

Comparative Analysis between DWT and WPD Techniques of Speech Compression

Preet Kaur¹, Pallavi Bahl²

¹(Assistant professor, YMCA university of science & Technology)

²(student, YMCA university of Science and Technology)

ABSTRACT: - Speech compression is the process of converting speech signal into more compactable form for communication and storage without losing intelligibility of the original signal. Storage and archival of large volume of spoken information makes speech compression essential and which improves the capacity of communications relatively of unlimited bandwidth. Discrete Wavelet Transform (DWT) and Wavelet Packet Decomposition (WPD) are the recent technique used to materialize the compression. In this paper, both the techniques are exploited, and a comparative study of performance of both is made in terms of Signal-to-noise ratio (SNR), Peak signal-to-noise ratio (PSNR), Normalized root-mean square error (NRMSE) and Retained signal energy (RSE) is presented.

Keywords: - Speech compression, DWT, WPD, PSNR

I. INTRODUCTION

Speech is an acoustic signal by nature and it is the most effective medium for face to face communication and telephony application[1]. Speech coding is the process of obtaining a compact representation of voice signals for efficient transmission over band-limited wired and wireless channels and/or storage. A Speech compression system focuses on reducing the amount of redundant data while preserving the integrity of signals. [2] Speech compression is required in long distance communication, high quality speech storage, and message encryption. Compression techniques can be classified into one of the two main categories: lossless and lossy. In lossless compression, the original file can be perfectly recovered from the compressed file. In case of lossy compression, the original file cannot be perfectly recovered from the compressed file, but it gives its best possible quality for the given technique. Lossy compression typically attain far better compression than lossless by discarding less-critical data. Any compression on continuous signal like speech is unavoidably lossy[3]. Speech compression plays an important role in teleconferencing, satellite communications and multimedia applications. However, it is more important to ensure that compression algorithm retains the intelligibility of the speech. The success of the compression scheme is based on simplicity of technology and efficiency of the algorithm used in the system.[3][4]

Various compression techniques have been used by researcher to compress speech signal [5]. In this paper, Discrete wavelet transform [6] and wavelet packet decomposition techniques are used to compress the speech signals. The

paper has been organized as follows: Section II talks about the speech compression techniques used i.e. Discrete wavelet transforms and the wavelet packet decomposition technique. Section III shows the compression methodology used in the experiment. In section IV, results and graphs are discussed and finally Conclusions are drawn in section V.

II. SPEECH COMPRESSION TECHNIQUES USED

This section deals with the speech compression techniques that we used in this experiment.

1.1 TRANSFORM METHOD

Transformations are applied to the signals to obtain information details from that signal. Fourier transform is time domain representation of signal and is not suitable if the signal has time varying frequency that is not stationary[7]. In particular, the wavelet transform is of interest for the analysis of non stationary signals, because it provides an alternative to the classical Short-Time Fourier Transform. In contrast to the STFT, which uses a single analysis window, the WT uses short window at high frequency and long windows at low frequencies.[8].

The Wavelet Transform (WT) is a mathematical tool for signal analysis. For certain applications, the WT has distinct advantages over more classical tools such as the Fourier transform. Two important features of the WT are its ability to handle nonstationary signals and its time-frequency resolution properties.[8]

1. Discrete Wavelet Transform

The signal is divided into two versions i.e. approximation coefficients and detail coefficients. The low pass signal gives the approximate representation of the signal while the high pass filtered signal gives the details

or high frequency variations. The second level of decomposition is performed on the approximation coefficients obtained from the first level of decomposition. [9]

Where, the original signal is represented by $x^0(n)$. Here $g(n)$ and $h(n)$ represent the low pass and high pass filter, respectively

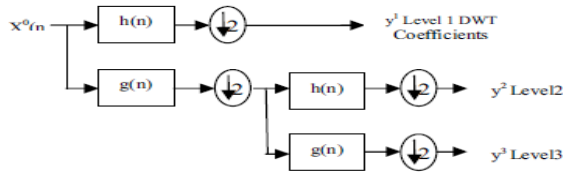


Fig.1 Filter bank representation of DWT decomposition

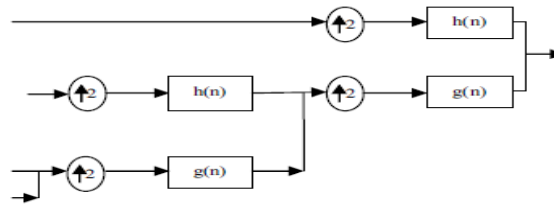


Fig. 2. IDWT filter bank representation

In order to reconstruct the original signal, at each level of reconstruction, approximation components and the detailed components are up by 2 and the detailed components are up sampled by 2, and then convolved which is shown in Fig. 2.

