Audio Compression in MPEG Technology

Manas Arora*,Neha Maurya**

*Electronics and Communication Department, BBDNIIT
**Electronics and Communication Department, BBDNIIT

Abstract- MPEG stands for MOVING PICTURE EXPERTS GROUP is a standard for video and audio compression for eliminating the noisy signals from the transmitted signals from the satellite. Audio compression is a basic method defined under MPEG-1 and MPEG-4 which by coding techniques compress audio signals to filter out undesired signals .This paper focuses on the MPEG technology, need and coding technique for audio compression.

Index Terms- *MPEG; Audio compression; ISO; IEC; bit rate; DCT; standards; MPEG encoder; audio signal; filter*.

I. INTRODUCTION

MPEG is an acronym for MOVING PICTURE EXERTS GROUP that was formed by ISO and IEC. It sets the standard for audio and video compression and its transmission. Today in the field of communication compression standards from MPEG-1 to MPEG DASH have been developed that give digital output. MPEG is a first international algorithm for digital audio compression that showed high fidelity.

Data compression is a technique in which data content of the input signal to system is compressed so that original signal is filtered out and unwanted or undesired signals are removed and true digital signal is obtained as output. Therefore when audio signals are used in the form of data it is termed as AUDIO COMPRESSION. In this compression technique reducing the bit rate which is required is an important task for perfect design of an audio system. High quality audio transmission and storage are the basic goals of audio compression.[1]

Audio compression involves transparent coding of hi-fidelity audio signals at the lowest possible bit rates. And the quality is maintained with the appropriate maintenance of standards for high fidelity. The compressed audio signals are characterized by a wide bandwidth with sampling rate of 44.1 kHz and a high resolution quantization of each sample of the signal at a rate of 16 bits per sample which uses pulse code modulation (PCM) [1]

The Audio compression basically utilizes MPEG 1 standard which is a generic audio compression standard. Audio compression in MPEG unlike vocal coders have coders known as encoders that do not assume the nature of source which generates audio signals but coder itself designs by hearing ability of the consumer at the output

II. MPEG TECHNOLOGY

The MPEG is a generic audio compression standard utilizing technology which compresses audio signal to make it audible to the human ear and remove the unwanted signals introduces in the communication path without assuming nature of audio source. In MPEG technology the perpetual parts of audio signals are maintained. MPEG technology also provides audio forward and reverse features by maintaining bit stream of the audio signal [5].

The MPEG technology runs by different standards which consist of different parts defining aspect of technology. These standards in technology specify Profiles and *Levels*. Profiles are set of tools that are available, and Levels define the range of appropriate values for the properties associated with them. The standards were revised by amendments MPEG technology has standardized the following compression formats [5]:

- MPEG-1 (1993)

Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s It was basically designed to allow moving pictures and sound to be encoded into the bit rate of a Compact Disc. It is used on Video CD and can be used for low-quality video on DVD Video. It was used in digital satellite/cable TV services, MPEG-1 down samples the images, as well as uses picture rates of only 24–30 Hz, resulting in a moderate quality.

- MPEG-2 (1995):

It transports video and audio standards for broadcast-quality television. MPEG-2 standard is considerably broader in scope than MPEG-1 It is also used on Blu-ray Discs

- MPEG-3:

MPEG-3 dealt with standardizing scalable and multi-resolution compression and basically intended for HDTV compressionMPEG-3 is different from MP3 [2,3], defined under MPEG-1 Audio Layer III.

- MPEG-4 (1998):

Coding of audio-visual objects (ISO/IEC 14496) is done. In this coding with additional complexity are made to achieve higher compression factors [2].It is closer to computer graphics applications. Defines compressed bit stream describes three-dimensional shapes and surface texture [2] MPEG-4 supports Intellectual Property MPEG-4 AVC (or MPEG-4 Part 10 or H.264). MPEG-4 AVC is used on HD DVD and Blu-ray Discs, along with VC-1 and MPEG-2.MPEG-4 involves compression schemes in Brazil (ISDB-TB) and JAPAN [6].

- MPEG-7 (2002):

Multimedia content description interface defined under ISO/IEC 15938.

- MPEG-21 (2001):

Multimedia framework defines under ISO/IEC 21000

MPEG is a standard for a multimedia framework and provides the intellectual property management and protection Moreover, recently MPEG international standards are increased for multiple[6]

MPEG technologies based on different applications are:

- MPEG-A.

