Available online at www.elixirpublishers.com (Elixir International Journal)

Computer Science and Engineering

Elixir Comp. Sci. & Engg. 59A (2013) 15793-15797

ISSN: 2229-712X

A study of MPEG series

S.Vetrivel[1] and G.Athisha[2]

[1]Mother Terasa College of Engineering

[2]PSNA College of Engg & Tech.

ARTICLE INFO

Article history:

Received: 11 December 2012;

Received in revised form:

10 June 2013;

Accepted: 14 June 2013;

Keywords

MPEG-1, ISO, IEC, DCT,

MPEG-2,

MPEG-4,

MPEG-7,

MPEG-21,

MPEG-A, MPEG-D.

ABSTRACT

The applications of audio and video compression are limitless and the ISO has done well to provide standards which are appropriate to the wide range of possible compression products. MPEG coding embraces video pictures from the tiny screen of a videophone to the high-definition images needed for electronic cinema. Audio coding stretches from speech-grade mono to multichannel surround sound. 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 standards. MPEG-7 is mainly used for object descriptions and MPEG-21 is for DRM (Digital Rights Management).

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-1 is used to deliver video and audio at the same bit rate as a conventional audio CD. MPEG-4 is a standard for the coded representation of visual information. This standard is a document that primarily defines two things, a coded representation that describes visual data in a compressed form and a method decoding the syntax to reconstruct visual information. 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 synchronize 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.

I. MPEG-1 (1992)

MPEG-1 video and audio compression standards were mainly devised for CD-ROM applications. MPEG-1 is currently the most compatible format in the MPEG family but does not support interlaced video coding. It typically operates at bit rates of 1.5 Mbit/s with a screen resolution of 352*288 pixels at 25 frames a second [1, 8].MPEG-1 coded bit stream 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.

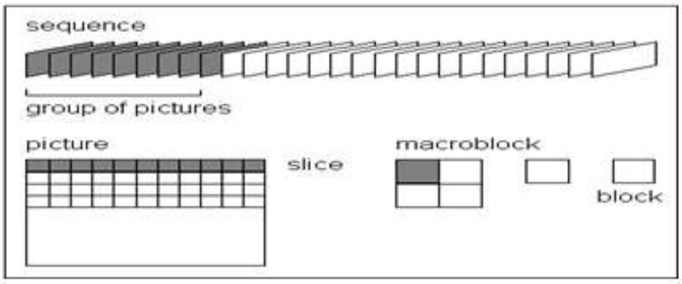


Fig. 1 -Example of temporal picture structure

Figure 1 above illustrates a possible combination of the three main types of pictures that are used in the standard. Page Layout

MPEG-2(1994)

The MPEG-2 standard is published in four parts. Part 1: Systems specifies the system coding layer of the MPEG-2. It defines a multiplexed structure for combining audio and video data and means of representation the timing information needed to replay synchronized sequences in real time. Part 2: video specifies the coded representation of video data and the decoding process required to reconstruct pictures. Part 3: Audio specifies the coded representation of audio data. Part 4: Conformance test.MPEG-2 was developed by ISO/IEC/JTC/SC29/WG11 and is known as ISO/IEC 13818. The MPEG-2 video coding standard is primarily aimed at coding of CCIRR-601 or higher resolution video with fairly high quality at challenging bit rates of 4 to 9Mbit/s. It aims at providing CCIR/ITU-R quality for NTSC, PAL, and SECAM, and also at supporting HDTV quality, at data rate above 10 Mbps, real-time transmission, and progressive and interlaced scan sources. It consists of 4 layers: GOP, Pictures, Slice, Macroblock, and Block.

Tele:

E-mail addresses: vetri76@gmail.com, gathisha@yahoo.com

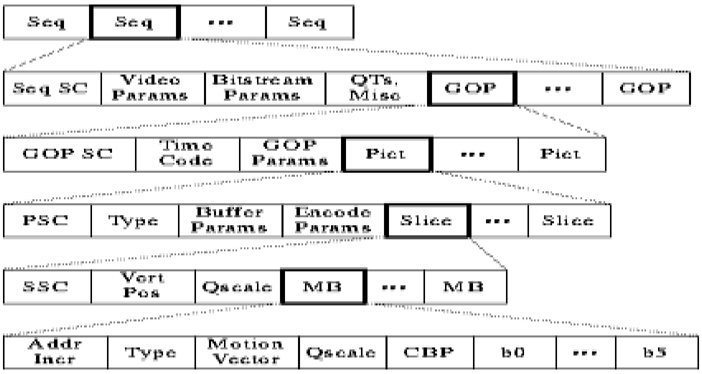


Fig 2. Video Stream Data Hierarchy

A. Video sequence:

Begins with a sequence header (may contain additional sequence header), includes one or more groups of pictures, and ends with an en-of-sequence code.

B. Group of Pictures (GOP):

A header and a series of one of more pictures intended to allow random access into the sequence.

