Volume 5, Issue 9, September 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Audio Compression Based on Isolation of Equal Adjacent Samples and DCT

Zinah S. Abduljabbar, Dhia A. Alzubaydi
[1,2]Dept. of Computer Science, College of Science,
University of Mustansiriyah, Baghdad, Iraq

Abstract— ***Audio compression becomes one of an important subject in the recent days. It addresses the problem of transmission requirements and the storage capacity, any compression system is satisfied by eliminating the redundant parts in the file. The purpose of this research is to design and implement an efficient lossy coding system based on discrete cosine transform (DCT). All files are used in this research are stereo type. Splitting process is applied on the two channels, framing each channel as preparing step to implement DCT, quantized the DCT coefficients by using appropriate quantization factor (Qf) and the main contribution in this proposed method is to isolate the equal adjacent samples in one slice with its location, so instead of sending two equals adjacent samples, only one of them is sent with its location. Differencing process is applied on the locations vector, run length encoding (RLE), one type of entropy coding, is applied in the last step in this system. Different sizes with different characteristics are used for audio file. Peak signal to noise ratio (PSNR) and compression factor (CF) are used to evaluate the performance of the system. Good results with a high quality for the reconstructed audio file achieved from the compression performance of the system and this obvious from the test results presented in the paper. CF is increased with increasing of frame size and increasing of Qf.***

Keywords— Audio Compression, DCT, RLE, Equal Samples in One Slice.

I. INTRODUCTION

Compression process is what reduces the redundancy exist in data representation and decreases the required storage capacity. It plays a major role in reducing the communication cost making use of available bandwidth [1]. It plays a vital part in most present communications; it is used when viewing an image, listening to mp3 and other applications. Compression can be defined as a process of re-encoding the digital data by using a special program called a codec for COmpressor / DECompressor and this program decrease the original file to the smaller version and then decompresses it to again present the data in a usable form [2].

An important feature of audio it can be lossy and is acceptable to lose features to which the ear is not sensitive. Audio codecs are used for coding either speech or music. Speech coding is a real-time communication while the gain in music coding is a high quality compression for efficient storage of files. Music has a very wide frequency range; therefor music codecs are expected to provide a higher fidelity experience which in effect, translates to higher sampling and data rate [3]. The most distinguished information is hidden in the frequency content of the signal where large signal mask others in the time domain. In this proposed system, discrete cosine transform (DCT) is used to transform the audio data to a form that lends itself well to compression. After transformation process, the produced coefficients are quantized to reduce the redundancy. Quantization process increases the number of contiguous samples in the two channels which have the same value, this feature can be exploited in compression process and at the last run length encoding (RLE) is applied.

II. AUDIO COMPRESSION SYSTEM

Audio signal contains redundant information in nature and the goals of compression process are to remove this redundancy, achieve a good compression and preserve the quality of the reconstructed audio file, Figure 1 shows the structure of the proposed system, the modules of the proposed system are described in the following sections.

A. *Transformation*

DCT has originated by [Ahmed et al.74]. This transform has been used and studied extensively due to its importance in many applications. DCT takes correlated input data and concentrates its energy in the first few transform coefficients, this characteristic make it a very useful in data compression [4]. There are many examples of lossy compression such as MP3 in audio compression and JPEG in image compression.

The forward DCT equation is (the mathematical representation of one dimensional DCT) [4]:

$$G_f = \sqrt{\frac{2}{n}} C_f \sum_{t=0}^{n-1} p_t \cos\left[\frac{(2t+1)f\pi}{2n}\right] \cdots\cdots (1)$$

$$C_f = \begin{cases} \frac{1}{\sqrt{2}}, f = 0, \\ 1, f > 0, \end{cases} \quad for\ f = 0, 1, \ldots, (n-1) \cdots\cdots (2)$$

Where n is a set of the input data values, p_t (audio samples) and the output is a set of n DCT transform coefficients, the first coefficient G_0 is the DC coefficient where all other transform coefficients are called the AC coefficients.

