

## Audio compression using transforms and high order entropy encoding

Zainab J. Ahmed<sup>1</sup>, Loay E. George<sup>2</sup>, Raad Ahmed Hadi<sup>3</sup>

<sup>1</sup>Department of Biology Science, College of Science, University of Baghdad, Baghdad, Iraq

<sup>2</sup>University of Information Technology and Communications, Baghdad, Iraq

<sup>3</sup>Department of Computer Science, College of Education, Al-Iraqia University, Baghdad, Iraq

### Article Info

#### Article history:

Received Oct 15, 2020

Revised Dec 12, 2020

Accepted Jan 19, 2021

#### Keywords:

Audio compression

Bi-orthogonal (tab 9/7) Wavelet

DCT

LZW

Quantization

Run-length

### ABSTRACT

Digital audio is required to transmit large sizes of audio information through the most common communication systems; in turn this leads to more challenges in both storage and archiving. In this paper, an efficient audio compressive scheme is proposed, it depends on combined transform coding scheme; it is consist of i) bi-orthogonal (tab 9/7) wavelet transform to decompose the audio signal into low & multi high sub-bands, ii) then the produced sub-bands passed through DCT to de-correlate the signal, iii) the product of the combined transform stage is passed through progressive hierarchical quantization, then traditional run-length encoding (RLE), iv) and finally LZW coding to generate the output mate bit stream. The measures Peak signal-to-noise ratio (PSNR) and compression ratio (CR) were used to conduct a comparative analysis for the performance of the whole system. Many audio test samples were utilized to test the performance behavior; the used samples have various sizes and vary in features. The simulation results appear the efficiency of these combined transforms when using LZW within the domain of data compression. The compression results are encouraging and show a remarkable reduction in audio file size with good fidelity.

*This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.*



### Corresponding Author:

Zainab J. Ahmed

Department of Biology Science

College of Science, University of Baghdad

Baghdad, Iraq

Email: zainabjahmed83@gmail.com

## 1. INTRODUCTION

Compression is a key mechanism applied in signal processing and has large significance because huge amounts of data are commonly transferred over a communication channel of a network [1]. Various types of information namely audio, video, images, and text is necessary to use the data compression [2]. Speech compression is a procedure that purposes to change the human speech in a coded frame so that it can be reestablished as the original signal to decrease the redundancy between neighboring samples and between the adjoining frames [3]. The idea of audio compression is to take up lower capacity space to lower transfer speed for transmission by implementing encodes audio data so that various compression methods have been prepared to apply this objective [4, 5]. It is possible to classify audio compression into two groups such as every other digital data compression; these groups are lossless compression and lossy compression [6]. The audio compression technique can be arranged into three functional groups; i) direct forms, ii) parameter extraction forms, and iii) transformation forms [7, 8]. The direct forms when the samples of the signal are directly handled to supply compression, parameter extraction forms when a preprocessor is employed to

extract some features that are later utilized to reconstruct the signal. While transformation forms, like discrete Fourier transform (DFT), discrete cosine transform (DCT), and discrete wavelet transform (DWT) [9] DCT and DWT process is utilized in the field of the audio signal. DCT is commonly used for signal compression especially when it has high correlation and can easily rebuild the signal with low fidelity error. DWT is appropriate for signal compression because of its localization characteristic over time-frequency space [10, 11].

Many methods have been suggested for the digital compression of audio signals utilizing DCT and wavelet transform. In [12] Kaur had studied the use of DCT and DWT to compress speech signal, he applied DCT on the signal, and then the encoded data is decoded by DWT. In general, digitized data can be represented by a lower number of bits when removing the existing. Also, Kaur indicated that wavelet filters can lead to good fidelity results when reconstructing the compressed speech signal. High-efficiency compression results performed in PSNR and MSE of the signal when using different filters with the hybrid model. For speech signal compression, this method can be very effectively applied. Drweesh and George [13] used the bi-orthogonal tab 9/7 wavelet filter to perform an effective audio coding scheme. The suggested coding scheme composed of five stages, these stages are the audio normalization, followed by transform coding using wavelet (tap 9/7), the wavelet coefficients are quantized using the progressive hierarchical scheme, a modified run-length encoding was applied to reduce long runs of the zeros numbers, and finally encoding using high order shift coding is applied. The last step of the decompression process is post-processing, this process is used to decrease the quantization of noise outcomes that happens when low energetic segments of the audio signal and can improve it. The achieved outcomes showed that the compression performance of the system is promising. The increase of several passes of wavelets made the CR is increased. The quality and the fidelity of the rebuild audio data considered improved if PSNR is lower than 38 Db when using the post-processing step. In [14] Kaur and Mehra performed compression of the audio signal by transform technology. Various transforms such as DWT and DCT are used. The test results showed that DCT is better in CR and SNR and PSNR is more than DWT, while DWT is more in MSE. For audio compression, the DWT is better than DCT. Viga and Chauhan [15] proposed a hybrid wavelet for speech compression. The percentage of energy to be kept are changing to implement different levels of compression, this is used as a threshold. The PSNR and MSE have been noticed their values by changing this threshold value from 99% to 99.9%. The results illustrated the hybrid wavelet performed better and notable development in PSNR for similar bit rates.

The main problem is the need for transferring information over the internet which required an additional storage device to handle growth in information and this will lead to an additional cost for equipment. Therefore, the project target is to improve the audio compression system based on transform coding and LZW. A combined transform coding scheme was suggested, the adopted system utilize both DWT and DCT to decompose audio signal. The output of the combined transform is a move to progressive hierarchical quantization followed by RLE to decrease the long runs for zero values, and finally, the LZW coding algorithm is performed. LZW is used because it, i) reduce the size of files having more repetitive data; ii) fast and easy entropy encoder to implement and iii) a lossless compression, all the contents of the file are remained during or after compression. So that the main contribution of this work has improved the compression for the audio signal by using the combined transforms, progressive hierarchical quantization, and run-length with LZW coding.

