International Journal of Research in Computer Science
eISSN 2249-8265 Volume 3 Issue 3 (2013) pp. 29-33
www.ijorcs.org, A Unit of White Globe Publications
doi: 10.7815/ijorcs. 33.2013.065

IMPROVING THE EFFICIENCY OF SPECTRAL SUBTRACTION METHOD BY COMBINING IT WITH WAVELET THRESHOLDING TECHNIQUE

G. R. Mishra[1], Saurabh Kumar Mishra[2], Akanksha Trivedi[3], O.P. Singh[4], Satish Kumar[5]

**Department of ECE, Amity School of Engineering & Technology, Amity University, Lucknow Campus, INDIA*
Email: [1]grmishra@gmail.com, [2]saurabhmishra18@gmail.com, [3]akanksha2trivedi@gmail.com

Abstract: *In the field of speech signal processing, Spectral subtraction method (SSM) has been successfully implemented to suppress the noise that is added acoustically. SSM does reduce the noise at satisfactory level but musical noise is a major drawback of this method. To implement spectral subtraction method, transformation of speech signal from time domain to frequency domain is required. On the other hand, Wavelet transform displays another aspect of speech signal. In this paper we have applied a new approach in which SSM is cascaded with wavelet thresholding technique (WTT) for improving the quality of speech signal by removing the problem of musical noise to a great extent. Results of this proposed system have been simulated on MAT LAB.*

Keywords: *Coefficient Thresholding, Musical Noise, SSM, Wavelet Coefficients, WTT.*

I. INTRODUCTION

The musical noise produced by SSM is a major drawback of this system, but there are so many methods that have been given for musical noise reduction. This paper proposed a new technique in which SSM is cascaded with WTT for musical noise reduction.

SSM requires a transformation of signal from time domain to frequency domain using FFT. In this method, a voice activity detector [1] is used for detecting the signal whether it is voiced signal or unvoiced signal. This method is based on the direct estimation of the short term spectral magnitude of speech signal during non-speech activity. Spectral subtraction method is successful in stationary or slowly varying noisy environment, otherwise the estimated noise is not correct and system generates musical noise [10]. On the other hand, if we transform a signal into wavelet domain it simply breaks the signal into low frequency and high frequency components with the help of low pass filter and high pass filter that yields the coefficients. In this method, a thresholding technique is used for signal de-noising that discards the coefficients below threshold level. WTT [7] has been successfully used for image de-noising but a very less attention has been paid for practical implementation of this technique in the field of speech signal. WTT can de-noise [2] a signal without noticeable loss because it reveals the aspects like trends, breakdown points, discontinuities in higher derivatives. In this paper we have cascaded [8] WTT with spectral subtraction method because both techniques use different approach for signal de-noising. First we applied SSM and then the output of SSM is given as input in WTT for better results. This new method will be very effective for military applications, real time noisy environments.

II. SPECTRAL SUBTRACTION METHOD (SSM)

A. Introduction

SSM is very popular and useful for acoustic noise suppression because of its relative simplicity and ease of implementation. This method is used for restoration of power spectrum or magnitude spectrum of a speech signal contains additive noise. In this method, a noise is added acoustically or digitally into the original speech signal and it becomes noisy speech signal. Then we take an estimation of the noise spectrum that updated from the periods during non-speech activity when only noise is present. The estimation of noise spectrum is subtracted from noisy signal and then we get an estimate of the clean reconstructed signal. Generally, spectral subtraction is effective for stationary or slowly varying noisy environments.

B. Mathematical Approach

Suppose speech signal $x(m)$ is corrupted by noise $n(m)$ that yields noisy signal

$$Y(m) = x(m) + n(m) \quad \ldots (1)$$

When windowing the signal

$$Y_w(m) = x_w(m) + n_w(m) \quad \ldots (2)$$

Fourier transform of equation (2) is as under

$$Y_w(e^{jw}) = X_w(e^{jw}) + N_w(e^{jw}) \quad \ldots (3)$$

Where $Y_w(e^{jw})$, $X_w(e^{jw})$and $N_w(e^{jw})$ are the Fourier transforms of noisy speech, original speech, and noise signals respectively.

