# Resource Conversion in MPEG-21 DIA

**Mariam Kimiaei Asadi[1] and Jean-Claude Dufourd[2]**

[1]Ecole Nationale Supérieure de Télécommunications, Paris, France
[2]Streammezo, Paris, France

## Abstract

MPEG-21 standard aims at defining elements needed to support a multimedia delivery chain. This paper introduces our approach for a media resource conversion framework in MPEG-21 DIA. We propose description tools for describing conversion parameters and preferences. We show that the metadata given in these descriptors, which are provided by either the author of the content or any other involving entity in the adaptation chain, helps the adaptation engine decide on the optimal nature, type and parameters of the adaptation that is to be applied to the media resource.

**Keywords**: multimedia content adaptation, media resource, resource conversion, Digital Item (DI).

## 1. Introduction

The demand, creation and then consumption of multimedia content are rapidly increasing. A large variety of devices, with different capabilities, are used to access to this multimedia content. Additionally, these access devices may be used in different contexts, i.e. in different environments, and by different end users, each having particular consumption preferences. The content providers may also have their own preferences on the adaptation of their content. A complete multimedia adaptation infrastructure is then needed to deal efficiently with all these issues. This motivated MPEG group in June 2000 to start to work on the definition of an enabling normative technology for multimedia applications: MPEG-21 "Multimedia Framework" [1].

MPEG-21 is a standard of ISO family that aims to define elements needed to support a multimedia delivery chain. It has a complete set of context describing tools in its seventh part: DIA (Digital Item Adaptation [2]), which cover the description of all different parameters of the usage environment. DIA also provides a complete support for a particular resource adaptation methodology (BSD-based adaptation), which works on the basis of direct bitstream modification and the XML [3] description of the resource bitstream. At this time, MPEG-21 does not provide a compete support for the usage of any other resource adaptation (also called resource conversion). This is where our work on resource conversion is situated within MPEG-21. We have participated in the establishment of an amendment to DIA that aims at providing the support for resource conversion in MPEG-21 DIA.

Our methodology for support of resource conversion in MPEG-21 is based on the description of conversion-related information. This metadata helps the adaptation engine choose the optimal form of the adaptation and the optimal values of the adaptation parameters. Through our proposed descriptors, the content author can express his preferences and hints for a certain conversion of a particular resource.

In the following section, we describe different types of resource adaptation and explain the missing features of MPEG-21 in the area of support of resource conversion. Section 3 gives a brief state of the art on research works on resource conversion in MPEG-21, while section 4 describes our approach on this subject. We finally conclude the paper in section 5, and present our perspectives on this work.

## 2. Media adaptation

Based on how a resource adaptor changes a media, we define two different main categories of single media adaptation: Direct Bitstream Modification and Resource Conversion. The former concerns resource adaptations tools that directly perform the modifications on the bitstream, and this, mostly by removing packages of data, while the latter concerns adaptation tools, which completely change the structure of the bitstream.

### 2.1. Direct Bitstream Modification

A variety of adapted versions can be retrieved from a single bitstream by performing simple editing-style operations such as data truncation and simple modifications. Media scaling of scalable resources will be categorized under this kind of resource adaptation. It can, therefore, be said that, the media (bitstream) has not been really *changed* but just modified. In order to provide interoperability, it is desirable that a processor

that is not aware of the specific bitstream coding format can be used for this task. For this, MPEG-21 DIA has defined a generic approach for manipulating bitstreams, by providing a method based on XML. A binary media resource consists of a structured sequence of binary symbols, this structure being specific to the coding format. A bitstream is defined as the sequence of binary symbols representing this resource. XML is used to describe the high-level structure of a bitstream; the resulting XML document is called a Bitstream Syntax Description (BS Description, BSD). This description, in most cases, will not describe the bitstream on a bit-per-bit basis, but rather addresses its high-level structure, e.g., how the bitstream is organized in layers or packets of data. With such a description, it is then possible for a resource adaptation engine to transform the BS Description, for example with an XSLT style sheet, and then generate back an adapted bitstream.

## 2.2. Resource Conversion

We define Resource Conversion to be a kind of digital resource adaptation that changes the structure of the Bitstream, by decoding, manipulating and re-encoding the media.