## 2. THE PROPOSED METHOD

Since the use of proper high entropy encoding is a vital step with any lossy compression scheme (i.e.; as shown in previous literature), so this work is dedicated to high entropy encoding combined with the combined transform coding scheme to compress the audio signal. The audio compression schemes have been applied with four connected stages; these stages are; i) Preprocessing, ii) Transform coding, iii) Quantizer, and iv) Entropy encoder. Firstly, preprocessing is applied to preamble the audio data. Secondly, a suitable transform (DWT and DCT) is implemented. Thirdly, the outcomes are passed through progressive hierarchical quantization operation to remove the existing psycho-visual redundancy, noting that the quantization is one time after applying both transforms. Finally, quantization values are coded using the LZW coding method. The structure of the system model is in Figure 1, the stages of the system are explained with details in the next sections.

### 2.1. Preprocessing stage

This preprocessing stage is necessary for the organization of the audio data and makes later stages of the system run effectively.

### 2.1.1. Load audio data

The header data is read to get the basic file and signal specification information (i.e., number of samples, number of channels, sampling rate, and sampling resolution). Also, the audio file (with WAVE) format is loaded as an array of unsigned bytes when the sample resolution is (8 bit/sample), and as an array of signed integers if the sample resolution is (16 bit/sample).

### 2.1.2. Normalization

Normalization is a process to uniform audio data range whether they are 8 or 16-bit sample resolution, for  $[-1, 1]$ . The normalization operation performed by one of the [13]:

$$W_n = \begin{cases} \frac{W(i) - 127.5}{127.5} & \text{if Sampling resolution} = 8 \text{ bit} \\ \frac{W(i)}{32768} & \text{if Sampling resolution} = 16 \text{ bit} \end{cases} \quad (1)$$

where  $W(i)$  is the  $i^{\text{th}}$  loaded audio data value.

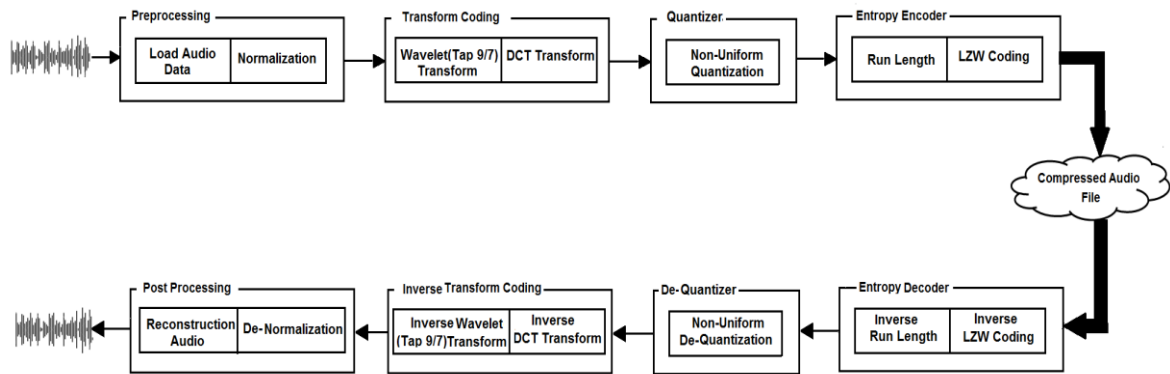


Figure 1. The structure of the developed system model (encoding unit and decoding unit)

## 2.2. Transformation coding stage

The transformation coding is the transformation of data to a more expressive shape for audio data. The audio signal is transformed in this stage from the time domain to the frequency domain; the bi-orthogonal (tap 9/7) wavelet transform and discrete cosine transform (DCT) is used to achieve this process.

### 2.2.1. Bi-orthogonal (tap 9/7) wavelet transform

It is a portion of the group of symmetric bi-orthogonal wavelet cohen\_daubechies\_feauveau (CDF). It holds more complexity than the other methods and a more accurate wavelet method. It includes for low pass filter nine coefficients while for high pass filter seven coefficients [16]. This transform is determined by performing lifting steps then the scaling steps, the lifting scheme is accomplished by using a sequence of phases, it can be declared in three phases: Split phase, predict phase, and update phase [17]. The bi-orthogonal transform is successively decomposed the original signal into low (approximation) frequency and high-frequency components. For high-frequency coefficients, no analysis is performed any further. The approximation signal is then classified into new approximation and detailed signals [18]. The bi-orthogonal (9/7) wavelet decomposition is performed using [19, 20]:

Equations (2)-(5) for "lifting" phase:

$$Y_{2n+1} = X_{2n+1} + a(X_{2n} + X_{2n+2}) \quad (2)$$

$$Y_{2n} = X_{2n} + b(Y_{2n-1} + Y_{2n+1}) \quad (3)$$

$$Y_{2n+1} = Y_{2n+1} + c(Y_{2n} + Y_{2n+2}) \quad (4)$$

$$Y_{2n} = Y_{2n+1} + d(Y_{2n-1} + Y_{2n+1}) \quad (5)$$

Equations (6), (7) for "scaling" phase:

$$Y_{2n} = \left(\frac{1}{k}\right) Y_{2n} \quad (6)$$

$$Y_{2n+1} = -kY_{2n+1} \quad (7)$$

The coefficient {a, b, c, d, and k} values are showed in Table 1:

Coefficient	Value
A	- 1.586134342
b	- 0.05298011854
c	0.8829110762
d	0.4435068522
k	1.230174105

