AN OVERVIEW OF MPEG FAMILY AND ITS APPLICATIONS

S.Vetrivel, M.Gowri, M.Sumaiya Sultana
Department of Computer Applications
Chettinad College of Engineering and Technology
Karur, Tamilnadu, India 639114
E Mail: vetri76@gmail.com

Dr G.Athisha,
Professor/ HoD,
Department of ECE
PSNA College of Engineering and Technology

Abstract: This paper presents an overview of the video compression standards related to the MPEG family. MPEG-7 and MPEG-21 are specially covered including its latest standard. MPEG-7 is mainly used for object descriptions and MPEG-21 is for DRM (Digital Rights Management).

Keyword: MPEG-1, MPEG-2, MPEG-4, MPEG-7, MPEG-21, MPEG-A, MPEG-D

I.INTRODUCTION

MPEG is the “Moving Picture Experts Group”, working under the joint direction of the international Standards Organization (ISO) and the International Electro Technical Commission (IEC). This paper will provide an overview of the recent standards in the MPEG family. MPEG-7 is developed for Multimedia content description interface ,it uses XML to store metadata, and can be attached to timecode in order to tag particular events, or synchronise lyrics to a song. MPEG-21 is an open framework for multimedia delivery and consumption. It can be used to combine video, audio, text and graphics. The other latest version in MPEG like MPEG-A, MPEG-D is also discussed in this paper.

II. MPEG-1 (1992)

MPEG-1 is currently the most compatible format in the MPEG family but does not support interlaced video coding. MPEG-1 typically operates at bitrates of 1.5 Mbit/s with a screen resolution of 352*288 pixels at 25 frames a second [1, 8].

MPEG-1 coded bitstream has been designed to support a number of operations including random access, fast search, reverse playback, error robustness, and editing [1].

A number of techniques are used to achieve a high compression ratio. The first is to select an appropriate spatial resolution for the signal. The algorithm then uses block-based motion compensation to reduce the temporal redundancy. The difference signal, the prediction error, is further compressed using the discrete cosine transform (DCT) to remove spatial correlation and is then quantized.

Finally, the motion vectors are combined with the DCT information, and coded using variable length codes.
Figure 1 below illustrates a possible combination of the three main types of pictures that are used in the standard.

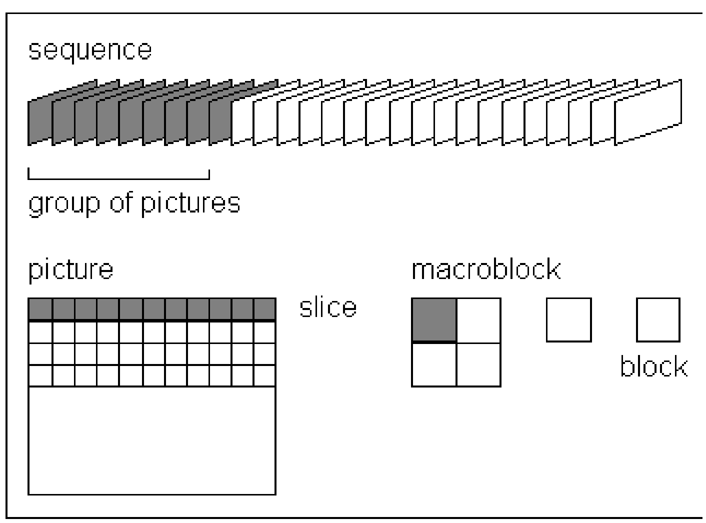

Figure 1 -Example of temporal picture structure.

A. *Application:*

It is basically designed to allow moving pictures and sound to be encoded into bitrate or a Compact Disc. It is used on Video CD SVCD and can be used for low-quality video on DVD video [1].

III. MPEG-2(1994)

A. Codec structure

MPEG-2 is aimed for high bitrate, high quality applications, seen as Digital TV broadcasting and DVD [6]. In an MPEG-2 system, the DCT and motion-compensated interframe prediction are combined, as shown in Fig. 2.

The coder subtracts the motion-compensated prediction from the source picture to form a 'prediction error' picture. The prediction error is transformed with the DCT, the coefficients are quantized and these quantized values coded using a VLC.

The coded luminance and chrominance prediction error is combined with 'side information' required by the decoder, such as motion vectors and synchronizing information, and formed into a bitstream for transmission. Fig.3 shows an outline of the MPEG-2 video bitstream structure.

