Journal of Information Engineering and Applications www.iiste.org
ISSN 2224-5782 (print) ISSN 2225-0506 (online)
Vol.4, No.4, 2014

Audio Compression Using DCT and DWT Techniques

Jithin James [1], Vinod J Thomas [2]

[1]PG Scholar, [2]Assistant Professor Department of Electronics and Communication Engineering

Vimal Jyothi Engineering College Chemperi, India.

Abstract

In today's world multimedia files are used, storage space required for these files is more and sound files have no option so ultimate solution for this is compression. Compression is nothing but high input stream of data converted into smaller size. Speech Compression is a field of digital signal processing that focuses on reducing bit-rate of speech signals to enhance transmission speed and storage requirement of fast developing multimedia. In many applications, such as the design of multimedia workstations and high quality audio transmission and storage, the goal is to achieve transparent coding of audio and speech signals at the lowest possible data rates. Therefore, the transmission and storage of information becomes costly. However, if we can use less data, both transmission and storage become cheaper. Further reduction in bit rate is an attractive proposition in applications like remote broadcast lines, studio links, satellite transmission of high quality audio and voice over internet. This paper explores a transform based methodology for compression of the speech signal. In this methodology, different transforms such as Discrete Wavelet Transform (DWT), Fast Fourier Transform (FFT) and Discrete Cosine Transform (DCT) are exploited. A comparative study of performance of different transforms is made in terms of Signal-to-noise ratio (SNR) and Peak signal-to-noise ratio (PSNR). The mean compression ratio is also calculated for all the methods and compared. The simulation results included illustrate the effectiveness of these transforms in the field of data compression.

Keywords-DCT (Discrete cosine transform), DWT (Discrete wavelet transform), Quantization Compression Factor (CF), Signal to Noise ratio (SNR).

I . INTRODUCTION

Speech is very basic way for humans to convey information. The main objective of Speech is communication. Speech can be defined as the response of vocal track to one or more excitation signal. Huge amount of data transmission is very difficult both in terms of transmission and storage. Speech Compression is a method to convert human speech into an encoded form in such a way that it can later be decoded to get back the original signal. Compression is basically to remove redundancy between neighboring samples and between adjacent cycles. Major objective of speech compression is to represent signal with lesser number of bits. The reduction of data should be done in such a way that there is acceptable loss of quality.

II. TECHNIQUES FOR SPEECH COMPRESSION

Transform Coding

This is the coding technique that we have used for our paper. In this method the signal is transformed into frequency domain and then only dominant feature of signal is maintained. In transform method we have used discrete wavelet transform technique and discrete cosine transform technique.When we use wavelet transform technique, the original signal can be represented in terms of wavelet expansion. Similarly in case of DCT transform speech can be represented in terms of DCT coefficients. Transform techniques do not compress the signal, they provide information about the signal and using various encoding techniques compressions of signal is done. Speech compression is done by neglecting small and lesser important coefficients and data and discarding them and then using quantization and encoding techniques.
Speech compression is performed in the following steps.

- Transform technique
- Thresholding of transformed coefficients
- Quantization
- Encoding

Transform Technique

DCT and DWT methods are used on speech signal. Using DCT, reconstruction of signal can be done very accurately; this property of DCT is used for data compression. Localization feature of wavelet along with time frequency resolution property makes DWT very suitable for speech compression. The main idea behind signal compression using wavelets is linked primarily to the relative scarceness of the wavelet domain representation of signal.

