Performance analysis of Compression Schemes on Audio Signals

B.Satish Chandra[1], Dr.T.Sheela[2]

1. Satish Chandra. B
Research Scholar, Vinayaka
Mission Research Foundation
University, Salem, India
2. Dr T.Sheela, Assoc professor,dept of ECE ,VMKVEC ,Salem

Abstract— The new technique of this paper studies the result of compression constraints and schemes bestowed in a very new and versatile paradigm to attain high compression ratios and acceptable signal to noise ratios of audio signals. Compression parameters are computed for variable frame sizes of level five to seven separate riffle rework (DWT) illustration of the signals for various analyzing mother wavelet functions. Results are obtained and compared for world threshold and level dependent threshold techniques. The results obtained conjointly embrace comparisons with Signal to Noise Ratios, Peak Signal to Noise Ratios and Normalized Root Mean sq. Error.

Keywords—Audio Compression, DWT, Wavelets, Normalized Root Mean **sq.** Error.

I. Introduction

APPLICATIONS to audio compression involve real time cryptography of audio for mobile satellite communication, cellular phones, and audio for videophones or video group discussion system. alternative applications involve vocoding of audio signals for storage, synthesis and transmission [8]. The DWT of a given audio signal concentrates audio energy in few neighboring coefficients permitting natural compression. during this paper we have a tendency to introduce a versatile compression theme that uses wavelets and their transforms. the pliability of this new paradigm is earned by observing:

i. The analyzing ripple used
ii. Decomposition level
iii. Compression ratios
iv. Frame size
v. Measured parameters
vi. style of threshold used

In this paper, a versatile paradigm represented in Fig. five is introduced to compress audio signals. The signal is initial divided it into completely different size frames, that are then analyzed victimisation specific mother wavelets up to tier seven illustration using DWT.

The digits are the focus of compression, namely. Different compression parameters are calculated for these signals and compression ratios are derived for to different types of compression schemes, namely, the Level Dependent and Global Threshold techniques. The following section introduces wavelets and two of their related transforms, namely, the Continuous Wavelets Transform (CWT) and the DWT. While audio compression is discussed in Section 3 with some details, the implementation of the system and compression parameters used in this work are included in Section 4. The last two sections give a detailed discussion on the results obtained, and the conclusion of this paper.

II. Wavelets

A wavelet is a finite energy signal defined over specific interval of time [1]. The main interest in wavelets is their ability to represent a given signal at different. Wavelets are used to analyze signals in much the same way as complex exponentials (sine and cosine) used in Fourier analysis of signals. Unlike Fourier, wavelets can be used to analyze non- stationary, time-varying, or transient signals [9] [10]. This is an important aspect, since audio signals are considered to be non-stationary. A given signal is represented by using translated and scaled versions of a mother wavelet as it is explained below. They are also localized in time and frequency domains [9].

A. Continuous Wavelet Transform

The wavelet transform is a two parameter expansion of a signal in terms of a particular wavelet basis function [1]. Given $\Psi(t)$ called the mother wavelet; all other baby wavelets are obtained by simple scaling and translation of $\Psi(t)$. $\Psi a,t(t) = (1/\sqrt{a})\Psi[(t-b)/a]$.

Where a and b are the scaling and the translation parameter respectively. A nice approach to the CWT representation is first to inspect the Fourier transform represented mathematically by:

$$F(\omega) = \int s(t)e^{-j\omega t}dt$$

Replacing the complex exponential in the Fourier transform with $\Psi a,t(t)$ yields:

$$C(S,U) = \int \sqrt{a}\, \Psi[(t-b)/a]dt$$

In other words with wavelet transform, reference to frequency is replaced by reference to scale [8], [9][10].

B. Discrete Wavelet Transform

In our application the discrete wavelet transform is applied. By choosing scale and position based on power of 2, CWT is reduced to DWT without any loss in energy. The scaling parameter is discrete and dyadic, $a = 2^{-j}$. The translation is discretized with respect to each scale by using $\tau = k2^{-j}T$ [5]. $\Psi_{j,k}(t) = (2^{j/2})\ \Psi[(2^{j}t-kT)]$.

The integer k represents the translation of the wavelet function; it indicates time in wavelet transform. Integer j, however, is an indication of the wavelet frequency or spectrum shift and generally referred to as scale. The DWT transforms a discrete input signal vector into two sets of coefficients the approximation CA containing low frequency information and the detail coefficients CD containing high frequency information. Fig. 1 shows a level 2 DWT decomposition of an input signal s (t) [8] [9] [10].

