DVB-based MPEG-21 Digital Items for Adaptive Multimedia Streaming

Martin Prangl, Christian Timmerer, Klaus Leopold, and Hermann Hellwagner
Department of Information Technology (ITEC), Klagenfurt University, Klagenfurt, Austria
E-mail: *{firstname.lastname}@itec.uni-klu.ac.at*

Abstract - *In this paper, we present a prototype application that receives Digital Video Broadcast (DVB) TV streams on a PC, extracts semantic and syntactic metadata from the MPEG-2 Transport Stream (TS), maps them to MPEG-7 compliant descriptors, and packs these metadata and the associated multimedia content into standard compliant MPEG-21 Digital Items (DIs). In this form, i.e., as DIs, the DVB content can be searched and accessed by clients such as PCs and PDAs, and can be transported over IP networks. The DVB server can be used as a live source of multimedia content to demonstrate multimedia streaming (e.g., via RTP) and adaptation to diverse devices.*

Keywords - *DVB, metadata, MPEG-7, MPEG-21, Digital Items*

1. INTRODUCTION

Although video-on-demand services are becoming increasingly popular, the most widespread audio/video (AV) content is still produced/consumed in the broadcast community such as for/on television (TV). TV content is usually consumed with special purpose devices such as TV sets. Quite recently, Digital Video Broadcast (DVB)[1] found its way into users' living rooms, offering a vast amount of high quality digital radio and TV channels. Due to the fact that DVB content is digital, it can also be consumed on regular PCs with ease. However, the computers have to be equipped with special hardware DVB cards that are connected to antennas.

Having the TV channels available on the Internet would enable all devices, which implement the Internet Protocol (IP), to consume DVB content. Usually, these devices use network connections with different characteristics (e.g., available bandwidth, delay, delay variation) through which they can access the DVB content. Additionally, the terminal capabilities of these devices such as available memory, computational power, or resolution of the screen may be diverse as well.

Therefore, one major goal within the multimedia research community is the development of Universal Multimedia Access (UMA) [1] strategies and technologies, which enable users to consume any kind of multimedia content, anywhere, and anytime. To achieve this goal, the multimedia contents have to be adapted to meet the limitations of the user's terminal and network characteristics. Such multimedia adaptation could be, e.g., transcoding from one video format to another or scaling a video in the spatial domain such that it fits on the terminal's screen. Furthermore, the content must also be adapted such that a user has an informative experience; in other words, the "end point" of universal multimedia consumption is the end user and not the terminal. Universal Multimedia Experience (UME) [1] might include, e.g., insertion of subtitles into a video to allow deaf users to follow the spoken content in a video.

Multimedia adaptation is a key means for addressing the challenges of UMA. The basic idea is to transform the content before or even during transmission taking into account the usage environment, i.e., user and network characteristics, terminal capabilities, or characteristics of the natural environment in which the content is consumed. In order to support transparent and interoperable multimedia adaptation and consumption, metadata will become a key issue [2].

In this paper we present an approach and a prototype application that map metadata from DVB streams to MPEG-7 Multimedia Description Schemes (MDS) [3]. Furthermore, both the DVB content and its metadata are packed into MPEG-21 Digital Items (DIs) [4] providing an open and interoperable interface to the multimedia content which is accessible through IP-based networks such as the Internet.

The remainder of this paper is organized as follows. Section 2 provides a brief overview of the MPEG standards used in this work. The mapping of DVB-based metadata into the MPEG-7 format as well as the structure of the Digital Item is described in Section 3. In Section 4 we present our prototype implementation and Section 5 concludes the paper and provides future prospects.

2. MULTIMEDIA METADATA

MPEG-7 provides comprehensive XML-based description schemes (DSs) for describing and annotating multimedia content, both at a syntactic and at a semantic level. The former includes DSs for describing the low-level signal characteristics such

[1] http://www.dvb.org

as bit-rate or frame size of a video whereas the latter facilitates the construction of semantically meaningful descriptions, e.g., title or leading actor of a movie.

In MPEG-21, the transaction of so-called Digital Items among users is one of the core concepts. The MPEG-21 Multimedia Framework aims at achieving interoperable and transparent access to multimedia content. In principle, a Digital Item is a container format for (multimedia) resources and metadata within a standardized structure. A user is defined as any entity (including individuals, communities, organizations as well as software agents) that interacts in the MPEG-21 environment or makes use of Digital Items.