### 2.2.2. Discrete cosine transform (DCT)

Discrete cosine transform, first submitted by [18] in 1974, has become very significant in recent years. DCT has many performed because of its optimal accomplishment, it applied in the signal, image analysis, and applied especially in speech compression because of its optimal achievement. DCT transforms an input signal from the time domain to the frequency domain and its one-dimensional form is good for the examination of one-dimensional signals like speech signals [21, 22]. DCT is composed of (DC) and (AC) coefficients, where the first coefficient C (0) is named the DC coefficient and carries average signal value and the rest coefficients are indicated as the AC coefficients [23].

$$C(u) = \alpha(u) \sum_{i=0}^N s(i) \cos\left(\frac{u\pi(2i+1)}{2N}\right) \quad (8)$$

$$\alpha(u) = \begin{cases} \sqrt{1/N} & \text{if } u = 0 \\ \sqrt{2/N} & \text{if } u \neq 0 \end{cases} \quad (9)$$

where  $u=0\dots, N-1$  and  $C(u)$  is the  $u^{\text{th}}$  coefficient of the DCT, and  $s()$  represents a set of  $N$  audio input data values.

### 2.3. Quantization

Quantization is the operation of representing a big set of values with a much smaller set. An easy quantization scheme would be to represent all products of the source with the integer value nearest to it [24]. Quantization maps a group of continuous-valued data to a group of discrete-valued data. The main goal of this process is to decrease the data located in threshold coefficients. The result of this process produces sure that it makes the smallest amount of errors [25].

### 2.4. Entropy encoder

It is a shape of the lossless compression and it is relying on the information-theoretic mechanism. Some entropy encoding techniques are Huffman coding, Arithmetic coding, LZW, and RLE [26]. The Entropy encoder techniques used in this article are run-length encoding and LZW.

#### 2.4.1. Run-length

Run-length encoding is the easy format of redundancy elimination. It eliminates redundancy according to the reality that a string has repeated sequences or "runs" of a similar symbol. Two structures are used to encode the runs of the same symbol: a count indicating the number of iterated symbols and the symbol itself [27].

#### 2.4.2. LZW

LZW compression techniques is a table (i.e., dictionary) coding of strings of characters using single codes, it shows outperformance especially when the long strings show a high probability of occurrence. LZW

only inserts each new string to the dictionary encodes it using the index, without the need to make costly analysis for input text. The use of a single index is an output rather than the strings of characters then compression is achieved [28, 29]. This technique begins by initializing the dictionary to all the symbols in the alphabet. In common 8-bit symbols, the dictionary takes the first 256 entries (entries 0 through 255) before any data is input. Always the next input character will be found in the dictionary because the dictionary is initialized [23]. If "X" and "Y" are two strings that are found in the dictionary, the sequence of character "XY" is changed into the index of "X" then by the index of "Y". "X" greedy string matching algorithm is utilized for scanning the input, when the first character of "Y" is "z", then "Xz" cannot be a component of the dictionary. The main feature of the algorithm is due to the reality that "X" "z" is automatically inserted into the dictionary if "X" is matched but "X" "z" is not matched [30]. LZW compression is a good mechanism for decreasing the size of files containing more repetitive data. Always the decompression algorithm comes after the compression algorithm. In the decompression code, it doesn't require to convey the string table. Using the input stream as data, the table can be created again as it was during compression, this makes LZW is an efficient algorithm [31].

### 3. RESEARCH METHOD

The audio compression system is composed of two units: The first one is named "Encoding unit", and the second is named "Decoding unit". These two units have many stages, as shown in Figure 1.

#### 3.1. Encoding unit

Many stages find in this unit which is altogether responsible for reducing the data size of the required audio and produce a compressed stream of data that represents the audio. The performed stages for the encoding unit are given:

- The wave file is converted to an array of values that demonstrate the audio signal. The one-dimensional array is using as the data structure. Just data of the wave audio samples are loaded.
- The loaded audio data file is normalized to make sure that the system is not changed with the audio loudness levels variation. The normalization operation is performed by (1).
- The bi-orthogonal tap 9/7 wavelet transform is implemented on normalized data using as (2), (3), (4), (5), (6), (7), and then this transform decomposes data into two subbands (low and high). Each subband contains a specific kind of audio information. The transform operation may be repeated depending on the number of wavelet transform passes parameter (i.e.;  $N_{\text{passes}}$ ); where its value is predefined by the user.
- Establishing a set of pointers indicating the start (st) and end (ed) of each subband of the wavelet coefficients.
- Calculating the standard deviation (std) for each wavelet coefficient (i.e.;  $H1, H2 \dots H_{N_{\text{pass}}}, L$ ).

$$Avr = \frac{1}{N} \sum_{i=st}^{i=en} Wav(i) \quad (10)$$

$$std = \sqrt{\frac{1}{N} \sum_{i=st}^{i=en} (Wav(i) - Avr)^2} \quad (11)$$

where st and ed is a start and endpoint for each subband of the wavelet coefficients. N is the total number of samples for each subband of the wavelet coefficients.

- The wavelet coefficients of each subband partition into blocks. Then each block is operated separately until reaching the last block. The no. of blocks ( $N_b$ ) depends on the length of each subband of the wavelet coefficients and the size of blocks according to (12):

$$N_b = \frac{\text{length of each subband of the wavelet coefficients}}{\text{block size}} \quad (12)$$

- The data of each block (i.e.; each subband of the wavelet coefficients) is decomposed separately using the DCT transform. The (8, 9) are applying to get a set of wavelet-DCT coefficients.
- The outcomes of wavelet-DCT coefficients are real-valued, and to increase the compression, they must be quantized before compression. Progressive hierarchical quantization is utilized to quantize the transformed wavelet-DCT coefficients of each produced block by using this [13]:

a. For DC coefficient:

$$C_q(0) = \text{round} \left( \frac{C(0)}{Q_0} \right) \tag{13a}$$

where  $C(0)$  is the DC-coefficient,  $C_q(0)$  is the quantized DC coefficient,  $Q_0$  for the DC coefficient.

b. For AC Coefficient:

$$C_q(u) = \text{round} \left( \frac{C(u)}{Q_1(1 + \alpha(u - 1))} \right) \tag{13b}$$

where,  $Q_1$  for AC coefficient of each block,  $u$  (coefficient index) = 1, 2 ... $N_b$ ;  $\alpha$  is the progressive rate parameter.