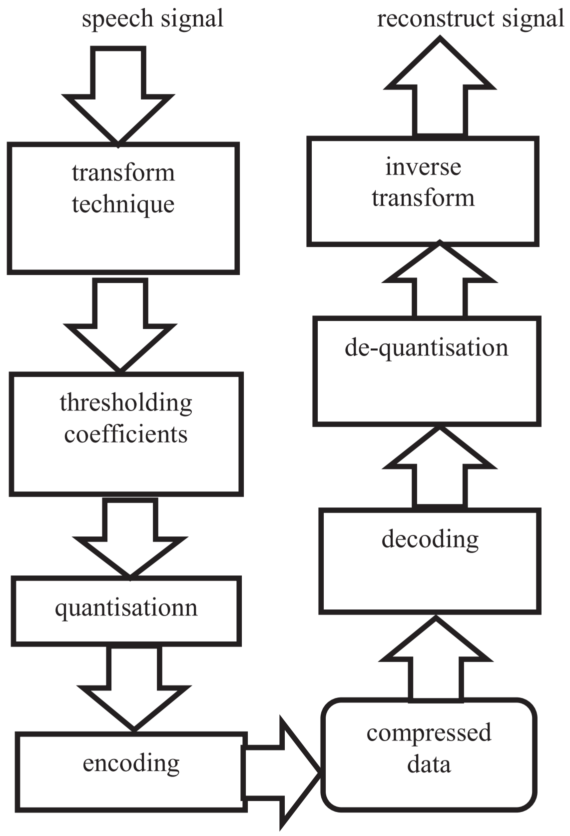

Fig.1- Block diagram of speech compression

1. Discrete Wavelet Transform

A discrete wavelet transform can be defined as a „small wave" that has its energy concentrated in time, and it provides a tool for the analysis of transient, non-stationary or time varying phenomenon. It has oscillating wave like property. Wavelet is a waveform of limited duration having an average value zero. They are localized in space. Wavelet transform provides a time-frequency representation of the signal. In DWT, the signal is decomposed into set of basic functions also known as „WAVELETS". Wavelets are obtained from a single MOTHER WAVELET by delay and shifting.

$$\varphi(t) = \frac{1}{\sqrt{a}}\varphi\frac{(t-b)}{a}$$

Where "a" is the scaling parameter and „b" is the shifting parameter.DWT uses multiresolution technique to analyze different frequencies. In DWT, the prominent information in the signal appears in the lower amplitudes. Thus compression can be achieved by discarding the low amplitude signals. Localization feature of wavelet along with time frequency resolution property makes DWT very suitable for speech compression. The main idea behind signal compression using wavelets is linked primarily to the relative scarceness of the wavelet domain representation of signal.

2. Discrete Cosine Transform

Discrete Cosine Transform can be used for speech compression because of high correlation in adjacent coefficient. We can reconstruct a sequence very accurately from very few DCT coefficients. This property of DCT helps in effective reduction of data.
DCT of 1-D sequence x (n) of length N is given by

$$X(m)=\left[\frac{2}{N}\right]^{0.5} C_m \sum_{m=0}^{N-1} X(n)\cos\left[\frac{(2n+1)m\pi}{2N}\right]$$

Where m=0, 1, - - - - - -, N-1

The inverse discrete cosine transform is

$$X(n)=\left[\frac{2}{N}\right]^{0.5} \sum_{m=0}^{N-1} C_m * X(m)\cos\left[\frac{(2n+1)m\pi}{2N}\right]$$

In both equations Cm can be defined as

Cm = (1/2)1/2 for m= 0

= 1 for m≠ 0

It expresses a finite sequence of data points in terms of sum of cosine function oscillating at different frequencies.They are very common encoding technique for audio track compressions. It is very similar to DFT, but the only difference is that the output vector is approximately twice as long as the DFT output. They are used in JPEG image compressions, MJPEG and many other video compressions.

A)Thresholding

After the coefficients are received from different transforms, thresholding is done. Very few DCT coefficients represent 99% of signal energy; hence Thresholding is calculated and applied to the coefficients. Coefficients having values less than threshold values are removed.

B) Quantization

It is a process of mapping a set of continuous valued data to a set of discrete valued data. The aim of quantization is to reduce the information found in threshold coefficients. This process makes sure that it produces minimum errors. We basically perform uniform quantization process.

