EXPERIMENTS ON TRANSMITTING MPEG-4 CONTENT OVER MPEG-2 TRANSPORT STREAM

S. M. Tran, L. Konyha, B. Enyedi, Cs. Szombathy, and K. Fazekas

Budapest University of Technology and Economics, Faculty of Electrical Engineering,
Department of Broadband Infocommunications and Electromagnetic Theory, Hungary
Email: sony@mht.bme.hu

ABSTRACT

The paper describes our experiments on packing the MPEG-4 contents with the delivery layer of MPEG-2 standard. First, we create various MPEG-4 contents, expanding from a simple up to a complex multimedia scene enhanced with interactivity. These information contained in several elementary streams are then split into MPEG-2 transport stream packages. The transport stream is transferred to a terrestrial RF communication channel using DVB-T technology. At the other end of the channel, the MPEG-4 contents are retrieved and evaluated. The goal of our experiments was to investigate the effect of using MPEG-2 delivery layer on the quality of MPEG-4 contents, especially on interactivity-enabled multimedia scenes that highly deviate from the conventional MPEG-2 contents, containing only audio video streams. This study introduces a stable background for a possible further improvement in the integration of MPEG-2 and MPEG-4 contents.

1. INTRODUCTION

The boom of the Internet together with the rapid development in digital storage capacity makes multimedia applications more and more popular. As a feedback, this new technology and the related applications expose some new requirements on the quality of multimedia applications. Among many others, emphasis shall be put on improving the compression of image / video data, on the availability of composing different media data and on the support of interactivity between users and applications. The recently published MPEG-4 standard brings efficient solutions to these new challenges. As a member of the MPEG family, the MPEG-4 standard inherits and improves all the virtues of its predecessors. Going further than offering a tool-set for efficient transmitting and / or storing a huge amount of digital image / video, the standard promotes the concept of multimedia scene supported by the definition of BInary Format for Scenes (BIFS). BIFS enables the construction of complex multimedia scene graph where the nodes of the graph are the media objects possessing audio and video data. The intermediate role of BIFS between media data and the final (composed) content provides a flexible way to manipulate various types of information in an MPEG-4 scene, such as scheduling, coordinating in temporal and / or spatial domain, synchronization, processing interactivity, etc.

During the design process of the MPEG-4 standard, it became clear that the specification of the standard should not follow the structure of its predecessor MPEG-2, which defines both a content representation format for audio and video data and a way to multiplex the coded data with signaling information into one serial bitstream. Today there are a number of different transport options for digital multimedia data. Most prominently, such transport options are provided by MPEG-2 itself and by the protocols that form the Internet. Other options in the telecommunication domain include Asynchronous Transfer Mode (ATM) and H223 videoconferencing multiplex defined by ITU-T.

The Digital Video Broadcasting (DVB) Project is an industry-led consortium of more than 300 broadcasters, manufactures, network operators, software developers, regulatory bodies and other institutions from more than 35 countries, committed to design a global MPEG-based standard for the delivery of digital television and data services. The success of the DVB project in the cable and satellite channels step by step expands to the low cost but highly interfered communication environment: the terrestrial channel. The investigation of transporting MPEG-4 data with DVB-T technology (i.e. the terrestrial transmission method) is twofold: it does not only demonstrate the protection capability of DVB for highly compressed data (the MPEG-4 content), but also reveals some possible shortcomings of the MPEG-4 standard when using external delivery layers.

The paper is structured as follows. The next section gives a short review on the standard way to deliver MPEG-4 content in MPEG-2 streams. In Section 3, we present our experimental scheme to evaluate the performance of the proposals. We close our paper with the conclusions and perspectives in the last section.

2. BACKGROUND OF TRANSPORTING MPEG-4 OVER MPEG-2

Due to the widespread availability of the MPEG-2 infrastructure, especially in broadcasting environments, it was important to make sure that the MPEG-4 data can be easily embedded in there. Indeed, the MPEG-2 design team was farsighted enough to leave room for bitstream syntax extension that could be exploited for an MPEG-4 over MPEG-2 amendment to MPEG-2 systems, which has subsequently become part of the second edition of the MPEG-2 system standard. In this paper, we discuss only the delivery of MPEG-4 data within MPEG-2 transport streams (TSs), which can be directly fed to DVB chains for broadcasting. According to the second edition of the MPEG-2 standard, transport of MPEG-4 data over MPEG-2 system comes in two ways:

Stream-based: the MPEG-4 audio or MPEG-4 video streams are simply added to an MPEG-2 frame. This makes it possible to design services that use MPEG-4 compression instead of MPEG-2 compression. However, interactivity (using a clickable button, moving objects around the screen, *etc.*) cannot be realized in this mode; or, more precisely, the standardized MPEG-4 system tools supporting these purposes are not available. The fact that the inserted media streams are actually MPEG-4 coded is expressed by the newly defined *stream-type* value carried in the Program Map Table (PMT). According to the MPEG-2 standard, each elementary stream is associated with a unique *stream-type* value, which is used as the first level of identifying the stream itself. Furthermore an MPEG-4_video_descriptor or MPEG-4_audio_descriptor is attached to the MPEG-4 stream also in the MPEG-2 system tool (PMT data structure). They convey the profiles and level information about the stream ─ that is, about the precise coding tools used, which enable the proper decoding process.