The wavelet-DCT coefficient ( $H_1$ ) takes the initial values of the quantization step ( $Q_0$  and  $Q_1$ ) which are predefined by the user. The quantization step values ( $Q_0$  and  $Q_1$ ) are not static and increase consistently with each sub-band coefficient. These values are calculated according to the following equation to the other wavelet-DCT coefficients (i.e.,  $H_2 \dots H_{N_{\text{pass}}}$ ,  $L$ ).

$$Q_{0\_new} = Q_{0\_old} \left( \frac{\text{std}_{new}}{\text{std}_{old}} \right) \times \text{thr} \tag{14a}$$

$$Q_{1\_new} = Q_{1\_old} \left( \frac{\text{std}_{new}}{\text{std}_{old}} \right) \times \text{thr} \tag{14b}$$

where,  $Q_{0\_old}$  and  $Q_{1\_old}$  are the previous quantization steps for the previous wavelet-DCT coefficients.

$Q_{1\_new}$  and  $Q_{1\_new}$  are the new quantization steps for the current wavelet-DCT coefficients.

$\text{std}_{old}$  is the previous standard deviation for the previous wavelet-DCT coefficients.

$\text{std}_{new}$  is the new standard deviation for the current wavelet-DCT coefficients.

$\text{thr}$  is a predefined threshold value.

- The outcomes of the quantization must be rounded to the close integer. Long runs of zero symbols are coded to generate sequences of small integer numbers. The solution to this problem is to apply RLE that replaces the long runs of zeros symbols with run-length count and non-zero symbols, it decreases the physical size of repeated consecutive values of quantized DCT coefficients. Two records are created; the first record contains the first item in a sequence followed by the run length for zero and non-zero values, this called runs record. The second record contains the non-zero values called values record. Figure 2 illustrates the run-length encoding. Example data:

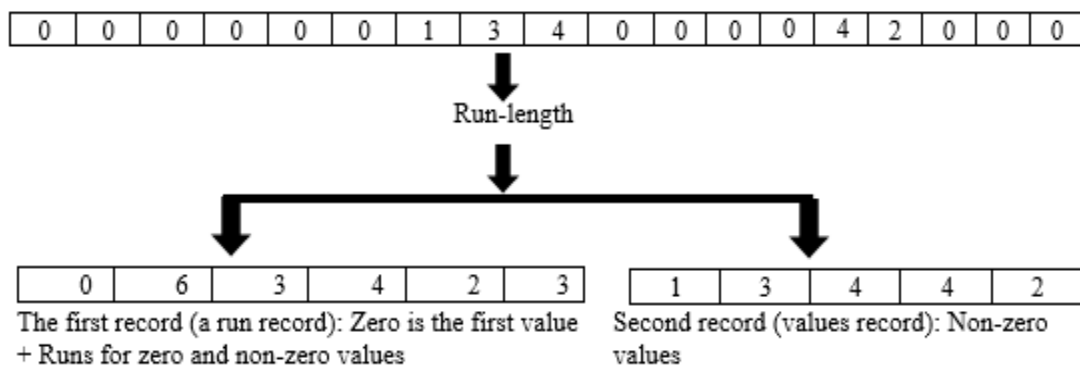


Figure 2. Example of run-length encoding

- This step maps the sequence elements in the record of the value (non-zero values) to be positive numbers to make the coding process of the next step simpler. The mapping step is to make each positive value is appeared as an even number, while each negative element value appears as a positive odd number. This conversion can be applied using the simple mapping [32]:

$$m_{\text{out}}(i) = \begin{cases} 2V(i) & \text{if } V(i) \geq 0 \\ -2V(i) + 1 & \text{if } V(i) < 0 \end{cases} \quad (14b)$$

- The run and mapping values records are merged to one record to increase the compression ratio because LZW generally performs best on files that hold high redundancy of symbols. Apply the LZW compression algorithm to encode the final record.

### 3.2. Decoding unit

In this unit similar, but inverted sequences, of processes, are built. It is applied to the compressed audio data to get the reconstructed version of the digital audio signal. The decoding unit is built by reversing the steps made to compress audio data. The decoding unit stages are: i) LZW decoding, ii) mapping to negative, iii) inverse run-length encoding, iv) de-quantization, v) inverse DCT, vi) inverse bi-orthogonal wavelet, vii) mapping data according to sample resolution, where in 8-bit mapping to the byte range [0, 255], and in 16-bit mapping to the integer range [-32768, 32767].

## 4. RESULTS AND DISCUSSION

The effectiveness of audio compression was evaluated using six different tested audio files to assess the performance of the proposed audio compression scheme. In this set of tests, the metrics used for determining the effectiveness of the compression process are the compression ratio (CR) and peak signal to noise ratio (PSNR). Figure 3 presents the waveform patterns of the audio test samples. Table 2 shows the characteristics of the used six audio files. All these samples are wave file format (wav) with the mono channel and PCM data format.