There exist quite numerous resource adaptors, which do not need the Bitstream Description Syntax. These resource adaptors are defined to be working on the basis of Resource Conversion. We consider three main types of Resource Conversion: Transmoding, Transcoding and Transforming. Transmoding is defined as a kind of resource adaptation that changes the modality of the original resource, such as video summarization to a slideshow. Transcoding is any format change of a resource while staying in the same modality, such as BMP to GIF. Transforming is defined as any parameter change of the original resource, while staying in the same modality and format, such as an image resizing or a text translation.

We consider a Resource Conversion to be an atomic process. For example a transmoding conversion is a pure modality change; it changes only the modality and is not supposed to change any other parameter such as resolution. To realize a composed conversion, the atomic conversions can then be cascaded (AND) and paralleled (OR).

MPEG-21 DIA provides a complete support for adaptations of Direct Bitstream Modification type; (g)BSD-based resource adaptation. However at this time, MPEG-21 DIA does not provide a complete support for Resource Conversion. This is where our work on resource conversion is situated within DIA.

# 3. MPEG-21-based approaches for resource conversion

Research works, contributing to MPEG-21, have been done to find necessary support for offline adaptation by modality change, called offline modality conversion [4]. The proposed solutions are limited to the support of static decision taking and adaptation and do not permit to take into account the metadata that is needed for online adaptation. [5] proposes an approach for multi-step adaptation of multimedia resources. The methodology is based on semantic descriptions of transformation steps, which are exploited by a classical state-space manner. The proposed framework relies on descriptions of the resource itself (MPEG-7 [6]), the usage environment of the resource (MPEG-21), as well as declarative descriptions of the transformation tool. The implemented prototype (a simple video resource adaptation engine) employs a knowledge-based engine for finding and executing the needed adaptation sequences. The adaptation module of [7] uses transcoding, transmoding and real-time media modification tools. These adaptations are performed based on content and context descriptions, respectively given by MPEG-7 and MPEG-21 DIA descriptors. The approach also uses some transcoding hints, expressed on MPEG-7.

# 4. Resource Conversion framework

In this section we describe our approach for realization of an MPEG-21 Resource Conversion framework.

Before going through the description of the methodology of our approach, let us define two basic concepts: XDI and CDI. XDI, conteXt Digital Item, is a DI or a fragment of a DI that contains the context information. It is presented in DIDL [8] language and contains the expression of context, commonly given in DIA UED descriptors. It does not contain or does not reference any media resource. CDI, Context Digital Item, is a DI or a fragment of a DI that contains the media resources and the content-related information. It is presented in DIDL.

Having defined these two concepts, we now proceed to describe the principle elements of our approach that are: a) context description, b) content description, c) description of the parameters of conversions and d) the expression of the conversion preferences of the author and the end user.

The context description is provided via DIA UED descriptors. We chose to provide the description of the resource in MPEG-7 descriptors. Figure 1 provides an MPEG-7 resource description for an mp4 video, coded

```
<mpeg7:MediaFormat>
 <mpeg7:Content href="urn:mpeg:cs:ContentCS:2001">
                  <mpeg7:Name>video</mpeg7:Name>
 </mpeg7:Content>
 <mpeg7:FileFormat
              href="urn:mpeg:cs:FileFormatCS:2001">
                <mpeg7:Name>mp4</mpeg7:Name>
 </mpeg7:FileFormat>
 <mpeg7:VisualCoding>
  <mpeg7:Format
       href="urn:mpeg:cs:VisualCodingFormatCS:2001">
           <mpeg7:Name>MPEG-4 Visual</mpeg7:Name>
  </mpeg7:Format>
  <mpeg7:Frame width="240" height="144" rate="24" />
 </mpeg7:VisualCoding>
</mpeg7:MediaFormat>
```

Fig. 1: MPEG-7 Resource description of a MPRG-4 video.

in MPEG-4 with 24 frames per second and with dimension of 240*144.

