Enhancement of Speech Compression Technique Using Wavelet Transforms With Parallel Processing and Fusion

S.Kamesh[1], V.Sailaja[2], K.Jyothi[3]

[1]M.Tech (DECS), ECE Department, Godavari Institute of Engineering & Technology, Rajahmundry, India

[2] Professor , ECE Department, Godavari Institute of Engineering & Technology, Rajahmundry, India

[3]Associate Professor, ECE Department, Godavari Institute of Engineering & Technology, Rajahmundry, India

***Abstract* — The Discrete Wavelet Transform is the most powerful and new signal compression technique which uses multi-resolution analysis for analyzing speech signal. Here we are doing parallel speech compression using two different wavelet transforms like Haar wavelets and Bi-orthogonal wavelet transformation (Bior). The resultant of components which are in high frequency bands are fusioned together and difference signal are used to reconstruct the speech which will gives the enhancement in speech signal with less loss speech data in high frequency portion.**

Keywords- Speech signal compression; Haar wavelet; Bi-orthogonal wavelet; Discrete approximation of Meyer.

I. INTRODUCTION

Since speech signal contains a large number of redundant information, how to compress high quality speech maintaining transparent quality at low bit rates is still a very important topic. To reduce redundancy and make full use of the human's auditory masking effect by using a variety of source coding techniques, not only can compress the coding rate by many times, but also has the ability to regain high intelligibility and acceptability of speech signals. Therefore, a speech compression system focuses on reducing the amount of redundant data while preserving the integrity of signals. The different transformation of speech signals to the time-frequency and time-scale domains for the purpose of compression aim at representing them with the minimum number of coding parameters. [1, 2]

Speech is non-linear random process. Wavelet transform devotes a lot to deal with time-varying, non-stationary signal, for it has excellent resolution in both time and frequency domain. Wavelet transform with detail of signal, decomposes the high-frequency, and the signal was decomposed to the time-frequency space which has a certain correspondence with critical band of speech. The result of

wavelet transform is called "wavelet coefficients". The coefficients conversion of wavelet can classify into two types,

Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT) [3, 4]. This article investigates the improvement speech compression technique based on the model wavelet transform which is focused in the frequency lossless.

There are some related researches works which used wavelet transform for speech compression. In [5], it explores the DWT as a tool for Hindi speech recognition. It studies the recognition of isolated words in Hindi Language speech. The mother wavelets are selected to use 3 families, Daubehies (db), Coiflets (coif) and Discrete Meyer Wavelet (dmey). It is found that Daubechies 10, 5-level decomposition and the Discrete Meyer wavelet give comparable performance, while the Daubechies 8, 3-level decomposition provides the poorest performance. In [6] presents reliable measures of speech compression by detecting the end points of the speech signals prior to compressing them.

Two different compression schemes used are the Global threshold and the Level method is tested with the Signal to Noise Ratios (SNR), Peak Signal to Noise Ratios (PSNR) and Normalized Root Mean Square Error (NRMSE) parameter measures. In [7] applies the Wavelet Packet Transform (WPT) to process speech signal to obtain optimal wavelet tree to allocate the dynamic bits, and then uses the modified Set Partitioning in Hierarchical Trees (SPIHT) coding algorithm to compress the coefficients from the wavelet packet transform. It indicates that it can gain better high compression. However the quality signal reconstruction is still not perfect according to loss of frequency. Thus this article is concerned to compare the frequency lossless of the new model wavelet for speech compression.

This article is organized as follows. Section II describes the Discrete Wavelet Transform while Section III details the speech compression using parallel process. Section IV shows its simulations results. Finally, Section V concludes this work.

II. DISCRETE WAVELET TRANSFORM

The DWT is the most powerful and new signal compression technique which uses multi-resolution analysis for analyzing speech signals. Function of the DWT is a frequency scale adjustments and values shifting position discrete which able to calculate coefficients as the equation (1):

$$C_{a,k} = \int_R s(x).\frac{1}{\sqrt{a}}\Psi(\frac{x-k}{a}) \qquad (1)$$

Where, Cak is the coefficient which a is scale, k is shifting. ψ (x)is the mother wavelet function, which three families wavelet have selected for speech compression. These details are followed.

