Novel Approach of Implementing Psychoacoustic model for MPEG-1 Audio

Jagadeesh B .K[1] Dr B Siva Kumar[2]

[1]*Research Scholar, PRIST University, Thanjvur*
[2]*Prof and HOD, TCE Dept, Dr AIT, Bangalore*

ABSTRACT:*Audio compression is the technology of converting audio signal into an efficiently encoded representation that can later be decoded to engender a close approximation of the pristine signal. In this work we are investigating implementation of psychoacoustics utilizing different wavelets to compress high quality audio signal and maintaining transparent quality at low bit rates*
Most psychoacoustic models for coding applications utilize a uniform spectral decomposition to approximate the frequency selectivity of the human auditory system; however the equal filter properties of the uniform sub bands do not match the non uniform characteristics of the cochlear filters. For implementing this algorithm a design of psycho-acoustic model was developed following the model utilized in the standard MPEG-1 audio. The architecture is predicated on opportune Wavelet packet decomposition (DWPT) in lieu of Short Term Fourier Transformation (STFT) or FFT.

KEYWORDS: *Psycho-Acoustic, DWPT, STFT, FFT*

I. INTRODUCTION

Compression refers to a variety of ways to reduce the size of a data file. Audio compression is form of data compression. To obtain this different methods for compression have been designed; from the simplest one that consists to make a banal under sampling, to the most advanced that takes account of the sensitivity of the human auditory perceiver. These last methods, takes in account of the sedulous assiduousness of human aurally perceiving. The auditory perceiver presents in effect, some limit in auricularly discerning that let to eliminate some sound information not perceived in the pristine signal. However, the auditory perceiver is an organ of astronomically immense sensibility, presenting a high resolution and a great dynamic range of the signal: a lamentable filtering can lead to a loss of an aural quality. The MPEG/Audio is a standard for both transmitting and recording compressed ratio. The MPEG algorithm achieves compression by exploiting the perceptual inhibition of the human auditory perceiver. Audio compression algorithms are acclimated to obtain compact digital representations of high-fidelity audio signals for the purport of efficient transmission.
The main objective in audio coding is to represent the signal with a minimum number of bits while achieving transparent signal reproduction.

The majority of MPEG coders apply a psycho-acoustic model for coding applications utilizing filter bank to approximate the frequency selectivity of the human auditory system.Figure1(a) and Figure1(b) shows a diagram of the structure of a generic perceptual audio coder. Figure 1(a) shows the structure of the encoder, which has three main stages and a fourth bit stream formatting stage, and Figure1 (b) shows the decoder, which has three stages. The encoder operates on the input audio signal and outputs the encoded bit stream, and the decoder operates on the encoded bit stream and reconstructs the pristine signal. The three stages in the decoder, as a result, are reverse operations of three stages in the encoder. Namely, the signal analysis, quantization and encoding, and bit stream formatting stages of the encoder correspond to the signal synthesis, dequantization and decoding, and bit stream extraction stages of the decoder, respectively.

The extra stage in the encoder is the psychoacoustic model, which is not required in the decoder since the information is implicitly encoded as side-information. This means that perceptual coders are asymmetrical in that the encoder has a greater computational requirement than the decoder, which actually can be desirable in certain applications, where one \server" encodes the signal for many clients". The psychoacoustic model is based on many studies of human perception. These studies have shown that the average human does not hear all frequencies the same. Effects due to different sounds in the environment and limitations of the human sensory system lead to facts that can be used to cut out unnecessary data in an audio signal. Auditory masking is a well-known psycho acoustical phenomenon in which a weak signal is masked in the presence of a stronger masker signal. Exploiting this phenomenon in perceptual audio coding is achieved so that the original audio signal is treated as a masker for distortions introduced by lossy data.

he psychoacoustic model used in the perceptual audio coder is based on the Psychoacoustic Model 1from the MPEG-1 Audio Standard. The MPEG-1 Audio Standard describes two sample psychoacoustic models, the first being computationally simpler and suitable for coding at Higher bit rates and the second being more complex but also more reliable at lower bit rates. Due to the complexity associated with the construction of a psychoacoustic model, a Simplified version was considered .In this work we are implementing psychoacoustic model -1 using wavelet packets.

