Studies on Implementation of Harr and daubechies Wavelet for Denoising of Speech Signal

Mahesh S. Chavan
Department of Electronics and Telecommunication,
Rajarambapu Institute of Technology, Rajaramnagar
Sakharale (MS) India.
maheshpiyu@gmail.com

Nikos Mastorakis
Technical University,Sofia, Bulgeria
mastorakis4567@gmail.com

ABSTRACT

In hands free speech communication environments situation occurs that speech is superposed by background noise. Over the past few decades there is tremendous increase in the level of ambient environmental noise. This has been due to growth of technology. Noise is added by various factors like noisy engines, heavy machines, pumps, vehicles, over noisy telephone channel or using radio communication device in an aircraft cockpit. As speech is transmitted and received using various media it introduces distortions and have bandwidth constraints. These degradations lower intelligibility of speech message causing severe problems in downstream processing and user perception of speech signal. There has been a lot of research in speech denoising so far but there always remains room for improvement. The motivation to use wavelet as a possible alternative is to explore new ways to reduce computational complexity and to achieve better noise reduction performance. The wavelet denoising technique is called thresholding. It is divided in three steps. The first one consists in computing the coefficients of the wavelet transform (WT) which is a linear operation. The second one consists in thresholding these coefficients. The last step is the inversion of the thresholded coefficients by applying the inverse wavelet transform, which leads to the denoised signal. This technique is simple and efficient. In this paper wavlet is used as denoising algorithm. Performance of the Haar and Daubechies wavelets are experimentally evaluated.

Keywords
Harr wavelet, Speech Processing, Filtering, Daubechies wavelet.

INTRODUCTION

Speech is a very basic way for humans to convey information with the emotion of a human voice to one another. People use speech to communicate messages.[1] Human speech can be modeled as filter acting on excitation waveform. The vocal tract shape causes certain frequencies in the excitation to be amplified and attenuates other frequencies. The excitation takes the form of quasi periodic puffs of air, which causes the output speech to appear periodic. Speech can be divided into voiced and unvoiced. Voiced speech has a spectrum with energy concentrated at discrete frequencies i.e. at the fundamental frequency of the vocal folds and its multiples (harmonics). About one third of speech is completely a periodic (unvoiced) resulting from a random excitation that resembles white noise, caused by air rapidly passing through a narrow constriction in vocal tract.[2]. There often occur conditions under which we measure and then transform the speech signal to another form in order to enhance our ability to communicate with a band width of only 4 kHz. With the advent of the wonders of digital technology, the analog to digital converter samples the electrical speech e.g. 8000 samples per second for telephone speech, so that speech signal can be digitally transmitted and processed. During transmission and reception signals are often corrupted by noise which is unwanted signal. There are many forms of noise. One of the most common sources of noise is background noise which is always present at any location. Other types of noise include channel noise which affects both analog and digital transmission, quantization noise which results from over compression of speech signals, multi talker babble, reverberation noise or delayed version of noise are also present in some situations.[5] The additive background noise is random in nature and also uncorrelated with speech[11]. It present in various environment scenarios like offices, cars, city streets fans, factory environment, helicopters etc. Incas of additive background noise the assumptions made for developing enhancement methods are 1) Speech and noise signals are uncorrelated at least over a short time basis. 2) Noise is stationary or slowly varying over several frame of speech and 3) Noise can be represented as zero mean random processes [13]. In case of reverberation, reflections of speech from various objects will be mixed with the speech in a convolutive fashion. Thus degradation in case of reverberation is signal dependent, whereas, it is independent in case of additive background noise. Speech from other speakers may also get mixed with desired speaker's speech in an additive fashion. Since the characteristics of degradation are different in each case, degraded speech may need to be processed in different ways. Therefore an automated means of removing the noise would be an invaluable first stage for many signal processing tasks. Denoising has long been a focus of research and yet there always remains room for improvement. Speech enhancement in general various objectives to increase accuracy of speech recognition systems operating in noisy environments.[5]

Broad classification of speech enhancement methods is classified as time domain approach and transforms domain approach. Filtering performed directly on time sequences includes techniques such as LPC based digital filtering , kalman filtering and hidden markov model(HMM). In the transform domain approach noise attenuation is performed on transform coefficients. Transform can be Fourier transform (FT), Karhunen Loeve transform, discrete cosine transform (DCT) or wavelet transform (WT).Time domain filtering of the corrupted signal are simple methods originally employed however, this is only successful when removing high frequency noise from low frequency signals and does not provide satisfactory results under real world conditions.[5] The problem of denoising consists of removing noise from corrupted signal without altering it. Fourier domain was long been the method of choice to suppress noise [3].The classical methods based on spectral subtraction are effective for this purpose; however they introduce artificial noise and alter the original signal. While solving engineering problems the utilization of wavelets in signal and image processing has been found to be a very useful tool.

1.1 Wavelet Transform:

The Wavelet transform was inspired by the idea that we could vary the scale of the basis functions instead of their frequency. The fundamental idea behind wavelets is to analyze according to scale, instead of representing a function as a sum of weighted delta functions (as in the time domain), or as a sum of weighted sinusoids (as in the frequency domain), it represents the function as a sum of time-shifted (translated) and scaled (dilated) representations of some arbitrary function, which is called a wavelet. An advantage of wavelet transforms is that, Wavelet analysis allows the use of long time intervals for low-frequency information, and shorter regions for high-frequency information.[8] The Discrete Wavelet Transform (DWT) involves choosing scales and positions based on powers of two. So called dyadic scales and positions. The mother wavelet is rescaled or dilated.[4][5][6]

1.2 Signal Decomposition:

Starting with a discrete input signal vector s, the first stage of the DWT algorithm decomposes the signal into two sets of coefficients. These are the approximation coefficients cA1 (low frequency information) and the detail coefficients cD1 (high frequency information). The coefficient vectors are obtained by convolving s with the low-pass filter Lo_D for approximation and with the high-pass filter Hi_D for details. This filtering operation is then followed by dyadic decimation or down sampling by a factor of 2. Mathematically the two-channel filtering of the discrete signal s is represented by the expressions[7][8]

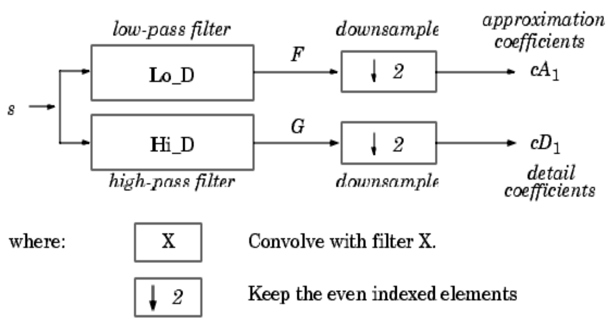


Figure 1. Filtering operation of the DWT

These equations implement a convolution plus down sampling by a factor 2 and give the forward fast wavelet transform. If the length of each filter is equal to 2N and the length of the original signal s is equal to n, then the corresponding lengths of the coefficients cA1 and cD1 are given by the formula:

The decomposition process can be iterated, with successive approximations being decomposed in turn, so that one signal is broken down into many lower resolution components. This is called the wavelet decomposition tree.[7][8]

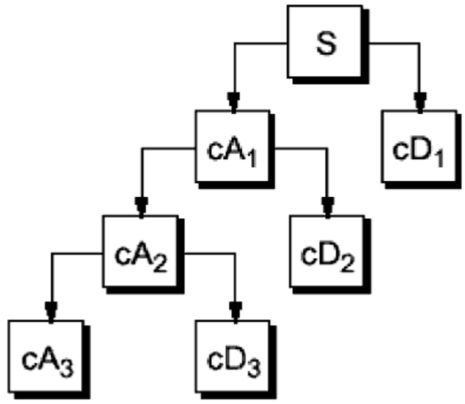


Figure 2. Decomposition of DWT coefficients

Since the analysis process is iterative, in theory it can be continued indefinitely. In reality, the decomposition can only proceed until the vector consists of a single sample. Normally, however there is little or no advantage gained in decomposing a signal beyond a certain level. The selection of the optimal decomposition level in the hierarchy depends on the nature of the signal being analyzed or some other suitable criterion, such as the low-pass filter cut-off.

The original signal can be reconstructed or synthesized using the inverse discrete wavelet transform (IDWT). The synthesis starts with the approximation and detail coefficients cA_j and cD_j, and then reconstructs cA_{j-1} by up sampling and filtering with the reconstruction filters.