A vital and important part of MPEG-21 with regard to UMA is Digital Item Adaptation (DIA) [5] which addresses, among others, the description of the usage environment as mentioned in Section 1. These description formats enable the development of interoperable adaptation engines, providing Digital Items which meet the constraints imposed by the usage environment where the multimedia content is expected to be consumed.

3. AUTOMATIC GENERATION OF MPEG-21 DIGITAL ITEMS FROM DVB STREAMS

Our approach is based on the concept that current DVB implementations provide, in addition to the AV content, also its associated metadata, both multiplexed into an MPEG-2 Transport Stream (TS) [6]. The metadata is provided through the Electronic Program Guide (EPG) which is encoded in the DVB System Information (DVB-SI) tables. Hence, this kind of metadata is referred to as semantic metadata which includes *title*, *publisher*, an *extended title or short abstract* as well as *start time* and *duration* of the respective television program. The second type of metadata is extracted directly from the MPEG-2 stream and, thus, is referred to as syntactic metadata. It contains information about the characteristics of the encoded AV stream such as the used *codec*, *spatial resolution*, *frame rate*, and *bit rate* of the video stream. For the audio the *codec* and *bit rate* is extracted as well. Additionally, the MPEG-2 Packetized Elementary Stream (PES) provides the *sample rate* and *number of audio channels*.

3.1. Semantic Metadata Extraction

Semantic information from the EPG stream is mapped to MPEG-7 MDS' `CreationInformation` DS and `MediaTime` data type, respectively, as shown in Document 1 and Document 2. Therefore, the plain text encoded EPG stream needs to be parsed and transformed into valid MPEG-7 descriptions as follows.

```
<CreationInformation>
 <Creation>
  <Title type="main">MacGyver</Title>
  <Abstract>
   <FreeTextAnnotation>Bei einer
internationalen Sportveranstaltung trifft
MacGyver ...
   </FreeTextAnnotation>
  </Abstract>
  <Creator>
   <Role
     href="urn:mpeg:mpeg7:cs:
     RoleCS:2001:PUBLISHER"/>
   <Agent xsi:type="OrganizationType">
    <Name>Kabel 1</Name>
   </Agent>
  </Creator>
 </Creation>
</CreationInformation>
```

Document 1 — MPEG-7 creation information extracted from DVB MPEG-2 PES.

```
<MediaTime>
 <MediaTimePoint>
   2004-12-03T15:14:00
 </MediaTimePoint>
 <MediaDuration>PT00H61M00S</MediaDuration>
</MediaTime>
```

Document 2 — MPEG-7 media time extracted from DVB MPEG-2 PES.

The `CreationInformation` DS describes information about the creation and production of multimedia content. In our case, the title from EPG is mapped to the `Title` element of MPEG-7, the extended title or abstract is provided through the `Abstract` element, and the name of the publisher and its role is contained within the `Creator` element.

The start time and duration of the DVB program are indicated by using the `MediaTimePoint` and `MediaDuration` elements, respectively. The `MediaType` data type specifies a notion of time encoded within the described media.

3.2. Syntactic Metadata Extraction

The syntactic metadata is directly extracted from the media stream and incorporated in an MPEG-7 description based on the `MediaInformation` DS as shown in Document 3. It is logically divided into three blocks. The first block provides general information about the content and the bit rate, i.e., the `Content` and `BitRate` element. The second and third blocks supply specifics regarding the video and audio coding format within the `VisualCoding` and `AudioCoding` elements.

Finally, both the semantic and syntactic metadata is combined within a single MPEG-7 description which is not shown in this paper due to space limits.

```
<MediaInformation id="S19.2E-0-12480-899">
 <MediaProfile><MediaFormat>
  <Content href="MPEG7ContentCS:2001">
   <Name>audiovisual</Name>
  </Content>
  <BitRate average="3500000"
    maximum="4000000">4000000</BitRate>
  <VisualCoding>
   <Format href="urn:mpeg:mpeg7:cs:
      VisualCodingFormatCS:2001:2"
      colorDomain="color">
    <Name xml:lang="en">MPEG-2 Video</Name>
   </Format>
   <Frame height="576" width="720"
      rate="25.0"/>
  </VisualCoding>
  <AudioCoding>
   <Format href="urn:mpeg:mpeg7:cs:
      AudioCodingFormatCS:2001:3.2">
    <Name xml:lang="en">
      MPEG-1 Audio Layer II</Name>
   </Format>
   <AudioChannels front="0" side="2"
      rear="0" lfe="0" track="0">2
   </AudioChannels>
   <Sample rate="48000.0" bitsPer="192"/>
  </AudioCoding>
 </MediaFormat></MediaProfile>
</MediaInformation>
```

Document 3 — MPEG-7 media information extracted from DVB MPEG-2 PES.

