International Journal of Computer Applications (0975 – 8887)
Volume 20– No.5, April 2011

# Noise Reduction of Speech Signal using Wavelet Transform with Modified Universal Threshold

Rajeev Aggarwal
Department of Elex.& Comm.
SSSIST, Sehore
M.P. India

Jai Karan Singh
Department of Elex.& Comm.
SSSIST, Sehore
M.P. India

Vijay Kumar Gupta
Department of Elex.& Comm.
IMS Engg.College Ghaziabad
U.P. India

Sanjay Rathore
Department of Physics
SSSIST, Sehore
M.P. India

Mukesh Tiwari
Department of Elex.& Comm.
SSSIST, Sehore
M.P. India

Dr. Anubhuti Khare
Department of Electronics
UIT,RGPV Bhopal M.P. India

## ABSTRACT
In this paper, Discrete-wavelet transform (DWT) based algorithm are used for speech signal denoising. Here both hard and soft thresholding are used for denoising. Analysis is done on noisy speech signal corrupted by babble noise at 0dB, 5dB, 10dB and 15dB SNR levels. Simulation & results are performed in MATLAB 7.10.0 (R2010a). Output SNR (Signal to Noise Ratio) and MSE (Mean Square Error) is calculated & compared using both types of thresholding methods. Soft thresholding method performs better than hard thresholding at all input SNR levels. Hard thresholding shows a maximum of 21.79 dB improvement whereas soft thresholding shows a maximum of 35.16 dB improvement in output SNR.

## General Terms
Thresholding, multi-resolution analysis, wavelet.

## Keywords
Discrete wavelet transform, hard thresholding, soft thresholding, signal to noise ratio.

## 1. INTRODUCTION
Speech denoising is a field of engineering that studies methods used to recover an original speech from noisy signals corrupted by different types of noises. Noise may be in the form of white noise, pink noise, babble noise and many other types of noise present in the environment. Over the last decades, noise removal from speech signals is an area of interest of researchers during speech processing. Wavelet methods are mostly used for speech denoising. The application areas of wavelet transform are 1D or 2D biomedical signal analysis, producing & analyzing irregular signals or images, wavelet modulation in communication channels, in video coding and forecasting etc. The fundamental idea behind wavelets is to analyze according to scale. Wavelet transforms can decompose a signal into several scales that represent different frequency bands, and at each scale, the position of signal's instantaneous structures can be determined approximately. Such a property can be used for denoising. Although without information on the signal to be analysed, wavelet select information by strongly reducing its quantity. Other way, wavelet is a small wave and wavelet transforms convert a signal into a series of wavelets and provide a way for analyzing waveforms, bounded in both frequency and duration. This allows signal to be stored more efficiently than by Fourier transform. Wavelet transform is preferred over Fourier Transform (FT) and Short Time Fourier Transform (STFT), since it provided multi-resolution.

Recently, various wavelet based methods have been proposed for the purpose of speech denoising. The wavelet split coefficient method is a speech denoising procedure to remove noise by shrinking the wavelet coefficients in the wavelet domain. The method is based on thresholding in the signal that each wavelet coefficient of the signal is compared to a given threshold; if the coefficient is smaller than the threshold, then it is set to zero, otherwise it is kept or slightly reduced in amplitude. Soft and Hard Thresholding are used for denoising the signals. Using wavelets to remove noise from a signal requires identifying which components contain the noise, and then reconstructing the signal without those components.

The paper is organized as follows :- Section 2 Wavelet and Multi-resolution, Section 3 Discrete wavelet Transform (DWT), Section 4 Multi-resolution analysis using filter bank, Section 5 Modified Universal Threshold, Section 6 Soft and Hard thresholding, Section 7 results and discussion, Section 8 conclusion.