For simplification purpose w (windowed) notation is dropped.

When multiplying both sides by their complex conjugates, we find

$$[Y(e^{jw})]^2 = [X(e^{jw})]^2 + [N(e^{jw})]^2 + \ 2[X(e^{jw})][N(e^{jw})]cosD_q \ \ldots (4),$$

Where, D_q stands for phase difference between speech signal and noise signal.

$$D_q = \angle X\left(e^{j\omega}\right) - \angle N\left(e^{j\omega}\right) \ \ldots.. (5)$$

We take expected value on both sides of equation (4)

$$E\{[Y(e^{jw})]^2\} = E\{[X(e^{jw})]^2\} + E\{[N(e^{jw})]^2\} + 2E\{[X(e^{jw})]\}E\{[N(e^{jw})]\}E\{\cos(D_q)\} \ \ldots\ldots (6)$$

1. *Power spectral subtraction:*

For power spectral subtraction it is assumed that $E\left\{\cos D_q\right\} = 0$, hence equation (6) becomes

$$E\{[Y(e^{jw})]^2\} = E\{[X(e^{jw})]^2\} + E\{[N(e^{jw})]^2\}$$

So, $[X(e^{jw})]^2 = [Y(e^{jw})]^2 - E\{[N(e^{jw})]^2$ .... (7)

2. *Magnitude spectral subtraction:*

For magnitude spectral subtraction it is assumed that $E\left\{\cos D_q\right\} = 1$, hence equation (6) becomes

$$E\{[Y(e^{jw})]^2\} = E\{[X(e^{jw})]^2\} + E\{[N(e^{jw})]^2\} + \ 2E\{[X(e^{jw})]\}E\{[N(e^{jw})]\}$$

$$E\{[Y(e^{jw})]\} = E\{[X(e^{jw})]\} + E\{[N(e^{jw})]\}$$

$$[X(e^{jw})] = [Y(e^{jw})] - E\{[N(e^{jw})]\} \ \ldots.. (8)$$

The procedure of spectral subtraction method has been shown below in figure 1.

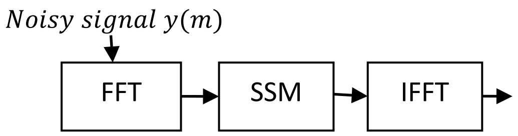

Figure1: Basic blocks of spectral subtraction method

III. WAVELET THRESHOLDING TECHNIQUE (WTT)

A. Introduction

SSM is effective for stationary or slowly varying noises, but in mobile communication, signal is definitely not stationary. So the next possible improvement in speech signal is to further decrease the problem of musical noise using WTT. In wavelet transform the output speech signal $\hat{x}(m)$ of spectral subtraction method has been taken as an input signal and that signal is divided up into low frequency and high frequency components. The output of LPF is known as approximation coefficients and the output of HPF is called detail coefficients. When we analyze approximation coefficients [9] at level 1 by using MATLAB command *sound* (cA1, Fs, bit depth) we can understand the speech with a low loss in the quality of signal. This shows that low frequency components contain essential information and that is why the output of LPF is called approximation coefficient. The output of HPF contains only high frequency non-essential information and is known as detail coefficient. For applying wavelet technique first we have to choose an appropriate mother wavelet and level of decomposition of the signal. Choosing a mother wavelet depends on the type of the signal we have to decompose. While speech de-noising our objective is to improve quality of the signal, so wavelet can be selected on the basis of energy conservation properties in the approximation coefficients [7]. By using Daubechies D20, D6, D4, D2 or Haarwavelets, more than 90% of the signal energy, level 1 approximation coefficients contains. For selecting a decomposition level, if the frame based input is applied, then frame size must be a multiple of 2^n, where n represents the decomposition level. In this paper, we have selected ‘Daubechies’ as a mother wavelet and decomposition level is 6.