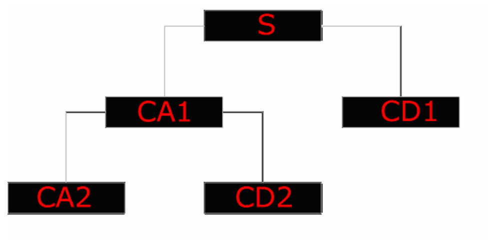

Fig. 1 Level 2 Decomposition

most ordinarily used moving ridges are classified into two classes: orthogonal and bi-orthogonal wavelet system. Orthogonal wavelets decompose signals into well behaved orthogonal signal areas. Biorthogonal moving ridges are a lot of sophisticated and are outlined supported a try of scaling and wavelet perform. The moving ridge of interest, the one utilized in this work, is that the Daubechies moving ridge family, during which do terribly distinctive compression properties, supposed for moving ridge coefficients.

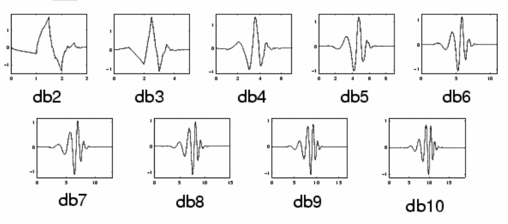

Fig. 2 Plots of Different Daubechies Orthogonal Wavelets

Fig. 2 shows several of the Daubechies orthogonal wavelets. One notes that db10 could be a abundant drum sander perform than the remainder. normally for the Daubechies orthogonal moving ridges dbi; as i increase the wavelet becomes drum sander. during this paper, the analyzing functions chosen were db4, db8, db10 and db20. The DWT may be computed victimisation octave band filter bank [6] [9]. The signal is split into 2 segments via a two-band filter bank, an occasional pass or lower resolution version, and a high pass one. The lower resolution version is then split once more, and so on. this can be illustrated in Fig. 1 and 3. The high pass filter power unit generates high frequency coefficients containing low energy; these are the detail coefficients of the signal indicated as CDi. Low pass filter phonograph record generates the approximation coefficients; selected by CA. Those coefficients contain most of the energy within the audio signal [10]. For multi resolution analysis CA are rotten A level more into detail and approximation coefficients. The output of the power unit filter is down sampled and fed into a noticeor to detect all coefficients below a specific threshold and replace them by a zero. The down sampling can retain N/2 of the signal constant those that are solely required. To reconstruct the signal we tend to apply the Inverse separate moving ridge rework (IDWT) illustrated in Fig. 4.

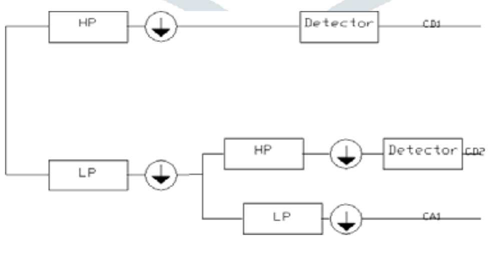

Fig. 3 DWT Decomposition (Analysis)

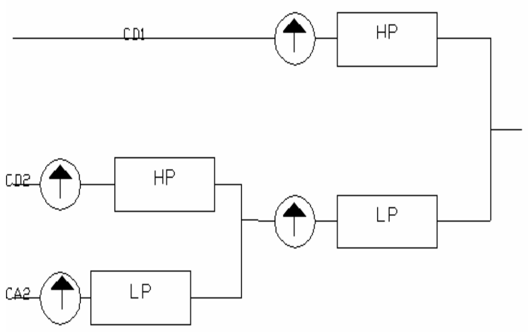

Fig. 4 IDWT Reconstruction (Synthesis)