## 2. WAVELET AND MULTIRESOLUTION
In time domain signal, the independent variable is time and the dependent variable is the amplitude. Most of the Information is hidden in the frequency content. By using wavelet Transform, we can get the frequency information which is not possible by working in time-domain. The analysis of a non-stationary signal using the Fourier Transform and Short Time Fourier Transform does not give satisfactory results. Better results can be obtained using wavelet transform analysis. In STFT, a fixed time-frequency resolution is used where as in wavelet transform, multi-resolution technique is used. One advantage of wavelet transform analysis is the ability to perform local analysis. Wavelet analysis is able to express signal appearance that other analysis techniques miss such as breakdown points, discontinuities etc.
[1] In multi-resolution analysis (MRA), signal has good time resolution and poor frequency resolution at high frequencies and

other way good frequency resolution and poor time resolution at low frequencies. It is more suitable for short duration of higher frequency and longer duration of lower frequency components. It is assumed that low frequencies appear for the entire duration of the signal, whereas high frequencies appear from time to time as short interval. This is often the case in practical applications.

## 3. DISCRETE WAVELET TRANSFORM

The continuous wavelet transform (CWT) performs a multi-resolution analysis by contraction and dilatation of the wavelet functions and discrete wavelet transform (DWT) uses filter banks for the construction of the multi-resolution time-frequency plane [1]. The DWT uses Multi resolution filter banks and special wavelet filters for the analysis and reconstruction of signals.

[2] In FT we know the information of frequency content of the signal, but we don't know at what times frequency components occur. So to avoid this problem we use Short term Fourier transform (STFT) and wavelet transform for analysis of signals like speech. If we want to choose transform from STFT & wavelet transform, we will give preference to wavelet transform because it analyses the signal at different frequency with different resolutions.

DWT provides sufficient information both for analysis and synthesis and reduce the computation time sufficiently. It analyse the signal at different frequency bands with different resolutions, decompose the signal into a coarse approximation and detail information. Human ear has better frequency resolution at low frequencies and lower frequency resolution at high frequencies. Decomposition of the signal is obtained by passing time domain signal through low pass and high pass filters.

## 4. MULTIRESOLUTION ANALYSIS USING FILTER BANK

Let us discuss the multi-resolution analysis upto Level 2. In figure 1, input noisy signal (Vn+1) is given to the high & low pass filter. Figure 1 shows the Two-level wavelet decomposition tree [3].

G is the high pass filter & H is the low pass filter. Wavelet analysis involves filtering and down-sampling. Output of these filters is decimated by 2 and gets the new coefficient named as Vn & Wn. Where Vn is Approximation coefficient and Wn is detail coefficient. The decomposition algorithm will decompose Vn into Vn-1 and Wn-1, as shown in figure 1.

Next, Wavelet thresholding is applied to the approximation coefficient (Vn-1) and detail coefficients (Wn,Wn-1) and this coefficient are again combined to reconstruct the noise free signal (where as practically it's not possible to remove the noise completely).

Process after decomposition or analysis is called synthesis where we reconstruct the signal from the wavelet coefficients. Wavelet reconstruction process consists of up-sampling and filtering. In reconstruction, components can be assembled back into the original signal without loss of information.

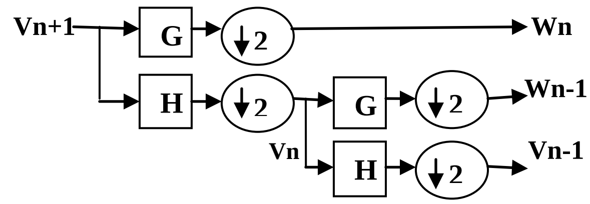



**Figure 1. Two Level Wavelet Decomposition tree**

Figure 2 shows the Two-level wavelet reconstruction tree.

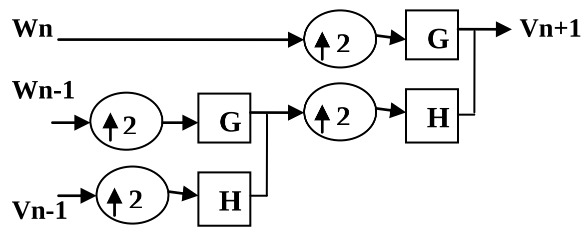



**Figure 2. Two Level wavelet reconstruction tree**

## 5. MODIFIED UNIVERSAL THRESHOLD

In this paper, we removed the babble noise from noisy signal which contain the noise contents of babble noise. Here we use the multi-resolution concept. Now, we want to find threshold value that will use to remove noise from noisy signal, but also recover the original signal efficiently. If the threshold value is too high, it will also remove the contents of original signal and if the threshold value is too low, denoising will not work properly.