Defined under ISO/IEC 23000 for the purpose for multimedia application formats [6]

- MPEG-B (2006):

Defined as systems technologies under ISO/IEC 23001for binary MPEG format for XML Fragment Request Units, Bit stream Syntax Description Language (BSD) [7].

- MPEG-C (2006):

Define as MPEG video technologies and under ISO/IEC 23002 for the accuracy requirements of implementation of integer-output 8x8 inverse discrete cosine transform[8]

- MPEG-D (2007):

Defined as MPEG audio technologies under ISO/IEC 23003 for MPEG Surround [9] SAOC (Spatial Audio Object Coding) and USAC (Unified Speech and Audio Coding)

- MPEG-E (2007):

It is defined for Multimedia Middleware. Under ISO/IEC 23004 [10] for multimedia application programming interface (API) and component models

- MPEG-V (2011):

Defined as media context and control and under ISO/IEC 23005[11] for information exchange with Virtual Worlds

- MPEG-M (2010):

It is a MPEG extensible Middleware (MXM) standard defined under ISO/IEC 23006 and defines MXM architecture and technologies, API and protocols for MPEG extensible middleware (MXM)[12,13].

- MPEG-U (2010):

It is rich media user interfaces which iis defined under ISO/IEC 23007.

- MPEG-H (2013):

It is for high efficiency coding and media delivery in heterogeneous environments.

- MPEG-DASH (2012):

It is latest information technology standard used for dynamic adaptive streaming over HTTP (DASH) and is defined under ISO/IEC 23009 and is used for media presentation, description and segment formats

III. AUDIO COMPRESSION

Audio compression enables efficient storage and transmission of data. There may be varying amounts of compression in an audio data. There may be different levels of system complexity and compressed quality of audio data. The recorded waveform is reduced with varying amounts for the purpose of transmission with or without loss. Therefore there are two types of compression lossy are lossless. The digital audio data is readily processed through mixing, filtering and equalization. This is now subjected to an encoder that uses fewer bits than that present in the original audio data. This results in reducing the transmission bandwidth of digital audio streams and storage size of audio files. [15]

Audio compression may be lossy or lossless Lossy compression is transparent to human audibility but lossless being have a compressing factor from 6 to 1[17,18].

At the decoder, bit stream unpacking of encoded bit stream is done. It is followed by frequency sample reconstruction. Finally frequency-to-time mapping produces decoded audio data stream. In the audio compression DCT is used under MPEG technology.

Discrete Cosine Transform (DCT) is a main transform or mapping method used in the MPES technology which maps the time domain into frequency domain A **DCT** expresses a finite sequence of data points in terms of a sum of cosine functions oscillating at different frequencies. The discrete cosine transform is a linear, invertible function F: $R^N \rightarrow R^N$ (where R denotes the set of real numbers), or equivalently an invertible $N \times N$ square matrix. Mathematically DCT is calculated from (1)

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\Pi\left(n+\frac{1}{2}\right)k}{N}\right]$$

$$k = 0,.......N-1 \tag{1}$$

The cosine graphs define the nature of frequency used in DCT transform and define the frequency of the audio signal. Basic cosine graphs used in the discrete cosine transform are:
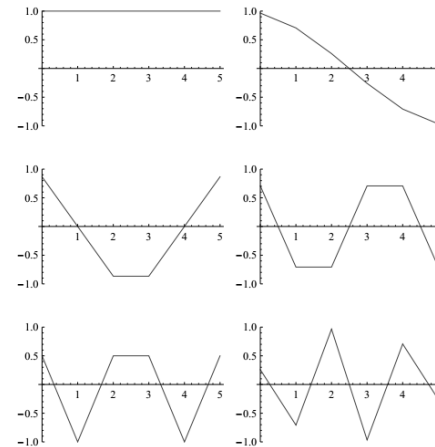
Figure 1: Cosine graphs used in DCT

The audio compression using DCT method is an efficient method by which we can change the frequency content of the signal and analyze the signals in different frequency ranges in a fixed frequency domain. DCT also effectively compresses the audio signals and maintain the audibility feature of audio signals

IV. NEED OF AUDIO COMPRESSION IN MPEG TECHNOLOGY

The basic MPEG standard for audio compression is MPEG-1 as video is not designed for DBS (Direct Broadcast Satellite) TV's. MPEG-1 audio compression supports both stereo and mono formats.

