RESEARCH PAPER | Engineering | Volume - 5 | Issue - 1 | Jan Special Issue - 2015 | ISSN - 2249-555X

A Modified Speech Enhancement Algorithm Based on the Discrete Wavelet Transform

KEYWORDS Discrete Fourier Transform, Minimum mean-square error, Noise estimation, Noise reduction, Signal to Noise Ratio, Speech enhancement, Spectral Subtraction, Wavelet Transform, Discrete Wavelet Transform, Wavelet Packet Transform

Kanu Patel	Vishnu Patel
Asst. Professor, EC Engineering, SPBPEC,Saffrony, Mehsana	Asst. Professor, EC Engineering, SPBPEC,Saffrony, Mehsana
Dipal Patel	**Hardik Patel**
Asst. Professor, EC Engineering, SPBPEC,Saffrony, Mehsana	Asst. Professor, EC Engineering, SPBPEC,Saffrony, Mehsana

ABSTRACT In an everyday life the speech communication is vivid uses for the hearing impaired and numerous other applications. Speech is a fundamental means of human communication. After over thirty years of research throughout the world, no perfect solution exists to this problem. The objective of our work is to implement a novel speech enhancement algorithm, which offers superior noise reduction over current methods. All speech enhancement systems suffer from distortion or residual noise due to imperfect noise removal. It is often necessary to perform denoising in speech processing system operating in highly noisy environment. Wavelet transform is one of the most promising techniques used in signal processing, due to its ability to decompose signals and to reduce noise having non- stationary characteristics. Wavelet Packet Transform is the advantageous over wavelet transform that it also resolves higher frequency component of the signal. In the present work wavelet thresholding and wavelet packet thresholding algorithm has been used to reduce the noise from the speech signal. A simple SURE threshold method is proposed to compute the optimum threshold value. Mean square error at different values of SNR is computed to evaluate the performance of the proposed method like traditional spectral subtraction, Weiner Filtering method, Spectral Subtraction with MMSE etc. The result obtained is compared with the other speech enhancement algorithms given in various reference papers. In comparison to other reference papers we get improved results in terms of SNR and MSE. Simulation is done on MATLAB platform.

I. INTRODUCTION
Enhancement of speech signals is required in many situations in which the signal is to be communicated or stored. Speech enhancement is required when either the signal or its receiver is degraded. For example, hearing impaired individuals require enhancement of perfectly normal speech with their individual hearing capabilities. Speech signals produced in a room generate reverberations, which may be quite noticeable when a hands-free single channel telephone system is used and binaural listening is not possible.

Whenever speech is recorded by a microphone, unwanted noise is also recorded. This noise depends on the environment and can range from anything such as computer fan noise, car engine noise to factory floor noise. The goal of any speech enhancement system is to suppress or completely remove the unwanted noise while maintaining the quality and/or intelligibility of the speech. This has been an ongoing area of research since it was first proposed in 1979 by Boll in [1].

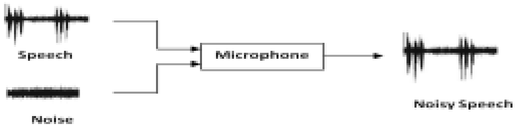

Fig. 1. Basic overview of additive noise

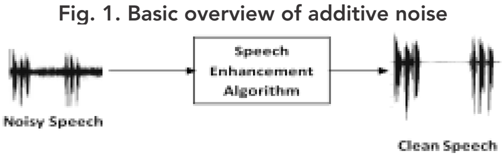

Fig. 2. Basic overview of a speech enhancement system

I. DENOISING BY WAVELET DECOMPOSITION PROCESS
The Wavelet Transform has recently gained a lot of popularity in the field of signal processing since it has the capability to provide both time and frequency information simultaneously, hence it gives a time-frequency representation of the signal [11].

A. Wavelet transform vs. Fourier Transform
The traditional Fourier Transform only provides spectral information about a signal and only works for stationary signals while many real world signals are non-stationary and need to be processed in real time [12]. The problem with Short Time Fourier Transform (STFT) can be attributed to the Heisenberg uncertainty principle which states that it is impossible for one to obtain the time instance at which frequencies exist but, one can obtain the frequency bands existing in a time interval. Also the resolution window used in STFT is of constant length whereas with Wavelet transform we can have multi resolution analysis.