[2] One of the first methods for selection of threshold was developed by Donoho and Jonstone [4] and it is called as universal threshold. Different universal threshold was proposed in [5]:

$$\mathrm{thr} = \sigma n\sqrt{2\ \mathrm{Log2}\ (N)} \qquad \text{-----------------} \ (1)$$

Where N denotes number of samples of noise and σn is standard deviation of noise. But threshold obtained by equation (1) is too high.

Again Universal threshold was proposed in [2] and modified with factor 'k' in order to obtain higher quality output signal:

$$\mathrm{thr} = k.\sigma n\sqrt{2\ \mathrm{Log2}\ (N)} \qquad \text{-----------------} \ (2)$$

During our research it is noticed that if we use two factors i.e. k & m, then new threshold value gives better results, especially to recover the original signal. We will discuss on this topic in detail in section 7.

Denoising algorithm scheme is showed in Figure 3. Inverse DWT is applied to get denoise time domain signal.

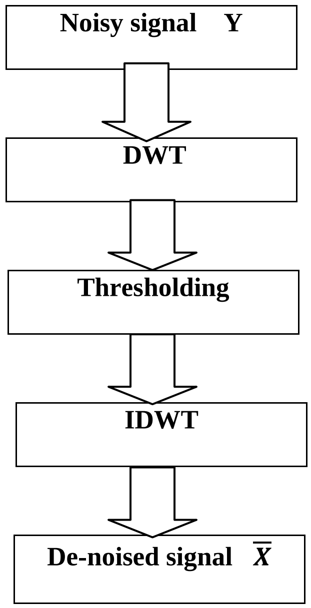



**Figure 3. Denoise Algorithm**

## 6. SOFT AND HARD THRESHOLDING

The soft and hard thresholding methods are used to estimate wavelet coefficients in wavelet threshold denoising. [1] Hard thresholding zeros out small coefficients, resulting in an efficient representation. Soft thresholding softens the coefficients exceeding the threshold by lowering them by the threshold value. When thresholding is applied, no perfect reconstruction of the original signal is possible.

Hard thresholding can be described as the usual process of setting to zero the elements whose absolute values are lower than the threshold. The hard threshold signal is x if x $\geq$ thr and is 0 if x $<$ thr, where 'thr' is a threshold value. 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. If x $\geq$thr, soft threshold signal is (sign (x). (x - thr)) and if x$<$ thr , soft threshold signal is 0. Hard thresholding is the simplest method but soft thresholding has nice mathematical properties and gives better denoising performance.

For example, as shown in figure 4, original signal has a line space with Z = (-1, 1, 100). This means line space generates a row vector Z of 100 points linearly spaced between and including -1 and 1. In hard and soft thresholding, threshold value 'thr' is 0.4

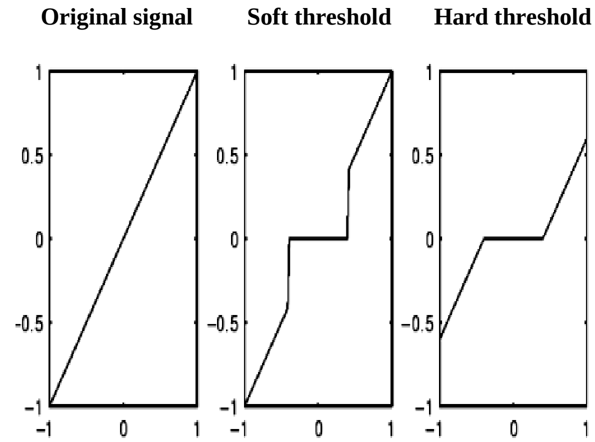



**Figure 4 Hard and Soft thresholding**

$$\mathrm{T_{Hard}}\ (\mathrm{x}) = \begin{cases} \mathrm{x} & |\mathrm{x}| \geq \mathrm{thr} \\ 0 & |\mathrm{x}| < thr \end{cases} \qquad \text{-------------} \ (3)$$