1) Picture:

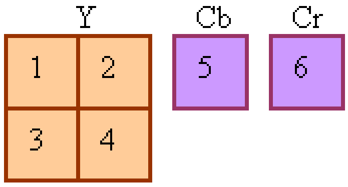
The primary coding unit of a video sequence. A picture consists of three rectangular matrices representing luminance (Y) and two chrominance (Cb and Cr) values. The Y matrix has an even number of rows and columns. The Cb and Cr matrices are one-half the size of the Ymatrix in each direction (horizontal and vertical).

2) Slice:

One or more “contiguous” macroblock. The order of the macroblocks within a slice is from left-to-right and top-to-bottom. Slice are important in the handling or errors. If the bitstream contains an error, the decoder can skip to the start of the next slice. Having more slices in the bitstream allows better error concealment, but uses bits that could otherwise be used to improve picture quality.

3) Macroblock:

The basic coding unit in the MPEG algorithm. It is 16x16 pixel segment in a frame. Since each chrominance component has one-half the vertical and horizontal resolution of the luminance component, a macroblock consists of four Y , one Cr, and one Cb block.



4) Block:

The smallest coding unit in the MPEG algorithm. It consists of 8x8 pixels and can be one of three types: luminance(Y), red chrominance (Cr), or blue chrominance (Cb). The block is the basic unit in intra frame coding.

A. 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.

B. 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 characterized 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 quantization noise compared to the original source signal. [2]

MPEG-4 (1998)

The different visions of the next generation multimedia applications are MPEG-4 and H.264/AVC. Among the two, the first is the vision of MPEG-4 standard which is a rich, interactive on-line world bringing together synthetic, natural, video, image, 2D and 3D ‘objects’.MPEG-4 is based on the foundation of MPEG-1 and 2 as can be seen in Fig. 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. 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. The MPEG-4 decoder is a hardware module optimized for FPGA technologies, making use of a limited number of logic resources and being able to decode a 4CIF (704*576) sequence in real time. The decoder can also be used to decode multiple streams simultaneously (up to 8)

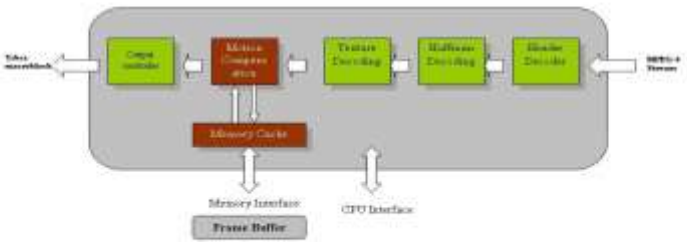


Fig. 3 Block Diagram

Figure 3 illustrates a simplified block diagram showing the internal modules and its interfaces. The core accepts the compressed stream contains headers. The decoded video data is organized in macroblock under YUV format. In that one macroblock is made of 4 luminance blocks which consists of 8*8 and 1 Cb block (8*8) and 1 Cr block (8*). The video data is output by the core through its video interface in macroblock raster scan order. The decoder has then make interface to a memory controller, that allows the connection to any custom memory controller. The core can be delivered with a standard

SRAM controller; a suitable SDRAM controller is separately available. The core has been optimized in order to minimize the amount and bandwidth of off-chip memory. A single frame needs to be stored and accesses that are reduced to 1 read and 1 write per input sample. The module constituting are depicted under Figure 3 are explained below:

Header decoding: The module decodes compliant MPEG-4 VOL and VOP headers.

Entropy decoding: The entropy decoder applies a Huffman decoding on both the motion vectors and the compressed pixels. This module uses pre-defined Huffman tables.

Texture decoding: This module decodes error frames when dealing with P-VOPs (motion compensated) or complete frames when dealing with I-VOPs.

Motion compensation: This module is bypassed for Intra-coded pictures (I-VOP). For P-VOPs, it combines the so-called error picture to the reference frame (stored in off-chip memory), using the movement information carried by the decoded motion vector. The motion compensation unit supports the definition of a single motion vector per macro block.

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 synchronize lyrics to a song, for example.It was designed to standardize: (1) a set of Description Schemes (short DS in the standard) and Descriptors (short D in the standard) (2) a language to specify these schemes, called the Description Definition Language (short DDL in the standard) (3) a scheme for coding the description. The combination of MPEG-4 and MPEG-7 has been sometimes referred to as MPEG-47.

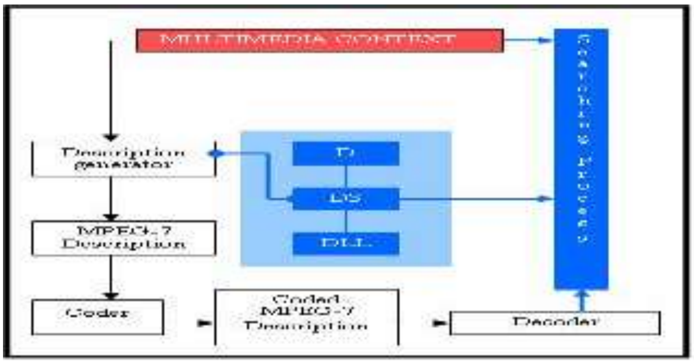


Fig. 4 Relation between different tools and elaboration process of MPEG-7

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.

