yi, H.-N. Li and X.-Y. Zhao, "Noise Smoothing for Structural Vibration Test Signals Using an Improved Wavelet Thresholding Technique", International Journal on Sensors, vol. 12, (**2012**), pp. 11205-11220.
[3] S. Vaidya and K. Shah, "Audio Denoising, Recognition and Retrieval by Using Feature Vectors", Journal of Computer Engineering, Journal of Computer Engineering, 2278-0661, 2278-8727, vol. 16, no. 2, Ver. II, (**2014**) March-April, pp. 107-112
[4] V. Harini, B. Sindhu, G. S. Kumari and H. Khan, "Block thresholding algorithm for enhancement of an audio signal corrupted by noise", IJRRAS, vol. 7, no. 2, (**2011**) vol. 7, no. 2.
[5] M. Alam, I. Islam, and M. R. Amin, "Performance Comparison of STFT, WT, LMS and RLS Adaptive Algorithms in Denoising of Speech Signal", IACSIT International Journal of Engineering and Technology (**2011**) June, vol. 3, no. 3.
[6] R. Sharma and V. P. Pyara, "A Comparative Analysis of Mean Square Error Adaptive Filter Algorithms for Generation of Modified Scaling and Wavelet Function", International Journal of Engineering Science and Technology, vol. 4, no. 4, (**2012**), pp. 1396-1401.
[7] C. Févotte, B. Torrésani, L. Daudet and S. J. Godsill, "Sparse Linear Regression with Structured Priors and Application to Denoising of Musical Audio", IEEE Transactions on Audio, Speech and Language Processing, vol. 16, no. 1, (**2008**), pp. 174-185.
[8] V. H. Nascimento, M. T. M. Silva, L. A. Azpicueta-Ruiz and J. Arenas-Garc´ıa, "On The Tracking Performance Of Combinations Of Least Mean Squares And Recursive Least Squares Adaptive Filters", (**2010**).

7.2. Book

[9] S. Haykin, "Adaptive Filter Theory", 3rd ed. Prentice Hall, (**2001**).

7.3. Chapter in a Book

[10] Y. Hu and P. Loizou, "Subjective evaluation and comparison of speech enhancement algorithms", Speech Communication, (**2007**), pp. 588-601.

7.4. Conference Proceedings

[11] N. Verma and A. K. Verma, "Real time adaptive denoising of musical signals in wavelet domain", Proc. Of IEEE International Conference on Engineering (NUiCONE), (**2012**) December, pp.1-5.
[12] J. Yu and D. C. Liu, "Thresholding Based Wavelet Packet Methods for Doppler Ultrasound Signal Denoising", IFMBE Proceedings Springer Verlag Berlin Heidelberg, vol. 19, no. 9, (**2008**), pp. 408-412.
[13] M. Al Mahamdy and H. Bryan Riley, "Performance Study of Different Denoising Methods for ECG Signals", Procedia Computer Science, vol. 37, (**2014**), pp. 325 – 332.
[14] G. Q. Xing, X. Coll, Xianning and Y. X. Zhang, "Analysis and comparison of RLS adaptive filter in signal De-noising", Proc of IEEE International Conference on Electrical and Control Engineering (ICECE), (**2011**), pp. 5754 – 5758.

Authors

T Jayasree, she completed her BE degree from Barathidasan University, Trichy in 1997 and ME from Govt. College of Technology, Coimbatore in 1999. She completed her Ph.D. degree from Anna University Chennai in 2011 in the area of signal processing. She has more than 15 years of teaching experience. Presently, she is working as an Assistant professor in the Department of ECE in Govt. College of Engineering, Tirunelveli, Tamilnadu, India. She has published many papers in National and International Journals and conferences. Her area of interest is signal processing applications in power quality, speech processing, Bio-medical signal processing.

G Renisha, she completed her BE degree from Anna University, Chennai in 2008 and ME from PSG College of Technology, Coimbatore in 2010. Presently, she is working as an Assistant professor in the Department of ECE in Govt. College of Engineering, Tirunelveli, Tamilnadu, India and pursuing Ph.D. in Anna University, Chennai. Her area of interest is signal processing applications in speech processing.