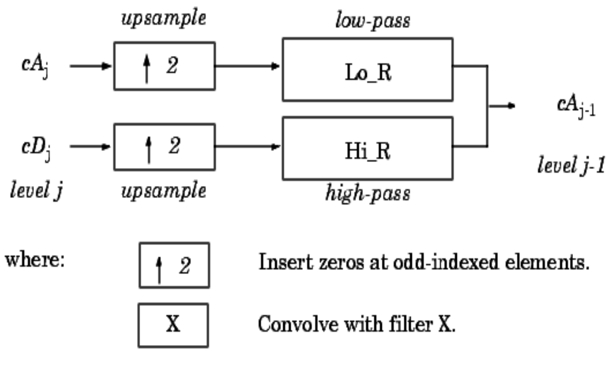


Figure 3. Wavelets reconstruction

The reconstruction filters are designed in such a way to cancel out the effects of aliasing introduced in the wavelet decomposition phase. The reconstruction filters (Lo_R and Hi_R) together with the low and high pass decomposition filters. For a multilevel analysis, the reconstruction process can itself be iterated producing successive approximations at finer resolutions and finally synthesizing the original signal.

1.3 Signal Denoising:

It has been seen that wavelets can remove noise more effectively than the traditionally used methods[4,6]. Use of wavelet transforms to denoise data is accomplished by applying a wavelet transformation to the noisy data, thresholding the resulting coefficients which are below some value in magnitude, and then inverse transforming to obtain a smoother version of the original data.

In this work the concept of Additive White Gaussian Noise (AWGN) is used. This simply means a noise, which has a Gaussian probability density function and white power spectral density function (noise distributed over the entire frequency spectrum) and is linearly added to whatever signal being analyzed. In the simplest model we suppose that,[10]

$$s\,(n) = f\,(n) + \sigma\, e\,(n) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3$$

Where time n is equally spaced. e(n) is a Gaussian white noise N (0,1) and the noise level σ .The de-noising objective is to suppress the noise part of the signal s and to recover f. The method is efficient for families of functions f that have only a few nonzero wavelet coefficients.

1.4 Denoising procedure:

The general de-noising procedure involves three steps. The basic version of the procedure follows the steps described below.

1. Decompose - Choose a wavelet, choose a level N. Compute the wavelet decomposition of the signal s at level N.
2. Thresholding: Threshold detail coefficients - For each level from 1 to N, select a threshold and apply soft or hard thresholding method to the detail coefficients.
3. Reconstruct - Compute wavelet reconstruction using the original approximation coefficients of level N and the modified detail coefficients of levels from 1 to N.

1.5 Soft and Hard Thresholding:

Hard Thresholding is the simplest method. Soft Thresholding has nice mathematical properties and the corresponding theoretical results are available.[9][10][12].

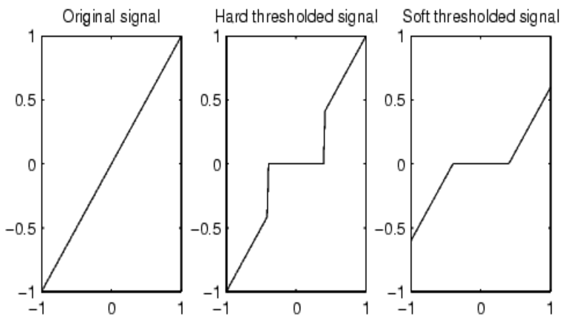


Figure 4. Signal, hard thresholding and soft thresholding

Let t denote the threshold.
The hard threshold signal x is x if |x| > t, is 0 if |x| t.
The soft threshold signal x is sign(x)(|x| - t) if |x| > t is 0 if |x| t.

Hard thresholding can be described as the usual process of setting to zero the elements whose absolute values are lower than the threshold. Soft thresholding is an extension of hard thresholding, first setting to zero the elements whose absolute values are lower than the threshold, and then shrinking the nonzero coefficients towards 0. The hard procedure creates discontinuities at x = ±t, while the soft procedure does not.

1.6 Choice of wavelet:

Choosing a wavelet that has compact support in both time and frequency in addition to significant number of vanishing moments is essential for an algorithm. Several criteria can be used in selecting an optimal wavelet function. The objective is to minimize reconstructed error variance and maximize signal to noise ratio (SNR).Optimum wavelets can be selected based on the energy conservation properties in the approximation part of the coefficients. Wavelets with more vanishing moments should be selected as it provides better reconstruction quality and introduce less distortion into processed speech and concentrate more signal energy in few coefficients. Computational complexity of DWT increases with the number of vanishing moments and hence for real time applications it cannot be suggested with high number of vanishing moments.

2 RESULTS AND IMPLEMENTATION OF WAVELET:

The speech signals used for the work are pronounced by male and female speakers, recorded using sound recorder facility with external microphone using mono channel. Samples used are 11000. Signal is sampled at sampling frequency Fs=8000 Hz, encoded using 16 bits, and degraded by additive Gaussian white noise. Signal is corrupted by 5db, 10 db and 15 db additive noises. Thus we have the noisy signal in required SNRs. In case of additive background noise the assumptions made for developing enhancement methods are (i) speech and noise signals are uncorrelated at least over a short-time basis, (ii) noise is either stationary or slowly varying as compared to speech, and (iii) noise can be represented as zero mean random process. The degradation level of additive background noise is normally specified by the measure called Signal to Noise Ratio (SNR) and is defined as the ratio of signal energy to noise energy. For evaluating performance of the method both objective and subjective tests are conducted. In objective test, SNR of signal after denoising is computed. Other two parameters used for comparing results are time required for reconstruction of signal and mean square error between clean signal and denoised signal. Haar and Daubechies wavelets are implemented on noisy signal and effort is made to remove additive white Gaussian noise from noisy signal. Signal is decomposed to level 4 and level 5.

Let s (n) is the clean speech, y (n) the noisy, sˆ (n) the enhanced signal and w (n) the noise then we have:

$$y\,(n) = s\,(n) + w\,(n) \text{-------------------------------} 5$$

SNR of denoised signal can be calculated as

$$SNR_{out} = 10 \log_{10} \frac{\sum S^2(n)}{\sum (S(n) - S^{\wedge}(n))^2} \text{--------------------------} 6$$

Minimizing mean square error (MSE) between the processed speech and the clean speech is a commonly used technique in the filtering algorithms. MSE is a valid distance measure between two speeches and it is computed directly as,

$$MSE = \frac{1}{N} \sum_{n=0}^{N-1} \left(\hat{s}(n) - s(n)\right)^2 \text{-------------------------------} 7$$

Speech signal pronounced by male speaker is as shown in figure and respective spectrogram of '.wav' file is as shown. Results are shown for level 4 decomposition of noisy signal with Haar wavelet implementation.

3 RESULTS AND DISCUSSION

From Table 1 it is observed that time elapsed and MSE are same for all Harr and Various Db Wavelets mentioned wavelets. Maximum SNR can be seen for Db6 and Db20. SNR using DB8 and Db18 are also approximately give same results for AGN noise using hard thresholding. In case of soft thresholding Db18 gives highest SNR as compared to others.

Haar wavelet gives worst results. If results of random noise using hard thresholding are compared then MSE is maximum for Db6, Db10 and Db20 for hard thresholding and minimum for Db2. Maximum SNR is at DB18. Db6 also performs better. For soft thresholding Db6 and Db20 gives good results, Haar is giving least SNR.

From Table 2 it is clear that in hard thresholding for AWGN, Db14 and Db16 are approximately gives same performance. Db16 requires little more time in reconstructing signal. Db20 takes maximum time to reconstruct among all wavelets and SNR is also very less as compared to others. Though Db20 takes long time for reconstruction it has maximum SNR than others for soft thresholding. For same SNR as Db20, Db18 is better as it takes less time than Db20.

For random noise of 10 dB, db20 is best but takes more time. Mean square error is same as that for Db10, 14, 16 and haar. When soft Thresholding is used Db6 and Db20 performs better and Db6 is preferable as it needs less time than Db20.

Table 3 represents results for 15 dB noise added to signal b1.wav and decomposed up to 5 levels. Haar wavelet needs maximum time to reconstruct signal but also has maximum SNR. Db20 gives same SNR as compared to Haar and also needs less time. For soft thresholding in AWGN db 20 gives maximum SNR and needs same time as Db10, 12, 14, 16, 18 which have comparatively less SNR. Db2 works better than others for random noise with hard thresholding. In soft thresholding results, it is observed that db2 and db6 are better. Out of these two wavelets db2 needs less reconstruction time.

As shown in Table no. 4 for AWGN, for hard thresholding Db8 and Db18 have approximately same SNR. Db18 requires little more time. From results it is clear that Haar needs least time and Db20 needs maximum time. Db18 also works well for soft thresholding. In case of random noise hard thresholding db20 is the best wavelet as it requires medium time to reconstruct and SNR value is also maximum. The Soft thresholding technique is best suited with Db16.

