Multimedia Standards

Peter L. Stanchev[1, 2]

[1] Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, Bulgaria
[2] Kettering University, Flint, USA
pstanche@kettering.edu

Abstract. The aim of this paper is to review some of the standards, connected with multimedia and their metadata. We start with MPEG family. MPEG-21 provides an open framework for multimedia delivery and consumption. MPEG-7 is a multimedia content description standard. With the Internet grow several format were proposed for media scenes description. Some of them are open standards such as: VRML1, X3D2, SMIL3, SVG4, MPEG-4 BIFS, MPEG-4, XMT, MPEG-4, LaSER, COLLADA5, published by ISO, W3C, etc. Television has become the most important mass medium. Standards such as MHEG, DAVIC, Java TV, MHP, GEM, OCAP and ACAP have been developed. Efficient video-streaming is presented. There exist a large number of standards for representing audiovisual metadata. We cover the Material Exchange Format (MXF), the Digital Picture Exchange (DPX), and the Digital Cinema Package (DCP).

Keywords: Multimedia Standards, MPEG, Scene Description, Interactive TV, Video-Stream Filtering, Audiovisual Media

1 Introduction

Today there are many technologies connected with multimedia presentations. Comprehensive review of the multimedia standards is given in [7]. Substantial research and standardization efforts have aimed at supporting Universal Multimedia Access [8] (Figure 1).

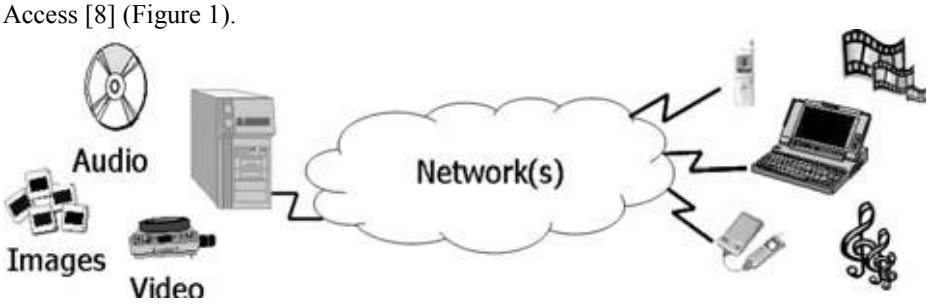

Fig. 1. Universal multimedia access

In 1988, the International Standards Organization (ISO) formed the Motion Picture Experts Group (MPEG) with the primary goal of developing coding techniques to achieve good quality audio and video with the compact disc as the target application. Through the years, the scope of the work evolved to include a standard for high-quality, efficient coded digital TV (MPEG-2), a standard for multimedia applications (MPEG-4), a standard for a multimedia content description interface (MPEG-7), and multimedia framework (MPEG-21). In 1996, MPEG was presented with the Emmy award for the MPEG-1 and MPEG-2 standards. Semantic Video and Image Retrieval is analyzed in [6].

With the Internet grow several format were proposed for media scenes description. Some of them are open standards such as: VRML1, X3D2, SMIL3, SVG4, MPEG-4 BIFS, MPEG-4, XMT, MPEG-4, LaSER, COLLADA5, published by ISO, W3C, etc. Others standards are proposed by Adobe like Flash or 3ds by Autodesk.

Television has become the most important mass medium. There is also a fast growing community for videos in the web (e.g. YouTube). Serious attempts are done to give the users interactive functions. For these purpose standards such as MHEG, DAVIC, Java TV, MHP, GEM, OCAP and ACAP has been developed. Efficient Video-Stream Filtering becomes important issue.

There exist a large number of standards for representing audiovisual metadata such as the Material Exchange Format (MXF), an example for a container format, the Digital Picture Exchange (DPX), used for image sequences in digital cinema production, the Digital Cinema Package (DCP), used to transport digital movies and associated metadata to cinemas and others.

2 Multimedia Standards Description

Some of the most used multimedia standards are discussed.