2. The Wavelet Packet Decomposition.

Wavelets packets have been introduced by coifman, meyer and wickenhauser.[10].The wavelet packet method is a generalization of wavelet decomposition that offers a richer range of possibilities for signal analysis.. In wavelet packet analysis each detail coefficient vector is also decomposed in to two parts using the same approach as in approximation vector splitting. This yields more than different ways to encode the signal. This offers the richest analysis . In the WPD, both the detail and approximation coefficients are decomposed in each level [10][11].

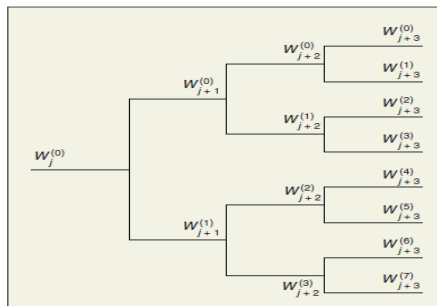


Fig.3: A binary tree representation of a Three-levels wavelet packet spaces

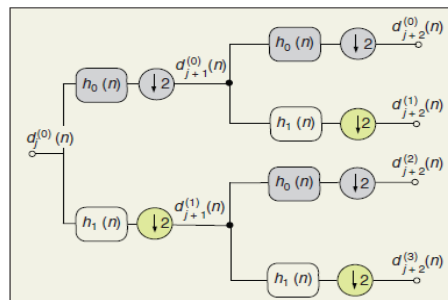


Fig.4: Wavelet packet filter bank analysis algorithm

III. COMPRESSION METHODOLOGY SPEECH COMPRESSION USING DWT/ WPD

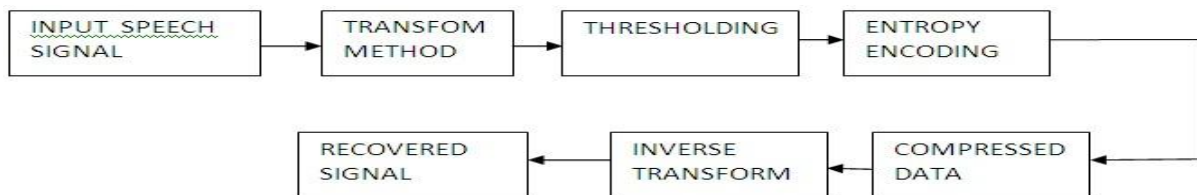


Fig.7: Block diagram of DWT/WPD.

Transform method:- Wavelets work by decomposing a signal into different resolutions or frequency bands. Signal compression is based on the concept that selecting a small number of approximation coefficients and some detail coefficients can accurately represent regular signal components.

Thresholding:- After calculating the wavelet transform of the speech signal, compression involves truncating wavelet coefficients below a threshold. The coefficients obtained after applying DWT on the frame concentrate energy in few neighbours. Thus we can truncate all coefficients with low energy and retain few coefficients holding the high energy value. The two thresholding techniques are implemented.

- 1) **Global Threshold :-** The aim of global thresholding is to retain the largest absolute value coefficients , regardless of the scale in the wavelet decomposition tree. Global thresholds are calculated by setting the percentage of coefficients to be truncated.
- 2) **Level Dependent Threshold :-** This approach consists of applying visually determined level dependent threshold to all detail coefficients. The truncation of insignificant coefficients can be optimized when such a level dependent thresholding is used. By applying this the coefficients below the level is made zero .

Entropy Encoding :- Signal compression is achieved by first truncating small-valued coefficients and then efficiently encoding them. We have used Huffman encoding to encode detail coefficients.

Inverse transform :- Inverse transform is applied to the decomposed compressed signal to recover the original signal.

Choosing the Decomposition Level

The DWT on a given signal, the decomposition level can reach up to level $L=2^k$,where k is the length of discrete signal. Thus we can apply transform at any of these levels. But infact ,the decomposition level depends on the type of signal being analyzed. In this paper , full length decomposition is obtained for signal and comparisons were made with level 6 and 7.

IV. RESULTS AND DISCUSSION

The coding of this paper is done in MATLAB 7. In this paper, we compared Discrete wavelet transform (DWT) and wavelet packet decomposition (WPD). A number of quantitative parameters can be used to evaluate the performance of the coder, in terms of reconstructed signal quality after compression scores. The following parameters are compared:

- Signal to Noise Ratio (SNR),
- Peak Signal to Noise Ratio (PSNR),
- Normalized Root Mean Square Error (NRMSE),
- Retained Signal Energy (RSE),
- Compression Ratio (CR).

