IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 17, Issue 6, Ver. V (Nov – Dec. 2015), PP 79-85
www.iosrjournals.org

Data Compression using Multiple Transformation Techniques for Audio Applications.

Arashpreet Kaur[1], Rajesh Mehra[2]
[1](M.E Scholar in Electronics and Communication Engineering, National Institute for technical teachers Training and Research, India)
[2](Associate Professor in Electronics and Communication Engineering, National Institute for Technical Teachers Training and Research, India)

Abstract: *As Multimedia Technology is growing day by day over past decades, therefore demand for digital information increasing rapidly. This digital information contains multimedia files like image files, audio files that require a large space so no other option than compression. In Compression high input stream of data is converted into small size. Data Compression for audio purposes is a field of digital signal processing that focuses on reducing bit-rate of audio signals to enhance transmission speed and storage capacity of fast developing multimedia by removing large amount of unnecessary duplicate data. The advantages of the compression technique are reduction in storage space, bandwidth, transmission power and energy. This paper is based on transform technology for compression of the audio signal. In this methodology, different transforms such as Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) are used. Mean compression ratio is calculated for DCT & DWT. Performance measures like peak signal-to-noise ratio (PSNR), signal-to-noise ratio (SNR) & normalized root mean square error (NRMSE) are calculated and then compared.*
Keywords: *Compression Ratio, DCT, DWT, NRMSE, PSNR, SNR.*

I. Introduction

Data compression is a technique in which data content of the input signal to system is compressed so that original signal is obtained as output and unwanted or undesired signals are removed [1]. Audio is an electrical representation of sound within the range of human hearing that specifically lies between 20 Hz to 20 kHz range of frequency which is detectable by the human ear. [2]The concept of Audio Compression is to compress the data in the form of audio so that it occupies less space for storing it. The need for audio compression is to accommodate more data in the available storage area so that the storage capacity can be enhanced. Due to less storage space occupancy, large amounts of data can be placed in the available memory. Therefore less storage room for information inhabitance, large amount of information can be transmitted with less transmission capacity[4] That implies compressed audio signal can be transmitted over the web with less transmission bandwidth at higher speeds. As speed is increased, audio files can be transferred and downloaded over the web faster with higher bit rates. Because of quick downloading and transferring of audio files, time delay is minimized. [1]

Signals compression is based on removing the redundancy between adjacent samples and/or between the adjacent cycles. In data compression, it is desired to represent data by as small as possible number of coefficients within an acceptable loss of visual quality.
Compression techniques has two main categories: lossless and lossy.
Compression methods can be classified into three functional categories:

- Direct Methods: The samples of the signal are directly handled to provide compression.
- Transformation Methods: such as Fourier Transform (FT), Wavelet Transform (WT), and Discrete Cosine Transform (DCT).
- Parameter Extraction Methods: A preprocessor is employed to extract some features that are later used to reconstruct the signal.[2],[5]

In this paper audio compression is carried out in two levels. In the first level a transform function (technique) like Discrete cosine transform, discrete wavelet transform are applied on audio signal which gives a result with a new set of data with smaller values. By applying transform technique compression ratio for each transform technique is obtained on different audio samples. Parameters like Signal to noise ratio (SNR), mean square error (MSE) are measured for the reconstructed audio obtained from DCT, WT these transform techniques. Second level is encoding. This step will present data in minimal form by using these encoding techniques. Compression ratios are also calculated.

II. Transformation Techniques

1. Discrete cosine transform

A discrete cosine transform (DCT) communicates a limited succession of information focuses as far as cosine functions wavering at distinctive frequencies[1] DCTs are critical to various applications in science and designing, from lossy compression of sound (e.g. MP3) and pictures (e.g. JPEG) (where little high-recurrence segments can be discarded), to spectral methods for the numerical arrangement of partial differential equations. The utilization of cosine instead of sine functions is critical for compression, since it turns out (as portrayed beneath) that less cosine functions are needed to approximate a typical signal, while for differential equations the cosines express a particular choice of boundary conditions. The DCT is similar to the discrete Fourier transform: it changes a sign or picture from the spatial area to the recurrence space.