The discrete wavelet packet transform can conveniently decompose the signal into an auditory critical band-like partition [2][11].Signal decomposition into critical bands resulting from
Wavelet analysis needs to gratify the spectral resolution requisites of the human auditory system.
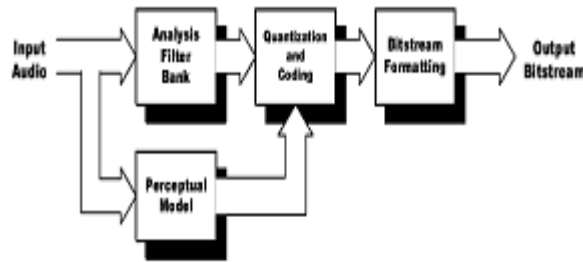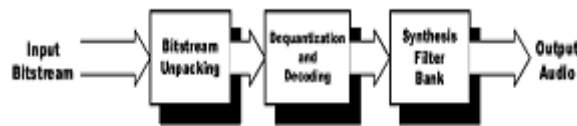

Figure 1(a).Encoder

Figure 1(b).Decoder

II. WAVELET PACKET TRANSFORM

The wavelet packet transform [4] constitutes a Solution that permits a finer an adjustable resolution of frequencies at high frequencies and gives a rich structure that allows adaptation to particular signals or signals classes [5]. Psychoacoustic model achieves an improved decomposition of the signal into 28 critical bands using the discrete wavelet packet transform (DWPT). This results in a spectral partition which approximates the critical band distribution much closer than before. Furthermore, the masking thresholds are computed entirely in the Wavelet domain.
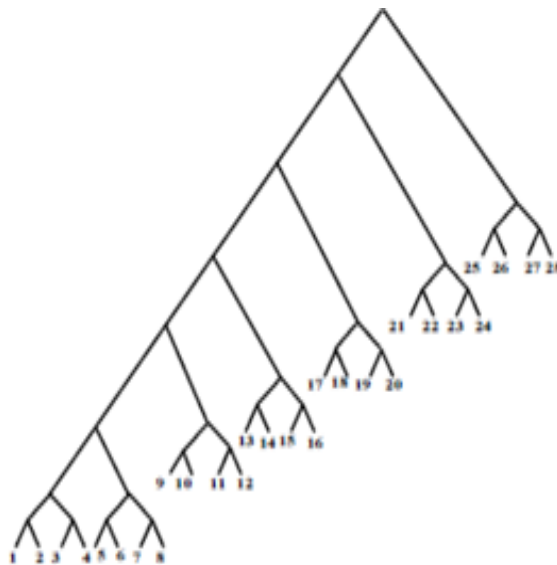

Figure 2.Wavelet Tree decomposition

Signal decomposition into critical bands resulting from wavelet analysis needs to satiate the spectral resolution requisites of the human auditory system. On the other hand, the cull of the wavelet substructure additionally is critical for meeting the required auditory temporal resolution, which ranges from less than 10ms at high frequencies to up to 100 ms at low frequencies [6]. Those constraints make those which are above the auditory perception absolute threshold. Individual masking threshold takes into account the masking threshold for each remaining component. Global masking threshold [3] is calculated by deducing tonal and non tonal components from the spectrum of the wavelet packet decomposition.

III. PROPOSED MODEL

In MPEG 1, psychoacoustic model plays a crucial role in audio compression. In audio compression technology, compression will be achieved by exploiting the auditory masking characteristic of human auditory perceiver. By applying psychoacoustic principles, it is possible to analyze the signal and computation of noise. Here, signal masking is a function of frequency. In this work, we are investigating implementation of psychoacoustics model-1 for MPEG-1 utilizing wavelet packet decomposition to compress high-quality audio signal. Most psychoacoustic models for coding applications utilize a uniform spectral decomposition to approximate the frequency selectivity of the human auditory system. For implementing this algorithm, a design of psychoacoustic model was developed following the model used in the standard MPEG-1 audio. In this work, . In this work, we have proposed we have proposed psychoacoustic model-
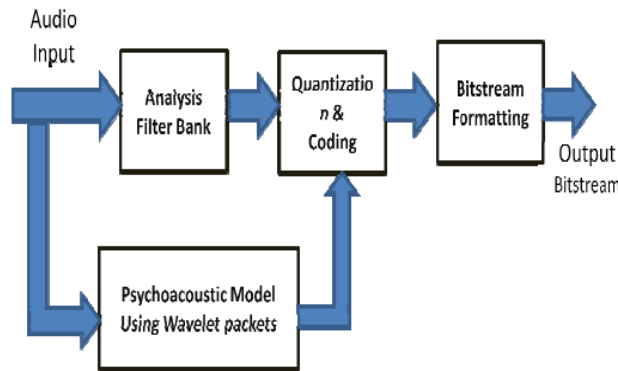

Figure.3. Perceptual Audio Encoder using wavelet based psychoacoustic model

implementation based on wavelet packet decomposition instead of conventional Fast Fourier transform Continuous Wavelet Transform. Figure 3 shows the system view of proposed method in which the implementation of psychoacoustic model will be done using wavelet packet decomposition techniques.