A. Haar wavelet (Haar):

Haar [3] is wavelet function which applies the easiest, using the least time and it has orthogonal property. The mother wavelet of Haar ψ(x) can be described as the equation (2).

$$\Psi(x) = \begin{cases} 1 & 0 < t < 1/2 \\ -1 & \frac{1}{2} < t < 1 \\ 0 & otherwise \end{cases} \qquad (2)$$

B. Bi-orthogonal wavelet (Bior):

Bior [3] or semi-orthogonal wavelet is only orthogonal to the shifted base function under different scale factor, but has no orthogonality in the same scale factor. The mother wavelet of Bior can be described as equation (3).

$$C_{a,k} = \int_R s(x).\tilde{\Psi}_{a,k}(x)dx \qquad (3)$$

C. Discrete approximation of Meyer wavelet (Dmey):

Dmey [3] wavelet is the discrete format of meyer wavelet function. Mayer's wavelet as shown in equation (4), is fundamentally a solvent method for solving the two-scale equation. Given a basis '□' for the approximation space, Mayer employed Fourier techniques to derive the DTFT of the two-scale education coefficients.

$$G_0(e^{jw}) = \sqrt{2}\sum_k \emptyset(2\omega + 4k\pi) \qquad (4)$$

III. SPEECH COMPRESSION USING DWT

Speech compression using discrete wave transforms (DWT) is shown in steps below.

- Choice of appropriate Wavelet:

The choice of the mother wavelet plays a very important role in designing high quality speech. Choosing the appropriate wavelet will maximize the SNR and minimizes the relative error. Here we selected bior1.1 wavelet for better results.

- Decomposition Level:

Wavelets work by decomposing a signal into different frequency bands and this task is carried out by choosing the wavelet function and computing the discrete wavelet transform (DWT). Choosing a decomposition level for the DWT Usually depends on the type of signal being analyzed.

- Truncation of Coefficients:

The coefficients obtained after applying DWT on the frame concentrate energy in few neighbors. Here we are truncating all coefficients with "low" energy and retain few coefficients holding the high energy value. Two different approaches are available for calculating thresholds.

- Global Thresholding:

The aim of Global thresholding is to retain the largest absolute value coefficients. In this case we can manually set a global threshold. The coefficient values below this value should be set to zero, to achieve compression.

- Level Dependent Thresholding:

This approach consists of applying visually determined level dependent thresholds to each de composition level in the Wavelet Transform. The value of the threshold applied depends on the compression. The task is to obtain high compression and an acceptable SNR needed to reconstruct the signal and detect it. Among these two, high SNR is achieved using global thresholding compared to level dependent thresholding.

- Encoding:

Signal compression is achieved by first truncating small valued coefficients and then efficiently encoding them. Another approach to compression is to encode consecutive zero valued coefficients, with two bytes. One byte to indicate a sequence of zeros in the wavelet transforms vector and the second byte representing the number of consecutive zeros.

IV. PERFORMANCE MEASURES

- Signal to Noise Ratio (SNR):

This value gives the quality of reconstructed signal. Higher the value, the better:

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

σ_x^2 is the mean square of the speech signal and σ_e^2 is the mean square difference between the original and reconstructed signals.

- Percentage of Zero Coefficients:

It is given by the relation: 100*(No. of Zeros of the current decomposition) / (No. of coefficients).

- Normalized Root Mean Square Error:

$$NRMSE = sqrt[(x(n) - r(n))^2 / (x(n) - \mu x(n))^2]$$

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

- Peak Signal to Noise Ratio:

$$PSNR = 10 \log_{10} NX^2 / ||X-r||^2$$

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.