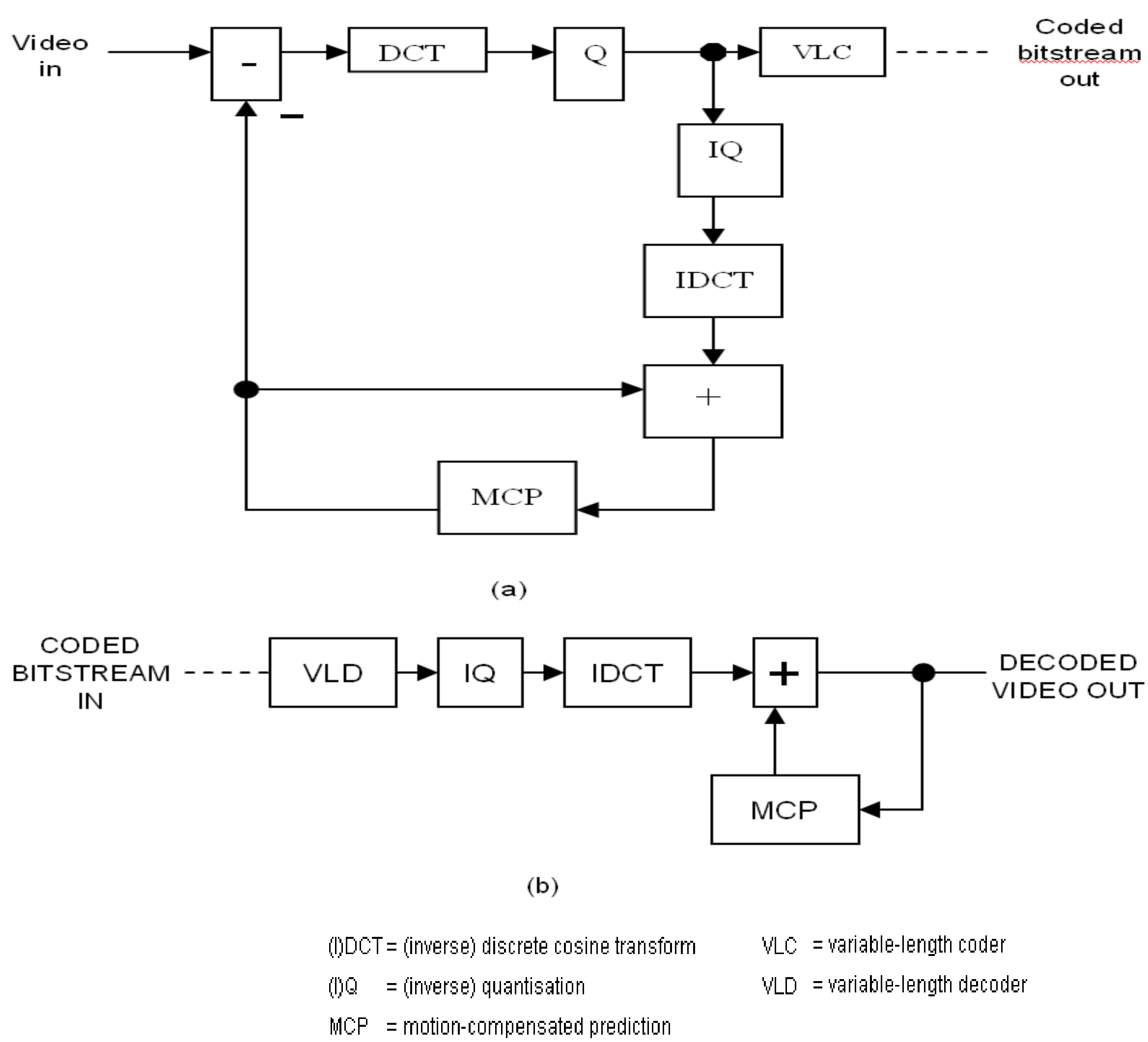

Fig. 2 - (a) Motion-compensated DCT coder; (b) motion compensated DCT decoder .

