Audio Compression Using Biorthogonal Wavelet Transform

Wafaa S. Ahmed

Abstract- The idea of audio compression is to encode audio data to take up less storage space and less bandwidth for transmission.. An effective wavelet-based audio compression algorithm is presented to provide highly efficient signal compression mechanism with acceptable human hearing perception. The basic engine that used in this method is the wavelet transform tap 9/7 followed by the quantization processes and used some of lossless compression techniques (run length encoder and shift coder) to obtain on good compression ratio with preserving the signal quality. The advantages of Wavelet transform are to provide characteristic of multiple resolution and global decomposition that are the significant features for the audio compression applications. The best compression ratio that obtains on it is (13.4) and the PSNR is (38.66) in the sound1, when the number level is (5) and quantization step value is (30). The main objective of this research is to construct a compression method that compress the audio files by using biorthogonal wavelet transform technique.

Keywords: Audio compression, Wavelet Transform 9/7 tap filter, Run length Encoder, Shift Coder.

———————————— ◆ ————————————

1 INTRODUCTION

AUDIO signal compression has found application in many areas, such as multimedia signal coding, high-fidelity audio for radio broadcasting, audio transmission for HDTV, audio data transmission/sharing through Internet, etc. High-fidelity audio signal coding demands a relatively high bit rate of 705.6 kbps per channel using the compact disc format with 44.1 kHz sampling and 16-bit resolution. For large amount of exchange and transmission of audio information through internet and wireless systems, efficient (i.e., low bit rate) audio coding algorithms need to be devised. Two major classes of techniques can be used in audio coding to reduce the coded bit rate. The first class takes advantage of the statistical redundancy in audio signal and applies some form of digital encoding, it is a lossless audio coding in which original audio signal can be perfectly recovered from the encoded audio signal [1], example Huffman ,Run length encoder, shift encoder. The second class employs some signal processing so that essential information and perceptually irrelevant signal components can be separated and later removed, it is a lossy audio coding in which original and reconstructed audio signal are not perfectly identical. This class includes techniques such as sub band coding, transform coding example(DCT, Wavelet Transform), critical band analysis, and masking effects [1].

Digital Signal Processing (DSP) techniques can be used to decrease the redundancy and irrelevancy contained in an audio signal. Audio coding is an important step towards delivering a high quality communications for multimedia and Internet. Digital audio compression allows the efficient storage and transmission of audio data [1].

One of the techniques that used in audio compression is wavelet transform, Wavelet compression is a form of predictive compression where the amount of noise in the data set can be estimated relative to the predictive function [2]. Wavelet coding is based on the idea that the coefficients of a transform decorrelates the sample values of an audio signal and can be coded more efficiently than the original samples themselves. Most of the important part of the information is contained by a smaller number of coefficients, and hence the remaining coefficients can be quantized coarsely or truncated to zero with little distortion in perception of coded audio signal. Because wavelet transforms are both computationally efficient and inherently local (i.e. their basis functions are limited in duration) [3]. Most modern compression techniques use a two-step process: First, a predictive compression function (such as wavelet transform) is applied. If the choice of the predictive compression function is good, the result will be a new set of data with smaller values and more repletion. Second, a coding compression step that will represent the data set in its minimal form (Huffman coding, run-length, shift coding) [2].

2 THE PROPOSED SYSTEM

The proposed system is shown in figure (1). Its consist of two main phases: compression phase and Decompression phase. In compression phase the audio signal is transformed from spacial domain to frequency domain and then compressed. In the second phase which is the reconstructed phase using the output of the first phase is used to reconstruct the original audio by using inverse transform operation on compressed phase.

The following sub-sections explain the action of each stage in the both phases of proposed system:

2.1 Compression and Decompression Phase

In the compression phase the audio file is transformed from spatial domain to frequency domain by using tap 9/7 wavelet transform and quantized the coefficients, after that applied two

compression techniques are run length encoder and shift coder to obtain on compressed audio file. This is shown in figure 1.
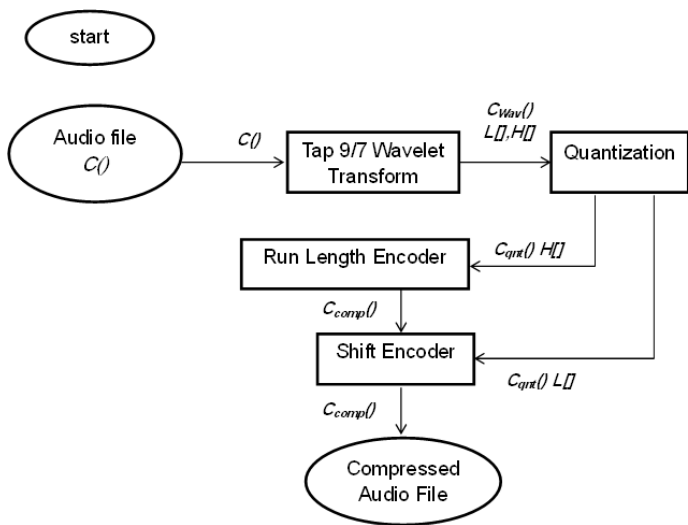

Fig 1: The block diagram of compression phase

 In the decompression phase the original audio file will reconstruct from compressed file by apply the same steps but in reversible. This is shown in figure 2.
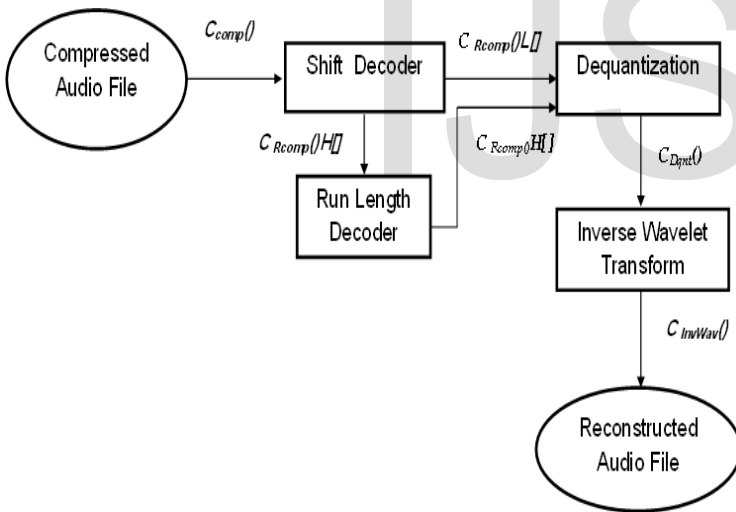

Fig 2: The block diagram of decompression phase

A. Wavelet Transform

The discrete wavelet transform (DWT) is a highly flexible family of signal representations that may be matched to a given signal and its well applicable to the task of audio data compression [4]. One dimensional discrete wavelet transform 1-D DWT and 1-D IDWT modules can be implemented by filter bank structures shown in figure(3) [5]. The 1-dimensional Wavelet Transform (1DWT) is applied on audio sample and applied horizontally to the row. 1-D DWT is the lifting based

wavelet transform implementation of filtering by the tap 9/7 [6]. The biorthogonal wavelets are perhaps the most widely used. These wavelets have symmetric scaling and wavelet functions, i.e., both the low pass and high pass filters are symmetric. The properties of these wavelets have made them very popular for compression applications. For high compression ratios more zeros are needed which can be obtained by using longer filters. But, if the filter is too long, ringing occurs and this destroys the quality. The tap 9/7 filter bank has rational filter length and yields good performance. 1-D DWT is applied to the finite length of audio samples [7]. Multiple level 1-D DWT is also recursively performed to low resolution sub-band coefficients to acquire audio transform bands [6].
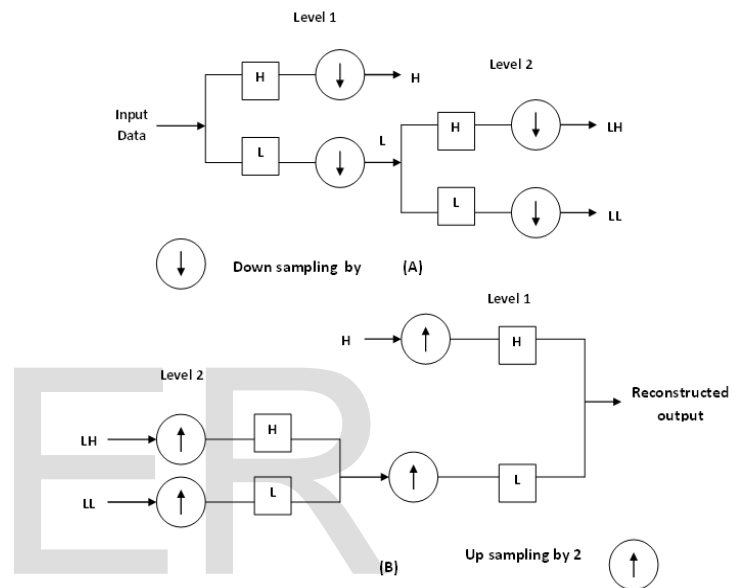

Fig 3: filter bank structure for No. level=2 (A) 1-D DWT (B) 1-D IDWT

In the discrete wavelet transform, a signal can be analyzed by passing it through an analysis filter bank followed by a decimation operation. This analysis filter bank, which consists of a low pass and a high pass filter at each decomposition stage . When a signal passes through these filters, it is split into two bands. The low pass filter, which corresponds to an averaging operation, extracts the coarse information of the signal. The high pass filter, which corresponds to a differencing operation, extracts the detail information of the signal. The output of the filtering operations is then decimated by two [3].

B. Quantization

A quantizer simply reduces the number of bits needed to store the transformed coefficients by reducing the precision of those values. Since this is a many-to-one mapping, it is a lossy process and is the main source of compression in an encoder. Quantization can be performed on each individual coefficient, which is known as Scalar Quantization (SQ). Quantization can also be performed on a group of coefficients together, and this is

known as Vector Quantization (VQ) [8]. Both uniform and non-uniform quantizers can be used depending on the problem at hand.

The forward and inverse uniform scalar quantization was done by using following equation:

$$C_{qnt}=\text{round}\left(\frac{C_{wav}}{Q}\right) \longrightarrow (\text{Forward}) \quad (1)$$

$$C_{Dqnt}=Q*C_{qnt} \longrightarrow (\text{Inverse}) \quad (2)$$

Where:

C_{wav}: is the original wavelet (detail) coefficients.

Q: is the quantization step value.

C_{qnt}: is the corresponding quantized coefficients.

C_{Dqnt}: is the inverse quantized coefficients

C. Entropy Encoder

An entropy encoder further compresses the quantized values losslessly to give better overall compression. It uses a model to accurately determine the probabilities for each quantized value and produces an appropriate code based on these probabilities so that the resultant output code stream will be smaller than the input stream [9]. The entropy encoders there used in this research are run-length encoder (RLE) and shift encoder.

It is important to note that a properly designed quantizer and entropy encoder are absolutely necessary along with optimum signal transformation to get the best possible compression.

I. Run Length Encoder

Data often contains sequences of identical bytes. By replacing these repeated byte sequences with the number of occurrences, a substantial reduction of data can be achieved. This is known as run-length coding [10]. In this research the run length coding is applied on the zero coefficients only that exist in the high band.

II. Shift Encoder

The proposed codec scheme is a variable length coding which gives a few numbers of bits to the short code word and many numbers of bits to the long code word. The main idea behind the shift coding algorithm is to find the max wavelet coefficient in the data set and optimize this coefficient to take a small numbers of bits. The other coefficient within the set is coded with the same number of bits [11].

The pseudo code for encoding and decoding scheme is as follow:

- Find the max no. in the set (positive wavelet coefficient).
- Compute the number of bits of the max called (n).
- Compute the histogram of the set coefficient.
- Optimize the total size of (short and long codeword).

Encoder:

```
P = 2^optimal codeword - 1
FOR I = 0 to NoCoef  DO
   IF Coef(I) < P THEN
        Output Coef(I), optimal codeword
   ELSE
       Output P, optimal codeword
       Output Coef(I) – P, n
   ENDIF
ENDFOR
```

Decoder:

```
P = 2^optimal codeword - 1
FOR I = 0 to NoCoef  DO
   Coef(I) = Get(optimal codeword)
   IF Coef(I) = P THEN
     Coef(I)=Coef(I) + Get(n)
      ENDIF
   ENDFOR
```

3 TEST MEASURES

In this research there are some measures used to test the results these measures are (PSNR) Peak Signal to Noise Ratio, (CR) Compression Ratio and (MSE) Mean Square Error [12].

$$\text{MSE}=\frac{1}{size} \sum_{i=0}^{size-1}(R_i - P_i)^2 \quad (3)$$

Where
R= Reconstructed audio.
P= Original audio.
Size= number of audio samples

$$\text{PSNR}=10 \log_{10} \frac{255^2}{\text{MSE}} \quad (4)$$

$$\text{Compression Ratio(C.R)}=\frac{Uncompressed\ \ File\ Size}{Compressed\ \ File\ Size} \quad (5)$$

4 THE RESULTS

There are some control parameters used in this research, such as quantization step (Q) and number of levels (Nolevel). These parameters effects on the compression ratio and mean square error and peak signal to noise ratio.

4.1 Quantization Step

To reduce the number of bits that needed to the transformed coefficients the quantization process must apply. In this research different scale quantization used into different wavelet subbands. The smallest band used small quantization value. With Sound1 has size (70.3 kB), Sound2 has size (40.2 kB) and Sound3 has size (74.9 kB).   Table(1) shows the effect of quantization step on the compression method.

TABLE 1

THE EFFECT OF QUANTIZATION STEP (Q) ON THE COMPRESSION METHOD.

Sample	No.level	Q	C.R	MSE	PSNR
Sound1	5	20	11.26	5.08	41.07
Sound1	5	25	12.35	6.91	39.73
Sound1	5	30	13.40	8.84	38.66
Sound2	5	20	10.67	6.54	39.97
Sound2	5	25	11.92	8.88	38.64
Sound2	5	30	13.11	11.19	37.64
Sound3	5	20	7.84	7.56	39.34
Sound3	5	25	8.91	10.53	37.90
Sound3	5	30	9.65	13.85	36.71

4.2 Number of Levels

The number of decompositions in the wavelet transform that effects on the compression. In this research the number of levels that used is 3, 4 and 5. Table(2) shows the effect of number level on the compression method.

TABLE 2

THE EFFECT OF NUMBER LEVELS (NO.LEVEL) ON THE COMPRESSION METHOD.

Sample	No.level	Q	C.R	MSE	PSNR
Sound1	3	25	5.36	2.53	44.08
Sound1	4	25	8.58	4.38	41.70
Sound1	5	25	12.35	6.91	39.73
Sound2	3	25	5.37	4.02	42.08
Sound2	4	25	8.11	6.11	40.26
Sound2	5	25	11.92	8.88	38.64
Sound3	3	25	4.61	4.90	41.22
Sound3	4	25	6.50	7.36	39.45
Sound3	5	25	8.91	10.53	37.90