- Compression Score:

It is the ratio of length of the original signal to the compressed signal.

Effects of Threshold:

In this experiment, there is a need to study the effects of varying threshold value on the speech signals in terms of SNR and compression score. For bior1.1 at level 3, the threshold value was slowly increased, and the corresponding values of the SNR and Compression score were recorded in Tables 1 & 2:

Table 1: Male

Threshold Values	SNR	Compression Score
0.1	10.8	5.64
0.5	10.49	6.23
1.0	9.87	7.75
1.5	9.83	7.93
3.0	9.5	7.95

Table 2: Female

Threshold Values	SNR	Compression Score
0.1	9.03	5.53
0.5	8.59	7.74
1.0	8.57	7.99
1.5	8.55	7.99
3.0	8.50	7.99

From the above tables, we can observe that after some particular threshold value, the Signal to Noise Ratio and Compression Factor do not change, because at this point all the detail coefficients are truncated to zero and remains only the approximation coefficients.

For a Speech Signal using Haar and Bior1.1 with threshold value 0.1, the corresponding waveforms are given below:

1. Male:

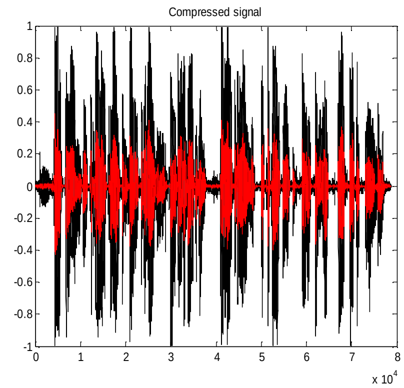

Fig. (a) Haar wavelet output

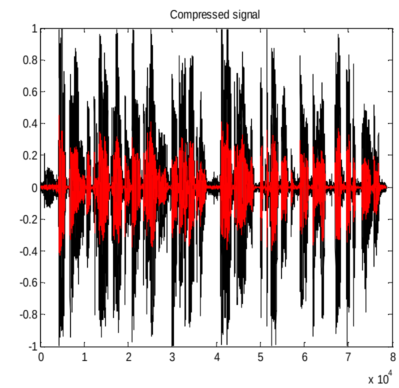

Fig. (b) Bior1.1 wavelet output

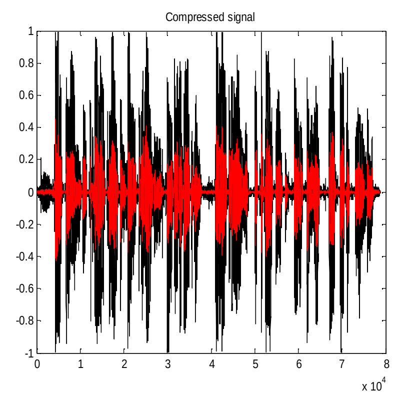

Fig. (c) Fusioned Signal output

- The performance measures are:
  Signal to Noise Ratio = 10.8034
  Peak Signal to Noise Ratio = 19.4911
  Normalized RMS Error = 0.4637
  Percentage of Zero Coefficients=85.35%