From Table 5 it is clear that Db6, Db8 and Db18 give same SNR results for AWGN hard Thresholding where Db 18 needs more time to reconstruct than other two. In db16 for soft thresholding has maximum SNR than Db8. Random noise results clearly show that Db18 gives good performance as compared to Db10 having same SNR value but takes more time for reconstruction. When soft threshold is used reconstructed signal has maximum SNR using Db12 with approximately zero MSE and needs same time as other wavelets having comparatively less SNR values

From results shown in Table 6 hard thresholding results for AWGN of 15 dB it seems that almost all wavelets have same or very close SNR values with approximately 0 MSE but take more or less time for reconstructing signal. Here Db8 takes maximum time than other wavelets. For soft thresholding method Db16 gives highest SNR with 0 MSE but takes more time for signal reconstruction.

Here Haar wavelet has least SNR with more reconstruction time for hard threshold and performs poor for both type of thresholding. When soft thresholding is used Db18 with 0 MSE but somewhat more time while reconstructing has maximum SNR than others.

Further discussion is about performance of Haar and different Daubechies wavelets on two signals b1.wav and j4.wav for level4 decomposition. For level 4 overall SNR values seems to be more than that of for level5. As signal energy is distributed in less number of levels

From results in Table 7, for level 4 decomposition of signal in AWGN and hard Thresholding, Db8 and Db16 both have good performance. In this case Db16 takes somewhat more time. Haar wavelet though needs less time for reconstructing as per results has lowest SNR. When soft Thresholding is used Db18 gives maximum SNR. In this case also it was found that Haar wavelet is giving poor performance. In case of random noise Db18 with 0.002 MSE has highest SNR. For soft thresholding Db 20 with 0.003 MSE and less time among all wavelets gives maximum SNR.

When hard Thresholding is used and signal is corrupted with AWGN, Db2 wavelet gives same MSE as other wavelets have good SNR. Using same method for random noise Db6 is found to have good results and less MSE than for AWGN.

For soft thresholding Db18 with comparatively more time gives good SNR. For random noise with soft thresholding it is observed that it takes same time to reconstruct with same MSE. Db16 performs better than Db6.

From results in Table 9 Haar takes less time to reconstruct and give same SNR as Db8, Db10, Db12. Db18 can be said to be good for AWGN when hard thresholding is used. Db18 has highest SNR among all wavelets. Db20 wavelet with hard thresholding and Db6 with soft thresholding performs well against random Noise.

Table 1: Results for B1.Wav input file for application of Haar and Db wavelet

Level 5 decomposition with addition of 5dB noise												
Wavelet	Additive White Gaussian noise						Random noise					
	Hard thresholding			Soft thresholding			Hard thresholding			Soft thresholding		
	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR
Haar	0.04	0.004	19.96	0.04	0.004	21.161	0.04	0.002	24.25	0.04	0.002	22.70
Db2	0.04	0.004	19.91	0.04	0.004	21.666	0.04	0.001	24.82	0.04	0.003	23.47
Db4	0.04	0.004	19.92	0.04	0.004	22.023	0.04	0.002	25.81	0.04	0.001	23.41
Db6	0.04	0.004	20.00	0.04	0.004	22.106	0.04	0.003	25.90	0.04	0.002	23.61
Db8	0.04	0.004	19.99	0.04	0.004	21.955	0.04	0.002	25.52	0.04	0.002	23.50
Db10	0.06	0.004	19.91	0.04	0.004	22.079	0.04	0.003	25.67	0.04	0.002	23.37
Db12	0.04	0.004	19.92	0.04	0.004	22.052	0.04	0.002	25.08	0.04	0.003	22.99
Db14	0.04	0.004	19.95	0.04	0.004	22.219	0.04	0.002	25.05	0.04	0.002	22.84
Db16	0.04	0.004	19.89	0.04	0.004	22.303	0.04	0.002	25.15	0.04	0.002	23.51
Db18	0.04	0.004	19.98	0.04	0.004	22.343	0.04	0.002	26.43	0.04	0.002	23.46
Db20	0.04	0.004	20.00	0.04	0.004	22.037	0.04	0.003	25.70	0.04	0.002	23.53

Table 2: Results for B1.Wav input file for application of Haar and Db wavelet

Level 5 decomposition with addition of 10 dB noise												
Wavelet	Additive White Gaussian noise						Random noise					
	Hard thresholding			Soft thresholding			Hard thresholding			Soft thresholding		
	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR
Haar	0.037	0.004	19.97	0.043	0.004	21.13	0.036	0.003	22.33	0.036	0.003	22.02
Db2	0.037	0.004	20.01	0.038	0.004	21.72	0.038	0.002	23.67	0.038	0.003	22.33
Db4	0.037	0.004	19.97	0.038	0.004	22.01	0.039	0.002	23.04	0.037	0.002	22.14
Db6	0.038	0.004	19.95	0.039	0.004	22.11	0.039	0.002	23.16	0.039	0.001	22.48
Db8	0.038	0.004	19.96	0.040	0.004	22.21	0.039	0.002	22.99	0.039	0.002	22.02
Db10	0.039	0.004	19.92	0.039	0.004	22.18	0.039	0.003	23.53	0.039	0.002	21.96
Db12	0.039	0.004	20.00	0.039	0.004	22.21	0.039	0.002	23.60	0.039	0.003	22.09
Db14	0.038	0.004	20.03	0.039	0.004	22.29	0.039	0.003	23.25	0.039	0.003	21.93
Db16	0.039	0.004	20.02	0.039	0.004	22.30	0.039	0.003	23.46	0.039	0.003	21.95
Db18	0.040	0.004	20.01	0.040	0.004	22.34	0.039	0.002	22.98	0.039	0.002	22.05
Db20	0.044	0.004	19.94	0.140	0.004	22.34	0.040	0.003	24.13	0.039	0.003	22.41

Table 3: Results for B1.Wav input file for application of Haar and Db wavelet

Level 5 decomposition with addition of 15 dB noise												
Wavelet	Additive White Gaussian noise						Random noise					
	Hard thresholding			Soft thresholding			Hard thresholding			Soft thresholding		
	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR
Haar	0.048	0.004	20.01	0.037	0.004	21.29	0.037	0.002	21.30	0.036	0.002	21.37
Db2	0.037	0.004	20.00	0.038	0.004	21.73	0.038	0.003	21.82	0.037	0.002	21.48
Db4	0.037	0.004	20.00	0.038	0.004	22.16	0.038	0.003	22.03	0.038	0.003	21.24
Db6	0.038	0.004	19.98	0.041	0.004	22.28	0.039	0.003	21.87	0.039	0.003	21.47
Db8	0.038	0.004	20.00	0.038	0.004	22.18	0.041	0.002	21.39	0.039	0.003	20.94
Db10	0.038	0.004	19.98	0.039	0.004	22.22	0.039	0.000	21.63	0.039	0.002	21.25
Db12	0.038	0.004	19.98	0.039	0.004	22.19	0.039	0.002	21.48	0.039	0.002	21.12
Db14	0.039	0.004	19.99	0.039	0.004	22.35	0.042	0.003	21.61	0.039	0.002	21.18
Db16	0.039	0.004	20.00	0.039	0.004	22.36	0.040	0.003	21.87	0.039	0.002	21.19
Db18	0.039	0.004	19.99	0.039	0.004	22.49	0.039	0.003	21.49	0.040	0.002	21.15
Db20	0.040	0.004	20.01	0.039	0.004	22.31	0.040	0.003	21.61	0.040	0.002	21.11

Table 4: Results for j4.Wav input file for application of Haar and Db wavelet

Level 5 decomposition with addition of 5 dB noise												
Wavelet	Additive White Gaussian noise						Random noise					
	Hard thresholding			Soft thresholding			Hard thresholding			Soft thresholding		
	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR
Haar	0.036	0.000	19.92	0.038	0.001	20.79	0.041	0.000	21.23	0.036	0.001	20.03
Db2	0.038	0.000	19.92	0.037	0.000	21.11	0.038	0.001	21.64	0.038	0.000	20.53
Db4	0.040	0.000	19.94	0.038	0.000	21.49	0.038	0.001	21.94	0.038	0.001	20.76
Db6	0.039	0.000	19.92	0.039	0.000	21.64	0.038	0.001	22.09	0.038	0.001	20.88
Db8	0.039	0.000	19.96	0.039	0.000	21.77	0.039	0.001	22.27	0.038	0.001	20.83
Db10	0.039	0.000	19.89	0.039	0.000	21.64	0.040	0.000	22.23	0.039	0.001	20.87
Db12	0.039	0.000	19.94	0.040	0.001	21.53	0.039	0.001	22.39	0.038	0.001	20.80
Db14	0.039	0.000	19.90	0.039	0.000	21.81	0.041	0.001	22.48	0.042	0.001	20.80
Db16	0.039	0.000	19.95	0.041	0.000	21.64	0.041	0.000	22.18	0.039	0.002	20.91
Db18	0.040	0.000	19.97	0.039	0.000	21.83	0.039	0.001	21.95	0.039	0.001	20.87
Db20	0.045	0.000	19.90	0.040	0.000	21.78	0.039	0.001	22.60	0.039	0.000	20.76

