Virtual Audio Processing Lab using GUI (Graphical User Interface)

Mohd Khairol Fitza Bin Mohamed Daud
Faculty of Electrical Engineering
Universiti Teknologi MARA Malaysia
40450 Shah Alam, Selangor, Malaysia
e-mail: krol_8705@yahoo.com

Abstract— **Recently, the wavelet transform has emerged as a cutting edge technology, within the field of audio compression. This paper studies the different effect of audio compression due to the different audio file, different wavelet families (Haar, Symlet and Daubechies) and different level of decomposition. An audio processing lab is implemented in Graphical User Interface (GUI) using MATLAB software. The audio compression coder is used to compress the signal. The results obtained were analyzed based on Signal to Noise Ratios (SNR) and Compression Ratios.**

Keywords – Wavelet Compression, GUI, SNR

I. INTRODUCTION

There is a wide range of applications for wavelet transforms. They are applied in different fields ranging from signal processing to biometrics, communication, broadcasting and medical image. One of the prominent applications is in the FBI fingerprint compression standard. Wavelet transforms are used to compress the fingerprint pictures for storage in their data bank. The previously chosen Discrete Cosine Transform (DCT) did not perform well at high compression ratios. It produced severe blocking effects which made it impossible to follow the ridge lines in the fingerprints after reconstruction. This did not happen with wavelet transform due to its property of retaining the details present in the data.

In Discrete Wavelet Transforms (DWT), the most prominent information in the signal appears in high amplitudes and the less prominent information appears in very low amplitudes. Data compression can be achieved by discarding these low amplitudes. The wavelet transforms enables high compression ratios with good quality of reconstruction. At present, the application of wavelets for image compression is one the hottest areas of research. Recently, the wavelet transforms have been chosen for the JPEG 2000 compression standard.

This paper has five sections. Section 1 is Introduction. This section will explain what this paper all about and the application of wavelet compression scheme. Basically, this paper is about doing research on audio compressing technique. Section 2 is about wavelet. This section explains more about wavelet and DWT. Section 3 is methodology. This section explains the steps in audio compression scheme, the coding used in the MATLAB software and the parameter that used in this project. Section 4 is result and discussion. This section shown the result obtained the experiment and discuss about it. Section 5 is the conclusion for all the experiments done in the project.

II. WAVELET

This technique called wavelet technique. Wavelet means 'small wave'. As its name, this technique is about analyzing signal with short duration and finite energy function. They transform the signal under investigation into another representation which presents the signal in more useful form. This transform called 'Wavelet Transform'. Unlike Fourier Transform, they have a variety of wavelet used for signal analysis.

A wavelet is a finite energy signal defined over specific interval of time. The main interest in wavelets is their ability to represent a given signal at different [1]. 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 [2] [3]. This is an important aspect, since speech 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

A. Discrete Wavelet Transform (DWT)

The Wavelet Transform (WT) is a technique for analyzing signals. It was developed as an alternative to the short time Fourier Transform (STFT) to overcome problems related to its frequency and time resolution properties. More specifically, unlike the STFT that provides uniform time resolution for all frequencies the DWT provides high time resolution and low frequency resolution for high frequencies and high frequency resolution and low time resolution for low frequencies [2][3]. In that respect it is similar to the human ear which exhibits similar time-frequency resolution characteristics. DWT is a special case of the wavelet transform that provides a compact representation of a signal in time and frequency that can be computed efficiently [4]. The DWT is defined by the following equation:

$$W(j,k) = \sum_j \sum_k x(k) 2^{-j/2} \psi(2^{-j} n - k) \quad (1)$$

where ψ (t) is a time function with finite energy and fast decay called the mother wavelet. In the pyramidal algorithm the signal is analyzed at different frequency bands with different resolution by decomposing the signal into a coarse approximation and detail information. The coarse approximation is then further decomposed using the same wavelet decomposition step [3]. This is achieved by successive highpass and lowpass filtering of the time domain signal and is defined by the following equations:

$$y_{high}[k] = \sum_n x[n] g[2k - n] \quad (2)$$

$$y_{low}\ [k] = \sum_n x[n] h[2k - n] \quad (3)$$

here y [k] *high*, y [k] *low* are the outputs of the highpass (g) and lowpass (h) filters, respectively after subsampling by 2. Because of the downsampling the number of resulting wavelet coefficients is exactly the same as the number of input points.