3.3. MPEG-21 Digital Item Structure

For each DVB channel we generate a so-called MPEG-21 Digital Item Declaration (DID) document which represents the Digital Item. Note that the complete Digital Item comprises the DID and its resource, i.e., the reference to the DVB stream. The proposed structure of the MPEG-21 DID is shown in Document 4. The DID consists of one `Item` element which in turn contains a `Component` element including the metadata and referencing the DVB stream. A component as defined in MPEG-21 binds a set of descriptors, i.e., metadata, to a resource. In particular, the MPEG-7 description which can be found as a child element of the `Descriptor` element is associated with the actual resource. Note that the DVB streams referenced by the two `Resource` elements are considered to be equivalent and are only differentiated by the Uniform Resource Identifier (URI) scheme, i.e., "http" refers to the HyperText Transport Protocol (HTTP) while "rtp" to the Real-time Transport Protocol (RTP).

Having now the DID including MPEG-7 metadata and the reference to the actual DVB stream, interoperability is guaranteed due to the fact that only normative descriptions are used. Thus, it becomes possible that any MPEG-21 capable peer may interact with this Digital Item, e.g., searching through the Digital Item repository for a certain DVB channel or adapting the Digital Item before streaming the content to a terminal where it is consumed. The latter becomes more and more important due to the heterogeneity of existing terminals and networks.

```
<did:DIDL>
 <did:Item>
  <did:Component>
   <did:Descriptor>
    <did:Statement mimeType="text/xml">
     <Mpeg7>
       <!-- MPEG-7 info here -->
     </Mpeg7>
    </did:Statement>
   </did:Descriptor>
   <did:Resource mimeType="video/mpeg"
     href="http://www.some-http-
location.com:3000/S19.2E-0-12480-899"/>
   <did:Resource mimeType="video/mpeg"
     href="rtp://www.some-rtp-
location.com:3000/S19.2E-0-12480-899"/>
  </did:Component>
 </did:Item>
</did:DIDL>
```

Document 4 — Proposed MPEG-21 DID structure for DVB-based multimedia content and metadata.

4. PROTOTYPE IMPLEMENTATION

This section describes our prototype implementation for the automatic generation of MPEG-21 Digital Items from video and audio channels broadcasted via DVB. The proposed architecture is depicted in Figure 1. The system is based on a Linux system using a standard PC with a built-in DVB-S card for receiving the DVB content via satellite. Furthermore, other technologies like DVB-C or DVB-T for receiving the content via cable or terrestrial antenna would be possible. For tuning and accessing the DVB streams, the *Video Disc Recorder (VDR)* toolkit[2] is used. For streaming of the channels over the network, a streaming server is needed. In our solution the VDR plug-in *streamdev* is used for this task. The channel is streamed in an MPEG-2 Packetized Elementary Stream (PES) format containing the video and audio track together with the synchronization information.

For the DID generation process, four main parts are needed: the EPG parser, the stream parser, the MPEG-7 generator, and the MPEG-21 generator.

The EPG parser is used to extract the provided semantic metadata from EPG which is encoded in the DVB System Information (DVB-SI) tables. A helpful tool to receive the DVB-SI data is *dvbsnoop*[3], however, in our solution we use VDR as set top box and this toolkit is able to receive the EPG data of the channels as well. After extraction, the useable information about title, abstract, start time and duration of the selected channel is passed to the MPEG-7 generator which generates the MPEG-7 description as described in Section 3.

