HIGH QUALITY AUDIO CODING AT LOW BIT RATE USING WAVELET AND WAVELET PACKET TRANSFORM

DR. D.C. DHUBKARYA [1] AND SONAM DUBEY [2]
Email at: sonamdubey2000@gmail.com,
Electronic and communication department
Bundelkhand Institute of Engineering and Technology, Jhansi (U.P.)

ABSTRACT

Audio coding is widely used in application such as digital broadcasting, Internet audio or music database to reduce the bit rate of high quality audio signal without comprising the perceptual quality.

In this paper work high quality audio codec at low bit rate using wavelet transform and improvement in reconstructed wave using post filtering has been proposed. The major issues concerning the development of audio codec are choosing optimal wavelets for audio signals, decomposition level in the digital wavelet transform and threshold criteria for coefficient truncation which is the basis to provide compression ratio for audio with suitable peak signal to noise ratio (PSNR), wavelet packet compression technique has also been used to compare the performance of audio codec using wavelet transform. After reconstructing the audio signal a post filtering technique is used to improve the quality of reconstructed audio signal. The proposed audio codec has been implemented in MATLAB 7.0 and various audio signals of different time duration have been tested. Result obtained show that the proposed codec improves quality of the reconstructed audio signal after post filtering.

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. 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 [1]. The basic task of high quality audio coding system is to compress the digital audio data in a way that [2] the compression is as efficient as possible and the reconstructed (decoded) audio sounds as close as possible to the original audio before compression.

Emerging digital audio applications in networks, wireless, and multimedia computers face serious shortfalls such as bandwidth limitations, and limited storage capacity.

These technologies have created a demand for high quality audio that can be transferred and stored at low bit rates. This creates a need of compression, whose role is to minimize the number of bits needed to retain acceptable quality of the original source signal [3].

Considerable interest has arisen in recent years regarding wavelet as a new transform technique for both image and audio processing applications. Like other transform coding techniques, 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 [4]. Because wavelet transforms are both computationally efficient and inherently local (i.e. their basis functions are limited in duration). First wavelet

based signal processing algorithm was given by David Marr in 1980's.

The Wavelet Transform provides a time-frequency representation of the signal. Better results can be obtained using wavelet analysis. Wavelet transform is breaking up of signals in shifted and scaled versions of original or mother wavelet. In wavelet transform every spectral component is not resolved equally as in STFT but analyzed at different frequency with different resolutions as shown in Fig 1.1. The basis functions of the wavelet transform are known as wavelets in   Fig 1.2. A wavelet is a waveform of effectively limited duration that has an average value of zero. In general, a wavelet is a small wave that has finite energy concentrated in time [5].
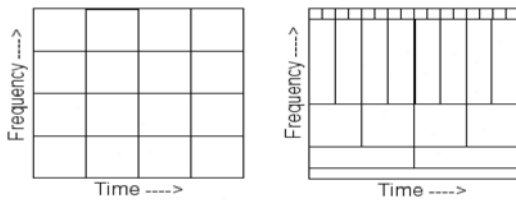

Fig 1.1: (a) STFT (b) Wavelet Multi resolution Analysis

Figure 1.2: Demonstration of a wavelet

They have their energy concentrated in time or space and are suited to analysis of transient signals. The wavelet transform uses a multi-resolution technique by which different frequencies are analyzed with different resolutions.

There are two types of wavelet transform. Continuous wavelet transforms (CWT) and discrete wavelet transforms (DWT). The main idea about the wavelet transform is the same in both of these transforms. However, they differ in the way the transformation is being carried out. But in this report transformation are done by discrete wavelet transform (DWT) and discrete wavelet packet transform (DWPT) because CWT computation may consume significant amount of time and resources, depending on the resolution required.

2. DISCRETE WAVELET TRANSFORM

The DWT, which is based on subband coding, is found to yield a fast computation of Wavelet Transform. It is easy to implement and reduces the computation time and resources required. In CWT, the signals are analyzed using a set of basis functions which relate to each other by simple scaling and translation. In the case of DWT, a time-scale representation of the digital signal is obtained using digital filtering techniques. The signal to be analyzed is passed through filters with different cutoff frequencies at different scales.