Most commonly used wavelets are categorized into two classes: orthogonal and bi-orthogonal wavelet system. Orthogonal wavelets decompose signals into well behaved orthogonal signal spaces. The coefficients of orthogonal filters are real numbers. Biorthogonal wavelets are more complicated and are defined based on a pair of scaling and wavelet function [4] [5].

The filters are of the same length and are not symmetric. The low pass filter, G_0 and the high pass filter, H_0 are related to each other by:

$$H_0(z) = z^{-N} G_0(-z^{-1}) \quad (4)$$

The two filters are alternated flip of each other. The alternating flip automatically gives double-shift orthogonality between the lowpass and highpass filters [1], i.e., the scalar product of the filters, for a shift by two is zero. i.e., ΣG[k] H [k-2l] = 0, where k,l€Z [4]. Also, for perfect reconstruction, the synthesis filters are identical to the analysis filters except for a time reversal. Orthogonal filters offer a high number of vanishing moments. This property is useful in many signal and image processing applications. They have regular structure which leads to easy implementation and scalable architecture.

B. Choice of Wavelet

The different families make trade-offs between how compactly the basis functions are localized in space and how smooth they are. The choice of the mother-wavelet function used in designing high quality coders is of prime importance. [1] Choice of wavelet has compact support in both time and frequency in addition to a significant number of vanishing moments is essential for an optimum wavelet compressor [2].

Several different criteria can be used in selecting an optimal wavelet function. The objective is to minimize reconstructed error variance and maximize SNR. In general optimum wavelets can be selected based on the energy conservation properties in the approximation part of the wavelet coefficients.

Wavelets with more vanishing moments provide better reconstruction quality, as they introduce less distortion into the processed signal and concentrate more signal energy in a few neighboring coefficients. [4] However the computational complexity of the DWT increases with the number of vanishing moments and hence for real time applications it is not practical to use wavelets with an arbitrarily high number of vanishing moments[3][4]. In this paper, the analyzing function chosen were Haar, Daubechies and Symlet wavelet family.

C. Different Level of Decomposition

Compress on wavelet compression scheme gave an effect to the level of decomposition to the compression ratio and SNR. The audio signal is split into two segments via a two-band filter bank, a low pass or lower resolution version, and a high pass one. The lower resolution version is then split again for a different level of decomposition. This is illustrated in Fig. 1 and 2.

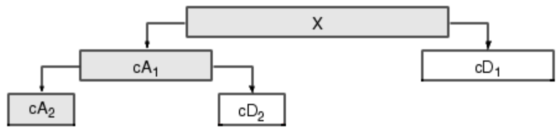

Fig. 1 Level 2 Decomposition

The highpass filter generates high frequency coefficients containing low energy; these are the detail coefficients of the signal indicated as cD_1. Lowpass filter generates the approximation coefficients; designated by cA_1. Those coefficients contain most of the energy in the audio signal. For multi resolution analysis cA_1 are decomposed a level further into detail, (cA_2) and approximation coefficients, (cD_2). The output of the highpass filter is down sampled and fed into a detector to detect all coefficients below with certain threshold and replace them by a zero. The down sampling will retain N/2 of the signal coefficient the ones that are only needed [5] [6]. N is referring to the present number of coefficient for particular level.

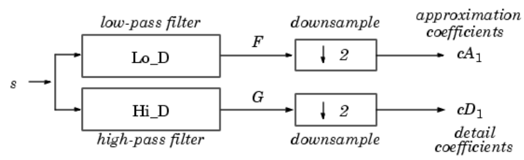

Fig. 2 Decomposition of one level

III. METHODOLOGY

A simple coder is designed to compress an audio signal based on wavelet compression scheme. The implementation algorithm used MATLAB and extends the program on GUI to analyze different type of wavelet and level of decomposition [7] [8]. The compression technique in MATLAB is described as Fig. 3:

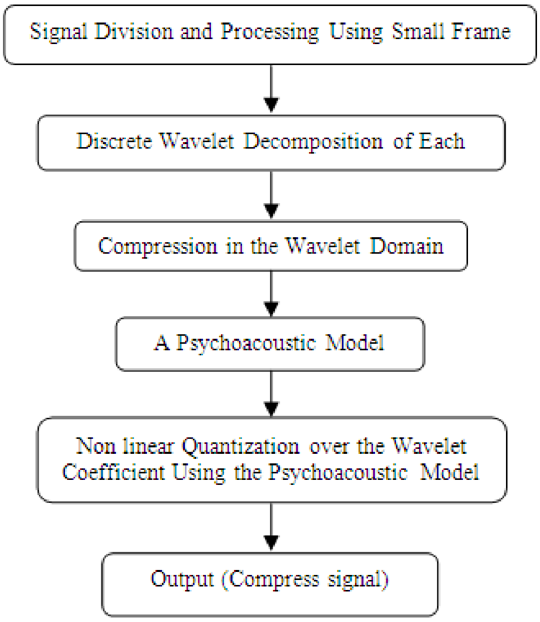

Fig.3 Wavelet Compression Coder

The psychoacoustic model is key component in the encoder. The function model is used to analyze the spectral content of the input audio signal and thereby compute signal to noise ratio for each subband in each three layers. This information coefficient in turns, used by quantizer-coder to decide how to apportion the available number of bits for the quantization of subband signals. These dynamic allocations of bits are performed as well as to minimize the audibility of quantization noise. Finally frame-packing unit assembles the quantized audio samples into decodable bit stream [3] [4].

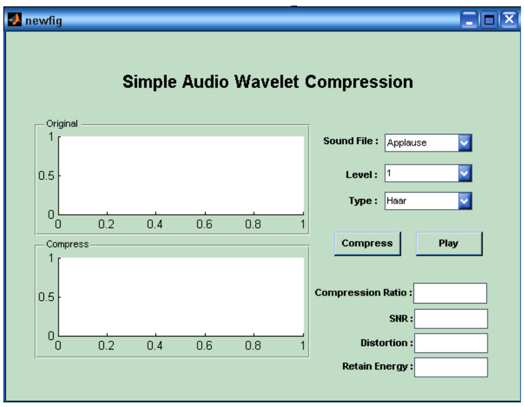

Fig. 4 MATLAB's GUI Layout

There are few coding commands were used in this project according to the designed coder. Wavelet toolbox in MATLAB software provides the simpler command in order to implement this simple wavelet audio compression. Several commands used for this project are [12] [13]:

i. *wevedec:*
   - Perform the multilevel one-dimensional wavelet analysis using specific wavelet.

ii. *ddencmp*
   - Give the default value for all procedure related to de-noising and compression using specific wavelet.

iii. *wdencmp*
   - Perform the de-noising or compression using specific wavelet.

iv. *compand*
   - Implement a µ-law compression for the input vector.

v. *quantiz*
   - Produce quantization index and quantize output value.

vi. *waverec*
   - Perform multilevel one-dimentional wavelet reconstruct using specific wavelet.

In this project, there are four parameter used, they are defined next along with their mathematical expressions [4] [9].

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

  SNR = 10*log (x^2/ (xc-x') ^2)

  Where x^2 is the mean square of the speech signal and (xc-x') ^2 is the mean square difference between the original and reconstructed signals.

- *Retained Signal Energy (RSE)*

  In the MATLAB coding, retain energy level gets from *PERFL2mean.*

- *Compression Score (CS)*

  CS = no. of coefficient / frame size

- *Compress Ratio (C)*

  C = length (x (n))/ length (xc)

  xc is the length of the compressed wavelet transform vector.

IV. RESULT AND DISCUSSION

The choice decomposition level and wavelet is important to obtained optimal result. Table 1, 2 and 3 has shown the analysis of level 1, level 3 and level 5 of decomposition. For every level of analysis, SNR, distortion and retained energy has been recorded.

Table 1 Level 1 Decomposition

Audio File	Type of Wavelet	SNR	Distortion	Retained energy
Applause	Haar	19.8680	2.899e-5	99.71
	Daubechies	20.7967	3.287e-5	99.70
	Symlet	20.4772	3.309e-5	99.71
Boing	Haar	19.4555	1.264e-4	99.79
	Daubechies	20.2280	1.063e-4	99.82
	Symlet	19.5184	1.267e-4	99.81
Camera	Haar	18.4243	1.939e-6	99.32
	Daubechies	16.4254	2.346e-6	99.26
	Symlet	16.9741	2.302e-6	99.26
Funky	Haar	18.8607	3.396e-5	99.62
	Daubechies	18.8476	3.385e-5	99.60
	Symlet	18.7040	3.396e-5	99.59

From the table 1, it shown for every type of mother wavelet used, it will give a certain SNR. It is because of every type of wavelet, has its own characteristic such as the vanishing moment and a different scale and position after decomposed from original signal.