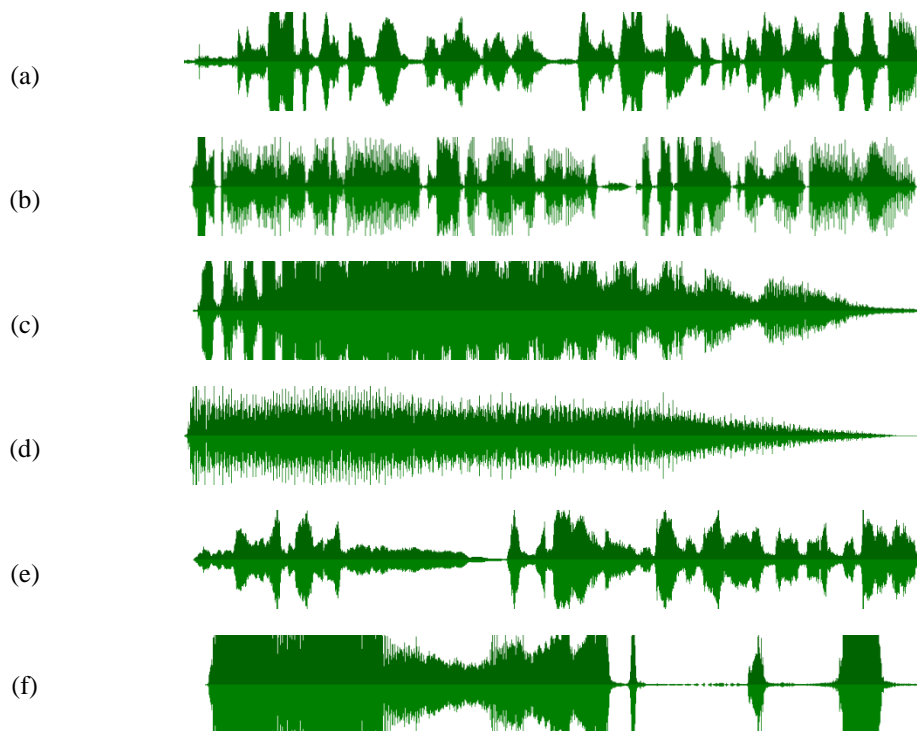


Figure 3. The waveform of the tested wave files, (a) test (1) audio signal, (b) test (2) audio, (c) test (3) audio signal, (d) test (4) audio signal, and (e) test (6) audio signal

Table 2. The characteristics of the audio test samples

Attributes	Audio samples					
	Test (1)	Test (2)	Test (3)	Test (4)	Test (5)	Test (6)
Sampling Rate (kHz)	44100	44100	44100	44100	44100	44100
Sample Resolution (bps)	8	8	8	16	16	16
Size (KB)	216 Kb	270 Kb	289 Kb	130 Kb	948 Kb	299 Kb
Audio Type	Dialog (Female)	Dialog (Male)	Laugh with Echo	Piano	Quran	Animal

The effects of control system parameters have been tested, they are; i) the number of wavelets passes ( $N_{pass}$ ), ii) quantization steps  $Q_0$  and  $Q_1$ , iii) alpha multiplication parameter ( $\alpha$ ), iv) sampling rate, v) sampling resolution, and vi) block size (BS), the vii) threshold value (Thr). The range values of the tested control parameters are illustrated in Table 3. High compression gain is the main concern in audio compression, while the main goal of this work is attaining good compression performance with a high-fidelity level. Table 4 shows the set of default values for control parameters; they were adopted to test the effectiveness of each system parameter.

Table 3. The range values of the control parameters

Parameter	Range
$N_{pass}$	[2,7]
$Q_0$	[0.01, 0.05]
$Q_1$	[0.06, 0.15]
$\alpha$	[0.01,0.05]
Sampling Rate	{11025, 22500, 44100}
Sampling Resolution	{8,16}
BS	[20, 200]
Thr	[0.5,1.5]

Table 4. The default values for the control parameters with the best PSNR and CR

	Audio Samples	PSNR	CR	$N_{pass}$	$Q_0$	$Q_1$	$\alpha$	BS	Thr
Audio samples with sample resolution= 8 bps and sample rate=44100 kHz	Test (1)	35.423	15.230	2	0.05	0.07	0.01	120	1.4
	Test (2)	36.070	15.370						
	Test (3)	36.360	10.693						
Audio samples with sample resolution= 16 bps and sample rate=44100 kHz	Test (4)	39.148	10.900	2	0.05	0.07	0.02	120	1.4
	Test (5)	40.814	14.918						
	Test (6)	36.025	9.9199						

The impact of the number of wavelets transforms passes ( $N_{pass}$ ) on PSNR and CR are presented in Figure 4, the results show that the increase of  $N_{pass}$  causes a high compression ratio while makes a decrease in PSNR value. Table 5 shows the effect of changing the sample rate on CR and PSNR with the same default control parameters; also Figure 5 shows the effect of changing the sample rate on results. The results show that the sample rate of 44100 kHz gives better results in CR and fidelity.

The impact of sampling resolution on the performance of the combined transform-based compression scheme shows in Figure 6. The results indicate that the significant impact of sampling resolution and more excellent compression results in the case of a 16-bit sampling resolution. In this set of tests, a conversion process from integer (16-bits) representation to byte representation is accomplished to get an audio (low-resolution) version. This conversion was implemented using a dynamic link library called "NAUDIO" belong to the C-sharp. The impacts of parameters  $Q_0$ ,  $Q_1$ ,  $\alpha$ , threshold (thr), and block size (BS) on PSNR and CR are shown in Figure 7 (see in Appendix), it is evident that the increase of these parameters leads to increase the attained compression ratio (CR) while decreasing the fidelity level (under the condition it is preserved to be above the acceptable level).