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, is commonly used in image compression. 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 [5].

Filters are one of the most widely used signal processing functions. Wavelets can be realized by iteration of filters with rescaling. The DWT is computed by successive low pass and high pass filtering of the discrete time-domain signal as shown in Fig 2.1. This is called the Mallat algorithm or Mallat-tree decomposition. In this figure, the signal is denoted by the sequence $x[n]$, where n is an integer. The low pass filter is denoted by G_0 while the high pass filter is denoted by H_0. At each level, the high pass filter produces detail information $d[n]$, while the low pass filter associated with scaling function produces coarse approximations $a[n]$.
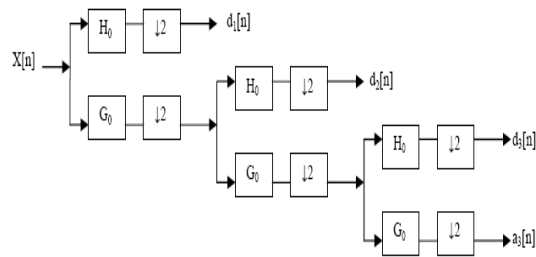

Figure 2.1: Three-level wavelet decomposition tree

At each decomposition level, the half band filters produce signals spanning only half the frequency

band. This doubles the frequency resolution as the uncertainty in frequency is reduced by half. In accordance with Nyquist's rule if the original signal has a highest frequency of ω, which requires a sampling frequency of 2ω radians, then it now has a highest frequency of ω/2 radians. It can now be sampled at a frequency of ω radians thus discarding half the samples with no loss of information. This decimation by 2 halves the time resolution as the entire signal is now represented by only half the number of samples. Thus, while the half band low pass filtering removes half of the frequencies and thus halves the resolution, the decimation by 2 doubles the scale. The filtering and decimation process is continued until the desired level is reached. The maximum number of levels depends on the length of the signal. The DWT of the original signal is then obtained by concatenating all the coefficients, a[n] and d[n], starting from the last level of decomposition.
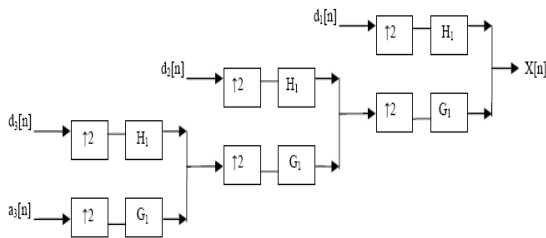

Figure 2.2 Three-level wavelet reconstruction tree

Fig 2.2 shows the reconstruction of the original signal from the wavelet coefficients. The approximation and detail coefficients at every level are upsampled by two, passed through the low pass and high pass synthesis filters and then added. This process is continued through the same
number of levels as in the decomposition process to obtain the original signal. The Mallat algorithm works equally well if the analysis filters, G_0 and H_0, are exchanged with the synthesis filters, G_1 and H_1.

3. DISCRETE WAVELET PACKET TRANSFORM

The wavelet packet method is a generalization of wavelet decomposition that offers a richer range of possibilities for signal analysis. In wavelet analysis, a signal is split into an approximation and a detail coefficient. The approximation coefficient is then itself split into a second-level approximation coefficients and detail coefficients, and the process is repeated. In wavelet packet analysis, the details as well as the approximations can be split. This yields

more than $2^{2^{n-1}}$ different ways to encode the signal. Fig 3.1 shows the level 3 decomposition using wavelet packet transform.
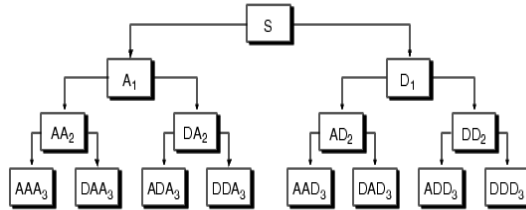