For the retained energy, level 1 decomposed will give the high retained energy compare to the retained energy for the higher level of decomposition. The higher the level of decomposition, more energy will be deducted because of for every level; the energy will be filtered to compress the signal. Retain energy important because of each energy brings along with the information data. If we filter too many energy, it also will reduce our quality of output signal. The objective is to compress the data, so energy need to be filtering but not too much until bother the quality of the output.

Table 2 Level 3 Decomposition

Audio File	Type of Wavelet	SNR	Distortion	Retained energy
Applause	Haar	18.3752	3.975e-5	99.60
	Daubechies	19.0229	4.115e-5	99.61
	Symlet	18.9225	4.104e-5	99.62
Boing	Haar	18.3446	1.347e-4	99.73
	Daubechies	17.9794	1.640e-4	99.74
	Symlet	17.8028	1.742e-4	99.72
Camera	Haar	16.8522	2.589e-6	98.96
	Daubechies	15.8051	2.948e-6	98.94
	Symlet	15.9667	2.927e-6	98.91
Funky	Haar	16.9810	4.214e-5	99.43
	Daubechies	17.0763	4.271e-5	99.43
	Symlet	16.8920	4.454e-5	99.42

In signal processing, distortion needs to be minimized as much as possible. It cause signal disturbing and cause the error in reconstruct signal. Reduction of errors increases the reconstruction signal quality. From the table above, it's shown that distortion in the wavelet audio coding is very small.

Table 3 Level 5 Decomposition

Audio File	Type of Wavelet	SNR	Distortion	Retained energy
Applause	Haar	18.3752	4.235e-5	99.58
	Daubechies	19.0229	4.326e-5	99.60
	Symlet	18.9225	4.304e-5	99.60
Boing	Haar	18.3446	1.671e-4	99.71
	Daubechies	17.9794	1.638e-4	99.72
	Symlet	17.8028	1.856e-4	99.71
Camera	Haar	16.8522	2.762e-6	98.89
	Daubechies	15.8051	3.116e-6	98.95
	Symlet	15.9667	3.100e-6	98.92
Funky	Haar	16.9810	3.737e-5	99.41
	Daubechies	17.0763	3.976e-5	99.46
	Symlet	16.8920	3.967e-5	99.44

For the lower level of decomposition, Haar has shows the better performance in SNR, and retained energy. But the goal of compression is to maximize the compression of signal, so the decomposition of higher level is recommended.

In order to get an optimum compression, the maximum level of decomposition recommended is level 5. Decomposition beyond level 5 will not give any gain or advantage except its only give worth output because for every decomposition, energy of output will deduct and if continue decompose, it will also cut the information of the signal. As a result, output produced obviously worth compare to original signal.

Table 4 Frame Size vs. Number of Coefficient

Audio File	Frame Size	Number of Coefficient		Compression Score	Compression Ratio
Haar	2048	Level 1	2048	1.0000	1
		Level 3	2048	1.0000	1
		Level 5	2048	1.0000	1
Daubechies	2048	Level 1	2066	1.0080	1
		Level 3	2103	1.0268	1
		Level 5	2140	1.0449	1
Symlet	2048	Level 1	2062	1.0068	1
		Level 3	2092	1.0215	1
		Level 5	2121	1.0356	1

Table 4 shows the related between frame size and number of coefficient for every type of wavelet. Consider here, the analysis of audio file 'boing.wave'. It shown here, by using the complex wavelet as the transform coefficient, it has expended the number of coefficient. This means, the ratio of the coefficient represent the signal become larger if used the complex wavelet as the transform coefficient.