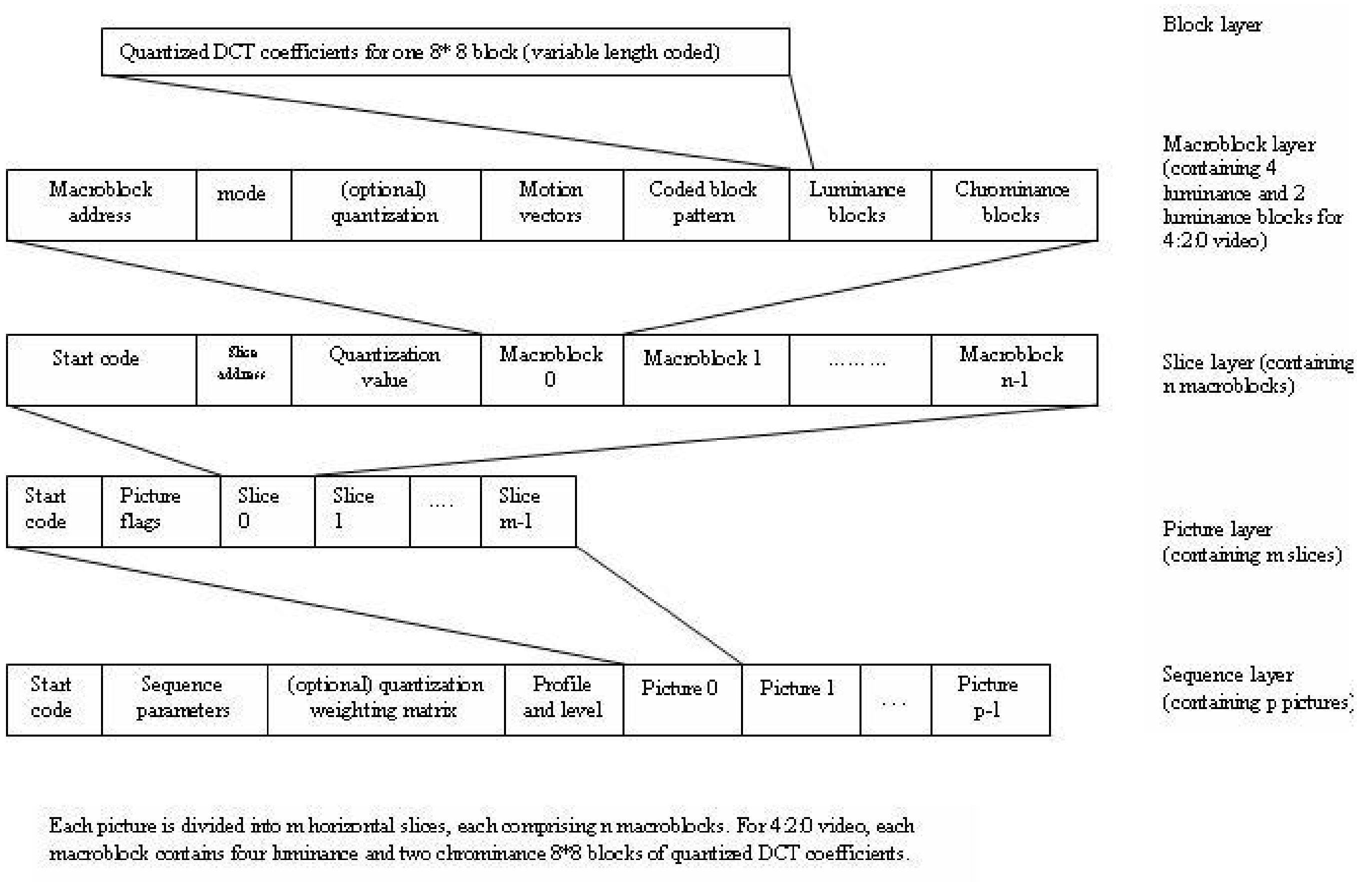

Fig. 3 - Outline of MPEG-2 video bitstream structure (shown bottom up).

In the decoder, the quantized DCT coefficients are reconstructed and inverse transformed to produce the prediction error. This is added to the motion-compensated prediction generated from previously decoded pictures to produce the decoded output.

In an MPEG-2 codec, the motion-compensated predictor shown in Fig. 2 supports many methods for generating a prediction

B. Details of non-scalable profiles:

Two non-scalable profiles are defined by the MPEG-2 specification.

The simple profile uses no B-frames, and hence no backward or interpolated prediction. Consequently, no picture reordering is required (picture reordering would add about 120 ms to the coding delay). With a small coder buffer, this profile is suitable for low-delay applications such as video conferencing where the overall delay is around 100 ms. Coding is performed on a 4:2:0 video signals.

The main profile adds support for B-pictures and is the most widely used profile. Using B-pictures increases the picture quality, but adds about 120 ms to the coding delay to allow for the picture reordering. Main profile decoders will also decode MPEG-1 video. Currently, most MPEG-2 video decoder chip-sets support the main profile at main level.

C. Details of scalable profiles:

The SNR profile adds support for enhancement layers of DCT coefficient refinement, using the 'signal to noise (SNR) ratio scalability' tool. The SNR profile is suggested for digital terrestrial television as a way of providing graceful degradation.