Scene-based: Complete MPEG-4 presentations with the attractive interactivity capability are retained in this mode. MPEG-4 content consists of the initial object descriptor and a variable number of streams, from the conventional audio-visual streams to the object descriptor streams, scene description streams, *etc*. Each of the MPEG-4 streams shall be contained in an SL-packetized stream and may optionally be multiplexed into a FlexMux stream, both defined in [4]. For transferring in an MPEG-2 Transport Stream, these SL-packetized streams and FlexMux streams shall be encapsulated either in PES packets or in *ISO_IEC_14496_sections* prior to Transport Stream packetization and multiplexing. Besides the *stream_id* and *stream_type* values, which are used to identify the nature of the associated streams in MPEG-2, the encapsulation also ensures the own addressing method of MPEG-4 based on *ES_IDs* ([4]). The assignment of the *ES_IDs* to streams is processed through the introduction of several new descriptors, such as *SL_descriptor*, *FMC_descriptor*. Another important descriptor, namely *IOD_descriptor* contains the Initial Object Descriptor, which is the key to construct the MPEG-4 scene from the *ES_IDs* referring to the underlying streams.

3. THE EXPERIMENTAL STRUCTURE FOR TRANSPORTING MPEG-4 OVER MPEG-2

Figure 1 shows the full scheme we use to transport MPEG-4 content over MPEG-2. The scheme contains the following functional units:

MPEG-4 coder: The very first section of transmission is to encode media data in an MPEG-4 compliant way. We applied software solution to compress three types of media data: audio, video and visual still texture. They are main material to build up various type of content. The visual and audio objects are encoded with the tools defined in the *Main profile* of MPEG-4 [5] [6].

Scene composition: this section makes our experimental scheme different from any transmission of MPEG-2 content. We deployed the new capability of MPEG-4 BIFS in a high abstraction level of an authoring tool to efficiently build up a complex multimedia scene. The second dialogue in Figure 1 is a snapshot of some settings available in our authoring tool. For instance, visual objects such as a video, a graphic or a text can be inserted into the scene; some objects can be coupled with compressed media data produced by an **MPEG-4 coder**, such as audio, video objects; the temporal and spatial position of the objects can be controlled. The so-called Script object ─ a key-important element for interactivity in MPEG-4 ─ can be added and specified with the assistance of the authoring tool. The current version of the authoring tool supports all BIFS features in the *Main2D* scene graph profile [4].

Stream combiner: this section is the combination of software and hardware implementation. Our own developed software is used in conjunction with the Rohde&Schwarz MPEG-2 application to generate a virtual MPEG-2 TS, which encapsulates the MPEG-4 content obtained from **Scene composition**. Both stream-based and scene-based encapsulations are available. The virtual stream, *i.e.* storage-format of the TS on computer disks is then converted physically into binary TS with the assistant of the Rohde&Schwarz MPEG-2 measurement generator [9].

DVB-T transmitter: In this section, the obtained TS is modulated and transmitted over terrestrial channel with the DVB-T technology. We use the Rohde&Schwarz TV

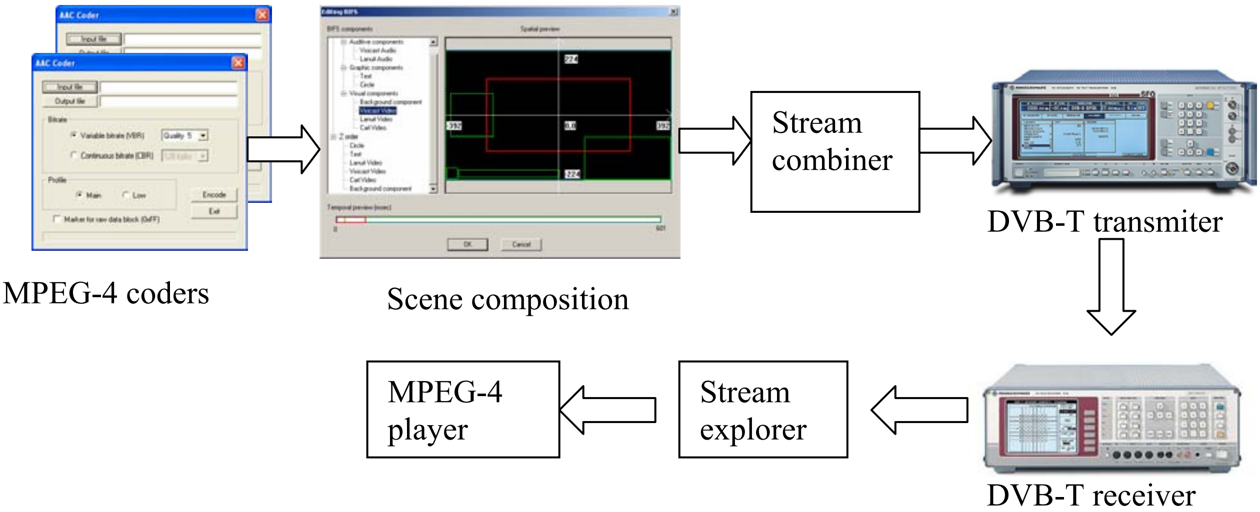

Figure 1: Scheme of encapsulating MPEG-4 content within MPEG-2 transport stream.