These are following steps that need to be performed to realize the implementation,
(a) Signal division and processing using small frames
(b) Pass the frame via analysis filter bank
(c) Simultaneously Apply frames to wavelets based Psychoacoustic model.
(d) Non linear quantization over the wavelet coefficient.
(e) Main output: Compressed Audio files.

IV. EXPERIMENTAL RESULTS

This new psychoacoustic model has been Implemented in reference coder based on the standard MPEG-1 audio. Table 1 indicates the type of the sound files used for the test, their duration as well as the compression ratio. The evaluation is based on the compression ratio (CR) and is defined as.
CR = Length wav File / Length of the compressed File

Table.1

Input Stream(.Wav)	*Duration(sec)*	*Compression Ratio*
Baby	7	1.45
Wolves	6	1.4
Classic	10	1.7
Voice	12	1.6

V. CONCLUSIONS

The amended psychoacoustic model predicated on wavelet packet takes an account of the critical bands and takes an account of the masking phenomenon. The essential characteristic of this model is that it proposes an analysis by wavelet packet transformation on the frequency bands that come more proximate the critical bands of the auditory perceiver.

The performance of DWPT coders for audio compression is evaluated for different signals and their compression ratios are compared

REFERENCES

[1].   T. Painter and A. Spanias,"Perceptual coding of digital audio," Proc. IEEE Trans.Sig, vol. 88, Apr. 2000, pp. 451-513
[2].   D.Sinha and A.Tewfik, Low Bit Rate Transparent Audio Compression using adapted wavelets, IEEE Trans.Sig.Proc, pp.3463-3479, Dec
[3].   1993.
[4].   Information technology-Coding of moving pictures and associated audio for digital storage media at up to about 1, 5 Mbit/s-Part3: Audio,
[5].   1999.
[6].   M. V. Wickerhauser, "Adapted wavelet analysis from theory to software," Wellesley, Massachusetts, 1994.
[7].   C.S.Burrus, R.A. Gopinath and H.Guo, "Introduction to wavelets and wavelets transforms: A Primer, "Prentice Hall, 1998.
[8].   M. Bosi and R. E. Goldberg, Introduction to Digital Audio Coding and Standards, Kluwer Academic Publishers, New York, NY, USA, 2003.
[9].   M. R. Zurera, F. L. Ferreras, M. P. J. Amores, S.M. Basc´ on, and N. R. Reyes, "A new algorithm for translating psycho-acoustic information to the wavelet domain," Signal Processing, vol. 81, no. 3, pp. 519–531, 2001.
[10].   B. Lincoln, "An experimental high fidelity perceptual audio coder," Project in MUS420 Win97, March 1998.
[11].   J. D. Johnston, "Transform coding of audio signals using perceptual noise criteria," IEEE Journal on Selected Areas in Communications,vol. 6, no. 2, pp. 314–323, 1988.
[12].   Ouni Kais," Contribution to the vocal signal analysis using knowledge on the auditory perception and Multiresolution time frequency representation of the speech signals" (in French), PhD Thesis on Electrical Engineering. National Engineering School of Tunis. February 2003.
[13].   B. Carnero and A. Drygajlo, "Perceptual speech coding and enhancement using frame-synchronized fast wavelet packet transform algorithms," IEEE Transactions on Signal Processing,vol. 47, no. 6, pp. 1622–1635, 1999.
[14].   M. R. Zurera, F. L. Ferreras, M. P. J. Amores, S.M. Basc´ on, and N. R. Reyes, "A new algorithm for translating psycho-acoustic information to the wavelet domain," Signal Processing, vol. 81,no. 3, pp. 519–531, 2001.
[15].   Jagadeesh B.K holds the Master degree from VTU, Karnataka, and currently working as Associate Professor, Department of Electronics and communication,HKBKCE ,Bangalore, Karnataka INDIA, Before entering in to teaching profession he has worked at several software industries like Toshiba, SASKEN communication and NXP semiconductors. He has extensively worked on Audio, Video codecs and Audio post processing algorithms on different DSP Platforms.
[16].   Dr B Sivakumar holds PhD degree from Anna University Tamilnadu, and currently working as Professor and head of the Telecom dept at Dr AIT, Bangalore.