The spatial profile adds support for enhancement layers carrying the coded image at different resolutions, using the 'spatial scalability' tool. Spatial scalability is characterised by the use of decoded pictures from a lower layer as a prediction in a higher layer. If the higher layer is carrying the image at a higher resolution, then the decoded pictures from the lower layer must be sample rate converted to the higher resolution by means of an 'up-converter'. The spatial profile is suggested as a way to broadcast a high-definition TV service with a main-profile compatible standard-definition service.

In MPEG-2 Transcoding differs from first generation coding, in that a transcoder only has access to a previously compressed signal which already contains quantisation noise compared to the original source signal. [2]

IV. MPEG-4 (1998)

This is based on the foundation of MPEG-1 and 2 as can be seen in Figure 2.The DCT transform is used along with similar quantization tables and entropy coders. The advances are with the use of multiple VLC tables and half pixel fractional motion estimation accuracy.
In the area of Audio, new tools are added in MPEG-4 Version 2 to provide the following new functionalities: [11]

- ***Error Resilience*** tools provide improved performance on error-prone transmission channels.
- ***Low-Delay Audio Coding*** tools support the transmission of general audio signals in applications requiring low coding delay, such as real-time bi-directional communication.
- ***Small Step Scalability*** tools provide scalable coding with very fine granularity, i.e. embedded coding with very small bitrate steps, based on the General Audio Coding tools of Version 1.
- ***Parametric Audio Coding*** tools combine very low bitrate coding of general audio signals with the possibility of modifying the playback speed or pitch during decoding without the need for an effects processing unit.
- ***Environmental Spatialisation*** tools enable composition of an "audio scene" with more natural sound source and sound environment modeling than is possible in Version 1.

MPEG-4 is an object oriented based image codec and actually uses the wavelet transform to represent textural information [8]. The steps involved in decompression are shown in figure and it should be noted that one of the aims of having a low complexity decoded has been met. MPEG-4 principally offers four error resilience tools.

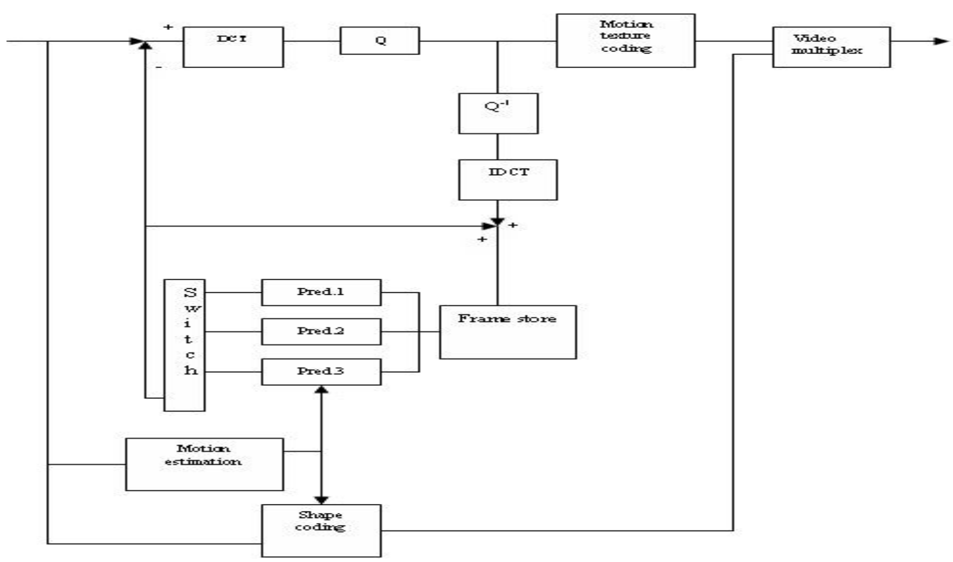

Figure 4. MPEG-4 Video Coder Basic Block Diagram

A. Application

MPEG-4 aimed at multimedia applications including streaming video applications on mobile devices [6].

IV. MPEG-7

MPEG-7 is a multimedia content description standard. It was standardized in ISO/IEC 15938 (Multimedia content description interface). This description will be associated with the content itself, to allow fast and efficient searching for material that is of interest to the user. MPEG-7 is formally called Multimedia Content Description Interface. The ultimate goal and objective of MPEG-7 is to provide interoperability among systems and applications used in generation, management, distribution, and consumption of audio-visual content descriptions. [3]
It uses XML to store metadata, and can be attached to timecode in order to tag particular events, or synchronise lyrics to a song, for example.

It was designed to standardize:

- a set of Description Schemes (short DS in the standard) and Descriptors (short D in the standard)
- a language to specify these schemes, called the Description Definition Language (short DDL in the standard)
- a scheme for coding the description

The combination of MPEG-4 and MPEG-7 has been sometimes referred to as MPEG-47.