IDCT is applied on the DCT coefficients in order to reconstruct the audio signal, the IDCT equation is:

$$p_t = \sqrt{\frac{2}{n}} \sum_{j=0}^{n-1} C_j\, G_j \cos\left[\frac{(2t+1)j\pi}{2n}\right], for\ t = 0, 1, \ldots, (n-1) \cdots\cdots (3)$$

B. Quantization

The process of representing a large-possibly infinite-set of values to a much smaller set is called quantization. The amount of compression gain and the loss incurred in the compression process are affected by the design of the quantizer [5]. So quantization can be exploited in the compression process by reducing the number of bits needed to store an integer value by reducing the precision of the integer.

C. Entropy Coding

Designing a variable bit length codes that eliminates the overall number of bits needed to transmit the coded signal are satisfied by using entropy coding [6].it is lossless because no information is lost, such as run length encoding (RLE).

The common data representations are redundant and one type of redundancy is runs of identical symbols. Audio files in nature have long runs, so RLE is used. A run-length encoder recognizes runs of the same symbol and changes each run with a token that includes the length and the symbol of the run. [7].

III. PROPOSED SYSTEM WORKFLOW

The proposed system consists of two major units; the first is encoding unit and the second one is the decoding unit.

A. Encoding Unit

This unit consists of the following stages:

1. Loading the waveform audio file, the selection of audio file was constrained on wave (.wav) files. This process includes reading the header of wave file (44 bytes long) that provided the specification of the file (i.e., number of channels, sampling rate). The audio signal data is loaded as an array, after that normalize the audio data to the range [-1,1] using the following equation:

$$WA = \frac{WA(i) - 2^{(nbits-1)}}{2^{(nbits-1)}} \cdots\cdots (4)$$

   Where:

   WA: the array of audio wave file.

   WA(i): the i^{th} elements of the aufio wave file.

   nbits: sampling resolution (8 bits, 16 bits).

2. Framing each channel into non overlapping parts; each one is called a frame. The frame size must previously be predefining. The Files used in this work have different sizes, so the size of the last frame may be less than the predefine size. Wave signal is changed through the time, so partitioning step is used to preserve sound quality and assures that the wave characteristics are same within each frame. Framing is performed logically and is a preparing step of transformation process.
3. Transform each frame by using DCT by applying equations (1and 2) in order to produce the DCT coefficients.
4. Uniform quantization is applied to quantize the DCT coefficients, the applied equation of quantization process is the following:

$$Q = \frac{DCT\ coefficient}{Qf} \cdots\cdots (5)$$

   Where Q is the quantized DCT coefficient, Qf is the factor of quantization process (the value of Qf can be adjusted to give the required results).

5. There is a large number of a contiguous samples have an equal value in the two channels of stereo wav file, the numbers of these samples which have an equal values are increased after quantization process. This can be exploited by isolated the adjacent equal samples in vector and save their locations, the outputs of this stage are four vectors: vector for isolated samples (SV), vector for locations of isolated samples (indV), the remaining from left channel (remL) and the remaining from right channel (remR).

The outputs from the above stage are four vectors, one of them is a vector of the stored locations which contains a consecutive value in nature, differencing process is applied on the vector of locations and this done by keeping the first value of the given vector in a new vector while the other values of the new vector are resulting from subtracting the preceding from current value of a given vector. Differencing process resulting a redundant occurrence of many sequences consist of long runs of ones.

6. In the proposed system, RLE is applied on two vectors:

The first vector which contains the isolated samples, this vector contains many sequences consist of long runs of zero resulting from the quantization process.

The second vector which contains the locations of the isolated samples, this vector contains many sequences consist of long runs of ones resulting from the differencing process, so RLE compress a runs by replacing it with a pair (length, run value). The run value in case (a) is equal to 0; where in case (b) is equal to 1.