Discrete cosine transform, linearly transforms information into frequency domain, so that the information can be represented by a set of coefficients. The advantage of discrete cosine transform is that, the vitality of the genuine information may be moved in just a couple low recurrence parts of DCT depending on the correlation present in the data. Equations (1) represents the D.C.T of 1-dimentional arrangement of length N.

$$\boldsymbol{y(k) = w(k)\sum_{n=1}^{N} x(n)\cos\left(\frac{\pi(2n-1)(k-1)}{2N}\right) \qquad k = 1, 2 \ldots\ldots N} \qquad \textbf{(1)}$$

Where

$$\boldsymbol{w(k) = \begin{cases} \frac{1}{\sqrt{N}} & k = 1 \\ \sqrt{\frac{2}{N}} & 2 \le k \le N \end{cases}} \qquad \textbf{(2)}$$

N is the length of x, and x and y are the same size. If x is a matrix, DCT transforms its columns. The series is indexed from n = 1 and k = 1 instead of the usual n = 0 and k = 0 because MATLAB vectors run from 1 to N instead of from 0 to N- 1. It often reconstructs a sequence very accurately from only a few DCT coefficients, a useful property for applications requiring data reduction.[12]

Wavelets: Wavelet is a part of waveform, whose average energy value approximately tends to zero. Generally, wavelet is a varying window length considered as a part of waveform. Wavelet transform are based on small wavelets with limited duration.

Equation (3) and (4) shows mathematical representation of wavelet [6]. Ψ(t) is a function called mother wavelet.

$$\int_{-\infty}^{\infty} |\Psi(\mathbf{t})|^{\boldsymbol{n}}\, \boldsymbol{dt} < \infty \qquad \textbf{(3)}$$

$$\int_{-\infty}^{\infty} |\Psi(\mathbf{t})|\, \boldsymbol{dt} < \infty \qquad \textbf{(4)}$$

2. Discrete Wavelet Transform

Jean Morlet introduced the idea of wavelet transform in 1982 and provided mathematical tool for seismic wave analysis [1],[4].

A discrete wavelet transform is characterized as a "small wave" that has its energy concentrated in time to give an apparatus to the examination of transient, non-stationary, or time-varying phenomena.[8] It has the oscillating wave like properties additionally can permit concurrent time and recurrence investigation. Wavelet Transform has risen as an effective mathematical tool in numerous regions of science and engineering, although in the field of sound and data compression. [1]

Wavelet transform decomposes a signal into a set of basic functions [12]. These basis functions are called “wavelets”.

$$\Psi_{\boldsymbol{a,b}}(\boldsymbol{t}) = \frac{\mathbf{1}}{\sqrt{\boldsymbol{a}}}\Psi\left(\frac{\mathbf{t-b}}{\mathbf{a}}\right) \qquad \textbf{(5)}$$

Where ‘a’ is the scaling parameter and ‘b’ is the shifting parameter. And parameter ‘b’ is translation parameter and it determines the time location of the wavelet. DWT decompose signal in to several n levels in different frequency bands. At each step there are two outputs, scaling and wavelet coefficients. The equations (6) and (7) represent the scaling and wavelet coefficients respectively.