MPEG-7 tools

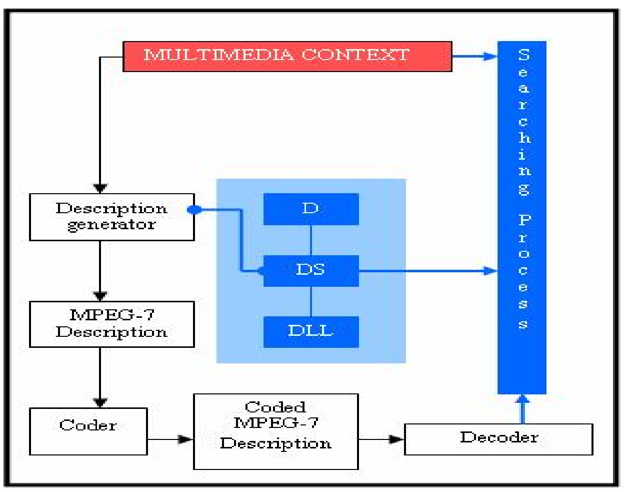

Figure-5 Relation between different tools and elaboration process of MPEG-7

A. MPEG-7 uses the following tools:

- ***Descriptor (D):*** It is a representation of a feature defined syntactically and semantically. It could be that a unique object was described by several descriptors.

- ***Description Schemes (DS):*** Specify the structure and semantics of the relations between its components, these components can be descriptors (D) or description schemes (DS).

- ***Description Definition Language (DDL):*** It is based on XML language used to define the structural relations between descriptors. It allows the creation and modification of description schemes and also the creation of new descriptors (D).

- ***System tools:*** These tools deal with binarization, synchronization, transport and storage of descriptors. It also deals with Intellectual Property protection [4].

There are many applications and application domains which will benefit from the MPEG-7 standard. A few application examples are:

- Digital library: Image/video catalogue, musical dictionary.
- Multimedia directory services: e.g. yellow pages.
- Broadcast media selection: Radio channel, TV channel.
- Multimedia editing: Personalized electronic news service, media authoring.
- Security services: Traffic control, production chains...
- E-business: Searching process of products.
- Cultural services: Art-galleries, museums...
- Educational applications.
- Biomedical applications.

B. Usage Environment [9]

The usage environment holds the profiles about user, device, network, delivery, and other environments. The system uses this information to determine the optimal content selection and the most appropriate form for the user.

The MPEG-7 user preferences descriptions specifically declare the user's preference for filtering, search, and browsing.

Traditional MPEG Systems Requirements: [11]

The fundamental requirements set for MPEG-7 Systems are described below.

Delivery: The multimedia descriptions are to be delivered using a variety of transmission and storage protocols. Some of these delivery protocols include streaming.

Synchronization: The MPEG-7 representation needs to allow a precise definition of the notion of time so that data received in a streaming manner can be processed and presented at the right instants in time, and be temporally synchronized with each other.

Stream Management: The complete management of streams of audio-visual information including MPEG-7 descriptions implies the need for certain mechanisms to allow an application to consume the content.

V. MPEG-21

One of the standards produced by the MPEG is MPEG-21 [4]. Its aim is to offer interoperability in multimedia consumption and commerce.MPEG-21 is an open framework for multimedia delivery and consumption. It can be used to combine video, audio, text and graphics. MPEG-21 provides normative methods for content identification and description, rights management and protection, adaptation of content, processing on and for the various elements of the content, evaluation methods for determining the appropriateness of possible persistent association of information.

Enabling access to any multimedia content from any type of terminal or network is very much in_line with the MPEG-21 standardization committee's vision, which is to achieve interoperable and transparent access to multimedia content. [4]

A .Consists of 12 parts/specifications

Part 1 - Vision, Technologies and Strategy
Part 2 - Digital Item Declaration
Part 3 - Digital Item Identification and Description
Part 4 - Intellectual Property Management and Protection
Part 5 - Rights Expression Language
Part 6 - Rights Data Dictionary
Part 7 - Digital Item Adaptation
Part 8 - Reference Software Part 9 - File Format
Part 10 - Digital Item Processing
Part 11 - Evaluation Tools for Persistent Association
Part 12 - Test Bed for MPEG-21 Resource Delivery

Three of these parts are directly dealing with Digital Right Management (DRM) [10].

Part 4. *Intellectual Property Management and Protection* (IPMP): provides the means to reliably manage and protect content across networks and devices.

Part 5. *Rights Expression Language* (REL): specifies a machine-readable language that can declare rights and permissions using the terms as defined in the Rights Data Dictionary.

Part 6. Rights Data Dictionary (RDD): specifies a dictionary of key terms required to describe users' rights.