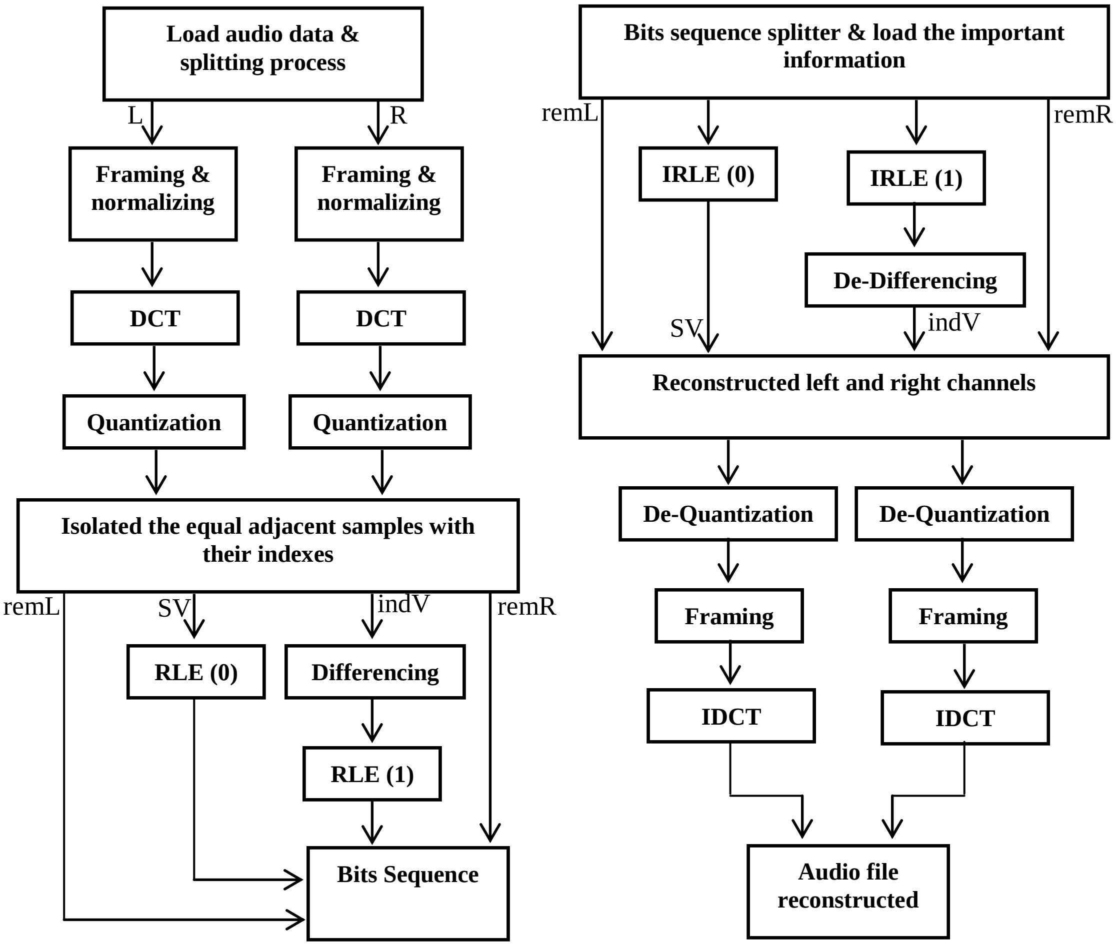

Fig. 1 The Proposed System Layout

B. Decoding Unit

Decoding unit consists of inverse steps of encoding unit; also the operations in each step are applied in reverse order. These steps are:

1. Decoding using RLE.
2. De-differencing process to reconstruct the index value of the isolated adjacent samples.
3. Reconstructed the left and the right channel.
4. De-Quantization.
5. Framing.
6. IDCT.
7. Reconstructed audio signal.