• **Signal to Noise Ratio:**

$$SNR = 10 \log_{10} (\sigma_x)^2 / (\sigma_e)^2$$

where $(\sigma_x)^2$ is the mean square of the speech signal, $(\sigma_e)^2$ is the mean square difference between the original and reconstructed signals

• **Peak Signal to Noise Ratio**

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

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

• **Normalized Root Mean Square Error (NRMSE)**

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

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

• **Retained Energy**

$$RSE = \frac{100 * \|x(n)\|^2}{\|r(n)\|^2}$$

$x(n)$ is the norm of the original signal and $r(n)$ is the norm of the reconstructed signal

• **Compression Ratio (CR)**

$$CR = \frac{Lengt h(x(n))}{Lengt h(r(n))}$$

Where $x(n)$ is the original signal and $r(n)$ is the reconstructed signal.

Speech compression is a way to representing a speech signal with minimum data values and favorable in case of storage and transmission. Two speech signal “good bye” and “wow” are compressed using different

Comparative Analysis between DWT and WPD Techniques of Speech Compression

wavelet and wavelet packet decomposition. Objective analysis of these two speech signals are done by evaluating the performance of parameters such as Compression Ratio (CR), Peak Signal to Noise Ratio (PSNR), Signal to Noise Ratio (SNR), Normalized Root Mean Square Error Rate (NRMSE) and Retained Signal Energy (RSE).

Table-1: Comparison between compression using wavelet transform and wavelet packet decomposition using different wavelets for speech signal “goodbye”.

	CR	SNR	PSNR	NRMSE	RSE
HAAR	1.1870	4.8822	19.0291	.5701	67.5075
WPD(HAAR)	1.3757	6.7175	20.8645	.4615	78.7065
DB2	1.2865	3.1380	17.2849	.6969	51.4488
WPD(DB2)	1.3596	5.9042	20.0512	.5068	74.3211
DB4	1.1363	4.0072	18.1541	.6305	60.2553
WPD(DB4)	1.3792	5.6226	19.7695	.5235	72.6008

Table 2: Comparison between compression using wavelet transform and wavelet packet decomposition using different wavelets for speech signal “wow”.

	CR	SNR	PSNR	NRMSE	RSE
HAAR	1.1675	4.384	15.4982	.6037	63.5579
WPD(HAAR)	1.3492	6.5057	17.6199	.4728	77.642
DB2	1.2851	3.0443	14.1585	.7044	50.3894
DB2(HAAR)	1.3469	6.1912	17.3055	.4903	75.9632
DB4	1.1326	3.0567	14.171	.7033	50.5319
DB4(HAAR)	1.3354	6.6056	17.7199	.4674	78.1507

As seen from the table1 and table 2, the performance of WPD is better than DWT. SNR obtained using DWT with HAAR as mother wavelet was found better than SNR obtained using DB2 as mother wavelet and SNR of DWT with DB4 as mother wavelet was found better than SNR obtained using DB2 as mother wavelet. CR of DWT with DB2 was found to be highest. No further enhancement was achieved with beyond level 6 decomposition. Table 3 and table 4 gives the comparison between compression using DWT with different wavelet and different thresholding techniques for speech signal “goodbye” and “wow” respectively. It can be seen from the table that for a particular wavelet, when global thresholding technique was used, the performance parameters were found better in comparison to hard thresholding technique.

Table 3: Comparison between compression using DWT with different wavelet and different thresholding techniques for speech signal “goodbye”.

	HAAR(hard threshold)	HAAR(global threshold)	DB2(hard threshold)	DB2(global threshold)	DB4(hard threshold)	DB4(global threshold)
CR	1.2982	1.3167	1.2865	1.3152	1.1363	1.1811
SNR	3.4558	3.7585	3.1380	3.3280	4.0072	4.0776
PSNR	17.6028	17.9055	17.2849	17.4749	18.1541	18.2246
NRMSE	.6719	.6489	.6969	.6818	.63255	.6254
RSE	57.8749	57.9130	51.4488	53.5270	60.2553	60.8947

Table 4: Comparison between compression using DWT with different wavelet and different thresholding techniques for speech signal “wow”

	HAAR(hard threshold)	HAAR(global threshold)	DB2(hard threshold)	DB2(global threshold)	DB4(hard threshold)	DB4(global threshold)
CR	1.3294	1.3391	1.2851	1.2852	1.1326	1.1360
SNR	3.3396	3.4573	3.0443	3.1443	3.0567	3.1218
PSNR	14.4538	14.5715	14.1585	14.2582	14.1710	14.2361
NRMSE	.6808	.6716	.7044	.6963	.7033	.6981
RSE	53.6506	54.8900	50.3894	51.5190	50.5319	51.2677

Fig 8 and fig 9 shows the comparison of speech signal “good bye and “wow” respectively, on the basis of SNR for different wavelet transform and wavelet packet decomposition. From the fig we can see that WPD gives better SNR as compared to DWT for both the speech signals..