B. Wavelet approach for musical noise reduction

Wavelet thresholding technique is very useful and a different technique for residual noise reduction. Residual noise come into existence because of variation in background noise, and that is why residual noise occurs during whole speech (including speech activity as well as non- speech activity). Using wavelet thresholding technique we are exploiting the fact that residual noise contains narrower peaks which are relatively high frequency components. More than 90% components of speech signal have values zero or near to zero that is clear from histogram representation. Here a threshold value is selected and all the coefficients are truncated that have values lower than threshold, so wavelet thresholding technique removes residual noise (also called musical noise in time domain) successfully to the great extent.

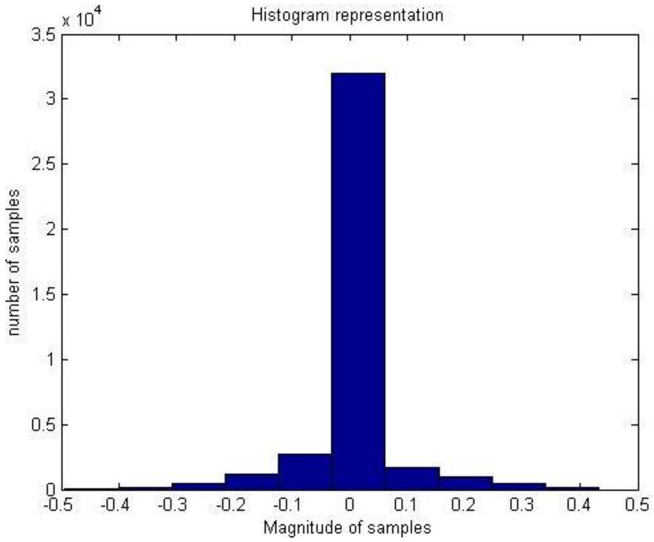

Figure2: Histogram representation

IV. THRESHOLDING OF COEFFICIENTS

After applying wavelet transform, input signal is decomposed into coefficients. Then we perform thresholding of coefficients for signal de-noising which is of two types, *hard thresholding* and *soft thresholding*. Generally hard thresholding is used for signal compression and soft thresholding is used for signal de-noising. Here we have used soft thresholding for de-noising the signals. Soft thresholding is an expansion of hard thresholding in which we first set to zero the elements whose absolute values are lesser than the threshold and then shrink the nonzero coefficients toward 0. After choosing soft thresholding, there are two types for finding a threshold value named *global thresholding* and *level dependent thresholding*. In global thresholding, a threshold value is set manually. For level dependent thresholding, we use Brige-Massart strategy [7] that yields a different threshold values for each level. To de-noise a signal we use a MATLAB command *wdencmp* that enables us to choose between global and level dependent thresholding. Coefficient thresholding discards the coefficient that has a value below the threshold and it results de-noised signal. In wavelet de-noising method we have taken $\hat{x}(m)$ as an input signal that is output signal of SSM. Steps involved in wavelet de-noising process are shown in figure 3.

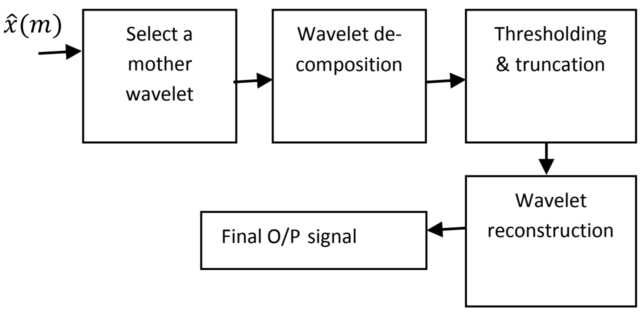

Figure3: Wavelet de-noising process

V. PERFORMANCE ANALYSIS OF PROPOSED SYSTEM

Performance analysis of this proposed system has been done in terms of Peak signal to noise ratio (PSNR) and Normalized root mean square error (NRMSE).