Test Transmitter to convert TS from the base-band form to the modulated one, which is compliant with DVB-T.

DVB-T receiver: it is the first functional unit at the other end of the terrestrial communication. We apply the Rohde&Schwarz TV Test Receiver to retrieve back the base-band form of the TS from the modulated radio-frequency. In parallel, we used the plug-in card DigiTV from Nebula Electrics Ltd. [10] for direct access to the base-band TS from computer, which facilitates the retrieval of MPEG-4 content for replaying with computer-based software player.

Stream explorer: Depending on the instrument we use in the **DVB-T receiver** section, software implementation of this unit exposes different functionalities. In case of the TV Test Receiver, the Stream Explorer provides an interface to the receiver (through serial or parallel port) to gain back the storage-format of the received TS. In case of the plug-in card, it creates low level of services to access the arriving data, which facilitate the continuous playback of the scene in higher level applications (MPEG-4 player). The output of this section is directed to a software player in the next step. Therefore Stream explorer also takes the responsibility for producing appropriate format of the data, which is acceptable for players. We reserve two modes to communicate with MPEG-4 players: MPEG-4 format and file and streaming format, for the offline and online playback respectively. The online playback should be understood as a continuous viewing as long as the sender transmits data.

MPEG-4 player: Our scheme ends with some persuasive MPEG-4 players running on computer. Currently we use the open source code VLC player from [12]. The virtue of open source code eases the integration of the **Stream explorer** and the **MPEG-4 player**. The shortcoming of VLC player is that it cannot decode the enhanced system information of MPEG-4 — i.e. there is no interactivity available with VLC. The plug-in software for Windows Media player and QuickTime Player from Envivio [11] overcomes this shortage.

4. CONCLUSION AND PERSPECTIVES

Although nowadays more and more applications exist making use of the efficient audio video compression defined in the recent MPEG-4 standard, most of the advanced features of the standard still remain unexploited — like, among many others, the encoding of arbitrary shaped video object, mesh animation, scene composition and interactivity. Our experiment is an effort to fill up this gap. We created several multimedia scenes enhanced with interactivity, which is impossible with the predecessor of MPEG-4. By experiments, we demonstrated the transfer of the advanced content on the current MPEG-2 platform, where the latter one can be considered as a widespread transport environment in the broadcasting domain.

Currently we examine the effect of noise and other kinds of interference occurring in the terrestrial RF channel on the quality of the MPEG-4 content. The fact that a TS comprises different types of data in its payload, may reduce the efficiency of the protection. This protection guarantees the quality of the content in the case of unmixed TS, where only MPEG-2 audio and video data are encapsulated. We are also working on building up an industrial scheme for transmitting MPEG-4 content. The goal is to enable real time communication and have expandable playback modules at the client-side, supporting various types of terminal as defined in the MPEG-4 standard.

REFERENCES

[1] F. Pereira, T. Ebrahimi, *The MPEG-4 Book*, Prentice Hall PTR, 2002.

[2] E. M. Schwalb, *iTV Handbook, technologies and Standards*, Prentice Hall PTR, 2003.

[3] ISO/IEC JTC1/SC29/WG11 N3050, ISO/IEC 13818-1/FDAM 7: Transport of ISO/IEC 14496 data over ISO/IEC 13818-1, January 2000.

[4] WD 3.0 ISO/IEC14496-1:2002, Information technology — coding of audio-visual objects: — Part 1: system, August 2002.

[5] ISO/IEC JTC1 / SC29 / WG11 N4350, Information technology — Coding of audio-visual objects — Part 2: Visual, July 2001.

[6] ISO/IEC JTC1/SC29/WG11 N3995, Information technology — Coding of audio-visual objects — Part 3: Audio, March 2001.

[7] EN 300 468 V1.3.1, Digital Video Broadcasting (DVB); Specification for Service Information (SI) in DVB systems, February 1998.

[8] ETR211, Digital Video Broadcasting (DVB); Guidelines on implementation and usage of Service Information (SI), August 1997.

[9] http://www.rohde-schwarz.com/, the website of Rohde&Schwarz company.

[10] http://www.nebula-electronics.com/information/info.asp?Code=0001, the website of Nebula Electronics Ltd.

[11] http://www.envivio.com, the website of Envivio company.

[12] http://www.videolan.org/vlc/, the website of a cross-platform media player and streaming server.