III. Wavelet Compression

In many applications, audio signals are both saved for later use or transmitted over a few media. In each instances, one is interested in decreasing the dimensions of the sign due to the cost, time and other advantages. the quest arises for a mechanism to compress the sign and gain lower bit quotes. special techniques were implemented each having its blessings and drawbacks. In fashionable, they may be categorized into two types, lossy and lossless. The Hufmann coding, as an instance, is considered to be lossless compression. The unique records will be completely restored without any changes. Lossy compression does now not completely retain the original signal; consequently some of the information is misplaced. For audio alerts, this loss is suitable for the reason that we're fascinated only in spotting the sign. Wavelets compression approach is taken into consideration to be lossy where the reconstructed signal isn't an exact healthy of the original sign. one of the most essential blessings of wavelet rework is that it concentrates audio formation (strength and notion) into a few neighboring coefficients [3]. also whilst making use of the DWT to a given audio sign many coefficients of small values (relying on level we pick out) are accordingly taken into consideration insignificant. The retained coefficients will nonetheless have the bigger percentage of strength in the sign. The method of compressing the digit indicators is mentioned within the subsequent phase which additionally contains the extraordinary mechanisms to be used all through the technique.

A. Data Base

The signals are chosen from the digit audio data base. The digits were spoken by different speakers and recorded in the studio at the Lebanese American University, Byblos. The small size studio is designed to minimize noise and equipped with a multi directional microphone made by Neumann to collect the audio signal with the best quality. The signals are then recorded and transformed into wave sounds (.wav). The tools used are the PRO-Tools Control 24 Dig-Design Device. This device is a computerized digital mixer.

B. 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 the discrete signal. Thus we can apply the transform at any of these levels. But in fact, the decomposition level depends on the type of signal being analyzed. For the processing of audio signals, decomposition up to scale 7 is adequate [7]. In this paper, level 5 DWT is obtained for every signal and comparisons were made with level 6 and 7 decompositions.

C. Choosing Appropriate Wavelets

The form of wavelet is of high importance for such experiments. It immediately impacts the signal to Noise Ratio (SNR) of the output signal. choosing the precise wavelet will maximize the SNR and minimizes the relative mistakes. As mentioned in advance Daubechies wavelets have proper compression assets for wavelet coefficients [1], giving higher SNR ratios. Wavelets with more vanishing moments provide better reconstruction great. Daubechies wavelets are developed with most regularity; the number of zero moments is maximized, leading to the high-quality wavelet circle of relatives for compression. the chosen members of this orthogonal Daubechies own family are db4, db8, db10 and db20.

D. Choosing the Frame Size

Dividing the audio signal in different frame sizes is used to examine their effect on the overall compression performance, since framing aims to improve the compression ratios effect. Framing aim to improve the compression ratios obtained. Three frame sizes are tested in this paper (20ms, 0.25s, and 0.5s). The frames obtained are analyzed separately being considered a vector in its own right.

E. Threshold Techniques

The coefficients obtained after applying DWT on the frame concentrate energy in few neighbors. Thus we can truncate all coefficients with “low” energy and retain few coefficients holding the high energy value. The two thresholding techniques are implemented according to the following algorithms [3].

1. Global Threshold

The global threshold technique works by retaining the wavelet transform coefficients which have the largest absolute value. For a given audio signal, the global threshold algorithm first divides the audio signal into frames of equal size **F**. Then the wavelet transform of each frame is computed. Usually with length **T** > **F**. These coefficients are sorted in an ascending order and the first **L** coefficients are retained. In practice, these coefficients along with their positions in the wavelet transform vectors are stored or transmitted [3][5]. For these reasons, 2.5***L** coefficients are used to represent the original **F** samples distributed as follows: 8 bits for the amplitude and 12 bits for the position leading to 2.5 bytes. The Compression Ratio **CR**, can then be defined by: **CR**=**F** / 2.5***L.**

Each frame is reconstructed by replacing the missing coefficients by zeros.

2. Level Dependent Threshold

The level-dependent threshold technique is derived from the Birge-Massart strategy [5]. This strategy works on selecting the retained wavelet coefficients as follows. Let J_0 be the decomposition level, m the length of the coarsest approximation coefficients over 2, and α be a real greater that

1. At level J_0+1 (and coarser levels), everything is kept. For level J from 1 to J_0, the K_J larger coefficients in absolute value are kept using this formula:

$$K_J = \frac{m}{(J_0 + 1 - J)^{\alpha}}$$

The value of α used is 1.5 as suggested in [5].

The value of the threshold applied depends on the compression ratio we want to achieve. The task is to obtain higher compression ratios and an acceptable SNR needed to reconstruct the signal and detect it. The signal is reconstructed by applying the Inverse Discrete Wavelet Transform IDWT as it is shown in Fig. 5.