Figure 3.1: Level 3 decomposition using wavelet packet transform

In wavelet packet analysis, an entropy-based criterion is used to select the most suitable decomposition of a given signal. This means we look at each node of decomposition tree and quantify the information to be gained by performing each split [6].The wavelets have several families. The most important wavelets families are Haar, Daubechies, Symlets, Coiflets , Biorthogonals, reverse Biorthogonal, Meyers, discrete approximation of Meyer wavelets, Gaussian, Maxican hat wavelets, Morlets, complex Gaussian, Shannons, frequency B-Spline wavelets and Complex Morlet wavelets.
Out of these wavelet families, Haar, Daubechies wavelets, Symlets, Coefilets and biorthogonal wavelet families are the most important wavelet families.

4. AUDIO CODING TECHNIQUE

The low bit rate audio codec using wavelet transform is shown in Fig 4.1. The major issues concerning the development of codec are choosing optimal wavelets, decomposition level and threshold criteria for coefficient truncation to provide low bit rate with suitable peak signal to noise ratio (PSNR). Wavelet packet coding technique has also been used to compare the performance of audio codec using wavelet transform.
After reconstructing the signal, there is always some coding error, which degrades the quality of reconstructed audio signal. The post filtering technique is used to enhance the perceptual quality of audio signal coded at low bit rates. In post filtering the reconstruction error of the coded audio signal is estimated and subtracted from the coded audio signal, so that the noise level in the coded audio signal is suppressed and hence better perceptual quality is achieved [7].
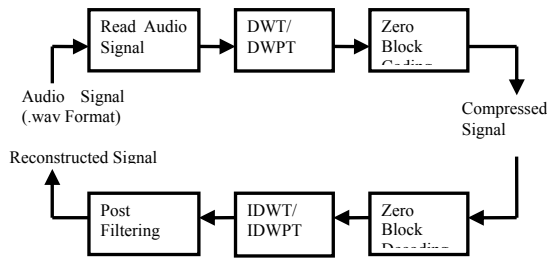
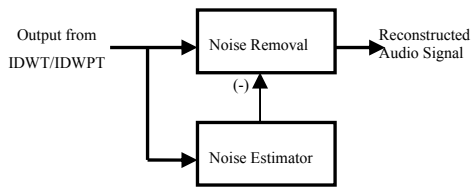**Figure 4.1: System Model to be implemented**

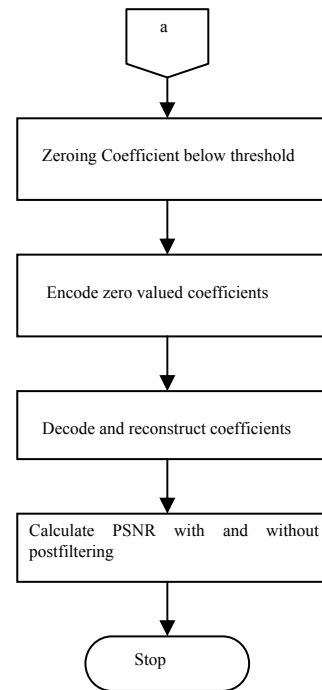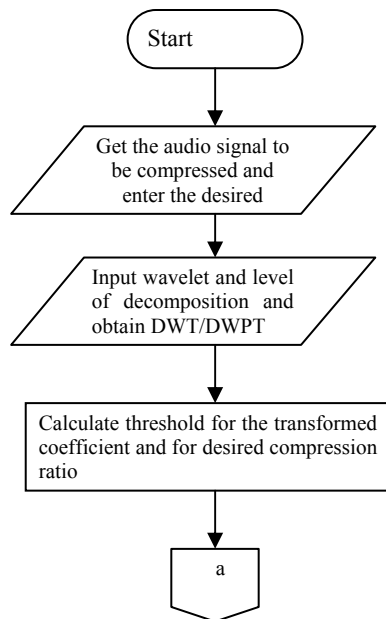**Figure 4.2 Basic Structure of Post filtering**

Figure 4.3 Flow Chart of Audio Codec

4.1 Comparison criteria between DWT and DWPT and between different levels of wavelets

When the level of information loss is expressed as a function of the original and processed audio signal, it is said to be an objective fidelity criteria. The following parameters fall under this criteria.