C) Encoding

We use different encoding techniques like Run Length Encoding and Huffmann Encoding. Encoding method is usedto remove data that are repetitively occurring. In encoding we can also reduce the number of coefficients by removing the redundant data. This helps in reducing the bandwidth of the signal hence compression can be achieved. the compressed speech signal can be reconstructed to form the original signal by decoding followed by dequantization and then performing the inverse-transform methods. This would reproduce the original signal.

III. WAVELET BASED COMPRESSION TECHNIQUES

Wavelets concentrate speech signals into a few neighbouring coefficients. By taking the wavelet transform of a signal, many of its‟ coefficients will either be zero or have negligible magnitudes. Data compression can then be done by treating the small valued coefficients as insignificant data and discarding them.

A.) Choice of wavelets

Choosing mother-wavelet function which is used in designing high quality speech coders is of prime importance.Choosing a wavelet having a compact support in time and frequency in addition to a significant number of vanishing moments is important for wavelet speech compressor. Different criteria can be used in selecting an optimal wavelet function. The objective is to minimize the error variance and maximize signal to noise ratio. Better reconstruction quality is provided by wavelets with more vanishing moments However the computational complexity of DWT increases with the number of vanishing moments. Hence it is not practical to use wavelets with higher number of vanishing moments.. Higher the number of vanishing moments, faster is the decay rate of wavelet coefficients. It leads to a more compact signal representation and hence useful in coding applications. However, length of the filters increases with the number of vanishing moments and the hence complexity of computing the DWT coefficients increases.

B.) Decomposition of wavelets

Wavelets decompose a signal into different resolutions or frequency bands. Signal compression is based on the concept that selecting small number of approximation coefficients and some of the detail coefficients can represent the signal components accurately. Choosing a decomposition level for the DWT depends on the type of signal being used.

C.) Truncation of coefficients

Compression involves truncating wavelet coefficients below threshold. Most of the speech energy is high-valued coefficient. Two different methods are available for the calculation of thresholds.

Global Thresholding- It takes the wavelet expansion of the signal and keeps the largest absolute value coefficient. In this we manually set a global threshold. Hence only a single parameter needs to be selected in this case.

Level Thresholding- It applies visually determined level dependent thresholds to each of the decomposition level in the wavelet transform.

D.) Encoding coefficients

Signal compression is achieved by first truncating small-valued coefficients and then encoding these coefficients. High-magnitude coefficients can be represented by storing the coefficients along with their respective positions in the wavelet transform vector. Another method for compression is to encode consecutive zero valued coefficient with two bytes. One byte indicates the sequence of zeros in the wavelet transforms vector and the second byte represents the number of consecutive zeros. For further data compression a suitable bit-encoding format can be used. Low bit rate representation of signal can be achieved by using an entropy coder like Huffman coding.

E.) Calculating threshold

Two different thresholding techniques are used for the truncation of coefficients i.e. global thresholding and level thresholding.