IV. Design and Implementations

The introduced device is depicted in Fig. five and implemented and simulated to examine its overall performance the use of Matlab®. a number of the functions used are, Wavdec, which computes the multi-degree decomposition of the signal and wavrec that reconstructs the sign from the coefficients acquired. two other essential capabilities are: wdencmp returns the coefficients after applying a determined threshold. It also computes the share of electricity retained and the proportion of truncated zeroes.

A. The Compression Parameters

In this paper, four compression parameters are used. They are defined next along with their mathematical expressions.

- **Signal to Noise Ratio (SNR)**

  $SNR = 10*\log(\sigma_x\text{^}2 / \sigma_e\text{^}2)$

  Where $\sigma_x\text{^}2$ is the mean square of the audio signal and $\sigma_e\text{^}2$ is the mean square difference between the original and reconstructed signals.

- **Peak Signal to Noise Ratio**

  $PSNR = 10*\log(NX^2 / ||x-r||^2)$

  Where 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 original and reconstructed signals.

- **Normalized Root Mean Square Error**

  $NRMSE = \text{sqrt}[(x(n)-r(n))^2/(x(n)-\mu_x(n))^2]$

  Where X(n) is the audio signal, r(n) is the reconstructed signal, and $\mu_x(n)$ in the mean of the audio signal.

- **Retained Signal Energy**

  $RSE = 100*||x(n)||^2 / ||r(n)||^2$

Where ||x(n) || is the norm of the original signal and ||r(n) || is the norm of the reconstructed one. For db orthogonal wavelets the retained energy is equal to the L^2-norm recovery performance.

The audio signals compressed are the digits “Zero’ and “eight”. They are depicted in Fig. 6 and Fig. 7 respectively along with their compressed versions. These compressed versions were obtained using “db8” and the size of frame is 0.02s. Different Compression Ratios (CR) were obtained. The compressed audio signal is still audible and you can still recognize the output signal. Different parameters were examined when simulating the code. The 8 KHz sampled signals are divided into frames (0.2ms, 0.25s, and 0.5s) and decomposed up to level 5. Each frame is decomposed separately. At this stage the threshold is applied on the coefficients to truncate whatever unnecessary. The obtained coefficients are then used to reconstruct the output compressed signal. Different results were obtained allowing efficient evaluations and comparisons of the used methods and parameters.

V. DISCUSSION

Average SNR, PSNR, and NRMSE are all measured given the frame size, the type of the mother wavelet ψ(t) and value of the threshold. Also, the percentage of Zeros (%Z) and the percentage of the Energy Retained (%ER) are included. Another set of experiment is done on the Digit zero. In these set of experiments, the threshold does not apply on the approximation coefficients. We are able to disregard more than 92% of the signal coefficients and still retain about 76 % of the energy in the signal using a frame size of 0.02ms and applying the global threshold. That is, keeping track of 8% of the frame to be reconstructed later. Furthermore, an increase in the length of the frame to 0.25s achieved better results. Less coefficients containing more energy and the SNR is maximized while the NRMSE is minimized. Changing the wavelet also has its effect on the compression ratio. If we use a smooth analyzing wavelet like db10 the percentage of the truncated coefficients decreased. However; they produced better SNR. Thus, these wavelets play the role of producing better SNR but less compressing ratios. On the other hand, un- smooth wavelets such as db4 lead to better compression ratios but resulting with low SNR.