2.1 MPEG Standards

The MPEG family includes MPEG-1, MPEG-2, MPEG-4, MPEG-7 and MPEG-21 [9]. MPEG-1 is a multimedia transmission standard that allows the combining and synchronizing of multiple media types (e.g. digital audio and digital video). MPEG-1 was primarily developed for CD-ROM multimedia applications. Because it is primarily focused on multimedia computers, there is no defined way to process interlaced video. The compression processes used in MPEG-1 systems can compress digital video up to 52 to 1.

MPEG-2 is a frame oriented multimedia transmission standard that allows the combining and synchronizing of multiple media types. MPEG-2 is the current choice of video compression for digital television broadcasters as it can provide digital video quality with a data rate of approximately 3.8 Mbps. The MPEG-2 digital video encoding processes can compression digital up to approximately 100 to 1. MPEG-2 became an international standard in 1996.

MPEG-4 is a digital multimedia transmission standard that was designed to allow for the transmission of digital video on packet data systems. A key feature of MPEG-4 is its ability to manage separate media components within image frames. MPEG-4 was originally aimed at medium-resolution video conferencing over low bandwidths at low frame rates but was later modified to define coding of audiovisual objects in applications such as Internet video, virtual reality games, interactive shopping and media databases. The main functionalities addressed are the ability to interact with objects in scenes, universal accessibility in terms of storage and transmission media, and improved compression compared to earlier standards. An example of MPEG-4 system is the ANIMATION system – a system for animation scene and contents creation, retrieval and display [5]. The object recognition mechanism in the ANIMATION system includes three steps: (1) Low level animation images analysis - element recognition based on the Attribute Relational Graphs (ARG); (2) Object recognition, based on production rules with degree of recognition; (3) Image interpretation, based on the Dempster-Shafer theory of evidence.

MPEG-7 is a standard that can be used to describe the characteristics and related information about digital media objects. The MPEG-7 system is an XML based language. The MPEG-7 system uses description definition language (DDL) to describe the characteristic objects using existing and custom created definitions. MPEG-7 was designed to standardize a set of Description Schemes and Descriptors where:

- Descriptors (D), define the syntax and the semantics of each feature (metadata element); and
- Description Schemes (DS), specify the structure and semantics of the relationships between their components that may be both Descriptors and Description Schemes. The main description schema is given in Figure 2.

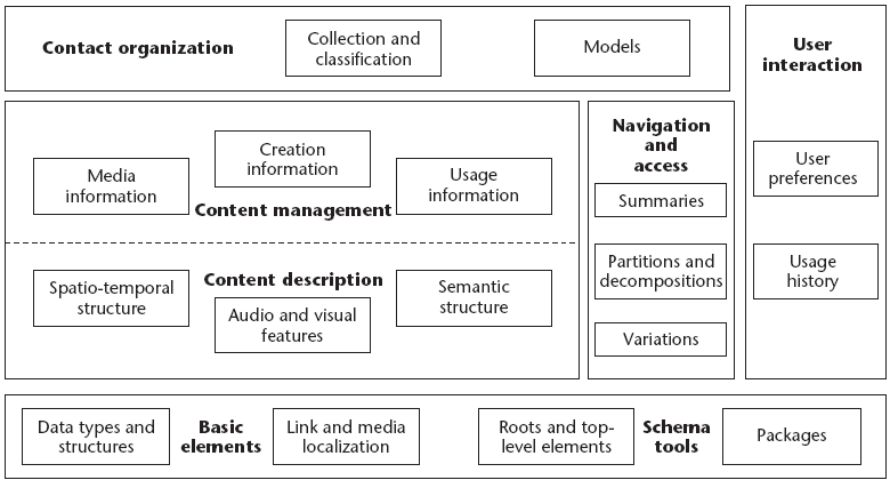

Fig. 2. The main description schema for MPEG 7

The mail description tools in MPEG 7 include:

- Basic Elements (used by the other Visual descriptors): grid layout, time series, 2D–3D multiple views, spatial 2D coordinates, and temporal interpolation;
- Color Descriptors: Color Space, Color Quantization, Scalable Color, Dominant Color, Color Layout, Color Structure, and Group-of-Frames/Group-of-Pictures Color;
- Texture Descriptors: Homogeneous Texture, Non- Homogeneous Texture (Edge histogram), and Texture Browsing;
- Shape Descriptors: Region-Based, Contour-Based, and 3D Shape;
- Motion Descriptors (for video): Motion Activity, Camera Motion, Parametric Motion, and Motion Trajectory;
- Location Descriptors: Region Locator and Spatial–Temporal Locator.

The audio MPEG-7 descriptors are given in Figure 3.

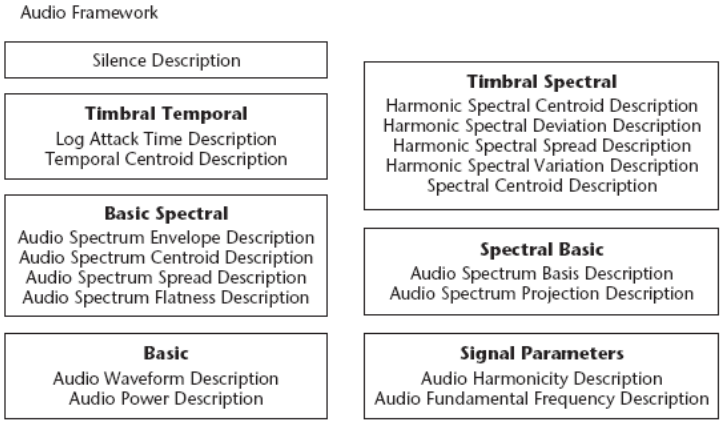

Fig. 3. Audio MPEG-7 descriptors

MPEG-21 aims to defining a normative open framework for multimedia delivery and consumption for use by all the players in the delivery and consumption chain. This open framework provides content creators, producers, distributors and service providers with equal opportunities in the MPEG-21 enabled open market. The streaming instructions facilitate the fragmentation of content-related metadata, the association of media and metadata fragments with each other, and the synchronized streaming and processing of those fragments.

Digital Items can be expressed using Digital Item Declaration Language (XML schema) and Digital Item = resources + metadata + structure, where resource is the individual asset(s); metadata are information about or pertaining to the Item and structure is the relationships between the parts of the Item. MPEG-21 digital identification declaration model is given in Figure 4.

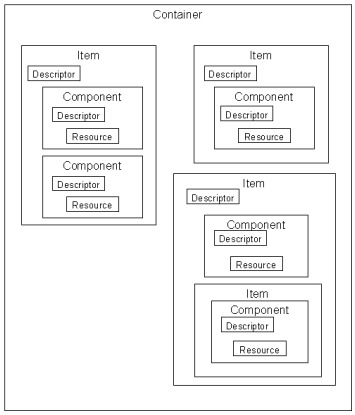

Fig. 4. MPEG-21 digital identification declaration model

2.2 Standards for Scene Description

There are several standards for scene description. Figure 5 gives a comparison on some of them [1].

	VRML	X3D	SMIL	SVG	MPEG-4
Standard name	Virtual Reality Modelling Language	eXtensible 3D	Synchronised Multimedia Integration Language	Scalable Vector Graphics	Coding of audio-visual objects
Organisation	ISO	ISO	W3C	W3C	ISO
Publication Year	1997	2005	1998-2007	2003	1998-2007
Media types	3D graphics and ref. to video / audio	2D/3D graph-ics and ref. to video / audio	Ref. to 2D graphics, video / audio	2D graphics, text and image	Still image, video, audio, 2D/3D graphics
Animation	Linear Interpolation	Linear Interpolation	Local animation	Not relevant	Linear and higher order interpolation
Compression	No	Yes	No	No	Yes
Streaming	Only for video / audio	Only for video / audio	Only for video / audio	Not supported	Video, audio, scene graph, graphics animation
Synchronisation	Only between animation tracks	Only between animation tracks	Not supported per frame	Not relevant	Full support
Interactivity	Yes	Yes	Yes	Yes	Yes