The second necessary module is the stream parser. In contrast to the EPG parser, it is used to extract the syntactic metadata from the PES. It includes the following coding information pertaining to the AV stream:

[2] http://www.cadsoft.de/vdr

[3] http://dvbsnoop.sourceforge.net

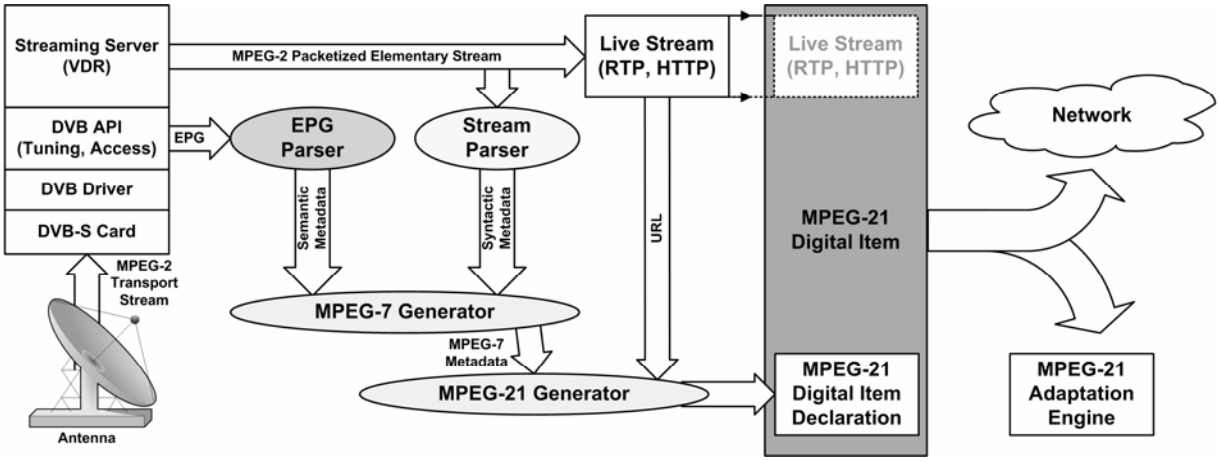

Figure 1 — *dvb2di* (DVB-to-Digital Item) server architecture.

- Video codec
- Spatial resolution
- Frame rate
- Video bit rate
- Audio codec
- Audio bit rate
- Audio sample rate
- Number of audio channels

For the extraction process of this data, helpful libraries like *dvbsnoop* and *ffmpeg*[4] are used. The resulting metadata form the second input of the MPEG-7 generator.

In order to produce the MPEG-7 description of the selected channel, the semantic and syntactic metadata need to be combined and translated into an XML-based format. This task is done in the MPEG-7 generator library which is based on a set of string parsing and conversion methods. For example, the input string *"mp2*" is converted to *"MPEG-1 Audio Layer II"*. The resulting converted values are then embedded in suitable MPEG-7 tags as explained in Section 3.

In order to produce the final DID, the last module required is the MPEG-21 Generator. First, the produced MPEG-7 description has to be encapsulated in an MPEG-21 `Descriptor` element. Second, the media stream URL has to be encoded. It is defined by the streaming plug-in and consists of the protocol type like http, hostname, port and the channel id. The final output is the DID structure like explained in Section 3.3. Together with the live stream, these outputs form the MPEG-21 DI which is available for further use, for example in MPEG-21 adaptation engines.

5. CONCLUSION

We have described an approach and a prototype implementation that receives and maps DVB streams to valid MPEG-21 Digital Items, containing media data and associated metadata. These DIs and, thus, the DVB content, can be accessed from computing devices like PCs and PDAs over standard IP networks. We will use this system as a live content source for experiments on streaming and adapting multimedia data to the characteristics of diverse end user devices and networks (Figure 2).

Figure 2 — DVB stream adapted to a PDA.

REFERENCES

[1] F. Pereira and I. Burnett, "Universal Multimedia Experiences for Tomorrow," *IEEE Signal Processing Magazine*, vol. 20, no. 2, Mar. 2003, pp. 63-73.

[2] F. Nack, "The Future in Digital Media Computing is Meta," *IEEE MultiMedia*, vol. 11, no. 2, Apr.-Jun. 2004, pp. 10-13.

[3] P. Salembier and J. R. Smith, "MPEG-7 Multimedia Description Schemes," *IEEE Trans. Circuits and Systems for Video Technology*, No. 6, 2001, pp. 748-759.

[4] I. Burnett, R. Van de Walle, K. Hill, J. Bormans, and F. Pereira, "MPEG-21: Goals and Achievements," *IEEE MultiMedia*, vol. 10, no. 6, Oct.-Dec. 2003, pp. 60-70.

[5] ISO/IEC 21000-7:2004, Information Technology — Multimedia Framework (MPEG-21) — Part 7: Digital Item Adaptation, 2004.

[6] ISO/IEC 13818-1:2000, Information Technology — Generic Coding of Moving Pictures and Associated Audio Information: Systems, 2000.

[4] http://ffmpeg.sourceforge.net/