$$x^{j+1}(n) = \sum_{i=1}^{2n} h(2n-i)x^j(n) \qquad (6)$$

$$y^{j+1}(n) = \sum_{i=1}^{2n} g(2n-i)x^j(n) \qquad (7)$$

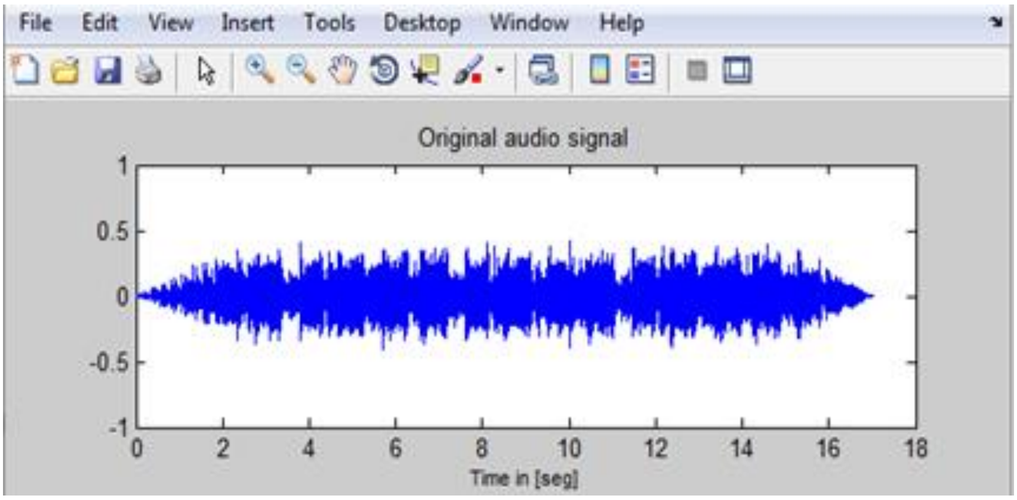
Wavelet transforms convert a signal into series of wavelets and they give an approach to breaking down waveforms in both frequency and time duration. Wavelet transform are a numerical intends to perform signal investigation when signal frequency varies with time [14], [15]. The Wavelet transform gives the time-frequency representation. That is, wavelet transforms give time and frequency information in the meantime, thus giving a time frequency representation of the signal [13]. The wavelet transform is processed independently for distinctive sections of the time-domain signal at different frequencies. Multi-resolution analysis: investigates the signal at different frequencies giving different resolutions. . MRA is intended to give good time resolution & poor frequency resolution at high frequencies and good frequency resolution & poor time resolution at low frequencies. Wavelet transforms are good for signal having high frequency components for short durations and low frequency components for long duration .e.g. images, video frames and speech signal.

III. Implementation Using Matlab

MATLAB could be a great tool that is employed to investigate audio signals which are read in wav format. Following commands are used for analyzing the audio signal

- wavread: it reads audio signal
- windowsize: defines window operate of transformation
- wavplay: it produces speech signal once transformation
- length: defines length of speech to be processed by
- reworking principle dct: performs distinct trigonometric function rework
- idct: performs inverse distinct trigonometric function rework[5],[7]

This is shown in figure (3) , that is design flow of DCT coder. After getting the reconstructed signal , the audio is further processed for calculating the peak signal-to-noise ratio (PSNR), signal-to-noise ratio (SNR) and normalised root mean square error (NRMSE). These values comparison is shown in table 1.
Here figure (1) is the original audio signal which is to be transformed. Figure (2) is the design flow for the DWT