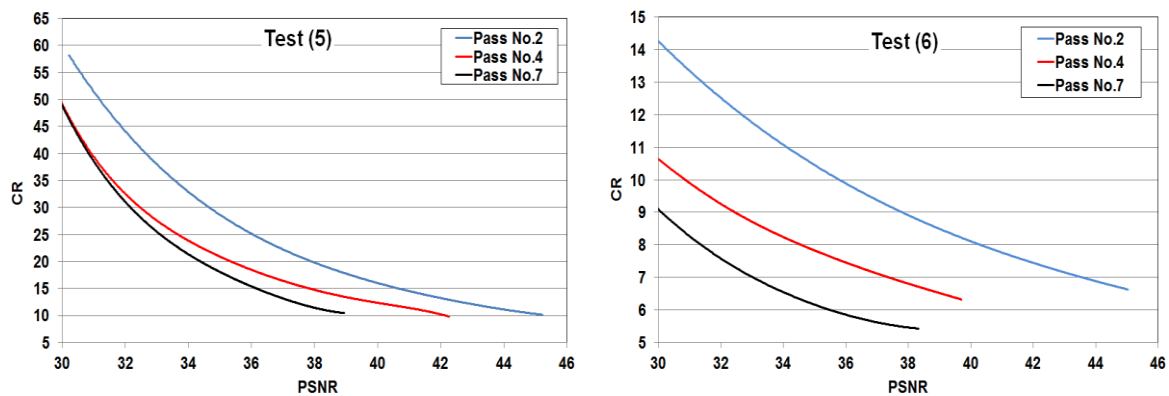


Figure 4. The effect of the number of wavelets passes  $N_{pass}$



Table 5. The effect of change sample rate on CR and PSNR with the same control parameters on test3

Sample Rate	CR	PSNR	$N_{pass}$	$Q_0$	$Q_1$	$\alpha$	BS	Thr
11025	4.229851	33.68311	2	0.01	0.05	0.01	120	1.4
22050	6.592916	35.53675						
44100	9.04161	38.04224						

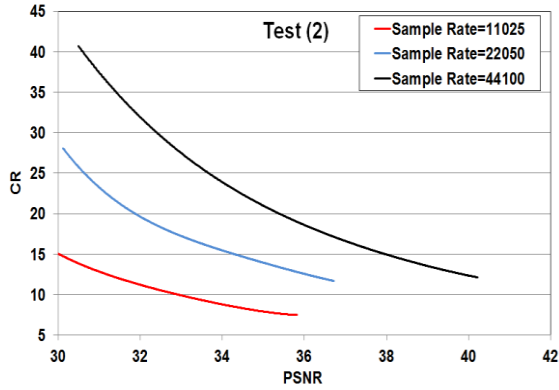


Figure 5. The effect of sampling rate on the relation between PSNR and CR

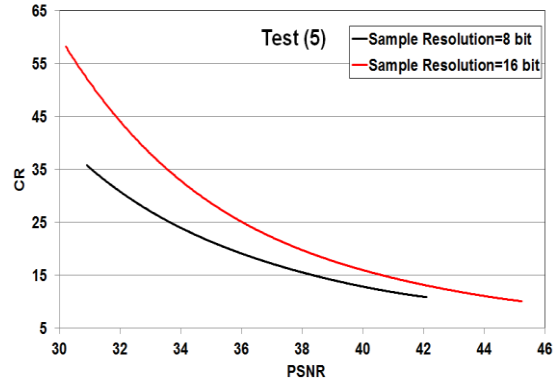
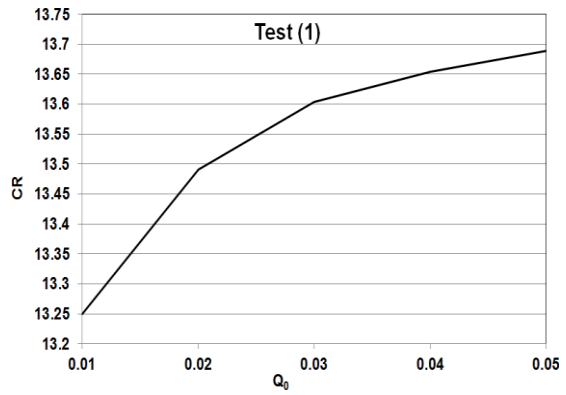
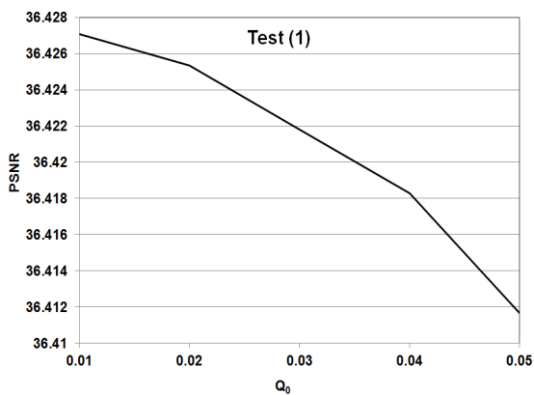


Figure 6. The effect of sampling resolution on the relation between PSNR and CR

5. CONCLUSION

In this study, compressive audio based on bi-orthogonal (tab 9/7), DCT transform, and LZW coding technique had been presented. The compressive audio system reduced the audio file size and avoided the need for much storage. The conducted test results indicated that the proposed system is promising and each control parameter that has significant effects on the system was analyzed from Figures 4, 5, 6, and 7. The following stimulated remarks are summarized: i) The proposed compression system using the LZW technique show acceptable the compressions ratio while preserving the audio quality as shown in Table 4, ii) The changing sample rates have an effect on CR and fidelity as shown in Table 5 and Figure 5; the sample rate 44100 kHz has better results in both compression and fidelity, iii) The increase of quantization steps leads to an increase in the attained compression ratio while decreasing the fidelity level as shown in Figure 7 in test results, iv) RLE decreased the physical size of repeated consecutive values, v) The system can be improved in the future using audio fractal coding as a compression tool (instead of wavelet transform coding and DCT) in the compressive audio scheme.

APPENDIX



(a)

Figure 7. The effect of ( $Q_0$ ), ( $Q_1$ ) and ( $\alpha$ ), threshold and block size on the relation between PSNR and CR, (a)  $Q_0$ , (b)  $Q_1$ , (c)  $\alpha$ , (d) thr, and (e) BS

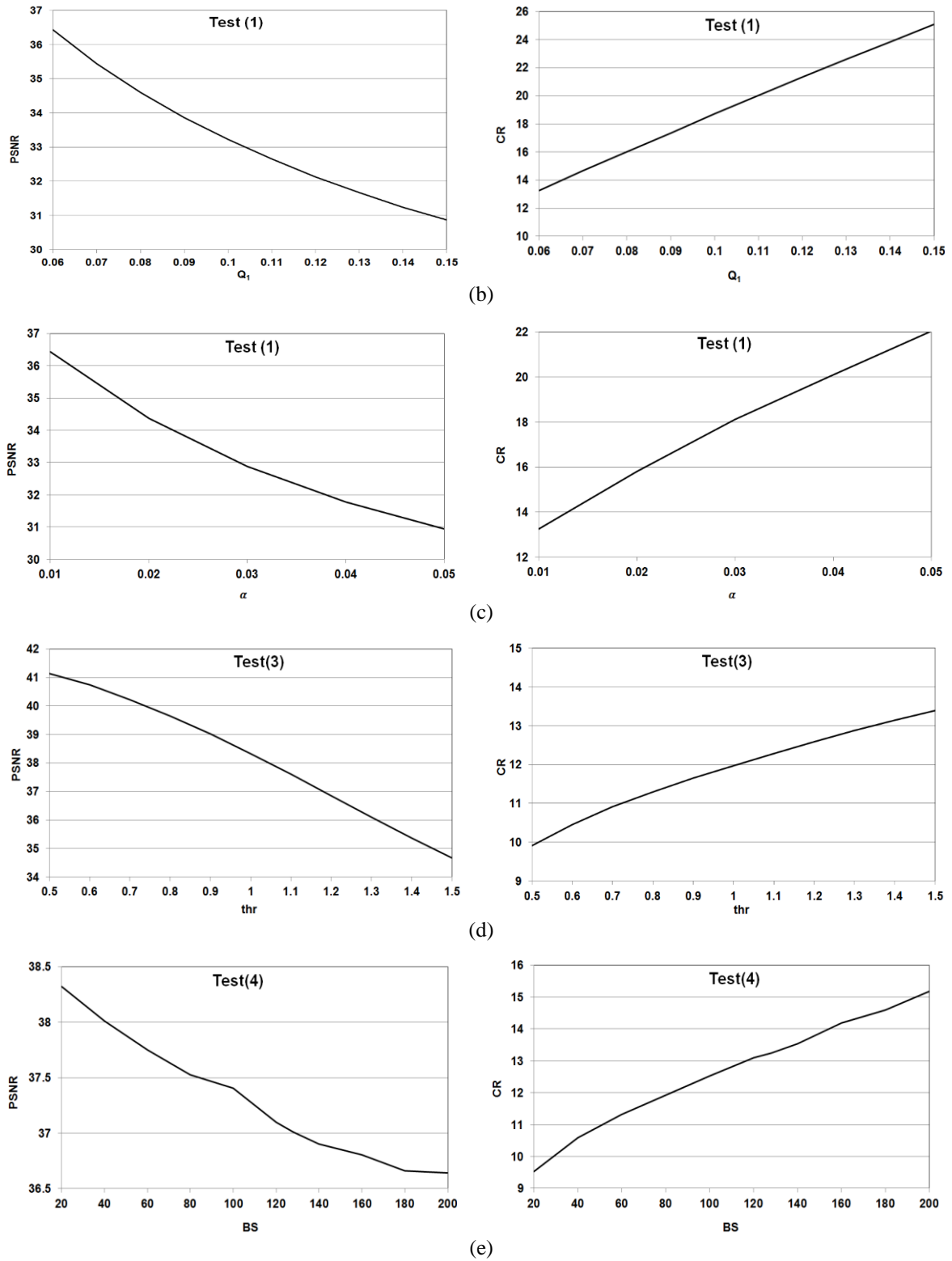


Figure 7. The effect of ( $Q_0$ ), ( $Q_1$ ) and ( $\alpha$ ), threshold and block size on the relation between PSNR and CR, (a)  $Q_0$ , (b)  $Q_1$ , (c)  $\alpha$ , (d) thr, and (e) BS (continue)

**REFERENCES**

[1] Salau, A. O., Oluwafemi, I., Faleye, K. F., and Jain, S., "Audio Compression Using a Modified Discrete Cosine Transform with Temporal Auditory Masking," *2019 International Conference on Signal Processing and Communication (ICSC)*, NOIDA, India, 2019, pp. 135-142.