PSNR has been evaluated using

$$PSNR = 10log_{10}\frac{NX^2}{\|x-r\|^2}$$

Where, N is the length of the reconstructed signal, X is the maximum absolute square value of signal x. $\|x-r\|^2$ is the energy of the difference between original and reconstructed signal.

And NRMSE has been evaluated using

$$NRMSE = \sqrt{\frac{(x(n)-r(n))^2}{(x(n)-\mu x(n)^2}}$$

Where, $x(n)$ is the speech signal, $r(n)$ is the reconstructed signal and $\mu x(n)$ is the mean of the speech signal.

For better results PSNR should be higher while value of NRMSE should be as low as possible.

We have taken a male spoken speech signal of 5 sec with 8 KHz sampling frequency and bit depth is 16, shown in figure4

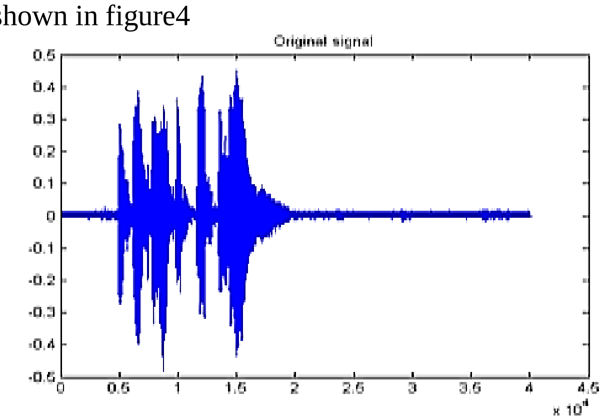

Figure 4: Original speech signal

After digitally added random noise in original speech signal, the noisy speech signal is shown in figure 5

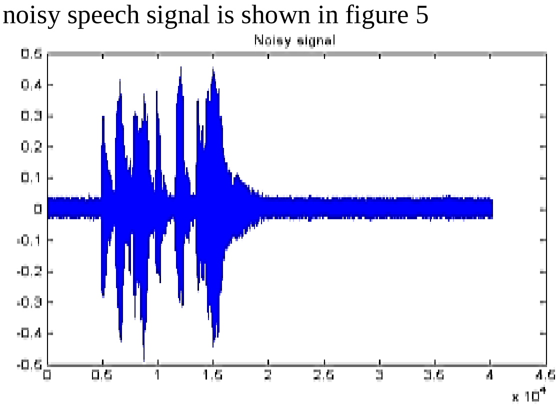

Figure5: Noisy signal

We applied SSM for signal de-noising and got reconstructed signal shown in figure 6.

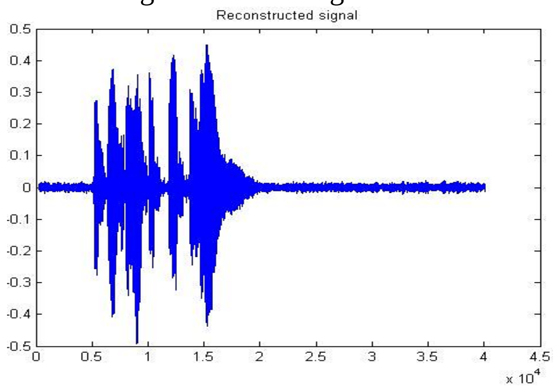

Figure 6: Output de-noised signal of spectral subtraction

After getting the output de-noised signal using SSM, we used command *sound (reconstructed signal, Fs, bit depth)* to hear the de-noised signal and got a great improvement in the quality of signal (PSNR and NRMSE of $\hat{x}(m)$ using SSM is 13.4981dB and 1.0818) but a little bit presence of noise still we can feel that is identified by musical noise. So we have used a new technique for reducing musical noise in which the reconstructed signal using SSM is taken as input signal for WTT. After transforming this signal into wavelet coefficients and applying thresholding respectively we got an output signal with reduced musical noise. This final output signal with reduced musical noise is shown in figure 7.