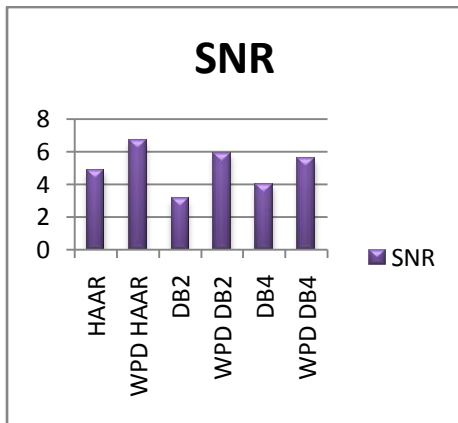


Fig.8. Comparison of speech signal “good bye” On the basis of SNR

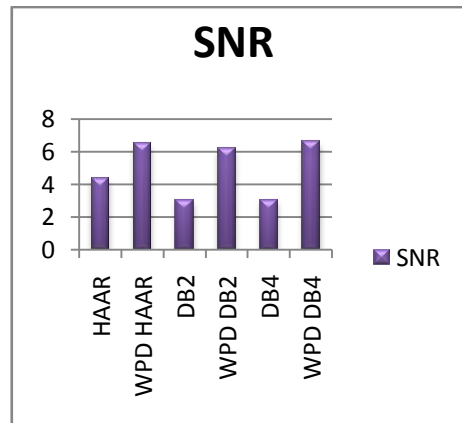


Fig.9. Comparison of speech signal “wow” On the basis of SNR

Fig 10 and fig 11, shows the comparison of speech signal on basis of PSNR and fig 12 and fig 13, compare the speech signal on basis of NRMSE.