Table 5: Results for j4.Wav input file for application of Haar and Db wavelet

Level 5 decomposition with addition of 10 dB noise												
Wavelet	Additive White Gaussian noise						Random noise					
	Hard thresholding			Soft thresholding			Hard thresholding			Soft thresholding		
	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR
Haar	0.036	0.000	19.97	0.036	0.000	21.00	0.037	0.001	20.87	0.036	0.001	20.09
Db2	0.037	0.000	19.97	0.038	0.000	21.24	0.038	0.000	21.10	0.040	0.001	20.00
Db4	0.039	0.000	19.98	0.038	0.000	21.60	0.037	0.000	21.22	0.038	0.002	20.39
Db6	0.039	0.000	20.00	0.038	0.000	21.73	0.039	0.000	21.16	0.039	0.000	20.19
Db8	0.039	0.000	20.00	0.039	0.000	21.84	0.038	0.000	21.48	0.039	0.001	20.21
Db10	0.038	0.000	19.97	0.039	0.000	21.78	0.045	0.001	21.73	0.041	0.001	20.28
Db12	0.043	0.000	19.97	0.039	0.000	21.92	0.038	0.000	21.49	0.039	0.000	20.56
Db14	0.042	0.000	19.99	0.039	0.000	21.77	0.039	0.001	21.23	0.039	0.001	20.47
Db16	0.038	0.000	19.97	0.040	0.000	21.91	0.039	0.001	21.57	0.041	0.001	20.29
Db18	0.041	0.000	20.00	0.039	0.000	21.82	0.039	0.001	21.74	0.039	0.001	20.39
Db20	0.040	0.000	19.97	0.040	0.000	21.75	0.040	0.001	21.31	0.039	0.001	20.49

Table 6: Results for j4.Wav input file for application of Haar and Db wavelet

Level 5 decomposition with addition of 15 dB noise												
Wavelet	Additive White Gaussian noise						Random noise					
	Hard thresholding			Soft thresholding			Hard thresholding			Soft thresholding		
	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR
Haar	0.036	0.000	20.00	0.036	0.000	20.97	0.072	0.000	19.85	0.036	0.001	19.57
Db2	0.038	0.000	20.00	0.038	0.000	21.32	0.037	0.000	20.14	0.038	0.000	19.74
Db4	0.038	0.000	19.99	0.039	0.000	21.61	0.038	0.001	20.21	0.038	0.000	19.82
Db6	0.039	0.000	20.00	0.039	0.000	21.68	0.038	0.002	20.12	0.038	0.001	19.65
Db8	0.045	0.000	19.99	0.039	0.000	21.85	0.042	0.001	20.26	0.039	0.000	19.81
Db10	0.040	0.000	20.00	0.039	0.000	21.85	0.039	0.002	20.28	0.039	0.000	19.68
Db12	0.040	0.000	19.99	0.039	0.000	21.82	0.039	0.001	20.66	0.038	0.001	19.85
Db14	0.039	0.000	19.99	0.040	0.000	21.87	0.039	0.001	20.79	0.039	0.000	19.83
Db16	0.039	0.000	20.00	0.041	0.000	21.90	0.041	0.001	20.61	0.039	0.002	19.79
Db18	0.042	0.000	19.99	0.039	0.000	21.79	0.044	0.001	20.09	0.040	0.000	19.96
Db20	0.042	0.000	20.00	0.039	0.000	21.83	0.041	0.001	20.58	0.039	0.001	19.81

Table 7: Results for B1.Wav input file for application of Haar and Db wavelet.

Level 4 decomposition with addition of 5 dB noise												
Wavelet	Additive White Gaussian noise						Random noise					
	Hard thresholding			Soft thresholding			Hard thresholding			Soft thresholding		
	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR
Haar	0.038	0.004	19.84	0.036	0.004	21.07	0.032	0.003	24.84	0.051	0.002	23.87
Db2	0.039	0.004	19.92	0.038	0.004	21.49	0.033	0.002	25.96	0.039	0.002	24.73
Db4	0.037	0.004	19.88	0.039	0.004	21.84	0.032	0.002	26.72	0.040	0.002	25.22
Db6	0.037	0.004	19.97	0.038	0.004	22.19	0.034	0.002	26.61	0.038	0.002	25.22
Db8	0.037	0.004	19.93	0.038	0.004	21.95	0.037	0.003	26.66	0.039	0.001	25.01
Db10	0.037	0.004	19.87	0.056	0.004	22.09	0.037	0.002	26.48	0.038	0.002	25.34
Db12	0.038	0.004	19.87	0.039	0.004	21.99	0.037	0.003	26.91	0.038	0.002	25.07
Db14	0.037	0.004	19.91	0.039	0.004	21.97	0.031	0.002	26.98	0.039	0.003	25.20
Db16	0.038	0.004	19.93	0.040	0.004	21.94	0.034	0.002	26.92	0.038	0.002	25.20
Db18	0.041	0.004	19.91	0.040	0.004	22.31	0.035	0.002	27.04	0.041	0.002	25.12
Db20	0.040	0.004	19.91	0.039	0.004	22.16	0.035	0.002	26.88	0.038	0.003	25.46

Table 8: Results for B1.Wav input file for application of Haar and Db wavelet.

Level 4 decomposition with addition of 10 dB noise												
Wavelet	Additive White Gaussian noise						Random noise					
	Hard thresholding			Soft thresholding			Hard thresholding			Soft thresholding		
	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR
Haar	0.035	0.004	19.98	0.036	0.004	21.24	0.031	0.003	23.04	0.036	0.001	23.70
Db2	0.036	0.004	20.02	0.038	0.004	21.76	0.033	0.002	23.64	0.037	0.003	25.00
Db4	0.036	0.004	19.93	0.037	0.004	22.05	0.033	0.003	24.48	0.037	0.002	24.93
Db6	0.037	0.004	19.95	0.040	0.004	22.19	0.034	0.003	24.72	0.038	0.002	25.52
Db8	0.037	0.004	19.94	0.071	0.004	22.20	0.034	0.003	24.29	0.039	0.002	25.23
Db10	0.038	0.004	19.97	0.038	0.004	22.19	0.034	0.002	24.24	0.038	0.002	25.34
Db12	0.037	0.004	20.00	0.039	0.004	22.12	0.036	0.002	23.72	0.038	0.002	25.04
Db14	0.038	0.004	20.01	0.041	0.004	22.30	0.034	0.003	24.18	0.040	0.002	25.04
Db16	0.038	0.004	19.96	0.041	0.004	22.33	0.034	0.002	24.55	0.038	0.002	25.54
Db18	0.044	0.004	19.98	0.043	0.004	22.46	0.035	0.002	24.27	0.039	0.002	25.25
Db20	0.040	0.004	19.93	0.039	0.004	22.26	0.036	0.003	24.32	0.039	0.002	25.34

Table 9: Results forB1.Wav input file for application of Haar and Db wavelet

Level 4 decomposition with addition of 15 dB noise												
Wavelet	Additive White Gaussian noise						Random noise					
	Hard thresholding			Soft thresholding			Hard thresholding			Soft thresholding		
	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR
Haar	0.035	0.004	20.00	0.036	0.004	21.23	0.031	0.002	21.41	0.036	0.002	24.02
Db2	0.036	0.004	19.98	0.037	0.004	21.74	0.034	0.003	22.00	0.037	0.002	24.71
Db4	0.037	0.004	20.01	0.044	0.004	22.17	0.034	0.003	22.27	0.039	0.002	25.21
Db6	0.037	0.004	19.98	0.038	0.004	22.24	0.038	0.002	21.99	0.037	0.002	25.46
Db8	0.039	0.004	20.00	0.038	0.004	22.22	0.034	0.002	21.97	0.038	0.001	25.23
Db10	0.038	0.004	20.00	0.040	0.004	22.23	0.034	0.003	22.34	0.038	0.003	25.05
Db12	0.038	0.004	20.00	0.038	0.004	22.19	0.036	0.003	21.94	0.038	0.002	25.15
Db14	0.040	0.004	19.99	0.039	0.004	22.31	0.035	0.003	22.20	0.038	0.001	25.27
Db16	0.038	0.004	19.99	0.040	0.004	22.44	0.036	0.002	22.30	0.042	0.003	25.08
Db18	0.039	0.004	20.00	0.039	0.004	22.48	0.035	0.002	22.19	0.038	0.002	25.37
Db20	0.039	0.004	19.97	0.039	0.004	22.30	0.036	0.002	22.36	0.040	0.003	24.97

Table 10: Results forJ4.Wav input file for application of Haar and Db wavelet.