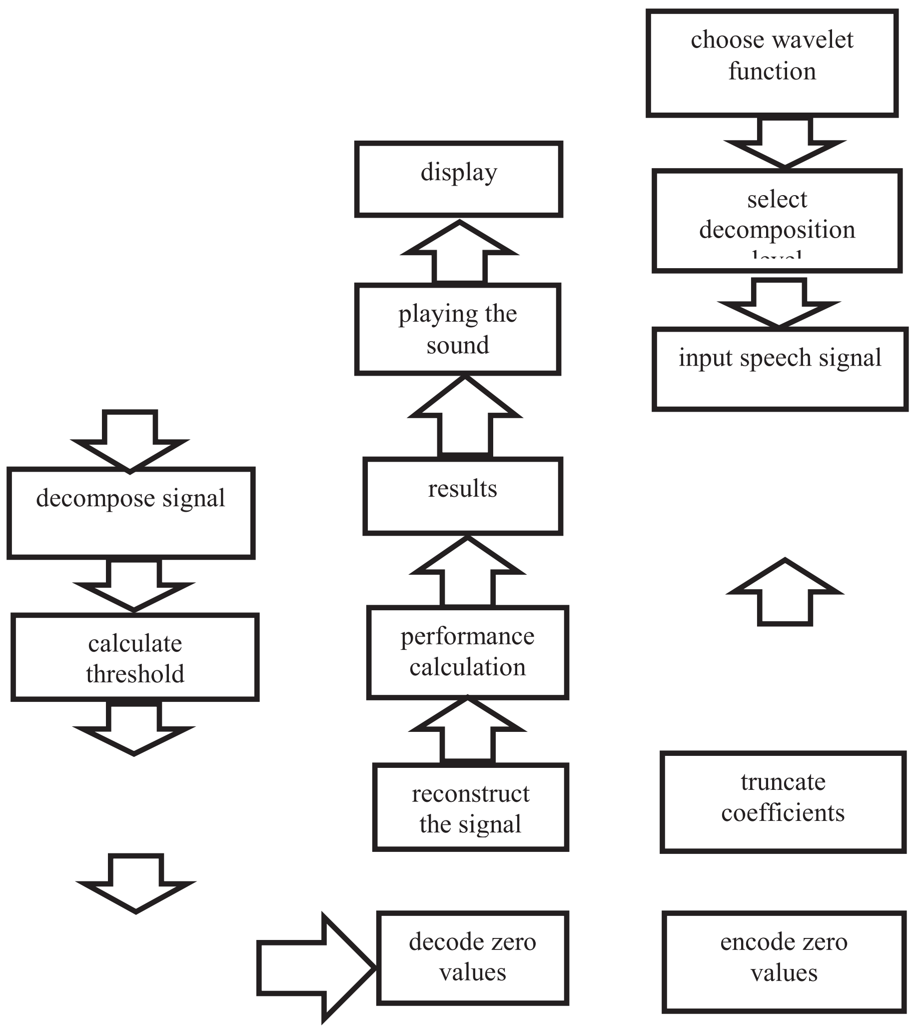

Fig.2 - Design Flow of Wavelet Based Speech Coder

F.) Encoding zero value functions

In this method, consecutive zero valued coefficients are encoded with two bytes. One byte specifies the starting string of zeros and the second byte keeps record of the number of successive zeros. This encoding method provides much higher compression ratio.

IV. DCT BASED COMPRESSION TECHNIQUE

The given sound file is read. The vector is divided into smaller frames and arranged into matrix form. DCT operation is performed on the matrix. DCT operation is performed and the elements are sorted in their matrix form to find components and their indices. The elements are arranged in descending order. After the arrangement has been done, a Threshold value is decided. The coefficients below the threshold values are discarded. Hence reducing the size of the signal which results in compression. The data is then converted back into the original form by using reconstruction process. For this we perform IDCT operation on the signal. Now convert the signal back to its vector form. Thus the signal is reconstructed. The fig.3 shows the DCT compression and decompression process.