5 CONCLUSIONS

The Discrete Wavelet Transform provides a multiresolution representation of signals. The transform can be implemented using filter bank. In this research, tap 9/7 biorthogonal filter was used, and a multiple level scale quantization which applies different scale quantization into different wavelet subbands. An effective algorithm of run length encoder and shift coding is also included to enhance the compression ratio. The performance of this method depended on some control parameters such as quantization step value and number of levels. From the tested results the compression ratio (C.R) and (PSNR) are depending on these parameters. The results shows the (C.R) and (MSE) increased when the quantization step value and number of levels are increased, but the (PSNR) will decrease. The best result shown in the table (1) in the sound1 when the number level is 5 and quantization step value is 30, the compression ratio (C.R) was 13.4 and the PSNR was 38.66. So the increase of number levels and quantization step value will effect on the quality of audio and lead to the distortion in the audio. Despite the differing sizes of the acoustic models used, but it did not affect the efficiency of the ratio compression. It was observed that whenever the sound was loud or uproarious, efficiency compression ratio decreased.

REFERENCES

[1] Chakresh, K.; Chandra, S.; Ashu ,S.; Bindu, T.; "Implementation of Audio signal by using wavelet transform" ; International Journal of Engineering Science and Technology; Vol. 2(10); pp. 4972-4977; 2010.

[2] Hatem, E.; Mustafa, J.; Mohammed, T.; " Speech Compression Using Wavelets"; Proceeding of the International Arab Conference on Information Technology; ACIT 2003, 2002.

[3] Dhubkarya, D.C; Sonam, D.; "High Quality Audio Coding at Low Bit Rate Using Wavelet and Wavelet Packet Transform"; Journal of Theoretical and Applied Information Technology; Vol 6; No.2; pp. 194-200; 2009.

[4] Iiya, P.; "Audio Compression using Wavelet compression Techniques"; Project Report; ECE 648; Purdue University; Spring 2005.

[5] Kavish, S.; Srinivasan, S.; "Vlsi Implementation of 2-D DWT/IDWT Cores Using 9/7-Tap Filter Banks Based On The Non-Expensive Symmetric Extension Scheme"; Design Automation Conference, 2002. Proceedings of ASP-DAC 2002. 7th Asia and South Pacific and the 15th International Conference on VLSI Design. Proceedings.

ASP-DAC (Asia and South Pacific - Design Automation Conference; pp. 435-440; 2002

[6] Yu-Shen, C.; Gen-Dow, H.; "Audio/Video Compression Applications using Wavelets"; Neural Networks, 2002. IJCNN '02. Proceedings of the 2002 International Joint Conference on; Vol (3); pp. 2214-2218; 2002.

[7] Deepika, S.; "Efficient Implementations of Discrete Wavelet Transforms Using FPGAs"; Msc. Thesis; Department of Electrical and Computer Engineering; College of Engineering; The Florida State University; 2003.

[8] Swastik, D.; Rasmi, R., S.; "Digital Image Compression Using Discrete Cosine Transform & Discrete Wavelet Transform"; Department of Computer Science and Engineering National Institute of Technology Rourkela Rourkela-769 008; Orissa; India; 2009.

[9] Gray, R., M.; Neuhoff, D., L.; "Quantization", IEEE Trans. Inform. Theory; Vol. 44; No.6; 1998.

[10] Ralf, S.; Klara, N.; "Multimedia Fundamentals: Media Coding and Content Processing"; Prentice Hall; Vol (1); Second Edition; 2002.

[11] Aree, A., M.; Loay, E., G.; " Intraframe Compression Using Lifting Scheme Wavelet-Based Transformation (9/7-Tap Filter)"; (JZS) Journal of Zankoy Sulaimani; Vol 11(1); Part A; pp. 53-60; December 2008.

[12] David, S.; "Data Compression the Complete Reference"; Third Edition, 2004.

_____

• *Wafaa Shihab Ahmed, Department of Shari' a, College of Education for Women, Al-iraqia University, Baghdad, Iraq. E-mail: flower_wa_sh_82@yahoo.com*