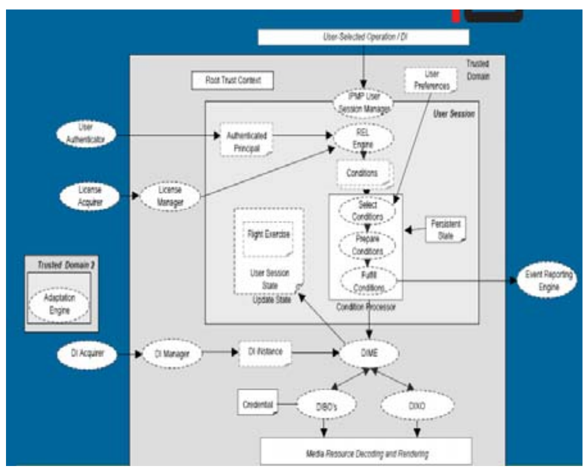

Figure-6 Architecture

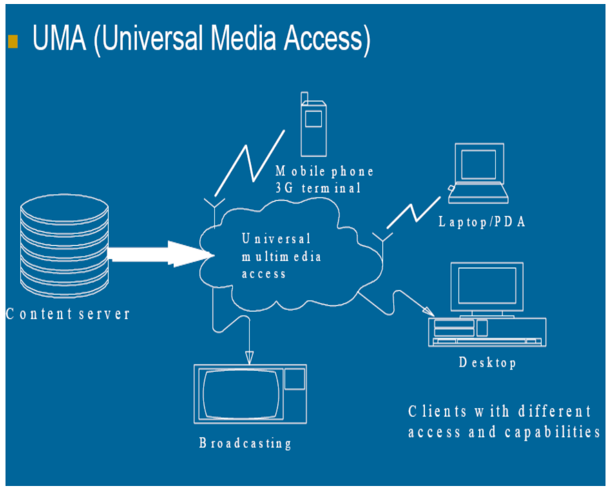

Figure-7 Use Case

B. MPEG-21 Benefits

Supports the creation, distribution and consumption of content that provides a richer user experience than previously possible except on a proprietary basis, MPEG-21 supports creation at all points in the distribution and consumption chain, Improves interoperability across applications, Opens the way for more user interaction with content, In the case of the REL and RDD, provide tools missing from MPEG2/4 IPMP.

VI. MPEG-A

MPEG-A supports a fast track to standardization by selecting readily tested and verified tools taken from the MPEG body of standards and combining them to form a MAF(Multimedia Application Format). This approach builds on the toolbox approach of existing MPEG standards. This means there is no need for time-consuming research, development and testing of new technologies. If MPEG cannot provide a needed piece of technology, then additional technologies originating from other organizations can be included by reference in order to facilitate the envisioned MAF. Hence, a MAF is created by cutting horizontally through all MPEG standards, selecting existing parts and profiles as appropriate for the envisioned application.

Consider Figure 8, which provides an illustration of this concept. MPEG standards are represented by the vertical bars on the right, and profiles are represented by the bold boxes. Non-MPEG standards or technologies are represented as vertical bars on the left. A particular MAF uses profiles from each technology (the various colored boxes) and combines them in a single standard. Ideally, a MAF specification consists of references to existing profiles within MPEG standards. However, if the appropriate profiles do not exist, then the experts can select and quantify the tools and profiles they believe are necessary to develop the MAF, which in turn provides feedback to the ongoing profiling activities within MPEG. It is also conceivable that the MAF process will help to identify gaps in the technology landscape of MPEG standards, gaps that may be mended subsequently by a new standardization campaign [5].