$$\mathrm{T_{Soft}}\ (\mathrm{x}) = \begin{cases} \mathrm{Sign}\ (\mathrm{x}).(\mathrm{x} - \mathrm{thr}) & \mathrm{x} \geq \mathrm{thr} \\ 0 & -\mathrm{thr} \leq \mathrm{x} < thr \\ \mathrm{Sign}\ (\mathrm{x}).(\mathrm{x} + \mathrm{thr}) & \mathrm{x} < -thr \end{cases}$$

------------- (4)

## 7. RESULTS AND DISCUSSION

We implemented babble noise removal algorithm in Matlab 7.10.0 (R2010a). Wavelet toolbox in Matlab has large collection of functions for wavelet analysis. Input of our simulation is noisy signal in Wave format which is sampled at sampling frequency of Fs=8000 Hz. Speech signal is corrupted by babble noise at 0dB, 5dB, 10dB and 15dB SNR levels.

This algorithm is very useful when we don't know about original signal (noise free). We only use the original signal just to compare the de-noised signal with the original speech signal. It is assumed that high amplitude DWT coefficients represent signal, and low amplitude coefficients represent noise. It is considered that some samples of noisy signal contain only noise, so we choose that samples for noise calculation.

It's very important to select the proper level in multi-resolution analysis. In multi-resolution analysis, the approximation signals (output of Low-pass filter & then decimation) are splits up to the certain level.

Figure 5 shows the Original speech signal and figure 6 shows the noisy signal.

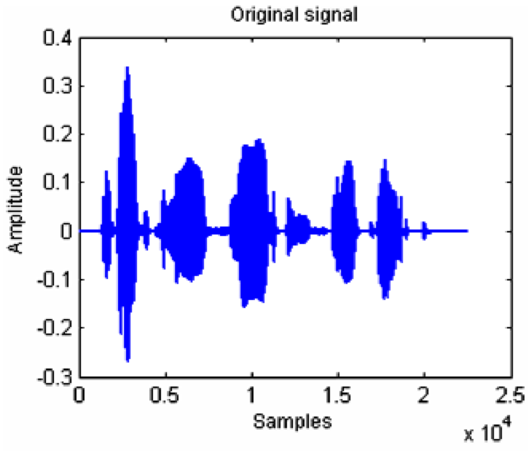



**Figure 5. Original speech signal**

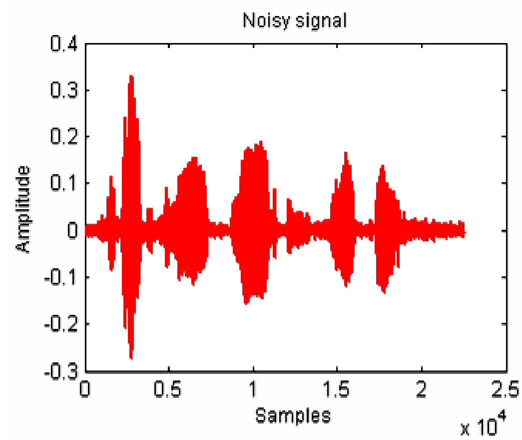



**Figure 6. Noisy signal**

We choose 5-level DWT and db5 wavelet. Improved threshold value is obtained by replacing threshold 'thr' (2) with

$$thr = m.k.\sigma n\sqrt{2\ Log2\ (N)} \qquad \text{----------------} (5)$$

Where, $0 < k < 1$ and $0 < m < 1$. 'N' denotes number of noise samples and σn is standard deviation of noise.

We use two factors 'k' and 'm'. It is found that if we fix one factor & vary other factor, than we can get different range of threshold value which gives improved results to recover the original signal especially for low level noise.