Basic question arises the need of audio compression? Answer lies within the requirement of appropriate bit rate for high quality audio. Bit rate is defined as number of sample per second which is given by R_{bits} calculates from (2)

$$R_{bits} = f_S \times n$$

(2)

The requirement of achieving high bit rate and clarity of audio signals basically defines the need of audio compression [14].

V. AUDIO COMPRESSION TECHNIQUE IN MPEG

An audio signal is a complex signal which has a wide spectrum and MPEG helps in perceptual phenomena for our ears. The large frequencies and smaller frequencies below masking threshold are inaudible to human ears. MPEG standard helps in defining standard and filtering audio signals in the available bandwidth and maintains signal to quantization ratio. In MPEG-1 audio compression is performed as given in figure 1 and involves two process.Filter bank which uses filters divides spectrum of incoming signals in sub bands .Quantizer quantizes the sub bands 15]
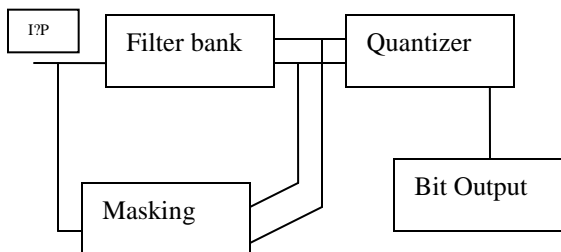

Figure 2: Block Diagram of MPEG-1 Encoder

The basic method by which audio compression in MPEG technology is done by MPEG encoder and decompression by MPEG decoder. Figure 3 shows block diagrams of the MPEG/audio encoder and decoder.[15]

 In MPEG compression the input audio stream passes through a filter bank which divides the input into multiple sub bands. The input audio stream is simultaneously also passed through a psychoacoustic model which is used to determine the signal-to-mask ratio of each sub band. The bit or noise allocation block utilizes the output of psychoacoustic model and a decision is made about the total number of code bits which are available for quantization process.

Code bits also help in removing quantization noise. The output of compressor is quantized audio samples and these formats of the data are transformed into a decodable bit stream.

The MPEG decompression of the audio signals simply reverses the formatting and then reconstruction of quantized code bits is done then reconstructs the quantized sub band values, which are finally transformed into the set of sub band values into a time-domain audio signal.[15]
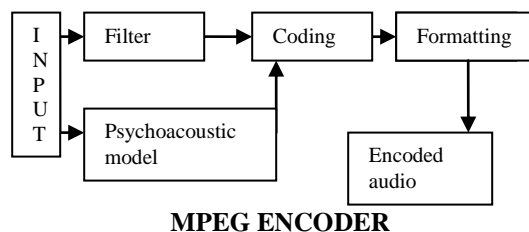
MPEG ENCODER

Figure 3.1: MPEG ENCODER R block diagram
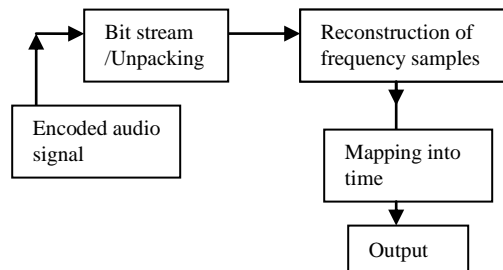
MPEG DECODER

Figure 3.2: MPEG DECODER block diagram

The MPEG audio standard has three distinct layers for compression.
Layer I is the basic algorithm

Layers II and III are enhancements of Layer I. Each successive layer improves the compression performance at the increasing cost of encoder and decoder complexity.

a) Layer 1. The Layer I algorithm uses the basic filter bank which is found in all layers. This filter bank divides the audio signal into 32 constant-width frequency bands. The filters are simple and provide time and frequency resolutions by which it is easily perceived by human ears The design of   Layer I has three concessions. First, the 32 constant width bands which do not accurately reflect the ear's critical frequency bands. Second, the filter bank and its inverse are not lossless transformations. Thirdly, adjacent filter bands have a significant frequency overlap [15]

b) Layer II. The Layer II algorithm is a simple enhancement of Layer I. It improves compression performance by coding data in larger groups. By the Layer II encoder forms frames of 3 by 12 by 32 that is total of 1152 samples per audio channel as compared to Layer I codes data in single groups of 12 samples for each sub band .Layer II removes stereo redundancy coding and there is one bit allocation The encoder encodes with a unique scale factor for each group of 12 samples to avoid audible distortion. The Layer II algorithm improves performance as it uses efficient code for representing the bit allocation, the scale factor values, and the quantized samples [15]

c) Layer III The Layer III algorithm is a much more refined approach.[16,17] though it is based on the same filter bank found in Layers I and II. Layer III compensates for some filter bank deficiencies by processing the filter outputs with a modified discrete cosine transform (MDCT).[15]

These al the three layers completely design the methods of audio compression in MPEG technology.