IV. PERFORMANCE MEASUREMENTS

To evaluate the performance of the proposed lossy compression system the fidelity criteria and compression gain (compression factor) must be defined. Fidelity criteria can be divided into two class objective and subjective fidelity criteria.

Peak signal to noise ratio (PSNR) is used in this proposed system as objective criteria, while the subjective criteria require definition of qualitative scale to evaluate the quality of the reconstructed audio file. The adopted Fidelity measures could be defined as follows [8]:

$MSE = \frac{1}{N}\sum_{n=1}^{N}(x_n - y_n)^2 \cdots\cdots (6)$

x_n is the original signal.

y_n is the reconstructed signal.

N is the total number of samples.

$$PSNR(dB) = 10\, log_{10} \frac{(2^{16}-1)^2}{MSE} \;....(7)$$

$$Compression\ factor\,(CF) = \frac{size\ of\ the\ input\ stream}{size\ of\ the\ output\ stream} \;...(8)$$

V. TEST RESULTS

This section is dedicated to evaluate the results of the suggested lossy compression system. Three audio samples are used to test the proposed approach and Table I shows the characteristics of these audio files.

Table I The attributes of the audio test samples

Audio Samples	Attributes			
	Sampling Rate (Hz)	Sample Resolution (bps)	Size (MB)	Duration (Sec)
Test1.wav	44100	16	2.47	14
Test2.wav	48000	8	1.34	14
Test3.wav	44100	8	1.23	14

Figure 2 presents the waveform patterns of these samples, the music in the three samples is same and the type of them is soft music but each file has different sample resolution and sampling rate from the other as present in the above table.

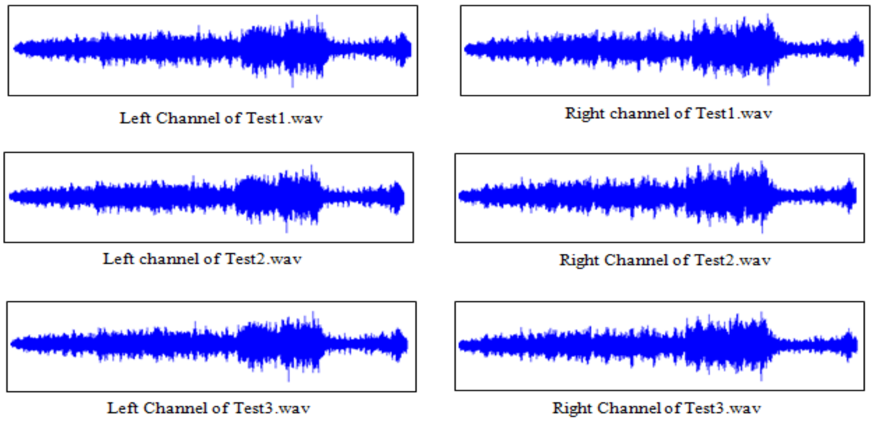

FIG. 2 The Waveform Patterns of The Audio Test Samples

The effects of the control parameter (Qf and frame size) have been investigated. When the value of Qf increased, CF increased also but the quality of audio file adversely affected (PSNR decreased). After comprehensive tests and by varying the value of Qf, 0.04 is the best value for it. Figures 3, 4, 5 and 6 show the effects of frame size on CF, PSNR, ET and DT respectively.