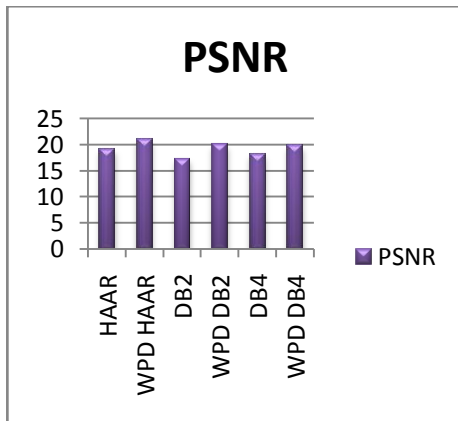


Fig.10: Comparison of speech signal “good bye” On the basis of PSNR

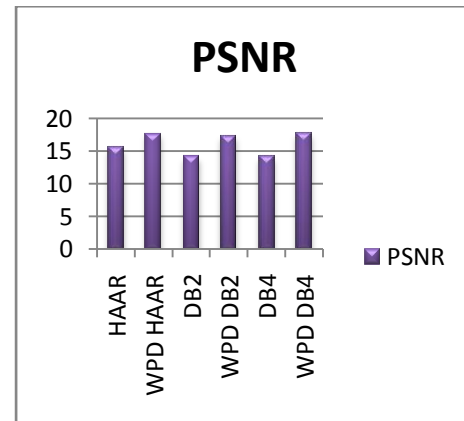


Fig.11: Comparison of speech signal “wow” On the basis of PSNR

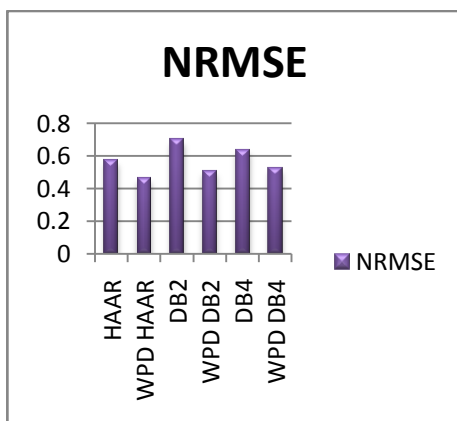


Fig.12: Comparison of speech signal “good bye” On the basis of NRMSE

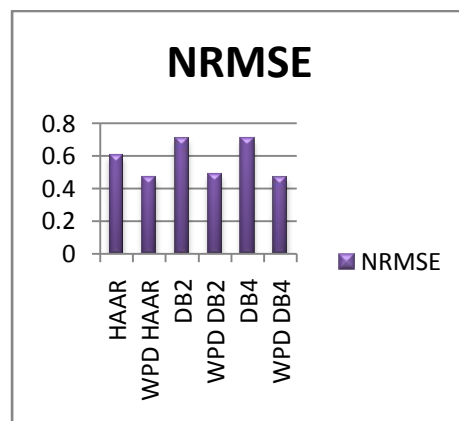


Fig.13: Comparison of speech signal “wow” On the basis of NRMSE

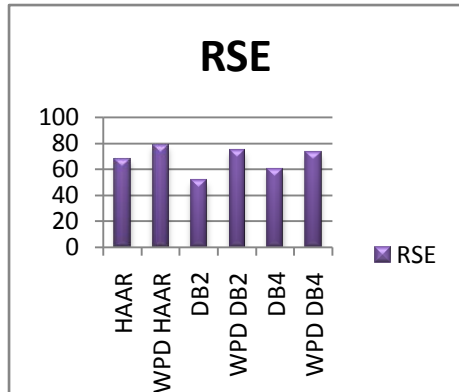


Fig.14: Comparison of speech signal “good bye” On the basis of RSE

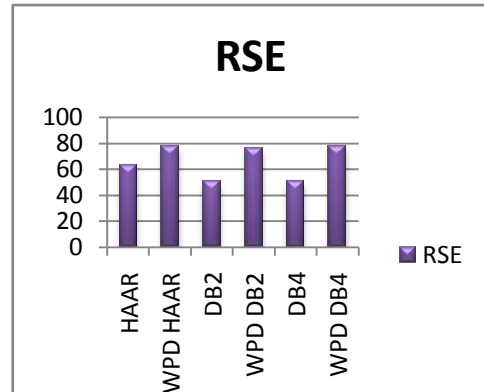


Fig.15: Comparison of speech signal “wow” On the basis of RSE

From fig 16 and 17, it can be observed, best CR in good bye speech signal is achieved with WPD DB4, it is comparable to WPD DB2 and WPD HAAR and best CR in “wow” is achieved with WPD DB2, which is comparable to WPD DB4 and WPD HAAR.