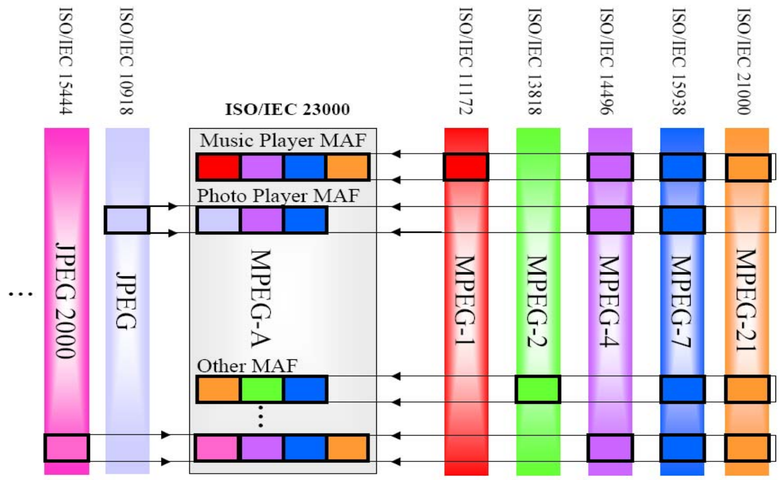

Figure 8: Conceptual Overview of MPEG-A

VII. MPEG-D SURROUND

MPEG Surround, is also known as Spatial Audio Coding (SAC) is a lossy compression format for surround sound that provides a method for extending mono or stereo audio services to multi-channel audio in a backwards compatible fashion. The total bit rates used for the (mono or stereo) core and the MPEG Surround data are typically only slightly higher than the bit rates used for coding of the (mono or stereo) core. MPEG Surround adds a Side-information stream to the (mono or stereo) core bit stream, containing spatial image data. Legacy stereo playback systems will ignore this side-information while players supporting MPEG Surround decoding will output the reconstructed multi-channel audio.

A. Perception of sounds in space

MPEG Surround coding uses our capacity to perceive sound in the 3D and captures that perception in a compact set of parameters. Spatial perception is primarily attributed to three parameters, or cues, describing how humans localize sound in the horizontal plane: Interaural level differences(ILD), Interaural time difference(ITD) and Interaural coherence (IC). These three concepts are illustrated in next image. Direct, or first-arrival, waveforms from the source hit the left ear at time, while direct sound received by the right ear is diffracted around the head, with time delay and level attenuation, associated. These two effects result in ITD and ILD are associated with the main source. At last, in a reverberant environment, reflected sound from the source, or sound from diffuse source, or uncorrelated sound can hit both ears, all of them are related with IC.

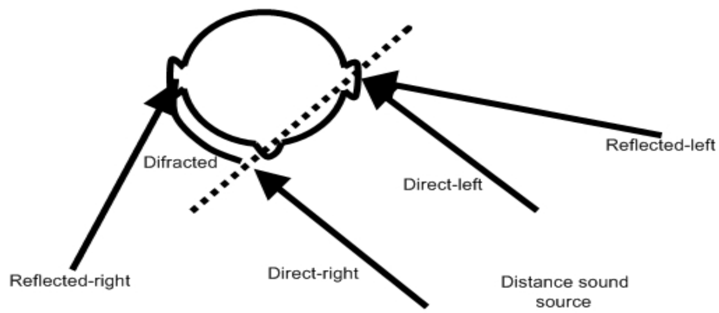

Figure-9 Humans localize sound in the horizontal plane

B. Description

MPEG Surround uses interchannel differences in level, phase and coherence equivalent to the ILD, ITD and IC parameters. The spatial image is captured by a multichannel audio signal relative to a transmitted downmix signal. These parameters are encoded in a very compact form so as to decode the parameters and the transmitted signal and to synthesize a high quality multichannel representation.