B. The Discrete Wavelet Transform
Calculating wavelet coefficients at every possible scale is a fair amount of work, and it generates an awful lot of data. So we choose scales and positions based on powers of two - called dyadic scales and positions - then our analysis will be much more efficient and just as accurate. We obtain just such an analysis from the discrete wavelet transform (DWT). An efficient way to implement this scheme using filters was developed in 1988 by Mallat.

In the continuous wavelet function

$$\Psi_{a,b} = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right) \quad (1)$$

Volume - 5 | Issue - 1 | Jan Special Issue - 2015 | ISSN - 2249-555X

By putting appropriate value of a and b (scale and shifting parameter), we get scaling and wavelet function for Discrete Wavelet Transform.

C. Wavelet Decomposition

Given a signal s of length N, the DWT consists of log2N stages at most. The first step produces, starting from s, two sets of coefficients: approximation coefficients cA1 and detail coefficients cD1. These vectors are obtained by convolving s with the low-pass filter LoF_D for approximation, and with the high-pass filter HiF_D for detail, followed by dyadic decimation.

More precisely, the first step is:
1. Decomposition step

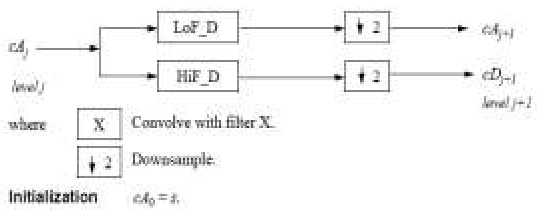

Fig. 3. Wavelet Decomposition with filters and downsampling

Low pass analysis channel: Filter and downsample:
Approximation vector,

$$cA_1 = (\downarrow 2)F = (\downarrow 2)\sqrt{2} * \mathrm{LoF\_D} \qquad (2)$$

High pass analysis channel: Filter and downsample:
Detail vecto

$$cD_1 = (\downarrow 2)G = (\downarrow 2)\sqrt{2} * \mathrm{HiF\_D} \qquad (3)$$

So the wavelet decomposition of the signal s analyzed at level j has the following structure like [cAj, cDj... cD1]. This structure contains for J=3, the terminal nodes of the following tree:

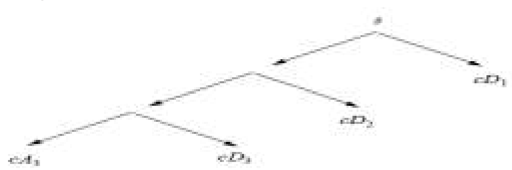

Fig. 4. Wavelet Decomposition Tree

Conversely, starting from cA_j and cD_j the IDWT reconstructs cA_{j-1} inverting the decomposition step by inserting zeros and convolving the results with the reconstruction filters.

2. Reconstruction step

• Before reconstruction apply any threshold technique to the detail coefficients to remove the noise from it.

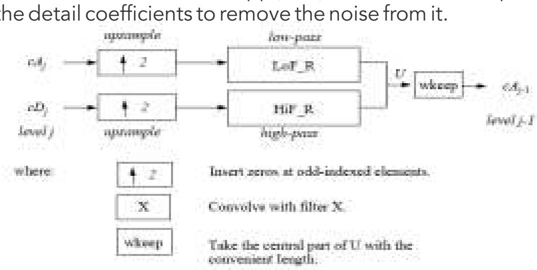

Fig. 5. Generation of coefficient

· As we know that at any level coefficient vectors cAj and cDj – cannot directly be combined to reproduce the signal. Because they were produced by downsampling, contain aliasing distortion, and are only half the length of the original signal. It is necessary to reconstruct the approximations and details before combining them.

· Let us denote h = LoF_R and g = HiF_R and focus on the one-dimensional case. We first justify how to go from level j to level j+1 for the approximation vector. This is the main step of the decomposition algorithm for the computation of the approximations.

$$A_k^{(j+1)} = \sum_n h_{n-2k} A_n^{(j)} \qquad (4)$$

This formula resembles a convolution formula.

- The details are calculated in the same way.
- In discrete wavelet transform (DWT) a signal s(t) limited to a scale J can be represented as:

$$s(t) = \sum_{k=-\infty}^{\infty} A_0[k]\phi_{0,k}(t) + \sum_{-\infty}^{\infty}\sum_{j=0}^{J-1} D_j[k]\psi_{j,k}(t) \qquad (5)$$

III.IMPLEMENTATION METHODOLOGY

We can define speech enhancement as away to improve the quality of speech, so that the resultant speech is better than the original one. This process of improving the quality of a speech by manipulating the noisy speech is done with MATLAB software. We worked in MATLAB for speech enhancement which removes the noise in the input speech. For that we use various speech files collected from internet open source database. We implemented the functional code for proposed work using MATLAB with Pentium core2duo window based PC. First step is to select the noisy speech or select only .wav speech file and manually add the various noises. Now apply the algorithm for speech enhancement.
Figure 6 shows the Block Diagram, followed in order to fulfill the aim of our research work. At last, we evaluated result in terms of SNR and MSE of the Denoised Signal. Also we compared with the major reference paper [14].
The block diagram shown in Figure 6 gives the actual implementation of the method proposed.

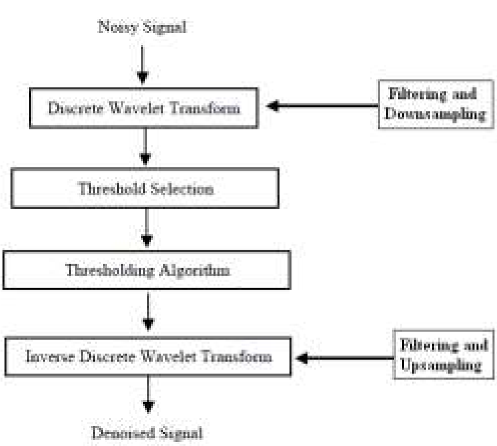

Fig. 6. Main algorithm work flow

The system described above is simulated on a MATLAB.

A. Discrete Wavelet transform

In case of Wavelet Packet Decomposition it offers a richer range of possibilities for signal analysis and which allows the best matched analysis to a signal. It provides level by level

Volume - 5 | Issue - 1 | Jan Special Issue - 2015 | ISSN - 2249-555X

transformation of a signal from the time domain into the frequency domain. In wavelet analysis, a signal is split into an approximation and a detail coefficient. The approximation coefficient is then itself split into a second-level approximation coefficients and detail coefficients, and the process is repeated. In wavelet packet analysis, the details as well as the approximations can be split. This yields more than different ways to encode the signal. The top level of the WPD tree is the time representation of the signal. As each level of the tree is traversed there is an increase in the tradeoff between time and frequency resolution. The bottom level of a fully decomposed tree is the frequency representation of the signal.

B.Entropy criteria (Number of Levels)[[16]]

Since the analysis process is iterative, in theory it can be continued indefinitely. In reality, the decomposition can proceed only until the individual details consist of a single sample or pixel. In practice, you'll select a suitable number of levels based on the nature of the signal or on a suitable criterion such as entropy.

C. Thresholding

Optimal de-noising requires a more suitable approach called thresholding. Before the reconstruction or synthesis process we apply one of the thresholding techniques. This involves discarding only the portion of the details that exceeds a certain limit. This is the portion where noise is removed from coefficients. Wavelet Packet analysis illustrates the use of Stein's Unbiased Estimate of Risk (SURE) as a principle [24] for selecting a threshold to be used for de-noising. This technique calls for setting the threshold T to

$$T = \sqrt{2\log_e(n\log_2(n))} \tag{6}$$

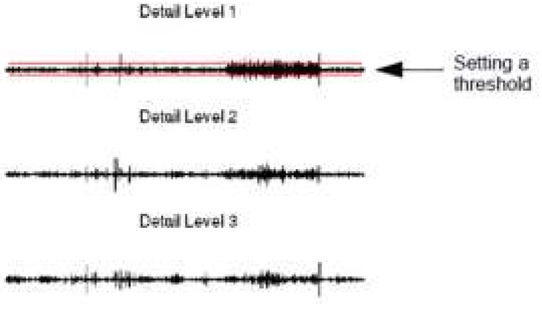

Fig. 7. Thresholding of Detail coefficients

D. Inverse Wavelet transforms

In this block we perform the synthesis process for denoising the signal, which includes two main functions, Filtering and Upsampling. Upsampling is the process of lengthening a signal component by inserting zeros between samples.

E.Performance Evaluation:

Performance evaluation tests can be done by subjective quality measures and objective quality measures. Objective measures provide a measure that can be easily implemented and reliably reproduced. Objective measures are based on mathematical comparison of the original and processed time domain signals. The majority of objective quality measures quantify time domain quality of the signal in terms of a numerical distance measure. The signal to noise ratio is the most widely used method to measure time domain signal quality. It is calculated as the ratio of the signal to noise power in decibels.

We apply our method to enhance the noisy signal. And the performance of our algorithm is compared with [14]. We evaluated our result in terms of SNR and MSE of the Denoised Signal which is implemented in MATLAB.

$$SNR = 10\log_{10}\left\{\frac{\Sigma(x)^2}{\Sigma(x-x_d)^2}\right\} \tag{7}$$

$$MSE = \frac{1}{N}\{\Sigma(x-x_d)^2\} \tag{8}$$

where x is speech signal, x_d is denoised signal (Enhanced signal) and N is the no. of samples of speech files.

IV. SIMULATION AND RESULTS

This topic contains the results, obtained after following the Speech Enhancement Algorithm. The results have been demonstrated in the form of comparison tables and graphs. After the comparison tables, a graphical representation has also been done for a quick analysis of results. All the techniques have been tested for all the assumed standard speech signals collected from internet open source database. Here we evaluate our result in terms of SNR and MSE of denoise signal. Also we compared it with other reference papers.

A. Comparison of SNR and MSE only for Babble Noise:

Here we compare the result of [14] with proposed work

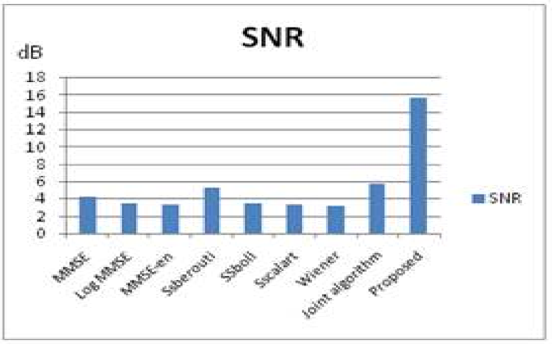

Fig. 8. Comparison in terms of SNR

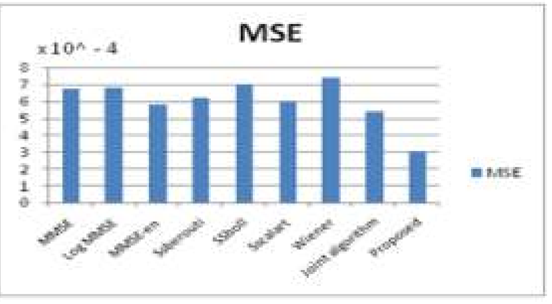

Fig. 9. Comparison in terms of MSE

TABLE I
Comparison in terms of SNR and MSE

References	Algorithm	SNR (dB) (Babble Noise 5 dB)	MSE
Proposed work	Wavelet Based Algorithm	15.57	0.0003
[10]	Joined Algorithm (spectral subtraction algorithm with Dereverberation)	5.8	0.0005

Volume - 5 | Issue - 1 | Jan Special Issue - 2015 | ISSN - 2249-555X

Table 1 shows the comparison of different speech enhancement algorithms given in [14] with speech signals corrupted by 5dB babble noise. In [14], Joined Algorithm (spectral subtraction algorithm with Dereverberation) for enhancement of the speech signal is applied. The graph (illustrates the comparison of SNR values obtained with the existing method and the proposed method with real time noises such as the babble noise. It can be observed that using wavelet packet transform in proposed algorithm, higher SNR can be achieved which in turn reduces the MSE.