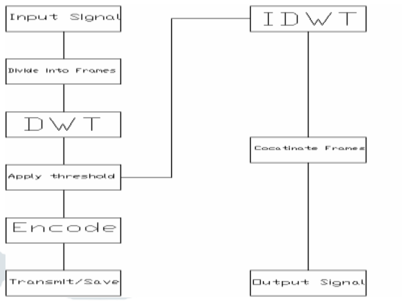

Fig. 5 The Introduced Flexible Paradigm

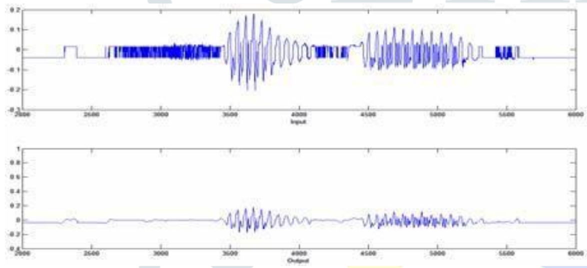

Fig. 6 Digit zero with compressed version. The CR = 7.65

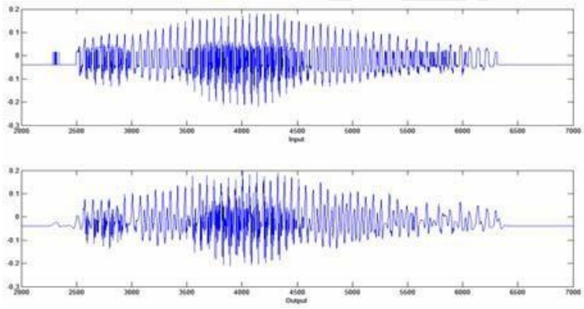

Fig. 7 Digit Eight with compressed version. The CR = 5.5

VI. CONCLUSION

In this paper, the performance of the separate moving ridge rework in press audio signals is tested and therefore the following points were discovered. High compression ratios were achieved with acceptable SNR. No more enhancements were achieved on the far side level five decomposition. The result of frame size and therefore the Level Dependent Threshold on the NRMSE is obvious whereas this measure remains nearly constant for all experiments with negligible changes. Increasing the frame size, absolutely affects the general performance in each threshold techniques used. Overall world threshold ends up in higher results than the amount dependent threshold technique within the case of SNR and Cr. This was the case with and while not framing and for each tested digits. it's worthy noting that we tend to couldn't pinpoint the most effective compression moving ridge.

Shri BOLISHETTI SATISH CHANDRA Ph.D research scholar from India and working as Associate professor ,in the Department of Electronics and communication , Mallareddy institute of Engineering and technology,Hyd. He studied B.Tech in Electronics and Communications engineering from Nagarjuna University, Guntur, AP and M.Tech in Systems and signal processing from JNTU college of engineering,JNTUH,Hyd,Telangana State .He is having 17 + years of work experience in Academics ,teaching and research. He participated and presented research papers in both national and international conferance,seminars and work shops also published 7 research papers in national and international peer review journals.

Dr.T.SHEELA obtained her Doctorate degree in VLSI Design in 2015 from Vinayaka Missions Research Foundation, Salem and Master degree from Anna University in 2008. She is an Associate Professor in Department of Electronics and Communication Engineering in VMKVEC, Salem. She is a member of IEEE and IEEE-AESS, and also a Life member of ISTE. She has published several papers in National and International Conferences and Journals. Her research interests include VLSI circuits, Digital circuit, Electronic circuits and optical communication.

REFERENCES

[1] Y. T. Chan “Wavelet Basics”, Kluwer Academic Publishers, 1995.

[2] DR. V.KRISHNANAIK “ Research Methods and Presentation” Lap labard publication, Germany in 2014

[3] C. Taswell, “Audio Compression with Cosine and Wavelet Packet Near-Best Bases”, ICASSP-96. Conference Proceedings, IEEE International Conference on Acoustics, Audio and Signal Processing, Vol.1, pp: 566 – 568, 1996.

[4] E. Fgee, W. J. Phillips, W.Robertson “Comparing Audio Compression Using Wavelets With Other Audio Compression Schemes”, IEEE, CCECE Conference Edmonton, Alberta, Canada. May 9-12, pp: 698- 701, 1999.

[5] N. M. Hosny, S. H. El-Ramly, M. H. El-Said, “Novel Techniques for Audio Compression Using Wavelet Transform”, ICM '99. 11th International Conference on Microelectronics, pp: 225 – 229, 1999.

[6] http://www.mathworks.com/academia MATLAB Wavelet Toolbox 2.

[7] O. Rioul and M. Vetterli, “Wavelets and Signal Processing”, IEEE Signal Process. Mag. Vol 8, pp. 14-38, Oct. 1991.

[8] J. I. Agbinya, “Discrete Wavelet Transform Techniques in Audio Processing,” IEEE Digital Signal Processing Applications Proceedings, IEEE, New York, pp: 514 – 519, 1996.

[9] J. F. Koegel Buford, “Multimedia Systems”, ACM Press, 1994.

[10] G. Strang and T. Nguyen, “Wavelets and Filter Banks”, Wellesley- Cambridge Press, 1996.

[11] J. S. Walker, “Wavelets and their Scientific Applications”, Chapman and Hall/CRC, 1999.