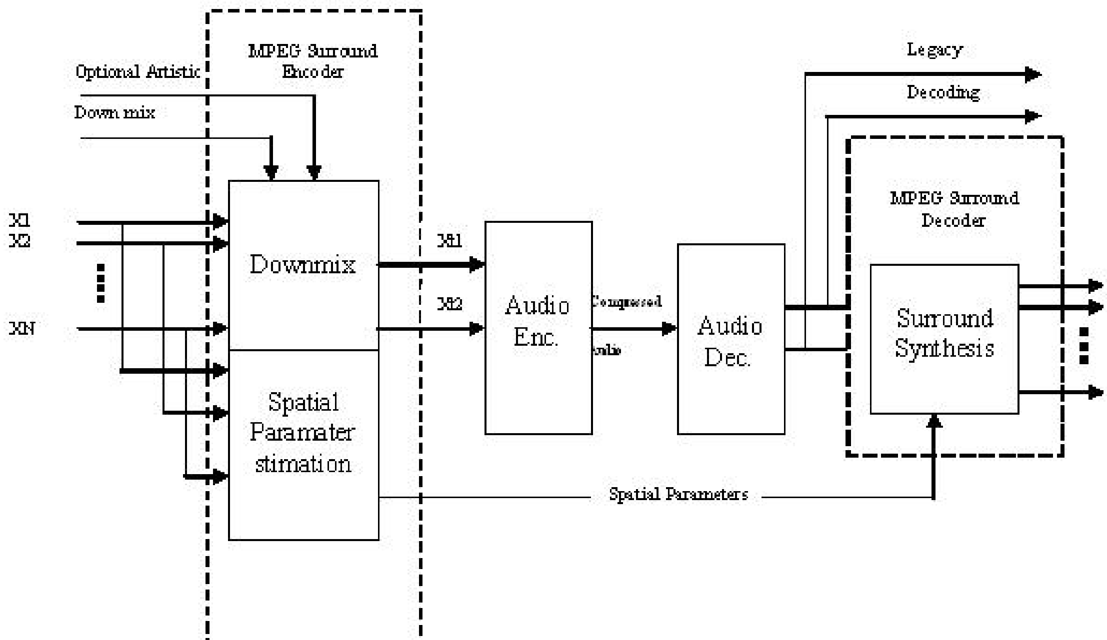

Figure-10 Block diagram of encoding and decoding MPEG Surround

MPEG Surround encoder receives a multichannel audio signal, x1 to xN where the number of input channels is N. The most important aspect of the encoding process is that a downmix signal, xt1 and xt2, which is typically stereo, is derived from the multichannel input signal, and it is this downmix signal that is compressed for transmission over the channel rather than the multichannel signal.

The encoder may be able to exploit the downmix process so as to be more advantageous. It not only creates a faithful equivalent of the multichannel signal in the mono or stereo downmix, but also creates the best possible multichannel decoding based on the downmix and encoded spatial cues as well. Alternatively, the downmix could be supplied externally (Artistic Downmix in before Diagram Block). The MPEG Surround encoding process could be ignored by the compression algorithm used for the transmitted channels (Audio Encoder and Audio Decoder in before Diagram Block). It could be any type of high-performance compression algorithms such as MPEG-1 Layer III, MPEG-4 AAC or MPEG-4 High Efficiency AAC, or it could even be PCM.

B. Legacy compatibility

The MPEG Surround technique allows for compatibility with existing and future stereo MPEG decoders by having the transmitted downmix (e.g. stereo) appear to stereo MPEG decoders to be an ordinary stereo version of the multichannel signal. Compatibility with stereo decoders is desirable since stereo presentation will remain pervasive due to the number of applications in which listening is primarily via headphones, such as portable music players.

MPEG Surround also supports a mode in which the downmix is compatible with popular matrix surround decoders, such as Dolby Pro-Logic.

C. Applications

Digital Audio Broadcasting**,** Digital TV Broadcasting**,** Music download service**,** Streaming music service / Internet radio

VIII. CONCLUSION

The MPEG family of standards has proven to be one of the most successful standards. The MPEG-1 is the basic compression that was used in CD and VCDS. With the enrichment of the various techniques, the MPEG standards are also developed. MPEG-7 will address both retrieval from digital archives as well as filtering of streamed audiovisual broadcasts on the Internet. MPEG-21 is mainly developed to adopt for all the distributed environment, it improves interoperability across applications. MPEG-A supports for the fast track of MAF. MPEG-D is a lossy compression format for surround sound that provides a method for extending mono or stereo audio services to multi-channel audio in a backwards compatible fashion . MPEG-U, MPEG-V, and MPEG-M are in under development.

REFERENCES

[1]. The ISO-MPEG-1 audio: A generic standard for coding of high-quality digital audio
[2].”Real-time transcoding of MPEG-2 video bitstreams”, Tudor, PN and Werner, OH, IEEE conference publication, 1, 1997.
[3]. “Overview of the MPEG-7 standard”, Chang, S.F. and Sikora, T. and Purl, A., IEEE Transaction of circuits and Systems for Video Technology, Vo.11, 6, 2001.
[4]. I. Burnett et al., “MPEG-21: Goals and Achievements,” *IEEE Multimedia*, vol. 10, no. 6, Oct.–Dec. 2003, pp. 60-70.
[5]. ISO/IEC JTC1/SC29/WG11/N7070, Busan, Korea, April 2005. “MAFs under Consideration”. [tobeaccessedthroughhttp://mpeg.chiariglione.org].
[6]. Ahmed et al.: “Video Transcoding: An overview of various Techniques and Research Issues”, IEEE Transactions on multimedia, Vol. 7, No.5, oct 2005
[7]. Belle L.Tseng et al: Using MPEG-7 and MPEG-21 for personalizing Video, IEEE Computer Society, 2004.
[8]. Maaruf ALI –The Latest Advances in Video compression and the MPEG family, University of Pitesti- Electronics and Computers Science, scientific Bulletin, No.8, Vol. 1, 2008.
[9].”MPEG-7 Systems Overview”, Olivier, A. and Philippe, S., IEEE Transaction on Circuits and Systems for Video Technology, 11, 6, 760--764, 2001
[10]. Interoperability between ODRL and MPEG-21 REL, J Polo, J Prados, J Delgado - Proceedings of the First International ODRL Workshop (Eds. Renato Iannella \& Susanne Guth, 22—23 2004,
[11]. “An Overview of MPEG-4 audio version 2”, Purnhagen H, Proc. AES 17th International Conference, 1999