Fig. 5. Comparison on difficult formats for scene description

2.3 Open Standards for Interactive TV

Around 1960, there were shows, where viewers called in and participated in quiz shows. In 1974, the teletext was developed in the United Kingdom. At the IFA (Internationale Funkausstellung Berlin), 1979 the Tele-Dialog was introduced. It was a televoting system which allowed viewers to participate in polls for TV shows by calling specific telephone numbers. Full Service Network by Time Warner, was launched in December 1994. This trail provided several interactive services, like video-on-demand, program guide, video games, and home-shopping to the customers. In 18 months only 65 people subscribed and therefore it was closed.

Levels of interactivity [4] can be described as: Level 1 - Basic TV, Level 2 - Call-In-TV, Level 3 - Parallel TV, Level 4 - Additive TV, Level 5 - Service on Demand, Communicative TV, and Level 7 - Fully Interactive TV. The Multimedia and Hypermedia Information Coding Expert Group (MHEG) [10], a subgroup of the International Organization for Standardization (ISO), published the MHEG standard in 1997. The MHP (Multimedia Home Platform) architecture is given in Figure 6.

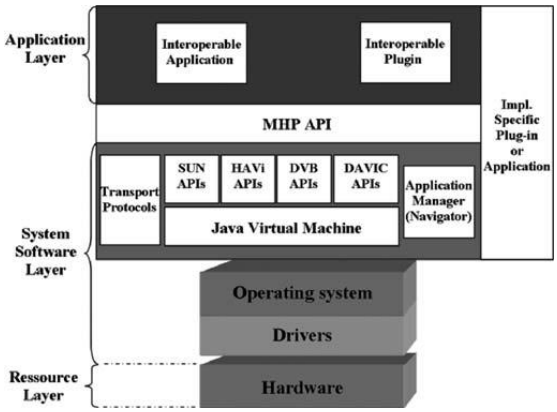

Fig. 6. The MHP architecture

2.4 Efficient Video-Stream Filtering

We define a scenario in which we presume a certain number of virtual sources such as channels of a digital video broadcaster [3]. Each source sends a video stream associated with an MPEG-7 stream that contains the description of this video stream. These streams are sent through a generic transmission channel (see Figure 7) to the receiver stations (e.g., set–top–boxes, digital multimedia recorders, media center PCs, etc.). The receiving station has a number of filters that select, within all the video streams that arrive to the station, only the video shots that are similar – within a certain range – to one of the filters.

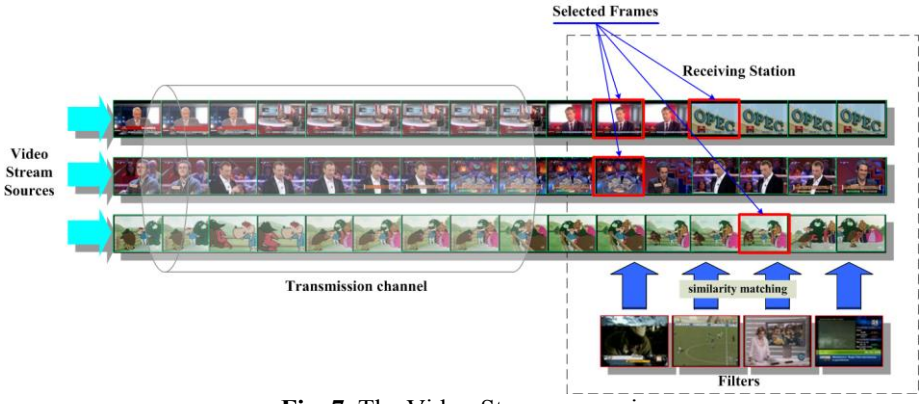

Fig. 7. The Video-Stream scenario

At the sender station we precompiling the similarity between some S-frames to reduce the number of similarity computations made at the receiver station between the query and the S-frame. In Figure 8 some of the results are shown.