- [2] L. A. Fitriya and T. W. Purboyo, "A Review of Data Compression Techniques," *International Journal of Applied Engineering Research*, vol. 12, no. 19, pp. 8956-8963, 2017.
- [3] zohra Chelali, F., Cherabit, N., Djeradi, A., and Falek, L., "Wavelet transform for speech compression and Denoising," *2018 6th International Conference on Multimedia Computing and Systems (ICMCS)*, Rabat, Morocco, 2018, pp. 1-7.
- [4] A. Tsegaye and G. Tariku, "Audio Compression Using DWT and RLE Techniques," *American Journal of Electrical and Electronic Engineering*, vol.7, no. 1, pp. 14-17, 2019.
- [5] Abd-Elhafiez, W. M., Gharibi, W., and Heshmat, M., "An efficient color image compression technique," *TELKOMNIKA Telecommunication, Computing, Electronics and Control*, vol. 18, no. 4, pp. 2371-2377, 2020.
- [6] Gunawan, T. S., Zain, M. K. M., Muin, F. A., and Kartiwi, M., "Investigation of Lossless Audio Compression using IEEE 1857.2 Advanced Audio Coding," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 6, no. 2, pp. 422- 430, 2017.
- [7] R. Kumar and H. Amhia, "A New Approach of Speech Compression by Using DWT & DCT," *International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering (IJAREEIE)*, vol. 3, no. 7, pp. 10762-10765, 2014.
- [8] H. Amhia and R. Kumar, "Integrated Approach for Speech Compression Using DWT & DCT," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 3, no. 8, pp. 11366- 11370, 2014.
- [9] Romano, N., Scivoletto, A., and Polap, D., "A Real-Time Audio Compression Technique Based on Fast Wavelet Filtering and Encoding," *2016 Federated Conference on Computer Science and Information Systems (FedCSIS)*, Gdansk, Poland, 2016, pp. 497-502.
- [10] Patil, M. V., Gupta, A., Varma, A., and Salil, S., "Audio and Speech Compression Using DCT and DWT Techniques," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 2, no. 5, pp. 1712-1719, 2013.
- [11] J. James and V. J. Thomas, "Audio Compression Using DCT and DWT Techniques," *Journal of Information Engineering and Applications*, vol. 4, no. 4, pp. 119-124, 2014.
- [12] H. kaur and R. kaur, "Speech compression and decompression using DWT and DCT," *International Journal Computer Technology & Applications*, vol. 3, no. 4, pp. 1501-1503, 2012.
- [13] Z. T. Drweesh and L. E. George, "Audio Compression Based on Discrete Cosine Transform, Run Length and High Order Shift Encoding," *International Journal of Engineering and Innovative Technology (JEIT)*, vol. 4, no. 1, pp. 45-51, 2014.
- [14] A. Kaur and R. Mehra, "Data Compression using Multiple Transformation Techniques for Audio Applications," *IOSR Journal of Computer Engineering (IOSR-JCE)*, vol. 17, no. 6, pp. 79-85, 2015.
- [15] R. Viga and S. S. Chauhan, "Speech Compression using Multi-Resolution Hybrid Wavelet using DCT and Walsh Transforms," *Procedia Computer Science*, vol. 132, pp. 1404-1411, 2018.
- [16] Beladgham, M., Bessaid, A., Abdelmounaim, M. L., and Abdelmalik, T. A., "Improving Quality of Medical Image Compression Using Biorthogonal CDF Wavelet Based on Lifting Scheme and SPIHT Coding," *Serbian Journal of Electrical Engineering*, vol. 8, no. 2, pp. 163-179, 2011.
- [17] Reddy, T. P. A., Reddy, P. L., and Kumar, T. V., "5/3 Lifting Scheme Approach for Image Interpolation," *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)*, vol. 12, no. 1, pp. 8-13, 2017.
- [18] S. M. Josepha and B. A. P., "Speech Coding Based on Orthogonal and Biorthogonal Wavelet," *Procedia Technology*, vol. 6, pp. 397-404, 2012.
- [19] G. Strang and T. Nguyen, "Wavelets and Filter Banks," *Wellesley-Cambridge Press, 2<sup>th</sup> Edition*, 1996.
- [20] H. A. Hadi and L. E. George, "EEG Based User Identification Methods Using Two Separate Sets of Features Based on DCT and Wavelet," *Journal of Theoretical and Applied Information Technology*, vol. 95, no. 22, pp. 6116-6126, 2017.
- [21] N. Ahmed, *et al.*, "Discrete cosine transform," *IEEE transactions on Computers*, vol. C-23, no. 1, pp. 90-93, 1974.
- [22] Novamizanti, L., Budiman, G., and Astuti, E. N. F., "Robust audio watermarking based on transform domain and SVD with compressive sampling framework," *TELKOMNIKA Telecommunication, Computing, Electronics and Control*, vol. 18, no. 2, pp. 1079-1088, 2020.
- [23] D. Salomon and G. I. Motta, "Handbook of Data Compression," *Springer, 5<sup>th</sup> Edition*, 2010.
- [24] K. Sayood, "Introduction to Data Compression," *5<sup>th</sup> edition*, United States of America, *Elsevier*, 2012.
- [25] Manohar, P. K. R., Pratyusha, M., Satheesh, R., Geetanjali, S., and Rajasekhar, N., "Audio Compression Using Daubechie Wavelet," *IOSR Journal of Electronics and Communication Engineering (IOSR-JECE)*, vol. 10, no. 2, pp. 41-44, 2015.
- [26] Z. M. Lu and S. Z. Guo, "Lossless Information Hiding in Images," *Zhejiang University Press, Elsevier*, 2017.
- [27] P. Havaladar and G. Medioni, "Multimedia Systems: Algorithms, Standards, and Industry Practices," *Course Technology, Cengage Learning, Canada*, 2010.
- [28] A. Banerji and A. M. Ghosh, "Multimedia Technologies," *Tata Mcgraw Hill Education*, 2010.
- [29] Saad, H. N., mushtaq Jafar, F., and Salman, H. A., "A new compression technique in MANET: compressed-LZW algorithm," *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, vol. 16, no. 2, pp. 890-896, 2019.
- [30] Z. N. Li and M. S. Drew, "Fundamentals of Multimedia," *Pearson Education International, United States of America*, 2004.
- [31] A. Makkar, *et al.*, "Improving LZW Compression," *International Journal of Computer Science and Technology (IJCST)*, vol. 3, no. 1, pp. 411-414, 2012.
- [32] A. H. Ahmed and L. E. George, "The Use of Wavelet, DCT & Quadtree for Images Color Compression," *Iraqi Journal of Science*, vol. 58, no. 1C, pp. 550-561, 2017.