We apply this threshold value to approximation coefficient & all detail coefficients. We use soft & hard thresholding separately. Finally use these new coefficients to reconstruct the signal. We found that soft thresholding results are more efficient than hard thresholding. Figure7 shows the reconstructed (denoise) signal. Figure 8 shows the noise in noisy signal and figure 9 shows the noise in denoise signal.

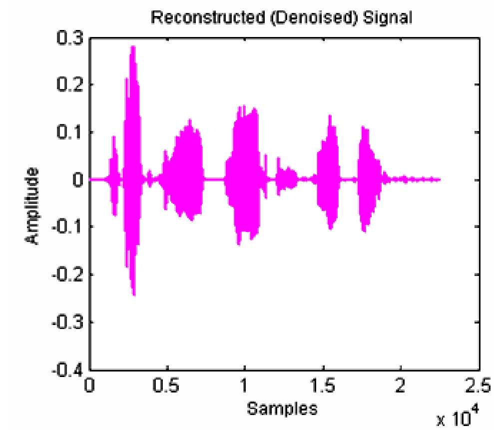



**Figure 7. Reconstructed signal for babble noise of 15 db using soft threshold**

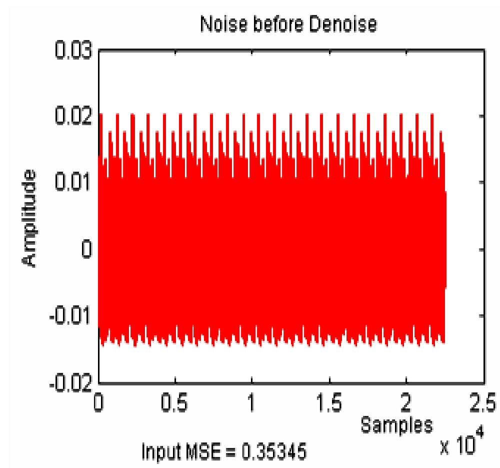



**Figure 8. Noise in noisy signal**

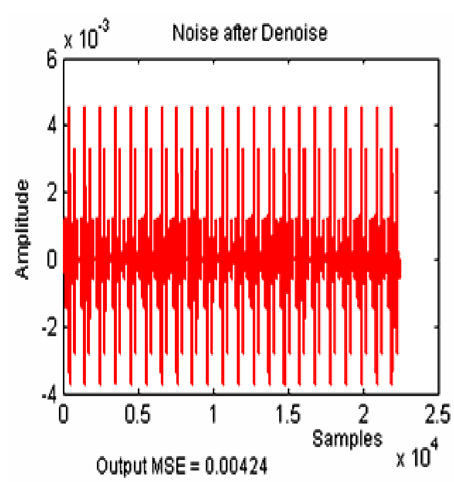



**Figure 9. Noise in de-noise signal**

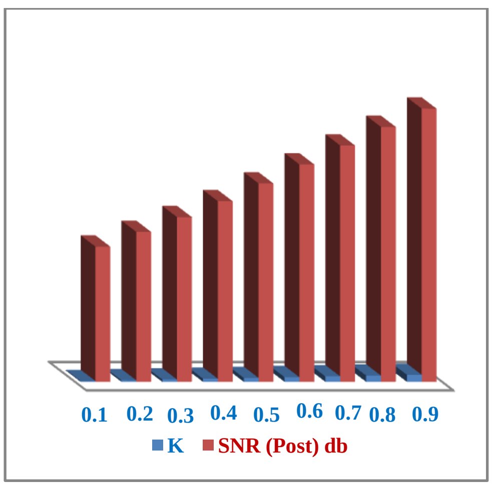
In example 1 as shown in table 1, when fix the value of factor m= 0.9 & varies the value of 'k' from 0.1 to 0.9, the value of Post SNR also increases.