We have defined XML schemas for the expression of conversion-related information. During several contributions and *Core Experiments*, we proposed these schemas to MPEG-21 DIA. They were not promoted to the standard, under the form described here, but gave birth to an activity –in which we participated– on resource conversion description in DIA. This opened an amendment to DIA, which finally considered a placeholder for the description of conversion parameters in DIA.

Our proposed descriptors contain the description of conversion preferences as well as the description of suggested/recommended generic parameters of a set of transmoding, transcoding and transforming conversions. Due to limited space, here we don't go thought the details of these generic parameters.

The objective is to express the metadata on the conversion of a resource. We are rather interested in the point of view of the author of the resource. Nevertheless, through this descriptor, any peer in the adaptation chain can express his preferences or hints on the parameters of a specific conversion. The author, based on his knowledge on the resources, may wish or need to express his preferences on conversion(s) of a particular content and provide some hints to facilitate, guide or enable the adaptation; this will be possible by using the proposed descriptors.

Through the proposed descriptor, are expressed the hints on conversion parameters. The hints include the descriptions of the most general parameters of a certain conversion for a particular resource. The considered parameters are generic, in other words they are based on no particular algorithm. We have considered a set of conversions, for which, we give the related and generic parameters on which the provider may wish to express his preferences or suggestions.

The structure of the proposed descriptors is extensible, meaning that other conversions and their generic parameters may at any time be added to it.

Concerning conversion preferences, we have considered two attributes for describing preferences (of the author) on a certain conversion of a particular resource: quality and priority. The quality attribute gives a hint at the underlying algorithm about the quality of adaptation, and is mapped to a quality factor of the underlying algorithm. Priority attribute is used to set priorities between different conversions of the same content. The general preferences of the end user are provided by DIA *ConversionPreference* in the corresponding XDI, which describes the context of the usage (terminal, network and user characteristics).

Figure 2 represents a conversion descriptor for a video-to-image transmoding. The conversion parameters, such as destination modality and image key frame, are given. We can imagine the same descriptor for a video-to-slideshow transmoding, with a list of key frames, and the values of duration and importance for each slide. The importance attribute then, would help the adaptation engine to choose the most important frames for a limited bandwidth. Figure 3 shows a conversion descriptor for a text translation transforming. Figure 4 demonstrates how we integrate resource and conversion descriptions into a DID (CDI) and how a composed conversion could be expressed. In this figure "rcd" is the namespace of our RCD (Resource Conversion Description) schema.

Figure 5 depicts the architecture of our MPEG-21 Resource Conversion Engine. The inputs are the CDI, the XDI and the original resource. The CDI contains the description of content in MPEG-7, description of conversion parameters and preferences, as well as references to original resources.

```
<Transmoding quality="1.0" priority="1">
 <TransmodingParameters
        xsi:type="VideoSummarizationParametersType">
   <To href="urn:mpeg:mpeg7:cs:ContentCS:2001">
              <mpeg7:Name>Image</mpeg7:Name>
   </To>
   <Slide importance="hight">
    <mpeg7:MediaTimePoint>
       T01:14:30:12F24
    </mpeg7:MediaTimePoint>
   </Slide>
 </TransmodingParameters>
</Transmoding>
```

Fig. 2: A video-to-image transmoding descriptor.

```
<Transforming priority="1" quality="0.8">
  <Parameters xsi:type="TextTranslationType">
    <Language
          href="urn:mpeg:mpeg21:cs:LanguageCS:2001">
                <mpeg7:Name>French</mpeg7:Name>
    </Language>
    <FontParameters fontSize="20">
        <FontStyle>BOLD</FontStyle>
        <FontFamily>ARIAL</FontFamily>
    </FontParameters>
  </Parameters>
</Transforming>
```