.**Fig 1:** audio signal

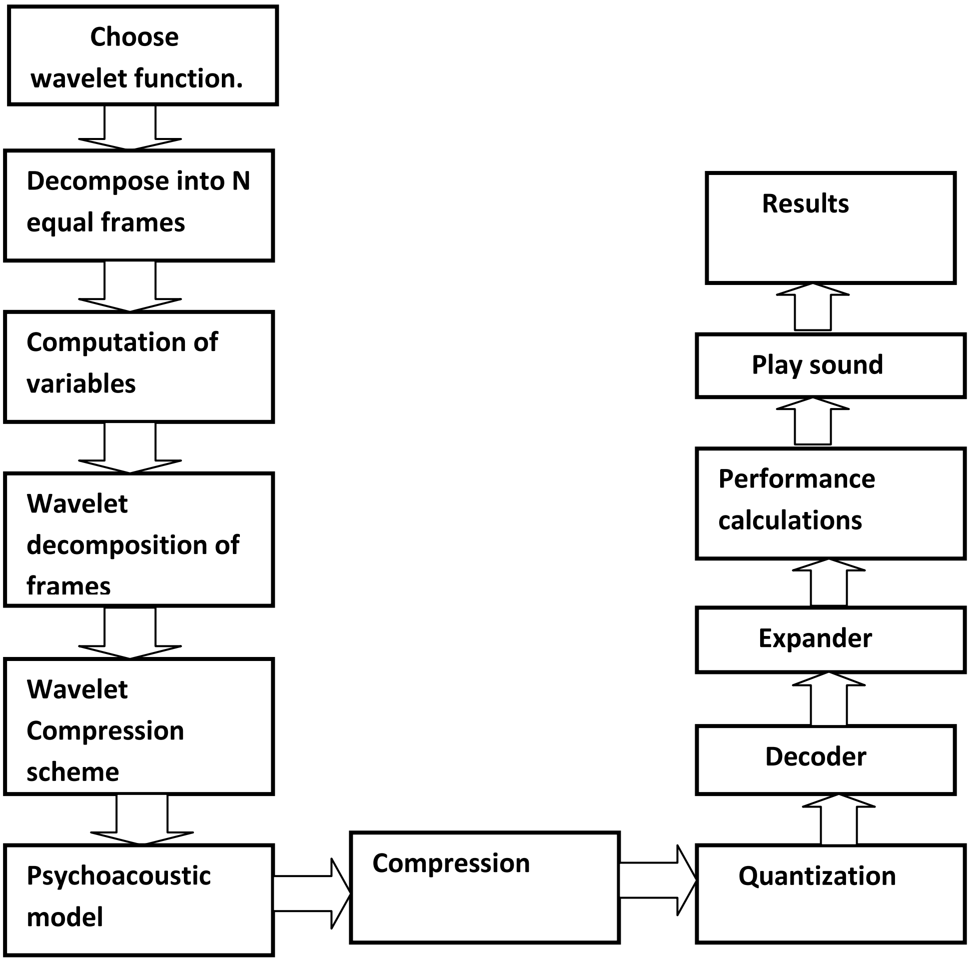


Fig 2 : Design flow of wavelet based audio coder

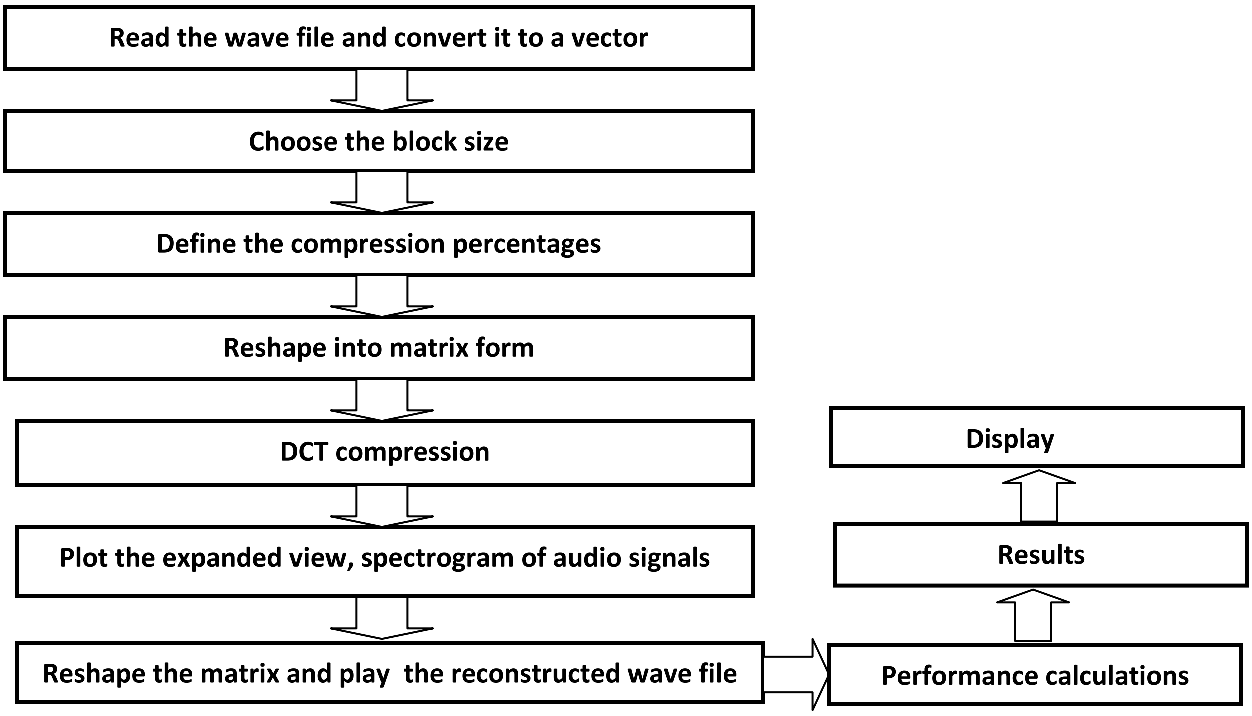


Fig 3: Design flow of DCT audio coder

IV. Output In Matlab

The output for the different waveforms is shown below. Figure (4) shows the audio signals in different compression ratios that is audio with different amplitudes. Figure(5) shows the portion of audio signal to be processed. Figure(6) represents the audio spectrograms while in Figure(7) compressed audio output using DWT is shown.