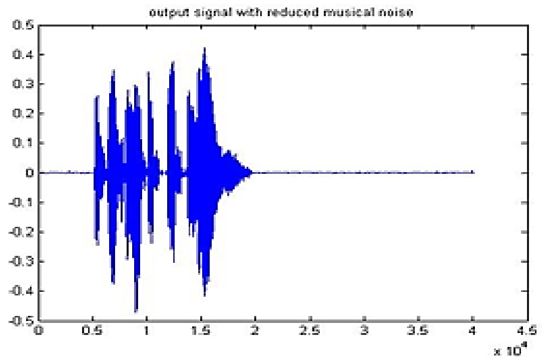

Figure 7: Signal of reduced musical noise using haar wavelet

Table (a)

Wavelet type	Decomp -osition level	Percentage Retained energy	PSNR in dB	NRMSE
Haar	6	83.7857	14.4836	1.0298
Db2	6	86.5747	14.3677	1.0357
Db4	6	87.9903	14.2931	1.0396
Db6	6	88.5790	14.2650	1.0411

PSNR using SSM is 13.4981dB, and NRMSE using SSM is 1.0818 and the PSNR and NRMSE values given in table (a) have been observed using proposed new system (SSM+WTT). So it's clear from PSNR and NRMSE values that there is a significant improvement in the speech signal by cascading SSM withWTT.

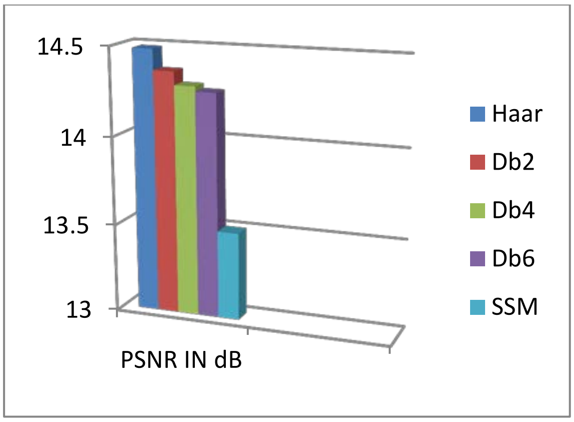

Figure 8: Performance evaluation based on PSNR

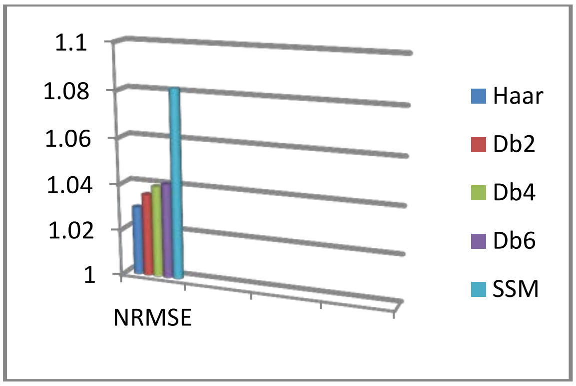

Figure 9: Performance evaluation based on NRMSE

VI. CONCLUSION AND FUTURE SCOPE

Musical noise is a problem of spectral subtraction method that has been eliminated using wavelet thresholding technique (WTT). In this paper we have proposed a new system (SSM+WTT) which combined SSM and WTT respectively and the efficiency of the proposed system is higher as compared to SSM. Result of this combined system is clear from the waveform shown in figure 7 and differences between PSNR and NRMSE values. Table (a) represents the type of mother wavelet, decomposition level, percent retained signal energy in de-noised signal, peak signal to noise ratio (PSNR) and NRMSE. Haar wavelet has highest PSNR and lowest NRMSE values. Results have been simulated on MATLAB.