1. PSNR
2. Compression Ratio
3. Bit Rate

4.1.1 psnr

Peak signal to noise ratio (PSNR) is one of the most important parameters used to estimate the quality of reconstructed audio signal with respect to original audio signal, it is given as

$$PSNR = 10 \log_{10}(NX^2/(x-y)^2)$$

Where N is the length of reconstructed audio signal, X is the maximum absolute square value of original audio signal. x and y are original and reconstructed audio signal respectively [8].

4.1.2 compression ratio

The compression ratio of a coder is usually defined as $R_{cr} = S/C$ (dimensionless)

Where S is the size of the source file and C is the size of the compressed file.

4.1.3 bit rate

Bit rate is used for indicating the transfer speed of a data or transfer speed in general. Bit Rate indicates the number of bits transmitted in one second.

5. SIMULATION AND RESULTS

We perform simulations in C language on MATLAB7.0. The performance of the audio codec is evaluated by considering different parameters such as decomposition levels, optimal wavelets and threshold value for wavelet coefficients to obtain low bit rate signal. PSNR is also calculated by varying above parameters which affect the quality of reconstructed signal. The audio signals have been tested for different wavelets (Haar, Daubechies, Symlets, Coiflets and Biorthogonal) function at level 3, 5, 7 and 9, of which Coiflets and Biorthogonal wavelets gives significant improvement in PSNR at level 5. Therefore results have been shown only for coiflets and biorthogonal wavelets for all tested signal. The input audio signals have been compressed at different threshold value for different bit rates and compression ratio.

The test signal 'audio3.wav'of size 4.86 MB is formed by converting the MP3 file of audio into wav file by 'Meda mp3 splitter' software. The 'audio3.wav' has 2552407 sampled data of 16 bits/sample with sampling frequency 44.1 kHz. The input signal has bit rate of 705.6 kbps, and 57 seconds long duration. Since the amplitude values of sampled data are in the range from [-1, +1], so the input audio has been tested at threshold value of 0.21, 0.175, 0.150, 0.115, 0.091 and 0.075 which provide the bit rate of 20, 30, 40, 60, 80 and 100 kbps respectively after encoding the signal.

The audio codec was tested for different wavelet functions with different level, of which only bior3.9 and coif5 gives significant improvement in PSNR for 'audio3.wav' as shown in Figure 5.1.

Figure 5.2 shows a graphical representation for the comparison of PSNR values at decomposition level 5 for coif5 wavelet using DWT for 'audio3.wav' with and without postfiltering. There is small improvement in PSNR with postfiltering.
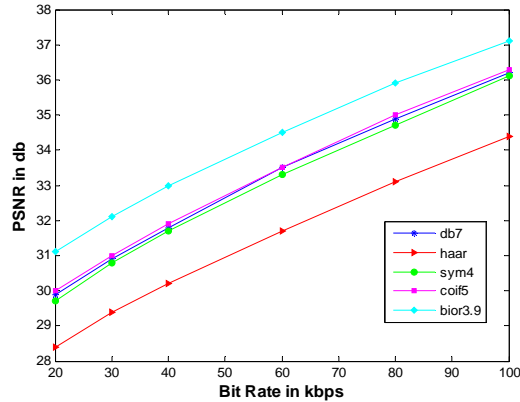

Figure 5.1: Graph between PSNR and bit rate using DWT for different wavelet at level 5 for 'audio3.wav'.

Figure 5.3 shows a graphical representation for the comparison of PSNR values at decomposition level 5 for coif5 wavelet using DWT and DWPT for 'audio3.wav'. There is some improvement in PSNR above 60 kbps when wavelet packet transform is used for level 5.