**Table 1. Results of noise removal algorithm for modified soft thresholding method (example 1)**

| Std. Dev. σ | k | m | Input MSE ($10^{-4}$) | Output MSE ($10^{-4}$) | SNR Pre (db) | SNR Post (db) |
|---|---|---|---|---|---|---|
| 0.006 | 0.1 | 0.9 | 0.3535 | 0.2038 | 15 | 17.40 |
| 0.006 | 0.2 | 0.9 | 0.3535 | 0.1253 | 15 | 19.30 |
| 0.006 | 0.3 | 0.9 | 0.3535 | 0.0772 | 15 | 21.20 |
| 0.006 | 0.4 | 0.9 | 0.3535 | 0.046 | 15 | 23.24 |
| 0.006 | 0,5 | 0.9 | 0.3535 | 0.026 | 15 | 25.53 |
| 0.006 | 0.6 | 0.9 | 0.3535 | 0.0141 | 15 | 27.99 |
| 0.006 | 0.7 | 0.9 | 0.3535 | 0.0077 | 15 | 30.42 |
| 0.006 | 0.8 | 0.9 | 0.3535 | 0.0042 | 15 | 32.81 |
| 0.006 | 0.9 | 0.9 | 0.3535 | 0.0024 | 15 | 35.16 |

Graph 1. shows the comparison between factor 'k' & Post SNR.

**Graph 1. K v/s SNR Post (db)**




In example 2 as shown in table 2, we fix the value of both factor 'k' & 'm' and analyse the post SNR at different SNR level of noisy signal. Mean square error (MSE) is mostly used for estimating signal quality. Graph 2 shows comparison between input MSE with Output MSE at different SNR level.

**Table 2. Results of noise removal algorithm for modified soft thresholding method (example 2)**

| Std. Dev. σ | k | m | Input MSE ($10^{-4}$) | Output MSE ($10^{-4}$) | SNR Pre (db) | SNR Post (db) |
|---|---|---|---|---|---|---|
| 0.036 | 0.8 | 0.9 | 12.6942 | 0.1071 | 0 | 13.90 |
| 0.018 | 0.8 | 0.9 | 3.1583 | 0.1013 | 5 | 16.99 |
| 0.010 | 0.8 | 0.9 | 1.0822 | 0.0290 | 10 | 23.46 |
| 0.006 | 0.8 | 0.9 | 0.3535 | 0.0042 | 15 | 32.81 |

Input MSE is defined as [2]:

$$\frac{1}{N}\sum_{i}(X_i - Y_i)^2 \quad \text{--------------------- (6)}$$

Where, $X_i$ is original signal and $Y_i$ is noisy signal.

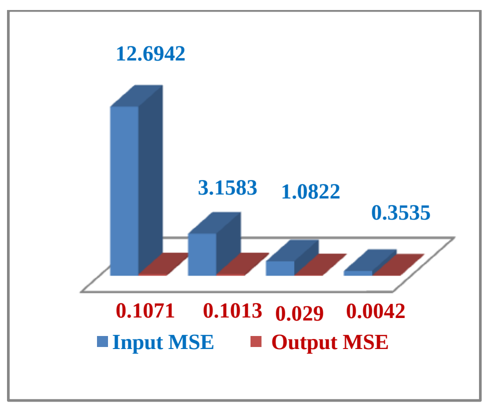
Output MSE is defined as:

$$\frac{1}{N}\sum_{i}(X_i - \overline{X}_i)^2 \quad \text{-------------------- (7)}$$

Where, $\overline{X_i}$ is reconstructed signal.

Denoising is successful if output MSE is lower than input MSE. Lower MSE means the closer match between two signals.

**Graph 2. Input MSE v/s Output MSE**




Another method to analyse the denoise signal is by Signal to Noise Ratio (SNR). SNR is used to quantify how much a signal has been corrupted by noise. It is defined as the ratio of signal power to the noise power corrupting the signal. Denoising is successful if Post SNR is higher than Pre SNR.