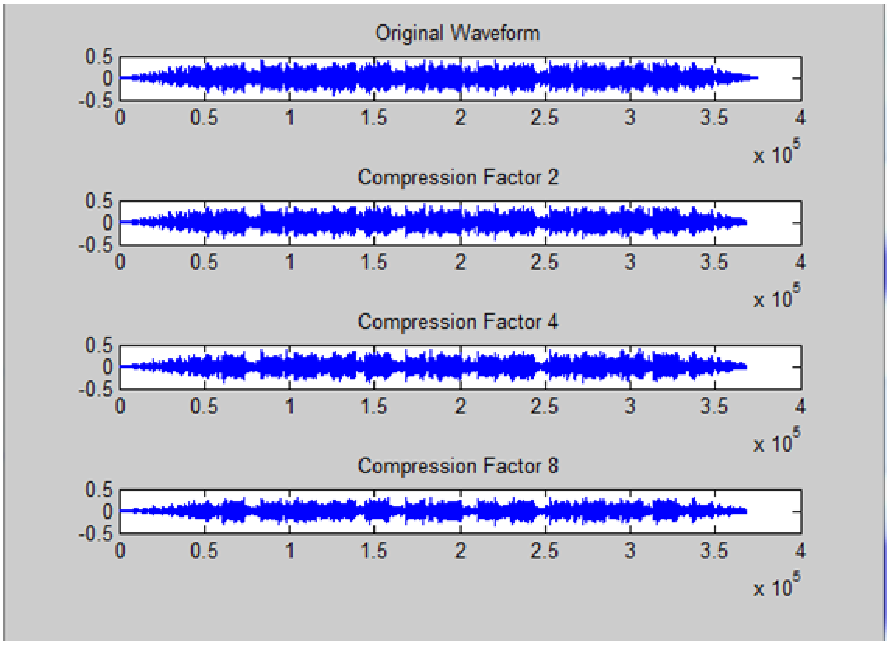


Fig 4: Audio signals with different amplitudes

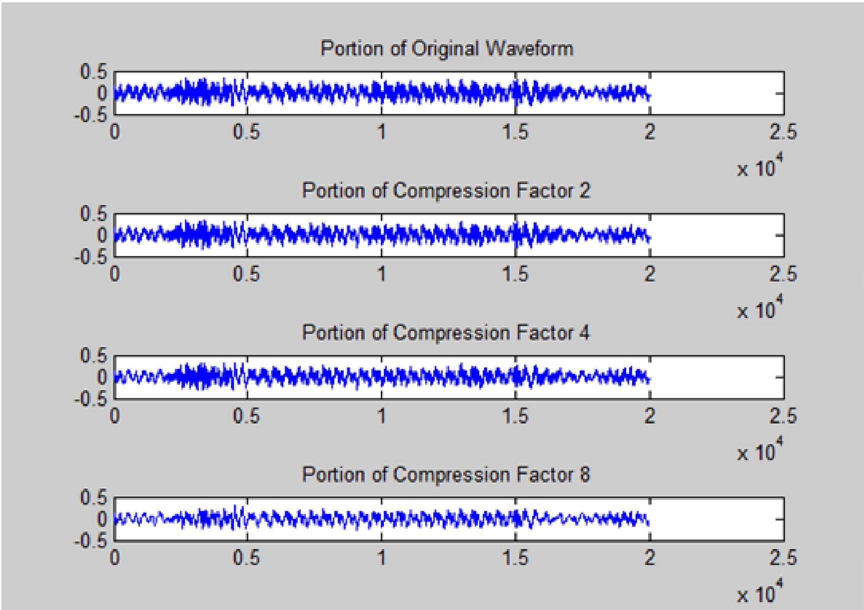


Fig 5: Portion of speech signal to be processed

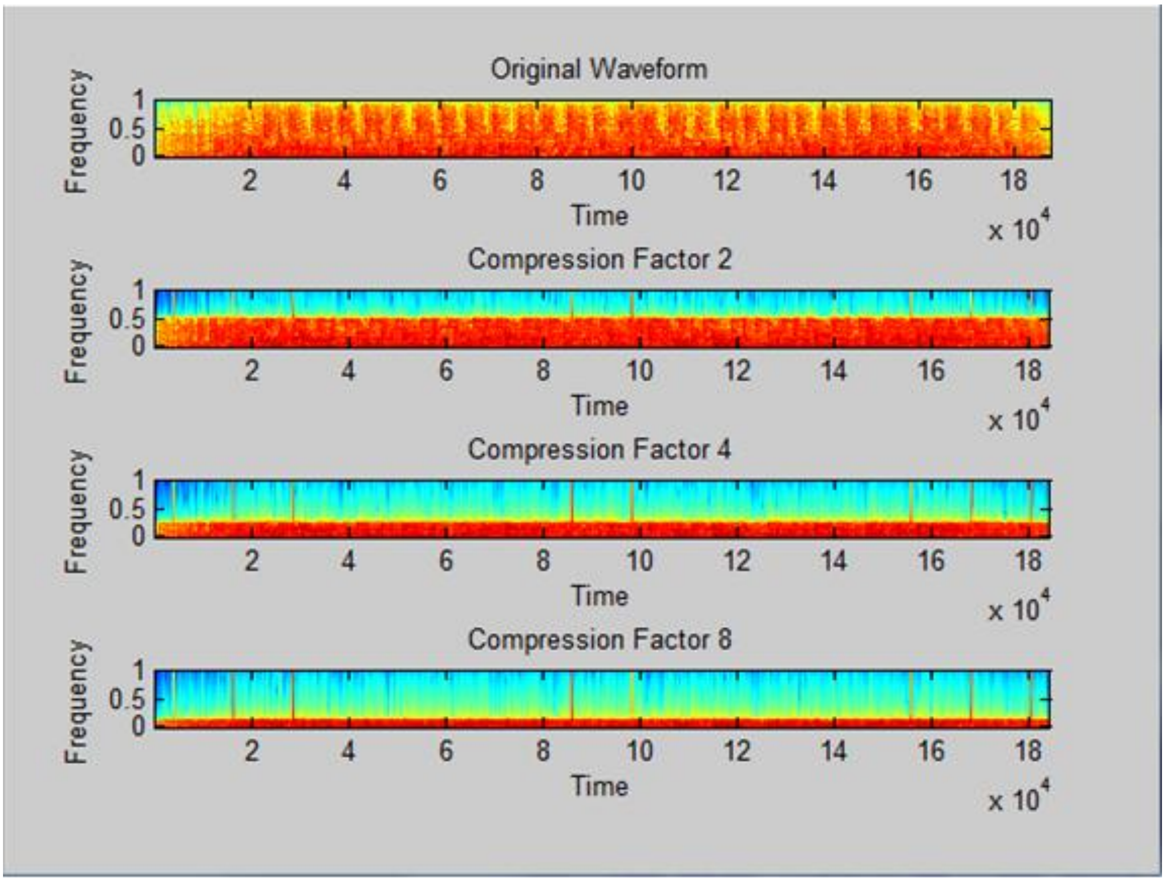


Fig 6: Audio Spectograms

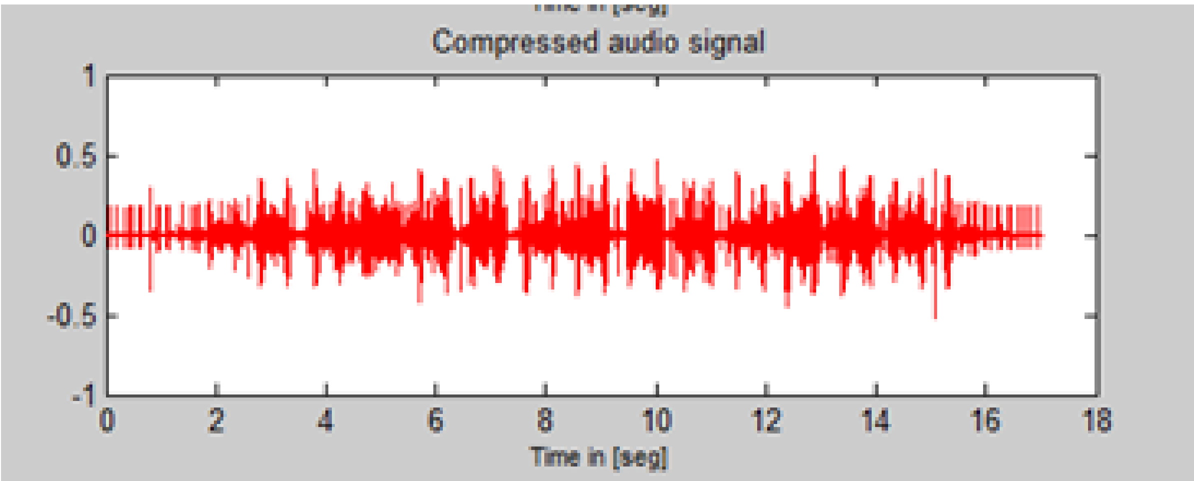


Fig 7: Compression by DWT

V. Performance Measures

For this audio compression method, based on transform techniques, the performance is measured in terms of Compression ratio, SNR, MSE.

1.Compression ratio: The definition for Compression ratio (C.R) is defined as ratio between Length of original signal and Length of compressed signal

$$C.R = \frac{\text{original audio length}}{\text{compressed audio length}}$$

2.Signal to noise ratio: It is a measure that compare the level of desired signal to the level of background noise.

$$\mathbf{SNR = 10\log_{10}\left[\left(\sigma_s^2 \middle| \sigma_e^2\right)\right]^2} \quad \textbf{(8)}$$

σ_s^2 is the mean square of audio signal and σ_e^2 is mean square difference between original and reconstructed audio signal.

3.Normalized Root Mean Square Error (NRMSE):

$$NRMSE = \sqrt{\frac{\sum_n (x(n) - x'(n))^2}{\sum_n (x(n) - u_x(n))^2}} \quad \textbf{(9)}$$

Here, X(n) is the speech signal, x"(n) is reconstructed speech signal and μ x(n) is the mean of speech signal.