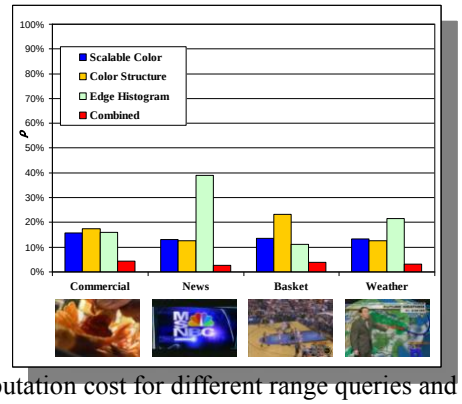

Fig. 8. Distance computation cost for different range queries and three different image descriptors combined

2.5 Standards in the Audiovisual Media

The main issues in the Audiovisual Media standards for audio visual media include [2]:

- EBU P Meta. The European Broadcasting Union (EBU) has defined P Meta as a metadata vocabulary for program exchange in the professional broadcast industry;
- Material Exchange Format. The Material Exchange Format (MXF) is a standard issued by Society of Motion Picture and Television Engineers (SMPTE), defining the specification of a file format for the wrapping and transport of essence and metadata in a single container.

- SMPTE Descriptive Metadata Scheme 1 (DMS-1), formerly known as Geneva Scheme uses metadata sets defined in the SMPTE Metadata Dictionary;
- SMPTE Metadata Dictionary (is a large thematically structured list of narrowly defined metadata elements, defined by a key, the size of the value and its semantics;
- Standard Media Exchange Framework (SMEF) the data model defined by the BBC to describe the metadata related to media items (media objects) and programs and parts thereof (editorial objects), down to the shot level;
- Controlled Vocabulary and Ontologies. Audiovisual content descriptions often contain references to semantic entities such as objects, events, states, places, and times.

3 Conclusion

The aim of the multimedia standards are to enable transparent and augmented use of multimedia resources across a wide range of networks, devices, user preferences, and communities, notably for trading of content.

Acknowledgments. This work was supported in part by Open Access Infrastructure for Research in Europe (OpenAIRE) EU project, the Bulgarian National Science Fund under the Project D002-308 "Automated Metadata Generating for e-Documents Specifications and Standards".

References

1. Granitzer M., Lux M. Spaniol M. (eds.), Multimedia Semantics - The Role of Metadata, Springer-Verlag Berlin Heidelberg (2008)
2. Information technology, Coding of audio-visual objects, Part 11: Scene descriptionand application engine, ISO/IEC 14496-11 (2005)
3. Falchi F., Gennaro C., Savino P., Stanchev P., Efficient Video Stream Filtering, IEEE Multimedia, January-March pp. 52—61 (2008)
4. Ruhrmann G. and Nieland J. Interaktives Fernsehen. Westdeutscher Verlag GmbH, Wiesbaden, Germany (1997)
5. Stanchev P., ANIMATION, System for Animation Scene and Content Creation, Retrieving and Viewing, SPIE, Vol 4672, pp. 86--94 (2002)
6. Stanchev P., Semantic Video and Image Retrieval - State of the Art and Challenges, ACMSE 2007, The 45th ACM Southeast Conference, Winston-Salem, North Carolina, USA, March 24 (2007)
7. Stanchev P., Multimedia Standards, The First International Conference on Advances in Multimedia, MMEDIA 2009, July 20-25, Colmar, France (2009)
8. Vetro A., Christopoulos C., and Ebrahimi T. Special Issue on Universal Multimedia Access. IEEE Signal Processing Magazine, 20(2), March (2003)
9. http://www.mpeg.org/
10. http://www.mheg.org/users/mheg/index.php