2. Female:

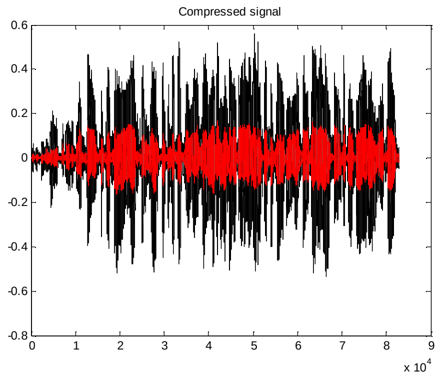

Fig. (a) Haar wavelet output

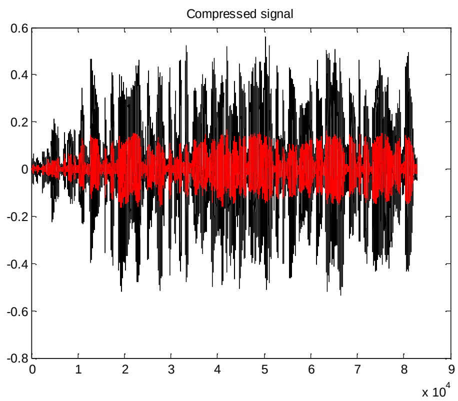

Fig. (b) Bior1.1 wavelet output

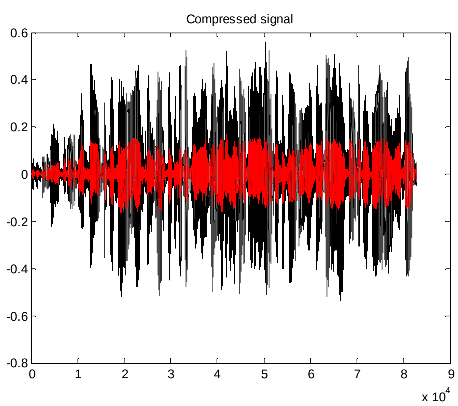

Fig. (c) Fusioned Signal output

- The performance measures are:
  Signal to Noise Ratio = 9.0323
  Peak Signal to Noise Ratio = 12.8232
  Normalized RMS Error = 0.5444
  Percentage of Zero Coefficients=85.27%

V. CONCLUSION

This article presents the frequency lossless comparison of the modern wavelet based on speech compression technique. The Haar wavelet (Haar), Bi-orthogonal wavelet (Bior) and discrete approximation of Meyer wavelet (Dmey), are investigated to test with different speech signals. The experiments show that the Bior provides the best performance in time and frequency domain when is compared to the all wavelet. This technique can also gain the compression rate as 2.67 to the original speech signal.

However Haar and Dmey do not provide the good performance because Haar and Dmey are unable to include a value changes immediately. Also the conversion coefficients of Haar and Dmey wavelet functions cannot be separate the data between low frequency and high frequency from each other effectively. Thus it wasn't suitable for the application on a compression and the noise removal.

REFERENCES

[1] S. M. Joseph, "Spoken digit compression using wavelet packet," International conference on signal and image processing, PP: 255-259, 2010.

[2] M. A. Osman and N. Al, "Speech compression using LPC and wavelet," 2nd International conference on computer engineering and technology, PP: V7-92 – V7-99, 2010.

[3] J. Y. Zhang and L. X. Gao, "Research on the selection of wavelet function for the feature extraction of shock fult in the bearing diagnosis," International conference on wavelet analysis and pattern recognition, PP: 1630-1634, 2007.

[4] R.S.Cai and Y.M.Guo, "Wavelet based multi feature V/UV speech classification algorithm," International conference wireless mobile and sensor networks, PP: 897 – 900, 2007.

[5] S. Ranjan. "Exploring the DWT as a tool for Hindi speech recognition," International journal of computer theory and engineering, PP: 642-646, 2010.

[6] Z. Dan and M. S. Qian, "Sppeech compression with best wavelet package transform and SPIHT algorithm," 2nd International conference on computer modeling and simulation, PP. 360-363, 2010.

[7] Kinsner and A.Langi, "Speech and Image Signal Compression with wavelets", IEEE Wescanex conference proceedings, IEEE, New York, NY-1993, PP. 368-375.

[8] Kornsing.S.; Srinonchat.J., "Enhancement Speech Compression Technique Using Modern Wavelet Transforms," Computer, Consumer and Control (IS3C), 2012 International Symposium on , vol., no., pp.393,396, 4-6 June 2012.

[9] E.B.Fgee, W.J.Philips, W.Robertson, "Comparing Audio Compression using Wavelets with other Audio Compression Schemes", IEEE Canadian Conference on Electrical and Computer Engineering, IEEE, Edmonton, Canada, 1999, pp.698-701.