Figure 5.4 shows a graphical representation for the comparison of PSNR values at decomposition level 5 for bior3.9 wavelet using DWT for 'audio3.wav' with and without postfiltering. There is small improvement in PSNR with postfiltering above 50 kbps. There is small improvement in PSNR with wavelet transform over wavelet packet transform above 30 kbps for level 5; at level 3 the value of PSNR is same for both wavelet transform as well as wavelet packet transform.
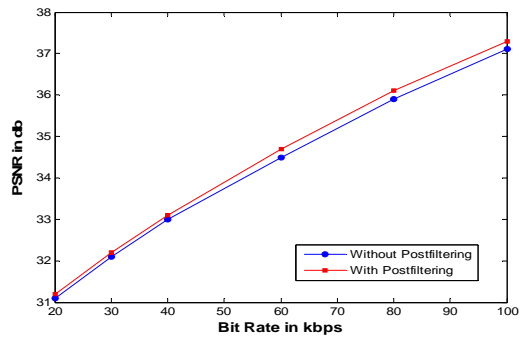

Figure 5.2: Graph between PSNR and bit rate for level 5 using coif5 wavelet for 'audio3.wav' with and without postfiltering
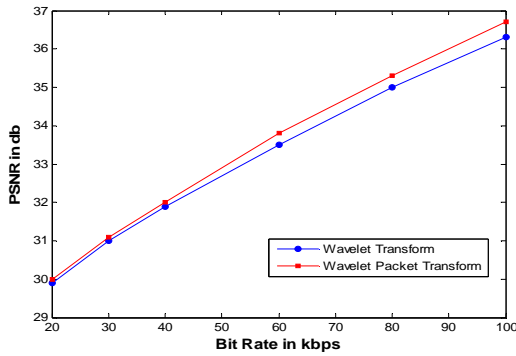
Figure 5.3: Graph between PSNR and bit rate using DWT and DWPT for level 5 using coif5 wavelet for 'audio3.wav'.

Figure 5.5 shows a graphical representation for the comparison of PSNR values at decomposition level 5 for bior3.9 wavelet using DWT and DWPT for 'audio3.wav'.
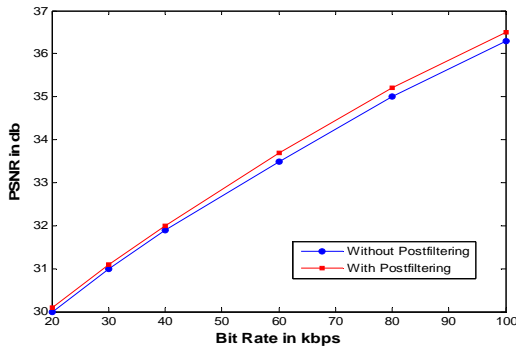

Figure 5.4: Graph between PSNR and bit rate for level 5 using bior3.9 wavelet for 'audio3.wav' with and without postfiltering
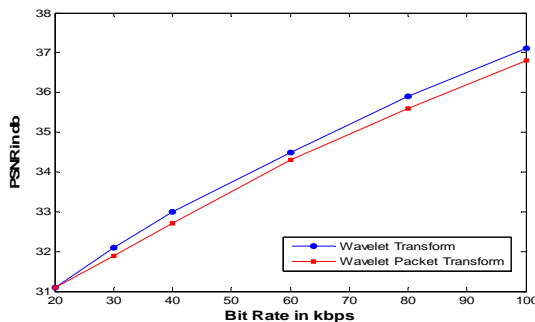

Figure 5.5: Graph between PSNR and bit rate using DWT and DWPT for level 5 using bior3.9 wavelet for 'audio3.wav'.