Fig. 3: A text translation transforming descriptor.

```
<DIDL>
 <Item>
  <Component >
   <Descriptor>
    <Statement mimeType="text/xml">
     <dia:Conversion>
      <dia:ConversionInformation>
       <dia:ConversionDescription
          xsi:type="rcd:TransmodingConversionType">
  <!-- text-to-image transmoding descriptor !-->
       </dia:ConversionDescription>
       <!-- AND !-->
       <dia:ConversionDescription
           xsi:type="rcd:TranscodingConversionType">
 <!-- image reformatting transcoding descriptor !-->
       </dia:ConversionDescription>
      </dia:ConversionInformation>
      <!-- OR !-->
      <dia:ConversionInformation>
       <dia:ConversionDescription
          xsi:type="rcd:TransformingConversionType">
<!-- text translation transforming descriptor !-->
       </dia:ConversionDescription>
      </dia:ConversionInformation>
     </dia:Conversion>
    </Statement>
   </Descriptor>
   <Descriptor>
    <Statement mimeType="text/xml">
 <!-- here goes the description of the resource !-->
    </Statement>
   </Descriptor>
   <Resource mimeType="text/text" ref="mytext.txt"/>
  </Component>
 </Item>
</DIDL>
```

Fig. 4: Integration and AND/ORing of conversions in a CDI.

The XDI contains the description of the context (user preferences, terminal and network characteristics) in DIA UED. The DID&DIA parsers parse the CDI and XDI and restore the metadata in an internal structure. The job of parsing and validating the RCD descriptors is also integrated into the DIA parser.

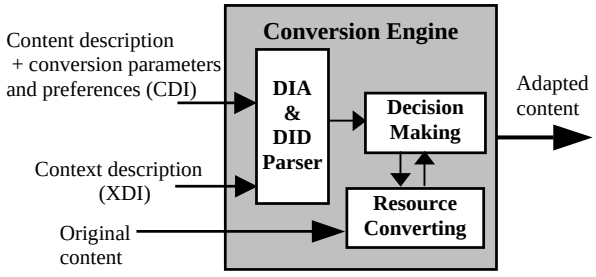


Fig 5. Architecture of Resource Conversion Engine

The Decision Maker uses this metadata to decide on the optimal type of conversion. It first verifies the description of the content and high-priority constraints such as terminal modality/format support and constructs a list of possible conversions. Then based on the preferences of the end user and the author, it chooses a conversion, for which it fetches the required parameters values in the description of conversion parameters. It then sends the chosen type of conversion, together with the values of necessary parameters, to the resource converter, which will convert the resource. The resource converter includes a set of free resource converters such as ImageMagick (http://www.imagemagick.org) and FFMPEG (http://ffmpeg.sourceforge.net) as well as a set of converters developed at ENST.

Using an on-demand real-time Resource Conversion engine in an adaptation system, removes the need for heavy authoring tasks. To be sure of the adaptability of his content, the author is no more obliged to provide different versions of his content. However, a client-server architecture for on-demand resource conversion becomes problematic when numerous adaptations are demanded at the same time by many user requests. This will cause high-delay response-time and or saturation of the server. A distributed architecture, using proxy servers, is obviously a better solution, which can improve the efficiency of the system face to such problem.

## 5. Conclusions and perspectives

In this paper, we presented our approach for support of Resource Conversion in MPEGE-21. The methodology is based on "static" and "hard" description of conversion parameters and preferences. As the continuation of this work we work on a non-static, "soft" and more efficient way for expressing the conversion parameters in MPEG-21 DIA. We believe this is feasible by using DIA *AdaptationQoS* tool for "dynamic" descriptions of conversion parameters.

## 6. References


[1] J. Bormans et al., "MPEG-21 Overview", ISO/IEC JTC1/SC29/WG11/N5231, 2002, Shanghai.

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

[3] Extensible Markup Language 1.0 , W3C Recommendation, 6 October 2000.

[4] Cong Thang T. et al., "Modality conversion for universal multimedia services", Proceedings of WIAMIS 2004, Lisbon, Portugal.

[5] D. Jannach et al.,"A Knowledge Based Approach for Multi-step Media Adaptation", Proceedings of WIAMIS 2004, Lisbon, Portugal.

[6] J. M. Martínez, "MPEG-7 Overview", ISO/IEC JTC1/SC29/WG11/N5525, 2003, Pattaya

[7] J. M. Martínez et al. " Integrating Metadata-Driven Content Adaptation Approaches", Proceedings of EWIMT 2004, London, UK.

[8] ISO/IEC 21000-2:2003, Information Technology —Multimedia Framework (MPEG-21)—Part 2: Digital Item Declaration, 2003.