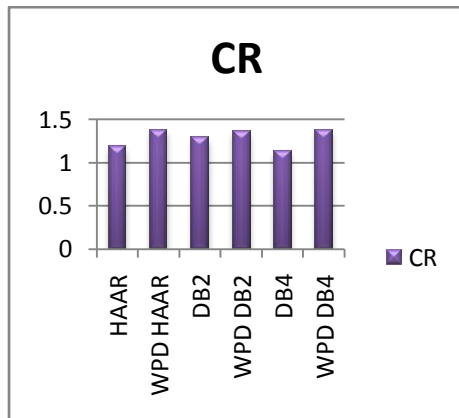


Fig.16: Comparison of speech signal “good bye” On the basis of CR

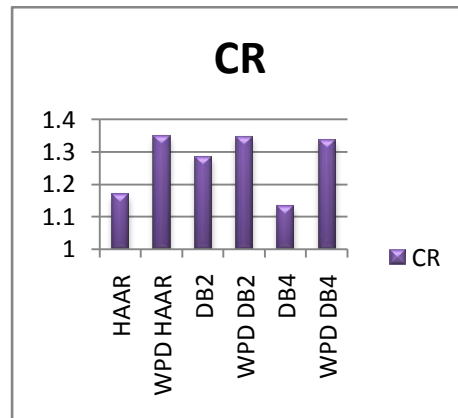
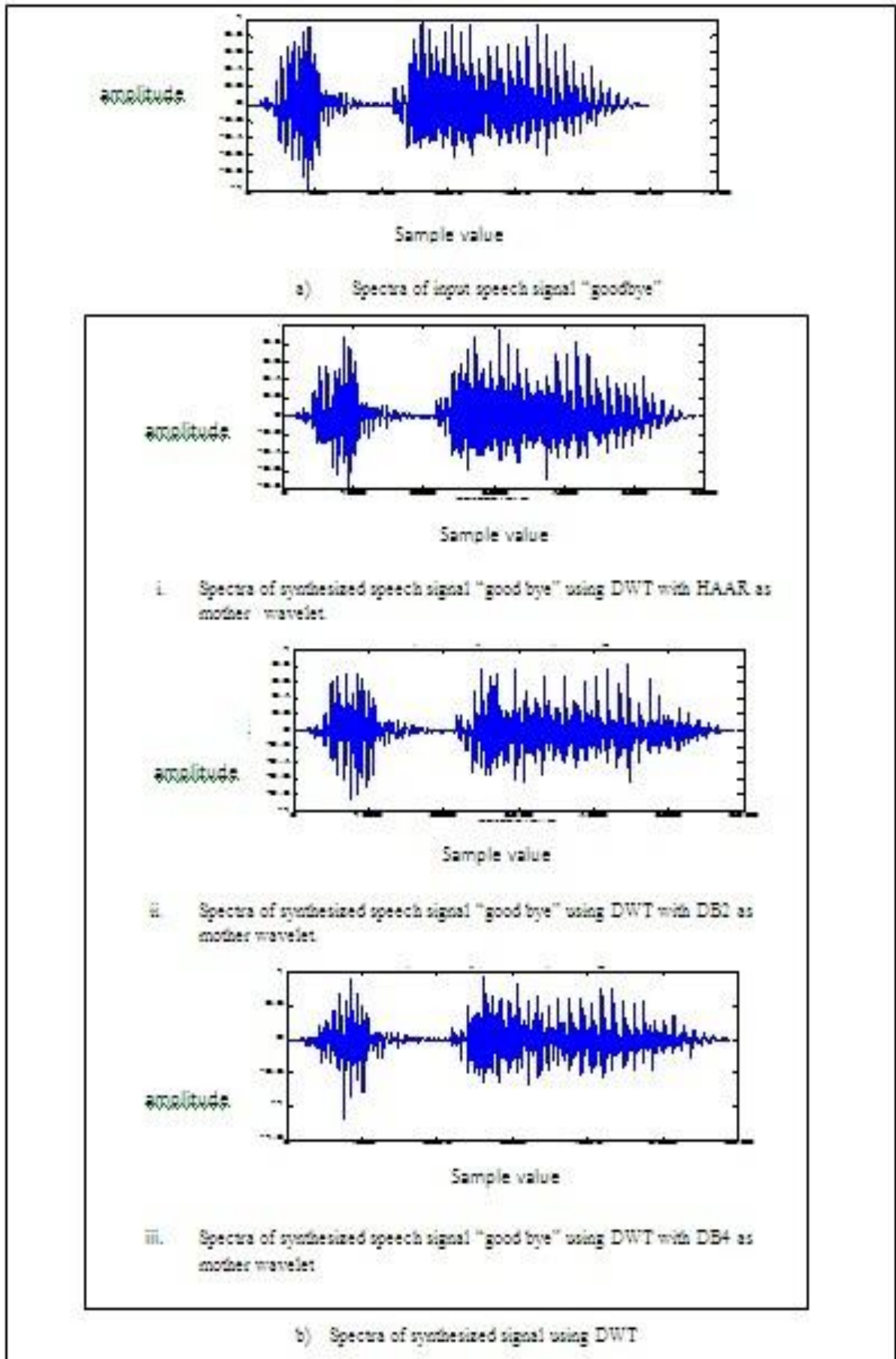