4.Peak Signal to Noise Ratio (PSNR):

$$PSNR = 10\log_{10}\frac{NX^2}{\|x - x'\|^2} \quad \textbf{(10)}$$

Where N is the length of reconstructed signal, X is the maximum absolute square value of signal x and ||x-x`||2 is the energy of the difference between the original and reconstructed signal.

Signal	CR	MSE	SNR(db)	PSNR(db)
funky	0.2639	0.02990	31.83	45.21

Table 1: Results of DCT based technique in terms of CF, SNR, PSNR & MSE

Signal	CR	MSE	SNR(db)	PSNR(db)
funky	0.0587	0.08	21.02	36.24

Table 2: Results of DWT based technology in terms of CF,SNR,PSNR & MSE

VI. Conclusion

In this paper a simple DWT & DCT based audio compression schemes are presented. These data compression is done by using MATLAB CODING. From the results shown above it is clear that DWT gives less compression ratio in comparison to DCT, while MSE for DCT is less. SNR and PSNR for DWT is less in comparison to DCT.DWT is better than DCT for audio compression.

Here audio is compressed in different factors in case of DCT by 2 , 4 and 8.

Acknowledgements

I am very thankful to my college and guide Dr. Rajesh Mehra for providing time to time help in studying the topic and providing me a background to understand it in deeper details.

References:

[1] M. Arora, N. Maurya, P. Pathak, and V. Singh, "SPEECH COMPRESSION ANALYSIS USING MATLAB," pp. 538–542, 2014.

[2] M. V Patil, A. Gupta, A. Varma, and S. Salil, "AUDIO AND SPEECH COMPRESSION USING DCT AND DWT TECHNIQUES," vol. 2, no. 5, pp. 1712–1719, 2013

[3] J. S. Jacaba, "AUDIO COMPRESSION USING MODIFIED DISCRETE COSINE TRANSFORM : THE MP3," no. October, 2001.

[4] H. S. Malvar, "LOSSLESS AND NEAR-LOSSLESS AUDIO COMPRESSION USING INTEGER- REVERSIBLE MODULATED LAPPED TRANSFORMS."

[5] H. kaur and R. kaur ,D. Cosine, "SPEECH COMPRESSION AND DECOMPRESSION USING DWT AND DCT," vol. 3, no. August, pp. 1501–1503, 2012

[6] M. Arora and N. Maurya, "AUDIO COMPRESSION IN MPEG TECHNOLOGY PEG IS AN ACRONYM FOR MOVING PICTURE EXERTS," vol. 3, no. 12, pp. 1–4, 2013.

[7] M.Siffuzzaman, M.R.Islam, and M.Z. Ali, "APPLICATIONS OF WAVELET TRANSFORM AND ITS ADVANTAGES COMPARED TO FOURIER TRANSFORM, by Journal of Physical Sciences, Vol. 13, pp:121-134, October-2009.

[8] Z. T. Drweesh and L. E. George, "AUDIO COMPRESSION BASED ON DISCRETE COSINE TRANSFORM , RUN LENGTH AND HIGH ORDER SHIFT ENCODING," vol. 4, no. 1, pp. 45–51, 2014.

[9] S. M. Ali, " INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN SPEEDING UP AUDIO FRACTAL COMPRESSION ," vol. 3, no. 6, pp. 86–92, 2013

[10] K.P. Soman, K.I. Ramachandran, "INSIGHT TO WAVELETS" second edition 2005, by Prentice Hall of India. ISBN- 81-203-2902-3

[11] Raghuveer M. Rao, Ajit S. Bopadikar, "WAVELET TRANSFORMS – INTRODUCTION TO THEORY AND APPLICATIONS". Pearson Education Asia. 1998 Pearson education, Inc. ISBN:81-7808-251-9.

[12] J. James and V. J. Thomas, "AUDIO COMPRESSION USING DCT AND DWT TECHNIQUES," vol. 4, no. 4, pp. 119–125, 2014.