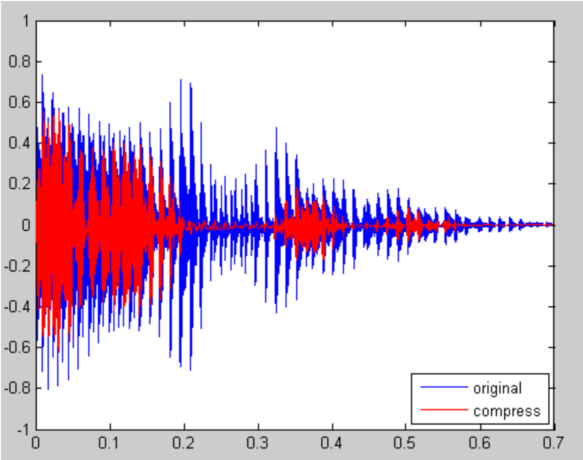

Fig. 5 Different between Original and Compressed Signal

Fig.5 has shown the different of original and compressed audio signal. More of the energy is reduce along with the deduction of signal output power and gave the minimum power require to deliver the information to human hearing scheme because human unable to detect the different between original signal and reconstructed signal after compression

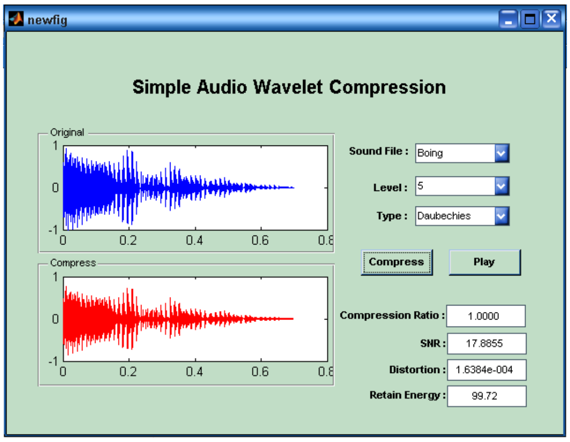

Fig. 6 Layout with compressed signal

V. CONCLUSION

We need the optimum level compression to give optimum output. The recommended level of decomposition is level 5[3]. Decomposition beyond level 5 will not give any gain or advantage except its only give worth output. From the result obtain, it shown that the Daubechiese wavelet family is the best compression family at level 5 of decomposition. In this paper, an audio compression scheme was implemented in GUI suing MATLAB. Although the performance of this system need to be further improved, this system has already shown some important significant features and obtained better SNR and signals quality with less distortion through the observation in GUI with better understanding on human hearing based on wavelet compression scheme that produced better SNR and compression ratio is shown in this project. Further research can be done to improve this system to give better audio signal quality.

REFERENCES

[1] Y. T. Chan, '*Wavelet Basics*', Kluwer Academic Publisher, 1995

[2] D. Sinha and A. Tewfik, '*Low Bit Rate Transparent Audio Compression using Adapted Wavelets*', IEEE Trans. ASSP, Vol. 41, No. 12, December 1993

[3] P. Srinivasan and L. H. Jamieson, '*High Quality Audio Compression Using an Adaptive Wavelet Packet Decomposition and Psychoacoustic Modeling*', IEEE Transactions on Signal Processing, Vol. 46, No. 4, April 1998.

[4] Andreas Spanias, Ted Painter, and Venkatraman Atti, *'Audio Signal Processing And Coding'*, John Wiley & Sons, Inc., Publication, 2007.

[5] Steven W. Smith,*' The Scientist and Engineer's Guide to Digital Signal Processing'*, California Technical Publishing, a Diego, California, second edition, 1990.

[6] Davis Yan Pan,*' Digital Audio Compression'*, Digital Technical Journal, Vol.5, No.2, Spring 1993

[7] L. Rabiner and R. Schafer, *'Digital Processing of Speech* Signals', Englewood Cliffs, NJ, Prentice- Hall, 1978.

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

[9] Stephen J. Chapman, *'MATLAB Programming For Engineers'*, Thomson 4th edition, Thomson Learning, part of the Thomson Corporation, 2008.

[10] Patrick Marchand, O.Thomas Holland.*'Graphics and GUIs with MATLAB'*, Chapman & Hall/CRC, Third edition, 2003.

[11] C. Grewin and T. Ryden, *'Subjective Assessments on Low Bit-rate Audio Codecs'*, Proceedings of the Tenth International Audio Engineering Society Conference, London (1991).

[12] MATLAB Toolbox.