Level 4 decomposition with addition of 5 dB noise												
Wavelet	Additive White Gaussian noise						Random noise					
	Hard thresholding			Soft thresholding			Hard thresholding			Soft thresholding		
	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR
Haar	0.035	0.001	19.91	0.036	0.001	20.76	0.031	0.000	22.05	0.037	0.001	21.54
Db2	0.039	0.000	19.94	0.037	0.001	21.31	0.033	0.000	22.65	0.037	0.000	21.92
Db4	0.037	0.001	19.94	0.037	0.001	21.31	0.037	0.000	22.77	0.037	0.001	22.16
Db6	0.037	0.001	19.90	0.038	0.001	21.68	0.033	0.001	23.10	0.045	0.000	22.28
Db8	0.037	0.001	19.87	0.038	0.001	21.78	0.032	0.001	23.63	0.040	0.001	22.57
Db10	0.038	0.001	19.92	0.039	0.001	21.54	0.034	0.001	23.05	0.041	0.000	22.34
Db12	0.037	0.001	19.87	0.038	0.000	21.71	0.034	0.000	23.31	0.039	0.000	22.55
Db14	0.040	0.001	19.95	0.038	0.001	21.68	0.055	0.000	23.18	0.038	0.001	22.71
Db16	0.042	0.001	19.95	0.044	0.001	21.69	0.038	0.001	23.56	0.040	0.000	22.75
Db18	0.039	0.001	19.90	0.039	0.001	21.78	0.052	0.001	23.05	0.040	0.001	22.62
Db20	0.041	0.001	19.94	0.039	0.000	21.54	0.039	0.000	23.57	0.039	0.001	22.56

Table 11: Results forJ4.Wav input file for application of Haar and Db wavelet.

Level 4 decomposition with addition of 10 dB noise												
Wavelet	Additive White Gaussian noise						Random noise					
	Hard thresholding			Soft thresholding			Hard thresholding			Soft thresholding		
	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR
Haar	0.035	0.001	19.96	0.036	0.001	20.90	0.033	0.001	21.07	0.035	0.000	21.534
Db2	0.039	0.002	19.97	0.037	0.001	21.29	0.032	0.000	21.56	0.040	0.001	21.978
Db4	0.037	0.001	19.97	0.037	0.001	21.49	0.034	0.000	21.97	0.038	0.000	22.285
Db6	0.038	0.001	19.98	0.038	0.001	21.70	0.035	0.001	22.21	0.040	0.001	22.405
Db8	0.038	0.001	19.97	0.038	0.001	21.86	0.038	0.002	22.07	0.039	0.000	22.515
Db10	0.038	0.001	19.98	0.038	0.001	21.72	0.034	0.000	22.12	0.038	0.001	22.404
Db12	0.038	0.001	19.99	0.039	0.001	21.70	0.035	0.001	22.05	0.039	0.000	22.531
Db14	0.038	0.001	20.00	0.039	0.001	21.83	0.038	0.000	22.19	0.040	0.001	22.626
Db16	0.040	0.001	19.99	0.039	0.001	21.85	0.038	0.002	22.55	0.039	0.001	22.641
Db18	0.040	0.000	19.99	0.039	0.001	21.89	0.039	0.001	22.36	0.039	0.001	22.600
Db20	0.039	0.001	19.99	0.039	0.001	21.83	0.040	0.001	22.39	0.039	0.002	22.643

Table 12: Results for J4.Wav input file for application of Haar and Db wavelet.

Level 4 decomposition with addition of 15 dB noise												
Wavelet	Additive White Gaussian noise						Random noise					
	Hard thresholding			Soft thresholding			Hard thresholding			Soft thresholding		
	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR	Time	MSE	SNR
Haar	0.035	0.001	20.00	0.036	0.001	20.97	0.032	0.000	20.22	0.035	0.001	21.59
Db2	0.039	0.001	19.99	0.038	0.001	21.29	0.033	0.001	20.56	0.037	0.000	21.95
Db4	0.037	0.001	19.98	0.037	0.001	21.60	0.033	0.002	20.67	0.045	0.000	22.21
Db6	0.039	0.001	19.99	0.040	0.001	21.73	0.034	0.002	20.98	0.038	0.001	22.34
Db8	0.037	0.001	19.99	0.038	0.001	21.80	0.034	0.001	21.06	0.041	0.001	22.61
Db10	0.038	0.001	19.99	0.039	0.001	21.74	0.034	0.001	20.77	0.045	0.001	22.67
Db12	0.038	0.001	19.99	0.039	0.001	21.89	0.035	0.000	20.86	0.041	0.000	22.51
Db14	0.038	0.001	19.99	0.041	0.001	21.87	0.038	0.001	21.20	0.039	0.000	22.49
Db16	0.040	0.001	20.01	0.039	0.001	21.87	0.039	0.001	21.17	0.039	0.000	22.60
Db18	0.040	0.001	19.99	0.041	0.001	21.82	0.038	0.001	20.67	0.040	0.001	22.58
Db20	0.039	0.001	20.00	0.039	0.001	21.79	0.038	0.000	21.00	0.039	0.001	22.65

For AWGN hard thresholding condition two wavelets namely Db14 and Db16 have maximum SNR. Both of these need more time as compared with others. In soft thresholding also Db8 and Db18 have same and higher SNR than others. Db8with very less reconstructing time gives highest SNR for hard and Db16 with 0 MSE and more time has higher SNR for soft in random noise. Db14 has maximum SNR and all other have similar SNR response with more or less time for reconstruction. Db16, Db18 need maximum reconstruction time and haar needs least time for hard thresholding of AWGN. Db12 to Db20 require same time to reconstruct signal out of which Db18 is better for soft thresholding. If hard thresholding is applied to signal corrupted by random noise Db16 performs better with MSE of 0.002. Db20 takes maximum and haar takes minimum time to reconstruct. For soft thresholding Db20 gives maximum SNR. Db16 also gives approximately close SNR with same reconstructing time and less MSE than DB20. From Table 12, for AWGN hard thresholding Db16 has highest SNR but needs long time to reconstruct signal. DB20 and haar have same SNR values where DB20 needs much more time than haar. Here haar is found to be performing well. For soft thresholding of AWGN Db16 gives maximum SNR. Db14, Db16 have same SNR which is approximately close to SNR of Db12. In case of hard thresholding for random noise Db14 has highest SNR value. Haar and Db12 have 0 MSE and all other wavelets have some error in reconstructed signal as compared to original signal. In soft threshold method, signal Db20 has better results than Db10 with maximum SNR but requiring more time than Db10.
Reconstructed signals which are denoised are as shown in following figures. Male and female speaker signals corrupted by white noise decomposed to level 4 and level 5 and reconstructed after thresholding the coefficients. Following figures indicate results for speech signal of male speaker with level 4 decomposition.