Figure 5.6 shows Original and reconstructed signal 'audio3.wav' using bior3.9 wavelet for level 5 at 20 and 40 kbps. At 20 kbps quality of reconstructed signal is not good due to higher compression ratio whereas at 40 kbps the quality of reconstructed signal is good and it is comparable with original audio signal.
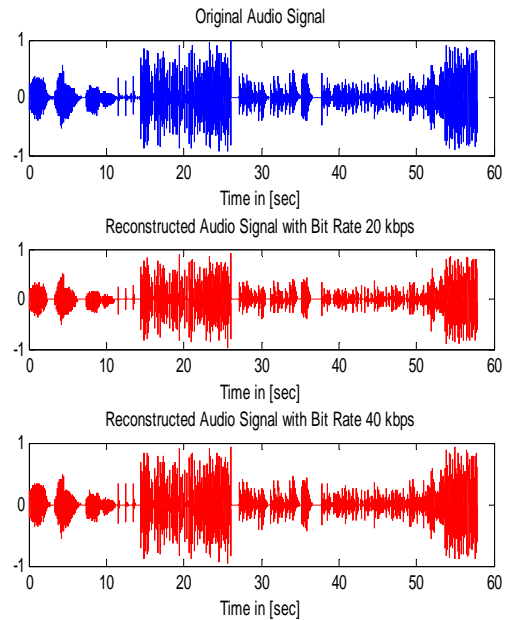

Figure 5.6: Original and Reconstructed 'audio3.wav'

The results show that wavelet packet transform improves the quality and PSNR of reconstructed audio signal for all the wavelets except bior3.9 over wavelet transform. Postfiltering also improves the quality of reconstructed audio signal. For all signals, level 5 gives better results. The comparative analysis of the results show that for good quality reconstructed signal, the bit rates of the proposed codec should be in the range of 40 - 60 kbps with PSNR values 30.8047 – 34.6779 dB respectively.

6. CONCLUSION

In this paper audio codec at low bit rate using wavelet transform and wavelet packet transform has been developed, which is simple yet effective compression technique. The codec successfully improves the quality of the reconstructed audio signal by using postfiltering at suitable bit rates. To test the codec several audio signals of different time durations have been used.

For the same compression ratio better PSNR values have been obtained with bior 3.9 wavelet. Thus bior3.9 wavelet has been chosen for the proposed codec. It has been observed that the optimum number of wavelet decomposition level is 5. Since high value of wavelet decomposition level require more computation time, moreover it does not improve the quality of the signal, whereas lower levels provide less compression ratio and less PSNR. So level 5 is preferred since it takes less computation time and provides a better compression ratio and PSNR. It shows that the wavelet packet transform instead of wavelet transform provides the improvement in PSNR of the reconstructed audio signal by 0.3 dB for all wavelets function except bior3.9.

It has been observed that postfiltering improves the quality of the reconstructed audio signal with wavelet packet transform as well as with wavelet transform.

8. REFERENCES

[1]	Yuan-Hao Huang and Tzi-Dar Chiueh, "*A New Audio Coding Scheme Using a Forward Masking Model and Perceptually Weighted Vector Quantization*", IEEE transactions on speech and audio processing, vol.10, no. 5, pp.325-331,July 2002,.

[2]	Karlheinz Brandenburg & Fraunhofer IIS Arbeitsgruppe, "*Low Bit Rate Audio coding - State of the Art Challenges and Future Directions*", Ilmenau Technical University Germany.

[3]	W. Kinser. "*Compression and It's Metrics for Multimedia*", Proceedings of the First IEEE International Conference on Cognitive Informatics (ICCE'02),pp.1-15,2002.

[4]	Boon-Lum Lim and Zi-Lu Ying, "*Performance Analysis of Audio Signal Compression Based on Wavelet and Wavelet Packet Transforms*", International Conference on Information, Communications and Signal Processing, Singapore, pp.735-739,9-12 September 1997 IEEE,.

[5]	Amara Graps, "*An Introduction to Wavelets*",published by the IEEE computer society, vol. 2, no. 2, pp.50-61 Summer 1995.

[6]	MATLAB7.0, www.mathwork.com.

[7]	Yu Rongshan. "*Improving Quality of Low Bit Rate Audio Coding by Using Short-Time Spectral Attenuation*", IEEE International Conference on Multimedia and Expo, pp.185-188, 2001.

[8]	Khaled N. Hamdy, Murtaza Ali and Ahmed H. Tewfik, "*Low Bit Rate High Quality Audio Coding With Combined Harmonic and Wavelet Representations*", Dept. of Electrical Engineering University of Minnesota, Minneapolis, pp.1045-1048, 1996 IEEE.