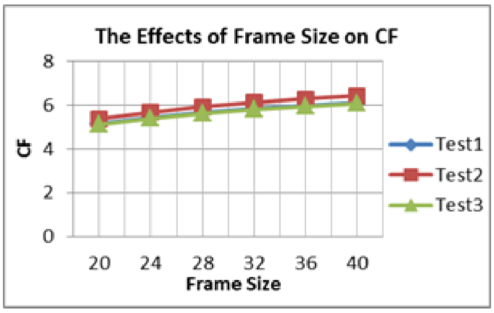

Fig. 3 The Effects of Frame Size on CF

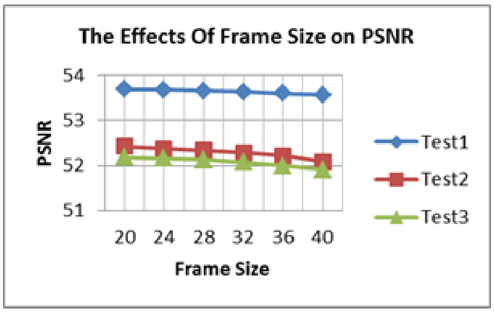

Fig. 4 The Effects of Frame Size on PSNR

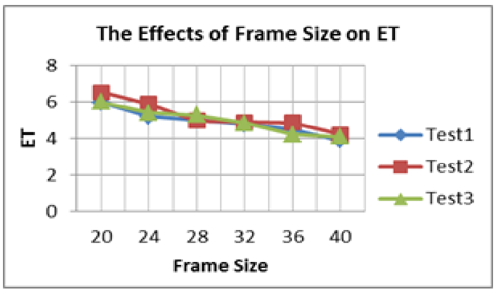

Fig. 5 The Effects of Frame Size on ET

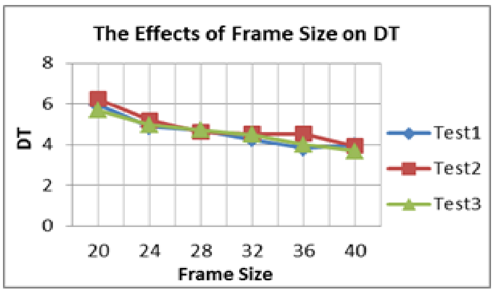

Fig. 6 The Effects of Frame Size on DT

As mentioned previously, framing process preserves the sound quality, so PSNR increased when the frame size is decreased since framing will redistribute the energy formally present along the frame and this can be seen in figure 4. Increasing in the size of the frame caused increasing in CF and decreased in ET and DT and this can be seen in figures 3, 5 and 6.

Table II shows the compression results when Qf equals to 0.04 and after a comprehensive test the best size for the frame equals to 32, from the results can be concluded the high compression gain can acquired from a file has a high sampling rate and this is obvious from the results listed in table II, where Test2.wav has a CF higher than the others file with a high quality for the reconstructed audio file.

Table II The results of compression process

Audio Samples	**CF**	**PSNR (dB)**	**ET (sec)**	**DT (sec)**
Test1.wav	5.8370	53.6361	4.7892	4.2432
Test2.wav	6.1302	52.2877	4.9764	4.5084
Test3.wav	5.8155	52.07	4.8360	4.4772

VI. CONCLUSIONS

In this paper, audio compression scheme based on isolation of equal adjacent samples and DCT has been introduced; from the test results presented above indicate that the introduced system is promising. The idea of isolation the equal adjacent samples in one slice improved the compression gain, the file which has high sampling rate can lend itself well to compression better than files which have less sampling rate.

REFERENCES

[1] M.Baritha Begum and Y. Venkataramani; “LSB Based Audio Steganography Based On Text Compression”, Elsevier Ltd, 2011.
[2] T.SAVAG and K.VOGEL; “An Introduction to Digital Multimedia”, by Jones and Bartlett Publishers, LLc, 2008.
[3] S. Hicsonmez, H. T. Sencar and I. Avcibas; “Audio codec identification from coded and transcoded audios”, Elsevier Inc, 2013.
[4] D.Salomon; “Data Compression the Complete Referance”, Springer-Verlag New York, 2004.
[5] K. Sayood; “Introduction to Data Compression”, Elsevier Inc, 2006.
[6] M. Bosi and R.Goldberg; “Introduction to Digital Audio Coding and Standards”, Kluwer Academic Publishers, Second printing, 2003.
[7] D. Salomon; “Variable-Length Codes for Data Compression”, Springer-Verlag London Limited, 2007.
[8] K. Sayood; “Introduction to Data Compression”, Elsevier Inc, 2012.