In future, if we use Wavelet Packet Transform instead of Wavelet transform with adaptive thresholding technique, the quality of reconstructed speech signal will be better.

VII. REFERENCES

[1] S. F. Boll, "Suppression of acoustic noise in speech, using spectral subtraction" *.IEEE. Acoustic.Speech, Signal Processing*, vol. ASSP-27, pp. 113-120, Apr. 1979. doi: 10.1109/TASSP.1979.1163209

[2] *Ing Yann Soon Soo Ngee Koh Cii Kiat Yeo*, "Wavelet For Speech De-noising", 1997 IEEE Tencon - Speech and Image Technologies for Computing and Telecommunications.

[3] Soltani Bozchalooi, Ming Liang, “A Combined Spectral Subtraction and Wavelet De-Noising Method for Bearing Fault Diagnosis”, IEEE Amercian Control Conference, pp 2533-2538, 2007. doi: 10.1109/ACC.2007.4282467
[4] Talbi Mourad, Cherif Adnene, “Simulation and comparison of noise cancellation technique in speech processing”, IEEE Electrotechnical Conference, pp 541-544, 2006. doi: 10.1109/MELCON.2006.1653158
[5] Wilfred N Mwema and Elijah Mwangi, “A Spectral Subtraction Method for Noise Reduction in Speech Signals”, IEEE AFRICON 4th, pp 382-385. 1996. doi: 10.1109/AFRCON.1996.563142
[6] Saeed V. Vaseghi “Advanced Digital Signal Processing and Noise Reduction”, Second Edition. John Wiley & Sons Ltd ISBNs: 0-471-62692-9 (Hardback): 0-470-84162-1 (Electronic). doi: 10.1002/0470841621
[7] Nikhil Rao, “Speech Compression Using Wavelets”, ELEC 4801 Thesis Project, School of Information Technology and Electrical Engineering, Qld 4108, October 18, 2001.
[8] WANG Guang-yan, ZHAO Xiao-qun, WANG Xia, “Musical Noise Reduction Based on Spectral Subtraction Combined with Wiener Filtering for Speech Communication”, IET International Communication Conference on Wireless Mobile and Computing, pp 726-729, 2009.
[9] Satish Kumar, O.P. Singh, G.R. Mishra, Saurabh Kumar Mishra, Akanksha Trivedi “Speech Compression and Enhancement using WaveletCoders”, International Journal of Electronics Communication and Computer Engineering Volume 3, Issue 6, ISSN (Online): 2249–071X, ISSN (Print): 2278–4209.
[10] James R. Hamilton, “Musical Noise”, British Journal of Aesthetics, Vol. 39, No. 4, October lag, pp 350-363. doi: 10.1093/bjaesthetics/39.4.350
[11] Ben Gold and Nelson Morgan. 'Speech and Audio Signal Processing'. John Wiley and Sons, 2000.
[12] Jr. J.R. deller, J. Hansen and J.G. Proakis, “Discrete-time processing of speech signals”, IEEE press, New York 2000.
[13] Chin-Teng Lin, “Single-channel speech enhancement in variable noise-level environment”, Systems, Man and cybernetics, Part A, IEEE Transactions, vol. 33 , no. 1 , pp 137–143, Jan. 2003. doi: 10.1109/TSMCA.2003.811115

How to cite

G. R. Mishra, Saurabh Kumar Mishra, Akanksha Trivedi, O.P. Singh, Satish Kumar, "Improving the Efficiency of Spectral Subtraction Method by combining it with Wavelet Thresholding Technique". *International Journal of Research in Computer Science*, 3 (3): pp. 29-33, May 2013. doi: 10.7815/ijorcs. 33.2013.065