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.

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.

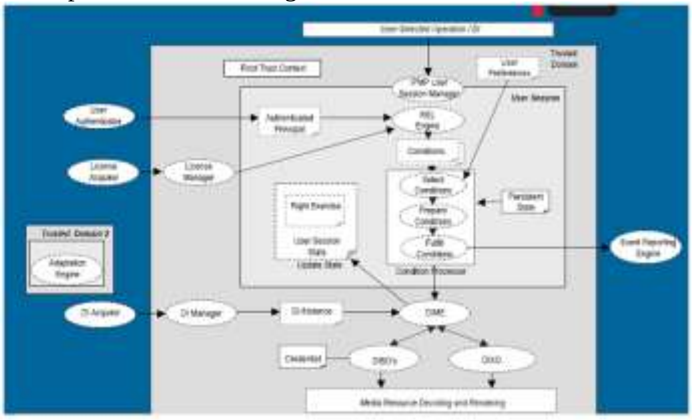


Fig. 5 Architecture

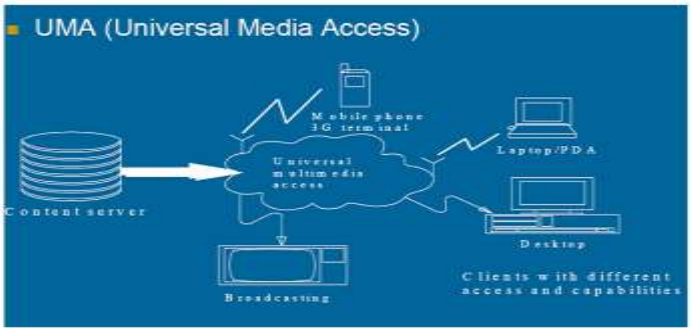


Fig .6 Use Case

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 echnology landscape of MPEG standards, gaps that may be mended subsequently by a new standardization campaigns.

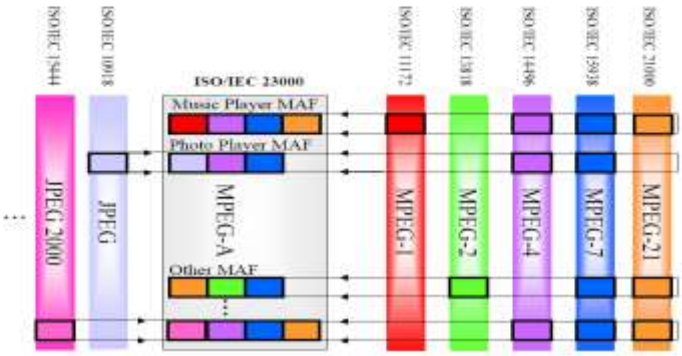


Fig 7. Conceptual Overview of MPEG-A

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.

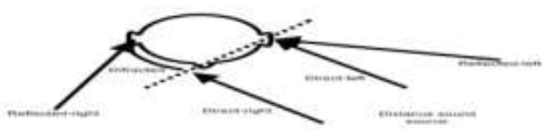


Fig: Humans localize sound in the horizontal plane

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

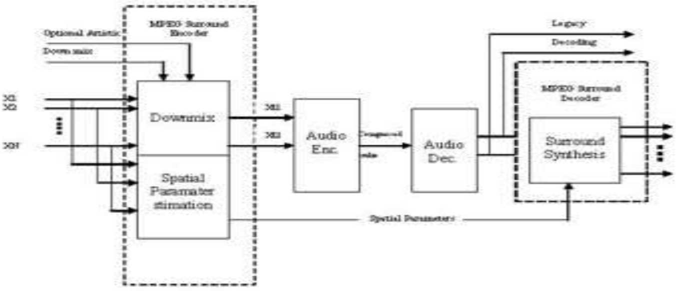


Fig . 9 Block diagram of encoding and decoding MPEG Surround

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.

Conclusions

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 high quality digital audio.

[2] Tudor, PN and Werner, OH: ”Real-time transcoding of MPEG-2 video bitstreams” , IEEE conference publication, 1, 1997.

[3] ”, Chang, S.F. and Sikora, T. and Purl, A :“Overview of the MPEG-7 standard, IEEE Transaction of circuits and Systems for Video Technology, Vo.11, 6, 2001.

[4] 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”. [to be accessed through http://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]J Prados, J Delgado, J Polo: ”Interoperability between ODRL and MPEG-21 REL, - Proceedings of the First International ODRL Workshop (Eds. Renato Iannella \& Susanne Guth, 22—23 2004.