VI. CURRENT TRENDS

MPEG technology is an ever developing technology in the field of communication which advances day by day as audibility is the chief medium used by the receiver in the communication .Currently MPEG are used in satellite communications, DBS TV's, TATSKY dish television network which uses MPEG-4,MPEG-1,MPEG-2 and MPEG-DASH. MPEG-D is being used in audio technologies like MPEG SURROUND, SAOC (Spatial Audio Object Coding) and USAC (UNIFIED SPEECH AND AUDIO CODING).MPEG dynamic adaptive streaming over HTTP (DASH) technology is an upcoming technology in the current era of communication which enhances and maintains fidelity of audio signals.

VII. CONCLUSION

MPEG is a basic standard for designing and compressing both audio and video signals which are then transmitted to the receiver end. MPEG 1 and MPEG 4 are standards dedicated to audio compression which by coding techniques compress audio signals and noisy or undesired signals are filtered.

ACKNOWLEDGMENT

We are thankful to our college for providing time to time help in studying the topic and providing us a background to understand it in deeper details.

REFERENCES

[1] Sinha D.,Tewhik A.H, "Low Bit Rate Transparent Audio Compression using Adapted Wavelets", IEEE TRANSACTIONS ON SIGNAL PROCESSING. VOL. 41. NO. 12. DECEMBER 1993

[2] The MPEG Handbook:MPEG-1,MPEG-2,MPEG-4,Fourth edition,2004,Watkinson J,Focal Press,p-p 4-5

[3] Salomon D, "Video Compression: Data compression" Springer ISBN 978-1-84628-602-5,2007

[4] Understanding MPEG-4,"MPEG-J White Paper"(http://mpeg.chiariglione.org/technologies/mpeg-4/mp04-j/index.htm).p-83,2005-2007.

[5] http://mpeg.chiariglione.org/

[6] ISO/IEC TR 23000-1:2007 – Information technology – Multimedia application format (MPEG-A) – Part 1: Purpose for multimedia application formats

[7] ISO. "ISO/IEC 23001-1:2006 – Information technology – MPEG systems technologies – Part 1: Binary MPEG format for XML

[8] ISO. "ISO/IEC 23002-1:2006 – Information technology – MPEG video technologies – Part 1: Accuracy requirements for implementation of integer-output 8x8 inverse discrete cosine transform

[9] ISO. "ISO/IEC 23003-1:2007 – Information technology – MPEG audio technologies – Part 1: MPEG Surround

[10] ISO. "ISO/IEC 23004-1:2007 – Information technology – Multimedia Middleware – Part 1: Architecture

[11] ISO. "ISO/IEC 29116-1:2008 – Information technology – Supplemental media technologies – Part 1: Media streaming application format protocols

[12] http://mpeg.chiariglione.org/working_documents.htm#MPEG-V

[13] Satellite Communications,Seventh Edition 2009,Agarwal D.C,Khana Publications,p-p 368-369

[14] Davis Yan Pan,"Digital Audio Compression", Digital Technical Journal Vol. 5 No. 2, Spring 1993

[15] K. Brandenburg and J. D. Johnston, "Second Generation Perceptual Audio Coding: The Hybrid Coder," Preprint 2937, 88th Audio Engineering Society Convention, Montreaux (1990).

[16] K. Brandenburg, J. Herre, J. D. Johnston, Y. Mahieux, and E. Schroeder, "ASPEC: Adaptive Spectral Perceptual Entropy Coding of High Quality Music Signals," Preprint 3011, 90th Audio Engineering Society Convention, Paris (1991)

[17] Pan, D "A tutorial on MPEG/audio compression" MultiMedia, IEEE (Volume:2 , Issue: 2),2002

[18] Gersho, A. "Advances in speech and audio compression" Proceedings of the IEEE Volume:82, Issue: 6 ,August 2002

AUTHORS

First Author – Manas Arora, M.Tech(Digital Communication) student, Babu Banarsi Das Northern India Institute of Technology, manasarora14@gmail.com

Second Author – Neha Maurya, M.Tech (Digital Communication) student, Babu Banarsi Das Northern India Institute of Technology, neha6maurya@gmail.com