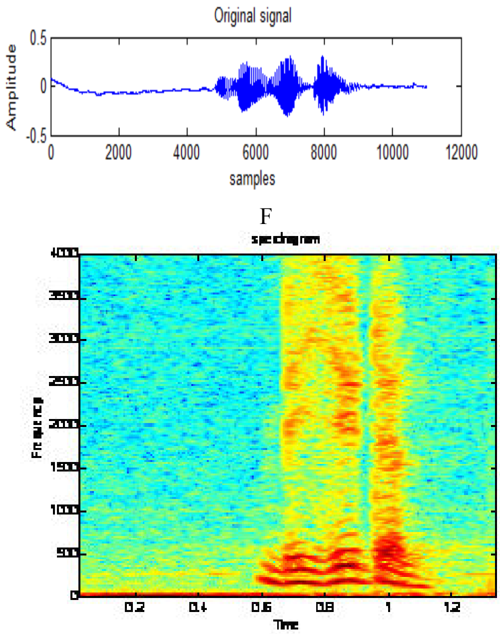


Figure 5: Original Speech Signal and Its Spectrogram

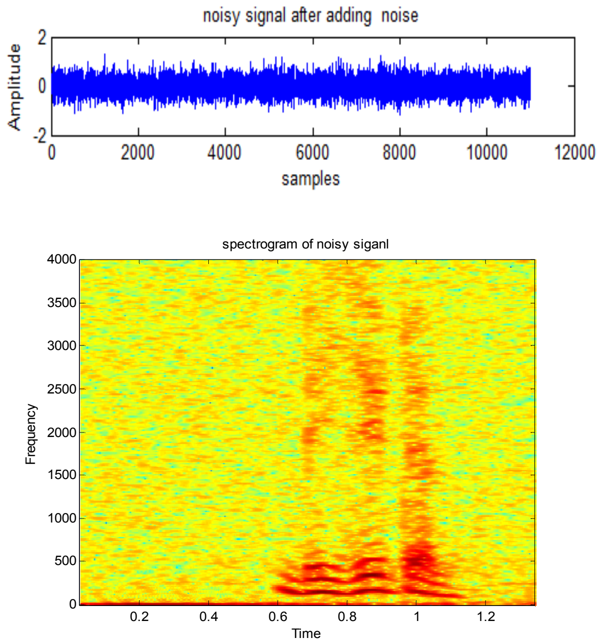


Figure 6 Noisy signal corrupted db 10 db AWGN and its Spectrogram

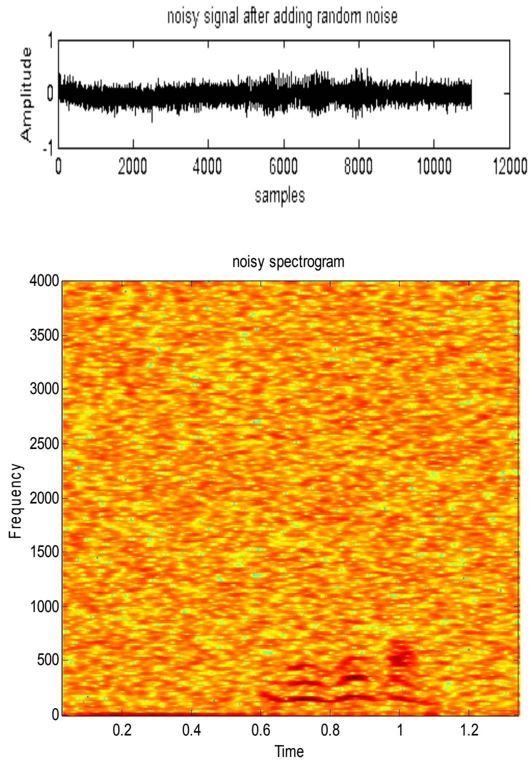
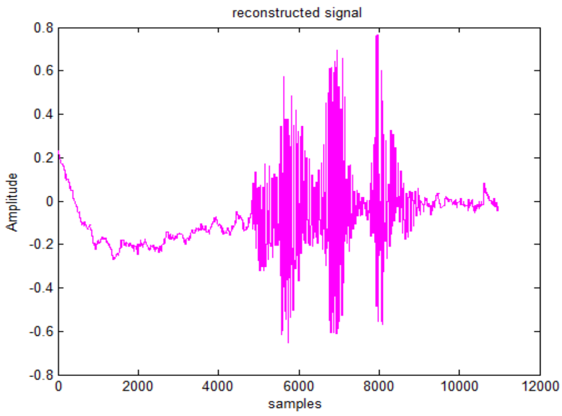


Figure 7: noisy signal corrupted by of 10 db random noise and Spectrogram of noisy signal.



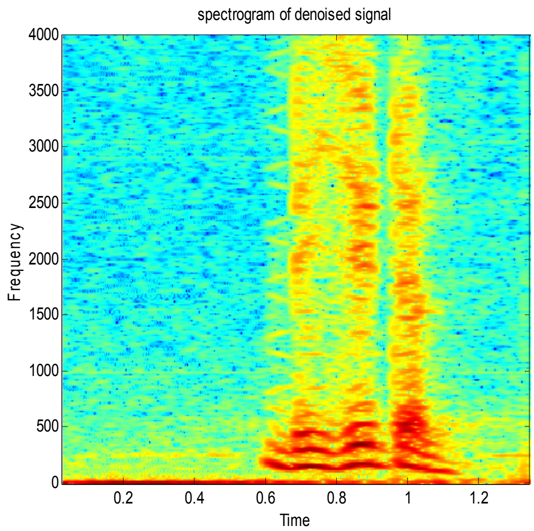
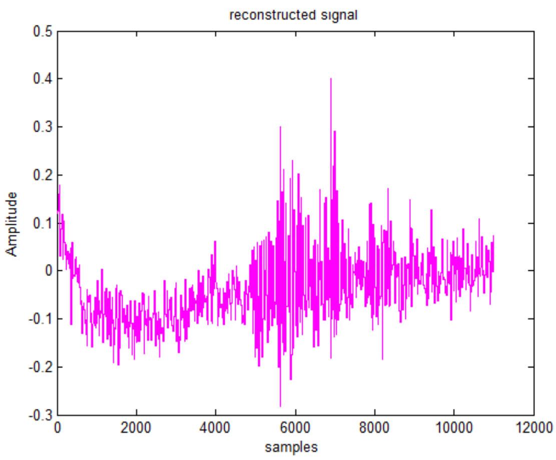


Figure 8: Reconstructed signal using Haar wavelet for 10 db AWGN.



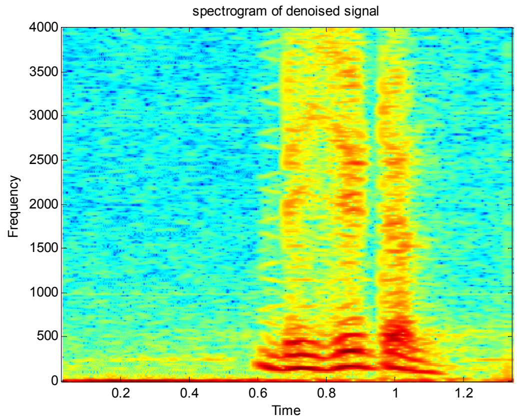
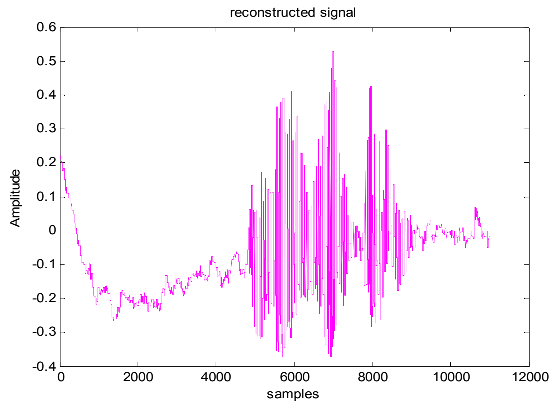


Figure 9: Reconstructed signal using Haar wavelet for random noise 10 db

From figure 8 it is observed that denoised signal is approximately same as of original signal with Haar wavelet for hard thresholding in additive white Gaussian noise. For random noise of 10 db for the same speech signal reconstructed signal has some noise which is tolerable as shown in figure 9. When same speech signal is denoised using soft thresholding, SNR of reconstructed signal improves. Reconstructed signal using soft thresholding is as shown in figure 10 and 11.



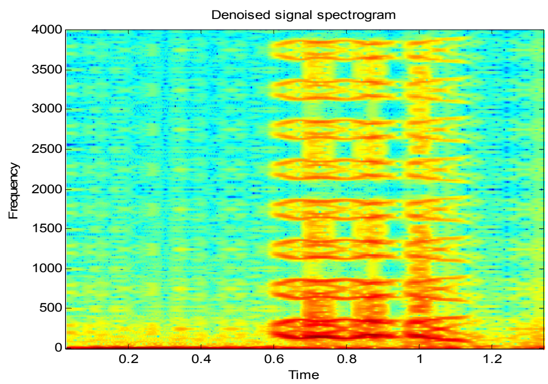


Figure 10: Reconstructed signal for AWGN of 10 db for soft thresholding

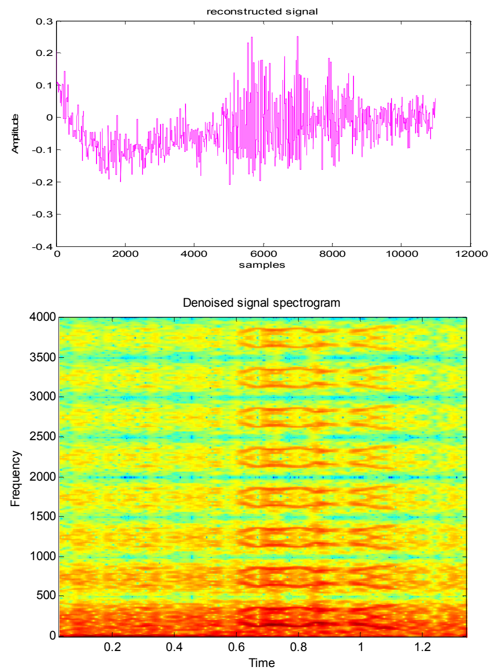


Figure 11: Reconstructed signal for Random noise of 10 db for soft thresholding

Following are the results of signal decomposed up to level5 with Haar wavelet implementation for hard as well as soft thresholding .

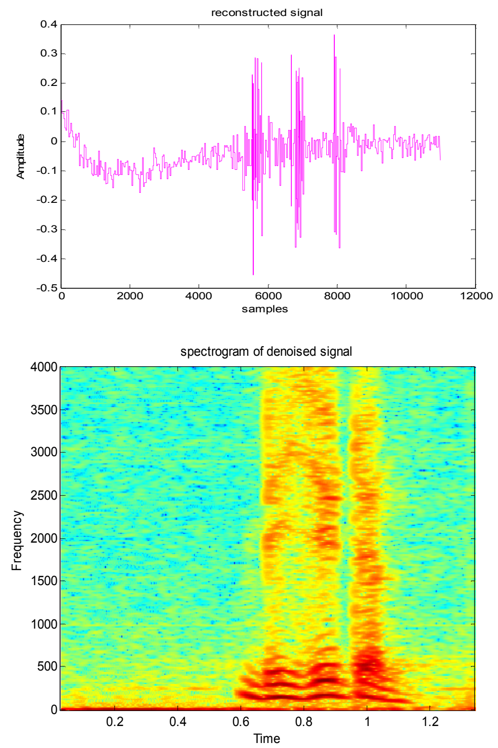


Figure 12: Reconstructed signal for random noise of 10 db using hard threshold

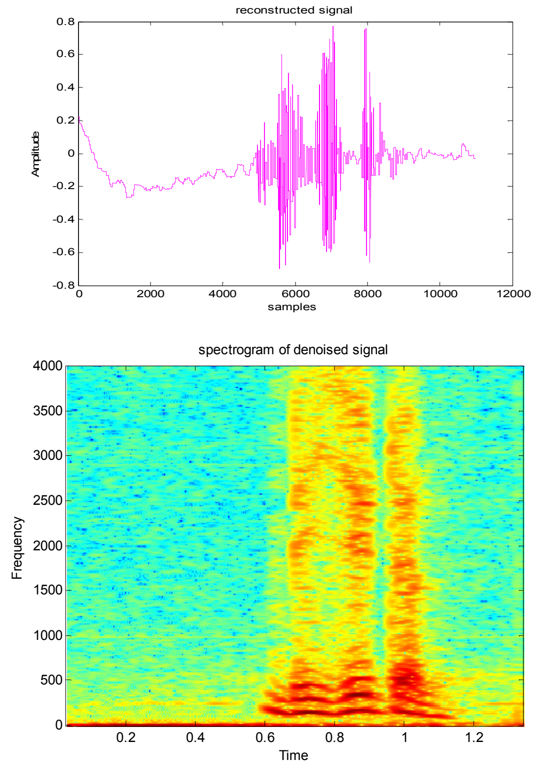


Figure 13: Reconstructed signal for AWGN 10 db

Following are the results of implementation of soft thresholding

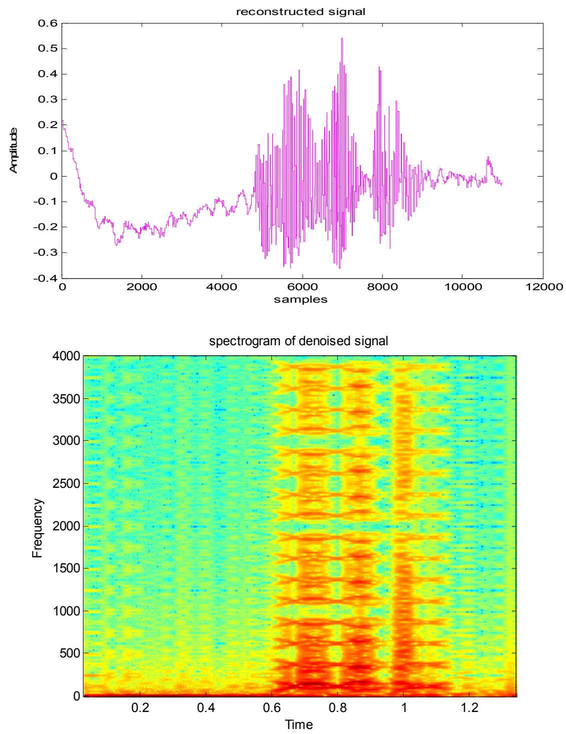


Figure 14: Reconstructed b1.wav for AWGN 10 db

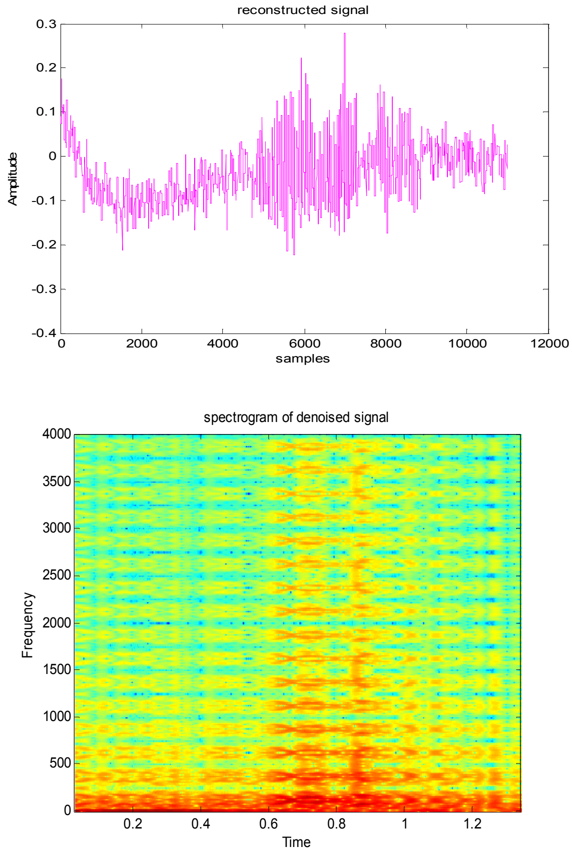
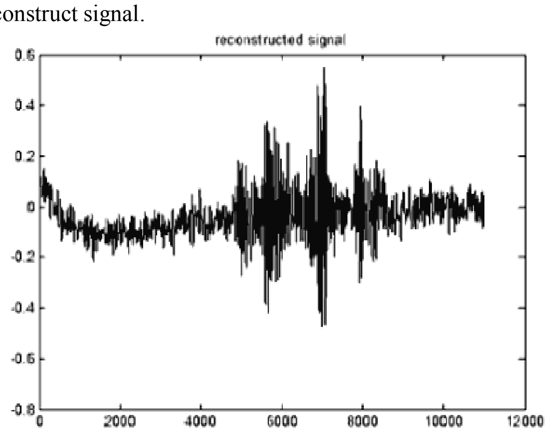


Figure 15: Reconstructed signal for Random 10 db

4. IMPLEMENTATION OF DAUBECHIES WAVELET

Following figures show signals reconstructed by hard thresholding and decomposed up to level 4 using db18 wavelet. Reconstructed signal seems to be close to the original signal. SNR is also around 24 dB. Signal was corrupted using random noise of 10 dB. It is observed that Daubechies wavelet needs more time to reconstruct signal.



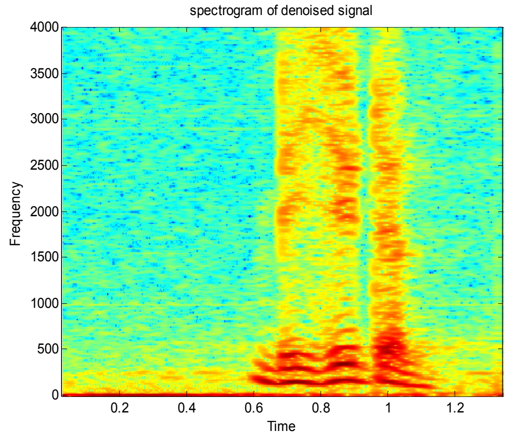


Figure 16: Reconstructed signal b1.wav using db18

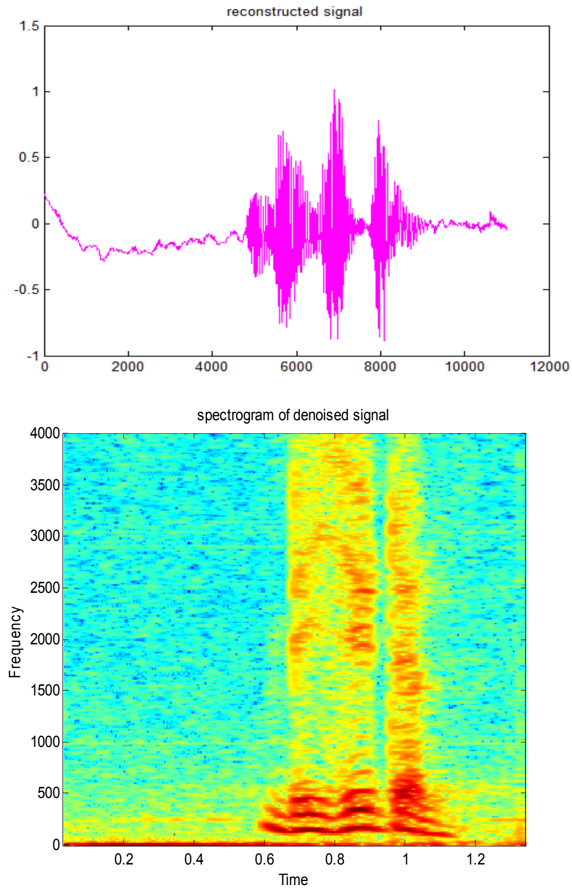


Figure 17: Reconstructed signal b1.wav using db18 for 10 dB AWGN

When the signal is denoised with soft thresholding, reconstructed signal is found to be similar to the original signal. Soft thresholding performs better than hard thresholding. Reconstructed signals are as shown in figure 18.