Fig.17: Comparison of speech signal “wow” On the basis of CR



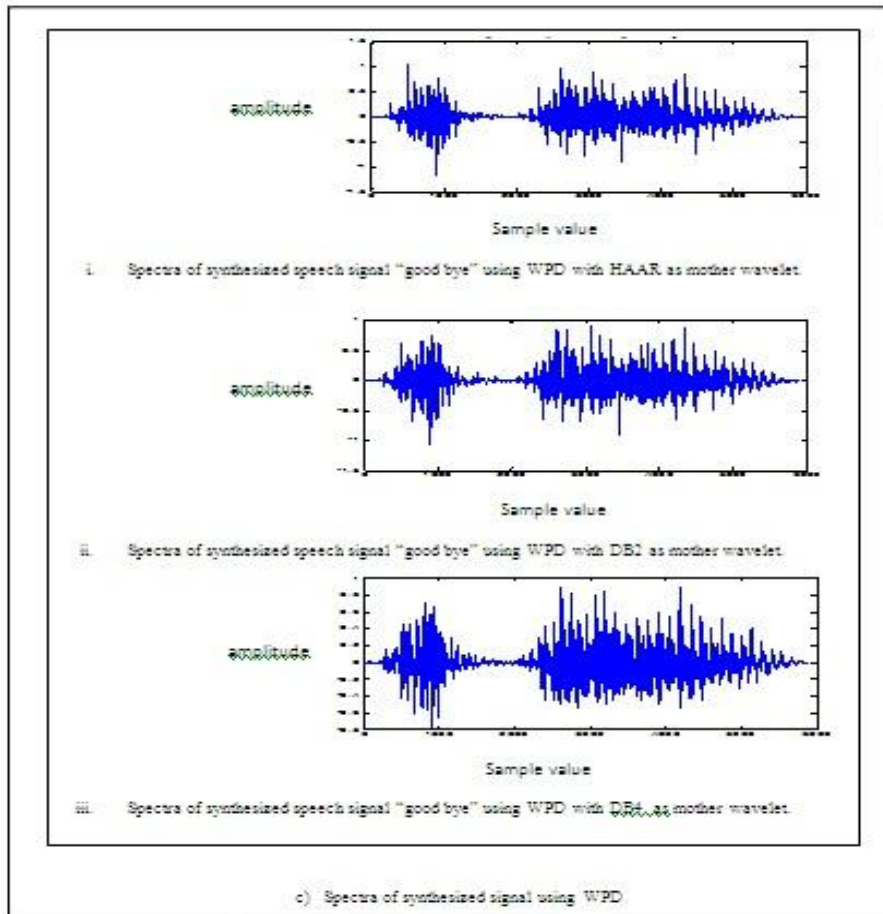
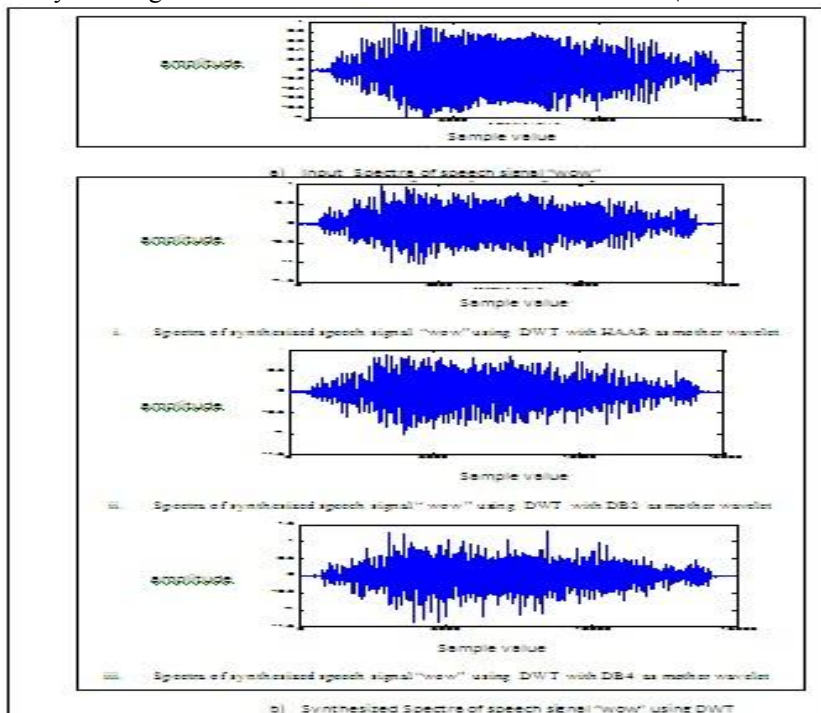


Figure 18:- Input and Synthesized spectra of speech "goodbye"

Figure 18 (a) shows the input spectra of speech signal "goodbye" and 18(b) and 18(c) shows the synthesized spectra of speech signal "good bye" using DWT and WPD with different mother wavelet. Figure 19 (a) shows the input spectra of speech signal "wow" and 19(b) and 19(c) shows the synthesized spectra of speech signal "good bye" using DWT and WPD with different mother wavelet.\