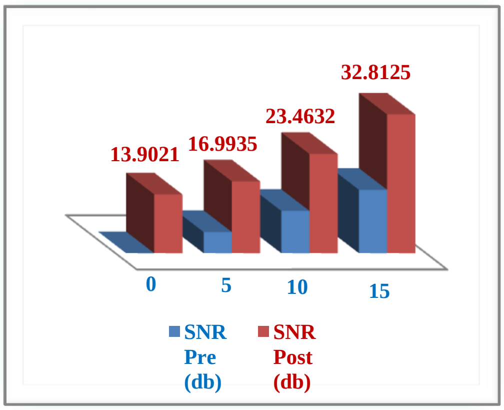
Signal to Noise Ratio is defined as:

$$SNR_{db} = 10 \, Log_{10}\left(\frac{P_{signal,db}}{P_{noise,db}}\right) = P_{signal,db} - P_{noise,db}$$

------------------- **(8)**

Graph 3 shows comparison between SNR (Pre) db with SNR (Post) db at different SNR level by using soft thresholding method.

**Graph 3. SNR Pre (db) v/s SNR Post (db)**



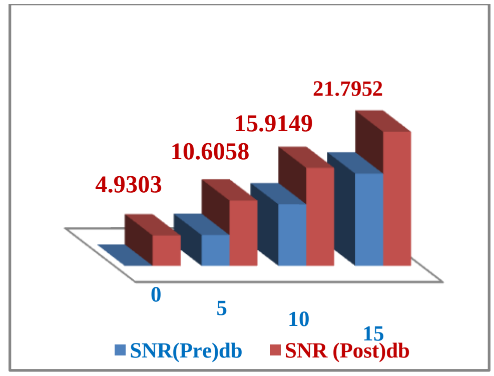

In example 3 as shown in table 3, hard thresolding method is used and found that results are not so good as compare to analysis done by using soft thresholding method.

**Table 3. Results of noise removal algorithm for modified hard thresholding method (example 3)**

| Std. Dev. σ | k | m | Input MSE $(10^{-4})$ | Output MSE $(10^{-4})$ | SNR Pre (db) | SNR Post (db) |
|---|---|---|---|---|---|---|
| 0.036 | 0.8 | 0.9 | 12.6942 | 3.4536 | 0 | 4.93 |
| 0.018 | 0.8 | 0.9 | 3.1583 | 1.0337 | 5 | 10.61 |
| 0.010 | 0.8 | 0.9 | 1.0822 | 0.2961 | 10 | 15.91 |
| 0.006 | 0.8 | 0.9 | 0.3535 | 0.0766 | 15 | 21.79 |

Graph 4 shows comparison between SNR (Pre) db with SNR (Post) db at different SNR level by using hard thresholding method.

**Graph 4. SNR Pre (db) v/s SNR Post (db)**




## 8. CONCLUSION

In this paper we used wavelet transform for denoising speech signal corrupted with babble noise. Speech denoising is performed in wavelet domain by thresholding wavelet coefficients. We found that by using modified universal threshold, we can get the better results of de-noising, especially for low level noise. During different analysis we found that soft thresholding is better than hard thresholding because soft thresholding gives better results than hard thresholding. Higher threshold removes noise well, but the part of original signal is also removed with the noise. It is generally not possible to filter out all the noise without affecting the original signal. We can analyse the denoise signal by signal to noise ratio (SNR) and mean square error (MSE) analysis.

## 9. REFERENCES

[1] Prof. Dr. Ir. M. Steinbuch, Dr. Ir. M.J.G. van de Molengraft, June 7 (2005), Eindhoven University of Technology, Control Systems Technology Group Eindhoven, "Wavelet Theory and Applications", a literature study, R.J.E. Merry, DCT 2005.53.

[2] MATKO SARIC, LUKI BILICIC, HRVOJE DUJMIC (2005), "White Noise Reduction of Audio Signal using Wavelets Transform with Modified Universal Threshold", University of Split, R. Boskovica b. b HR 21000 Split, CROATIA.

[3] Adhemar Bultheel, September 22 (2003), "Wavelets with applications in signal and image processing".

[4] D.L. Donoho, May (1992), Stanford University: "De-noising by Soft Thresholding", IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 41, NO. 3.

[5] Alexandru Isar, Dorina Isar, May (2003) : "Adaptive denoising of low SNR signals", Third International Conference on WAA 2003, Chongqing, P. R. China, 29-31, p.p. 821-826.

[6] Yuan Yan Tang, 2001, "Wavelet analysis and its applications": second international conference, Springer-Verlag,(U.K.)