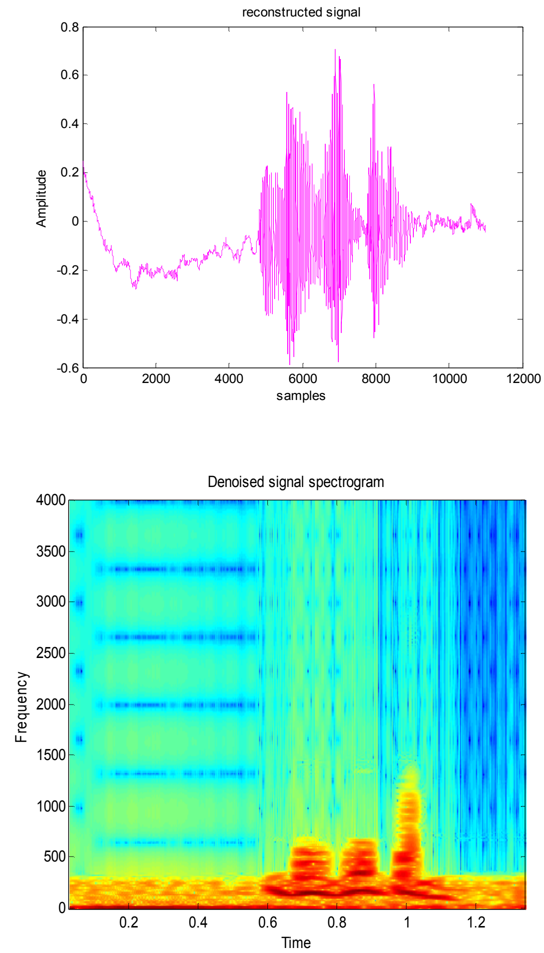
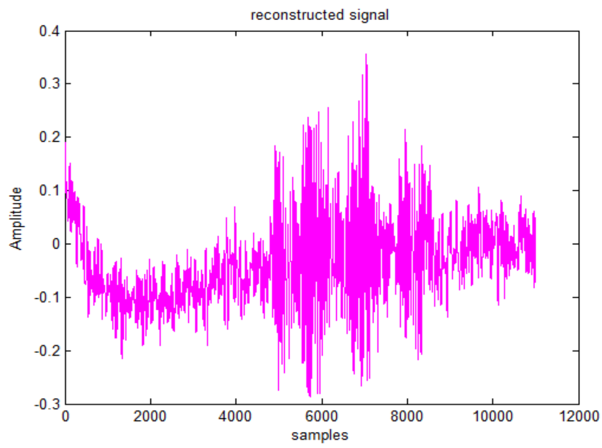


Figure 18: Reconstructed signal using db18 for AWGN

Figure 19 shows reconstructed signal which was corrupted with 10 db random noise. From figure it is clear that still noise is present in reconstructed signal and not totally removed, spectrogram also do not match to that of original and is quite is blurred & indicates presences of noise



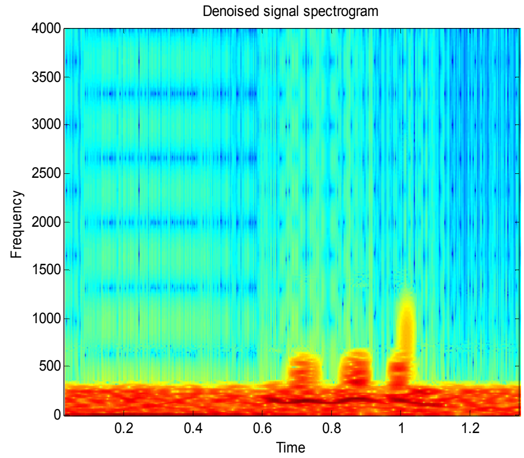


Figure 19: reconstructed b1.wav for random noise

5. PERFORMANCE EVALUATION USING SUBJECTIVE TEST

It is well known that SNR cannot faithfully indicate speech quality. For evaluation of performance of wavelets for speech enhancement one more criteria used is subjective test [5]. While conducting this test 15 listeners are selected in the age group of 18 to 22 years. Listeners are divided into group of five and they were asked to listen to the test material. Participants were not familiar with the test material. Listeners participated in two sessions. In the first session they were asked to carefully listen noisy signal and denoised signal. In second session they were asked to listen original signal and denoised signal. Throughout the subjective test input SNR was set to 10 dB.

For both AWGN and random noise, listeners preferred denoised signal with soft thresholding. For additive white Gaussian noise listener observed that original signal is superior than denoised signal and signal strength is also poor for later one. With implementation of wavelet transform denoised signal which is reconstructed from approximate and detail coefficients is found to be audible and approximately similar to that of original signal.

6. SUMMARY AND CONCLUSION

In the present work wavelet based speech Denoising algorithm is addressed. Wavelet Denoising is a non-parametric estimation method that has been proposed in recent years for speech enhancement applications. In this work both objective and subjective methods were used for evaluation of wavelets performance in speech Denoising confirms its superiority. Haar wavelet has comparatively good performance than some other wavelets from Db family. SNR for level 4 is better than for level5. If SNR obtained for speech signals in any noise with all type of wavelets is below 19 db then it is not properly denoised and still some noise components are remained. Therefore From results obtained it can be concluded that Haar wavelet is not suitable for speech Denoising application. As Haar is not smooth as compared to other wavelets it has limitations when applied to non stationary signal such as speech. Higher order Daubechies can be used and are found to be suitable for the work done. Also soft thresholding is better than hard thresholding. Depending upon application trade off is to be made between time required for reconstructing the signal by wavelet and signal to noise ratio. Use of a unique threshold for all wavelet bands is disadvantage for speech applications. The work carried out here can be extended to speech Denoising for sentences recorded in varies noisy environment also it can be extended for the real time signal Denoising or speech enhancement. In the work carried out noisy signal is directly decomposed without silence Detection therefore the work may be extended by separating existence of speech and absence of speech (Silence) and then computing thresholds separately. Wavelet packet transform can be implemented further to achieve good performance.

REFERENCES:

[1] S.Manikandan, "Speech enhancement based on wavelet denoising" Academic Open Internet Journal www.acadjournal.com, Volume 17, 2006.

[2] Hamid Sheikhzadeh, Hamid Reza Abutalbi, "An improved wavelet based speech enhancement system",Proceedings of 7th European Confrence on Speech Communication and technology, Alborg, Denmark, 3-7, Sept 2001, pp 1855-1857.

[3] Lallouani, M. Gabrea and C.S. Gargour, "Wavelet based Speech enhancement using different threshold based denoising algorithms", Canadian conference, Electrical and Computer Engg., May 2004.

[4] Amara Graps, "An introduction to Wavelets", IEEE Computation Science and Engineering, Volume 2, Issue 2, June 1995, page 50-61.

[5] Soon Ing Yann, "Transform based Speech Enhancement Techniques", PhD Thesis Nanyang Technological University 2003.

[6] Chris Perkins, Tobin Fricke, "Wavelets", University of California at Berkely, 1St December 2000, pp 1-18.

[7] Nikhil Rao, "Speech Compression using Wavelets", B.E Thesis, School of Information Technology and Electrical Engineering University of Queensland Oct 8th, 2001.

[8] Yan Long, Lin Gang and Guo Jun, " Selection Of The best Wavelet Base For Speech Signal", Proceedings of International Symposium on Intelligent Multimedia, Video and Speech Processing, October 20-22, 2004.

[9] Qin Linmei, Hu Guangrui and Li Chongni," A new speech enhancement method", Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and speech processing, May 2-4 2001, Hongkong.

[10] V. Balakrishnan, Nash Borges, Luke Parchment, " Wavelet Denoising and Speech Enhancement", Spring 2006.

[11] P Krishnamoorthy, Mahadeva Prasanna, "Processing Noise Speech for Enhancement", IETE Technical Review, Volume 24, No 5, Sept-Oct 2007,pp 351-357.

[12] D.L.Donoho, "Denoising by soft Thresholding", IEEE Trans. On Information Theory, Vol.41,no. 3, pp 613-627,1995.

[13] S. Boll, "Suppression of acoustic noise in speech using Spectral Subtraction", IEEE transactions on Acoustic Speech and Signal Processing Vol ASSP 27, April 1979, pp-113-120.

[14]