A. Comparison of Output SNR for different noise:
Here we compare the result of [15] with proposed work.

TABLE II
Comparison in terms of SNRs only

Ref.s	Algorithm	SNR (dB)		
		White Noise	Random Noise	Color Noise
Proposed work	Wavelet Based Algorithm	18.8655	9.6955	13.43
[11]	Spectral Subtraction	8.5699	2.2479	1.123
	Wiener Filter	12.2530	3.0670	9.23
	Kalman Filter	17.8932	5.45473	12.81

Table 2 shows the comparison of different speech enhancement algorithms given in [15] with speech signals corrupted by three different noises. In [15], it uses Kalman filter for speech enhancement and compares the other two algorithms shown in table 2. In proposed work by the use of wavelet packet transform the output SNR is improved in each case.

I. CONCLUSION

With wavelet packets we have a greater variety of options for decomposing the signal. The method presented is used for time as well as frequency analysis of time varying signals. From the results we conclude that the wavelet filtering find applications in the time domain analysis and synthesis era. Here we use Coiflet and Daubechies wavelet that improves the SNR of denoise signal as compared to other. The system has been tested with various sampling frequencies for time domain samples which gave satisfactory output. From the results we conclude that Discrete Wavelet Decomposition technique and wavelet packet transform with the use of different thresholding techniques improves the SNR and hence reduces the MSE as compared to other speech enhancement techniques.

REFERENCE

[1] Zhixin Chen, “Simulation of Spectral Subtraction Based Noise Reduction Method” (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No.8, 2011 | [2] The Artifacts of Spectral Subtractions, http://vocal.com/noise_reduction/spectral_subtraction.html | [3] S.F. Boll, “Suppression of acoustic noise in speech using spectral subtraction,” IEEE Transactions on Acoustics, Speech, Signal Processing, vol.ASSP-27, Apr. 1979 | [4] Audio Demonstration: Speech Enhancement for Electronic Hearing Aids http://www.ind.rwth-aachen.de/research/cochlear/audiodemo.html | [5] Y. Ephraim and H.L. Van Trees, “A signal subspace approach for speech enhancement,” IEEE Transactions on Speech and Audio Processing, vol. 3,July 1995 | [6] M. Dorbecker, S. Ernst, “Combination of Two-Channel Spectral Subtraction and Adaptive Wiener Post-Filtering for Noise Reduction and Dereverberation,” | [7] “Perceptual Evaluation of Speech Quality (PESQ),” recommendation ITU-TP.862, International Telecommunication Union, Feb. 2001. | [8] Wang, Zhanfeng; Li, Suping, "Discrete Fourier Transform and Discrete Wavelet Packet Transform in speech denoising," Image and Signal Processing (CISP), 2012 5th International Congress on , vol., no., pp.1588,1591, 16-18 Oct. 2012 | [9] Gilbert Strang, Truong Nguyen, “Wavelets and Filter Banks”, Wellesley, MA, Wellesley - Cambridge Press,1996 | [10] Deepa, D.; Prakarsha, C.; Shanmugam, A., "Single channel speech enhancement using spectral gain shaping method and dereverberation for digital hearing aid," Computer Communication and Informatics (ICCCI), 2012 International Conference on , vol., no., pp.1,6, 10-12 Jan. 2012 | [11] Mathe, M.; Nandyala, S.P.; Kumar, T.K., "Speech enhancement using Kalman Filter for white, random and color noise," Devices, Circuits and Systems (ICDCS), 2012 International Conference on , vol., no., pp.195,198, 15-16 March 2012 | [12] Ming Zhang; Kai-Cheng Li, "The application of wavelet energy entropy and LS-SVM to classify power quality disturbances," Wavelet Analysis and Pattern Recognition, 2009. ICWAPR 2009. International Conference on , vol., no., pp.154,159, 12-15 July 2009