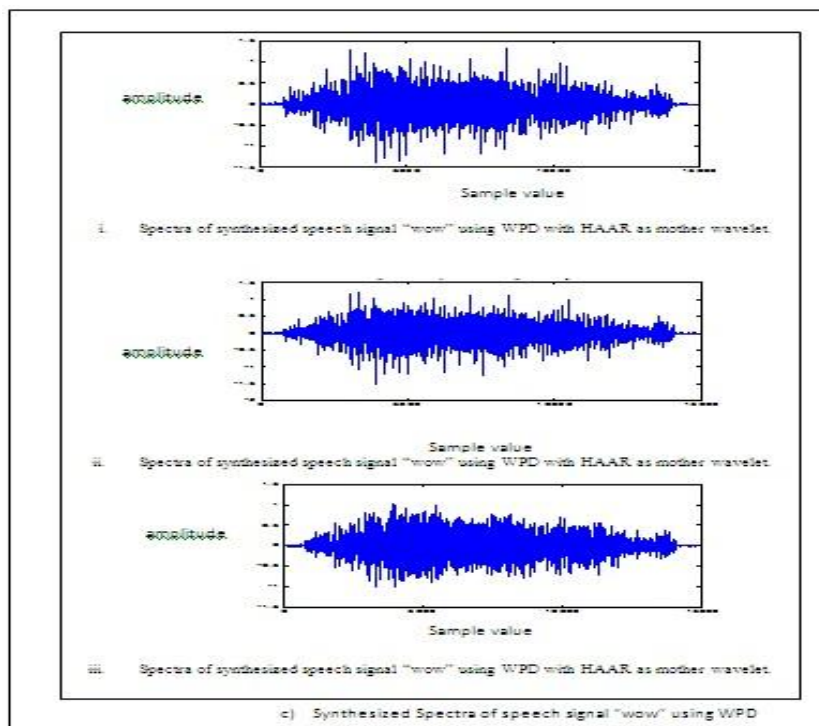


Figure 38. Input and Synthesized spectra of speech signal "wow".

V. CONCLUSION

In this paper, the performance of the, discrete wavelet transform (DWT) and wavelet packet decomposition (WPD) in compressing speech signals is tested and following points were observed. Wavelet packet decomposition gives better results than discrete wavelet transform. The results of wavelet packet decomposition for a particular mother wavelet were found to be better when compared with the results of wavelet transform. In both, DWT and WPD high compression ratios were achieved with acceptable SNR. It was observed that in DWT as we move from one family to another the Signal to Noise Ratio decreases and Compression Ratio increases as percentage of the truncated coefficients increases. And within a family the Signal to Noise Ratio increases. The reason behind this is that the number of vanishing moments increases as the order increases. Higher number of vanishing moments provides better reconstruction quality, thus better SNR value and Compression Ratio decreases. Overall global thresholding produces better results than hard thresholding in discrete wavelet transform and in WPD the results for global and hard thresholding found to be comparable.

REFERENCES

- [1] Shijo M Josepj ,and Babu Anto P "Speech Compression Using Wavelet Transform" IEEE-International conference on recent trends in information technology,ICRTIT 2011 MIT, Anna University ,Chennai,june 3-5,2011.
- [2] Jalal Karam "End Point Detection for Wavelet Based Speech Compression" International Journal of Biological and Life Sciences 4:3 2008
- [3] Dr .V.Radha , Vimala. C ,and M.Krishnaveni "Comparative Analysis of Compression Techniques for Tamil Speech Data sets" IEEE-International conference on recent trends in information technology,ICRTIT 2011 MIT, Anna University ,Chennai,june 3-5,2011.
- [4] Mahmoud A.Osman,Nasser Al, Hussein M.Magboub and S.A.Alfandi "Speech compression using LPC and wavelet" IEEE 2nd International conference on computer Engineering and Technology 2010.
- [5] Jalal Karam , and Raed Saad "The Effect of Different Compression Schemes on Speech Signals" world Academy of Science ,Engineering And Technology 18,2006.
- [6] Amara Graps "An introduction to wavelets" IEEE computational science and Engineering, summer 1995, vol 2, num 2, published by IEEE computer society,10662 Los Alamitos, CA 90720 ,USA
- [7] Ms P.M.Kavathekar/Mrs P.M.Taralkar,Prof U.L.B.ombale,Prof. P.C.Bhaskar " Speech compression using DWT in FPGA" international journal of scientific & Engineering Research ,Volume 2,Issue 12 December-2011
- [8] Olivier Rioul and Martin Vetterli " Wavelets ans signal processing" IEEE SP magazine,1991.
- [9] M.A.Anusuya and S.K.Katti " comparison of different speech feature extraction techniques with and without wavelet transform to kannada speeh recognition" International journal of computer applications(0975-8887) volume 26-No-4,july 2011.
- [10] Christian Gargour, Marcel Gabrea, Venkatanarayana Ramachandran, and Jean-Marc Lina "A short introduction to wavelets and their applications" IEEE circuits and systems magazine 2009.
- [11] Shijo. M. Joseph1, Firoz Shah A.2, and Babu Anto P.3 "Comparing Speech Compression Using Waveform Coding and Parametric Coding" International Journal of Electronics Engineering, 3 (1), 2011, pp. 35– 38.