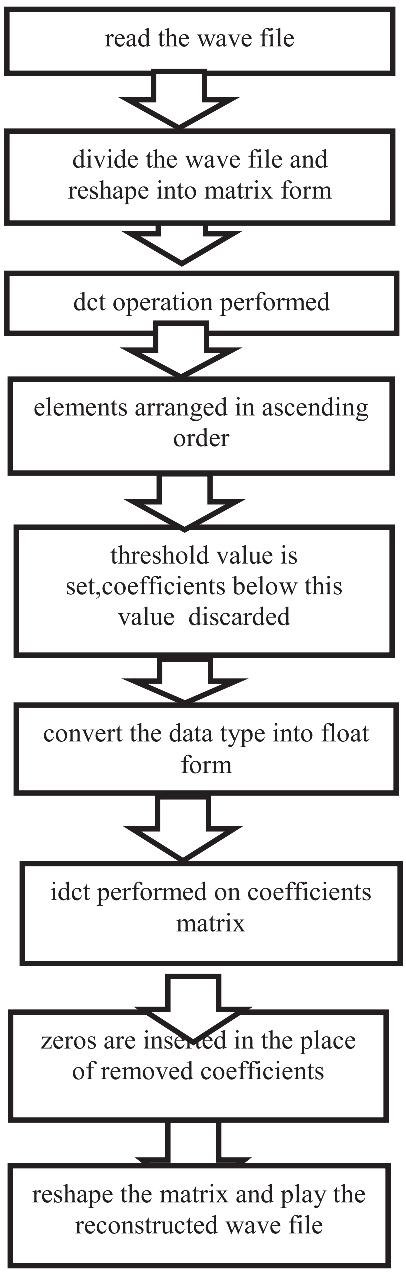

Fig.3- DCT based speech coder design flow

V. PERFORMANCE EVALUATION

To evaluate the overall performance of the proposed audio compression scheme, several objective tests were made.To measure the performance of the reconstructed signal, various factors such as compression factor and Signal to noise ratio are taken into consideration.

i) Compression Factor (CF)

We take into account all the values that would be required to completely represent the signal.

$$\text{compression factor} = \frac{\text{length of original signal}}{\text{length of compressed signal}}$$

ii) Signal to Noise Ratio (SNR)

$$\mathrm{SNR} = 10\log\left[\frac{\sigma x^2}{\sigma e^2}\right]^2$$

Where σx2 is the mean square of the speech signal and σe2 is the mean square difference between the original and reconstructed speech signal.

Signal	Compression factor (C.F)	SNR (dB)
funky2	2	76.6982
funky4	4	49.0264
funky8	8	30.1212

Table 1: Results of DCT based technique in terms of CF and SNR

VI. CONCLUSION

A simple discrete wavelet transform & DCT based audio compression scheme presented in this paper. It is implemented using MATLAB . Experimental results show that in general there is improved in compression factor & signal to noise ratio with DWT based technique. It is also observed that Specific wavelets have varying effects on the speech signal being represented.

REFERENCES

[1] HarmanpreetKaur and RamanpreetKaur, “Speech compression and decompression using DCT and DWT”, International Journal Computer Technology &Applications,Vol 3 (4), 1501-1503 IJCTA | July-August 2012
[2] Jalal Karam and RautSaad, “The Effect of Different Compression Schemes on Speech Signals”, International Journal of Biological and Life Sciences, 1:4, 2005
[3] O. Rioul and M. Vetterli, “Wavelets and Signal Processing”, IEEE Signal Process. Mag. Vol 8, pp. 14-38, Oct. 1991..
[4] HatemElaydi and Mustafi I. Jaber and Mohammed B. Tanboura, “Speech compression using Wavelets”, International Journal for Applied Sciences, Vol 2, 1-4,Sep 2011 Russel Lloyd Lim “Speech Compression using the Discrete WaveletTransform”
[5] Othman O. Khalifa, Sering Habib Harding & Aisha-Hassan A. Hashim “Compression using Wavelet Transform” in Signal Processing: An International Journal, Volume (2) : Issue (5).
[6] “A Tutorial of the Wavelet Transform” by Chun-Lin, Liu in February 23,2010. J.I. Agbinya, “Discrete Wavelet Transform Techniques in Speech Processing”, IEEE Tencon Digital Signal Processing Applications Proceedings, IEEE, New York, NY, 1996, pp 514-519.
[7] Ken C. Pohlmann “Principles of Digital Audio”, McGraw-Hill, Fourth edition, 2000.
[8] X. Huang, A. Acero & H-W. Hon “Spoken Language Processing: A Guide to Theory, Algorithm and System Development”, Pearson Education, 1st edition 2001

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

Recent conferences: http://www.iiste.org/conference/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar