MERL – A MITSUBISHI ELECTRIC RESEARCH LABORATORY
http://www.merl.com

MPEG-21 Digital Item Adaptation: Enabling Universal Multimedia Access

Anthony Vetro

TR-2004-002 January 2004

Abstract

The access devices of today are becoming increasingly sophisticated. For better or worse, these devices keep us connected at all times to our families, friends, and office. They allow us to give input on urgent matters, share our experiences and emotions, or just say hello. Thanks to multimedia, communication is much widespread and therefore more powerful. However, we face a serious problem of heterogeneity in our terminals, in our networks, and in the people who ultimately consume and interact with the information presented to them. This article describes work developed within the MPEG standardization committee to help alleviate some of the burdens confronting us in connecting a wide range of multimedia content with different terminals, networks, and users. Ultimately, this work will enable what we often refer to as Universal Multimedia Access (UMA).

IEEE Multimedia, January-March 2004

Editor: John R. Smith
IBM

MPEG-21 Digital Item Adaptation: Enabling Universal Multimedia Access

Anthony Vetro
Mitsubishi Electric Research Labs

The access devices of today are becoming increasingly sophisticated. For better or worse, these devices keep us connected at all times to our families, friends, and office. They allow us to give input on urgent matters, share our experiences and emotions, or just say hello. Thanks to multimedia, communication is much widespread and therefore more powerful. However, we face a serious problem of heterogeneity in our terminals, in our networks, and in the people who ultimately consume and interact with the information presented to them. This article describes work developed within the MPEG standardization committee to help alleviate some of the burdens confronting us in connecting a wide range of multimedia content with different terminals, networks, and users. Ultimately, this work will enable what we often refer to as Universal Multimedia Access (UMA).[1]

Background

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.[2] It's crucial to realize, however, that we can't achieve this vision with any one standard, but rather with a collection of technologies developed over the past decade, including scalable content representation formats and the description of multimedia content. Recognizing that there are still some missing elements in this picture, the MPEG-21 committee has proceeded to standardize tools that attempt to fill those gaps.

In this article, we focus on Part 7 of the MPEG-21 standard (ISO/IEC 21000-7), which we refer to as Digital Item Adaptation (DIA). At the time of this writing, the DIA specification is at the penultimate stage of Final Committee Draft;[3] final approval is scheduled for December 2003. Figure 1 shows the general DIA concept: A Digital Item is subject to both a resource adaptation and a descriptor adaptation engine, which together produce the adapted Digital Item. Note that the standard specifies only the tools that assist with the adaptation process, not the adaptation engines themselves.

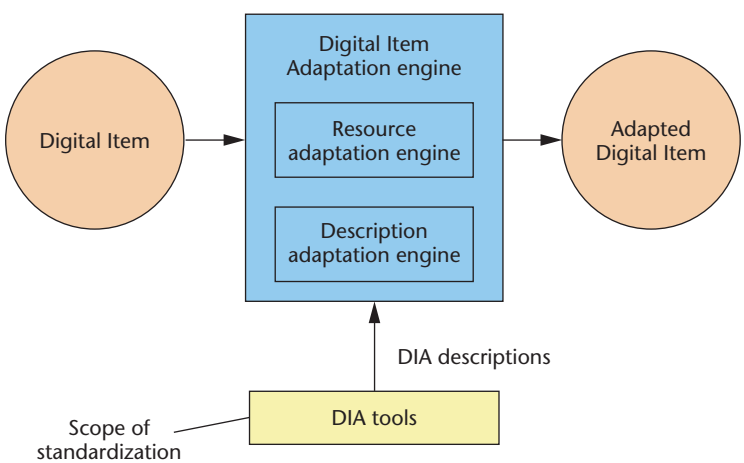

Figure 1. Digital Item Adaptation architecture.

Usage environment description tools

The usage environment description tools describe the terminal capabilities as well as characteristics of the network, User, and natural environment. (In the context of this standard, *natural envi-*

ronment pertains to the physical environmental conditions around a User such as lighting condition or noise level, or a circumstance such as the time and location.) More-specific information is outlined below, along with some potential uses for these descriptions.

Terminal capabilities

In addition to enabling media format compatibility, the terminal capabilities description lets Users adapt various forms of multimedia for consumption on a particular terminal. The following classes of description tools are specified as part of DIA:

- *Codec capabilities* specify the format that a particular terminal is capable of encoding or decoding, for example, MPEG-4 Simple Profile at Level 3. The specification is symmetric with the MPEG-7 description tools for media format.

- *Input-output capabilities* include a description of display characteristics, audio output capabilities, and various properties of several input device types.

- *Device properties* characterize power-related attributes of a device, as well as storage and data I/O characteristics.

Network characteristics

The specification considers two main categories: network capabilities and network conditions. One application for these descriptions is to enable multimedia adaptation for improved transmission efficiency. For instance, we can lower the delivery bandwidth of an audio stream if the available bandwidth is insufficient, or we can increase the rate of intra-coded blocks in a video stream if the packet loss rate is high.

- *Network capabilities* define a network's static attributes, such as the maximum capacity of a network and the minimum guaranteed bandwidth.

- *Network conditions* describe network parameters that tend to be more dynamic and time varying, such as the available bandwidth, error, and delay characteristics. The error is specified in terms of packet loss and bit error rates. Delay types include one-way and two-way packet delay, as well as delay variation.

User characteristics

We can classify the User characteristics specification into the following subcategories:

- *User info, usage preferences,* and *usage history* have imported description schemes (DSs) from MPEG-7 to describe a User's general characteristics as well as user preferences and usage history of Digital Items.

- *Presentation preferences* define preferences related to the means by which audiovisual information is presented or rendered to the User. For audio, the specification describes preferred audio power and equalizer settings. For visual information, the specification defines display preferences, such as the preferred color temperature, brightness, saturation, and contrast.

- *Accessibility characteristics* provide descriptions that enable Users to adapt content according to certain auditory or visual impairments. For audio, an audiogram specifies a person's hearing thresholds at various frequencies in the left and right ears. For visual impairments, the standard specifies the type and degree of color vision deficiencies.

- *Location characteristics* include a description of mobility and destination. *Mobility* describes a User's movement over time, particularly information about a User's angular changes and degree of random movement. *Destination*, as the name implies, indicates a User's destination. We can use these tools together to provide adaptive location-aware services.

Natural environment characteristics

Digital Item Adaptation (DIA) specifies the following natural environment description tools:

- *Location and time* refer to a Digital Item's location and time of usage. Both descriptions use MPEG-7 description schemes, in particular the Place DS and Time DS. Besides being standalone tools, we can also use location and time to specify User characteristics.

- *Audiovisual environment* describes audiovisual attributes that can be measured from the natural environment and affect the way content is delivered and/or consumed by a User in this environment. For audio, the specification describes the noise levels and a

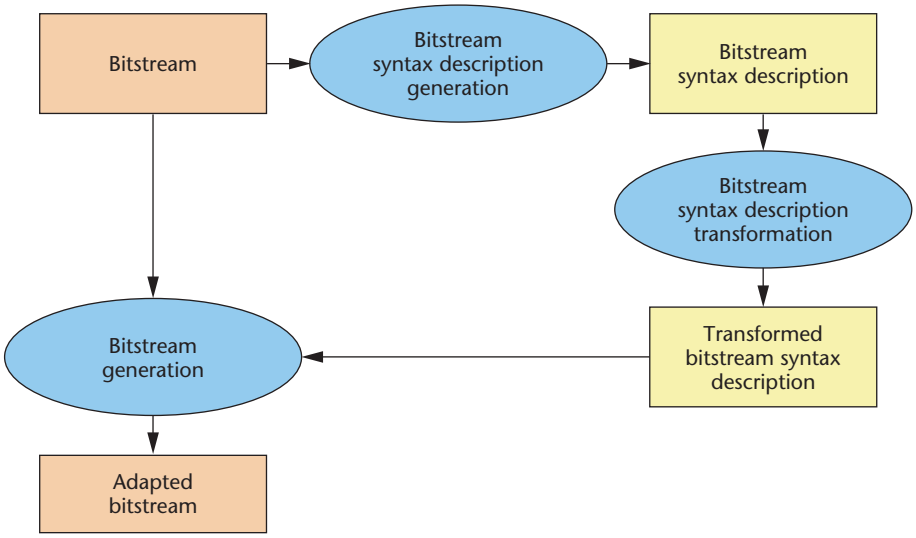

Figure 2. Bitstream syntax description adaptation architecture.

noise frequency spectrum. For the visual environment, the specification defines illumination characteristics that may affect the perceived display of visual information.

Bitstream syntax description

Suppose that an intermediate network node or proxy has no knowledge of any specific media formats. Wouldn't it be great if the node or proxy could perform media adaptation without having to understand the intricacies of every media format? This is precisely the motivation for the bitstream syntax (BS) description tools that DIA has standardized.

Figure 2 shows the high-level view of the adaptation process using BS description tools. As a first step, a User generates the BS description of the input bitstream. The description essentially details, in the Extensible Markup Language (XML), how data layers are organized in the bitstream. With this description, an adaptation engine can perform a transformation of the description in the XML space—for example, with an Extensible Stylesheet Language Transformation (XSLT) style sheet. The adaptation engine can then generate an adapted bitstream based on the transformed description. Although this process only allows for simple adaptation methods that truncate or remove data, this process could be applied to a wide variety of coding formats.

To support these operations, the MPEG-21 standardization committee has specified two tools as part of the DIA standard. The first tool, based on XML Schema, is the Bitstream Syntax Description Language (BSDL). With this language, we can design specific BS Schemas to describe the syntax of a particular coding format, which a generic processor could then use to perform the operations described earlier. The second tool—generic Bitstream Syntax (gBS) Schema—is used to describe any binary resource in a codec-independent manner. Additionally, the gBS Schema provides a means for Users to associate semantic labels with the syntactical elements being described. Also, the tool enables Users to describe a bitstream hierarchically so that groupings of bitstream elements can be formed to allow for efficient, hierarchical adaptations of the bitstream and addressing of elements.

Terminal and network QoS

Media adaptation is usually driven by constraints imposed by terminals and networks. Often, more than one adaptation operation could satisfy the given constraints. The set of terminal and network quality of service (QoS) tools defined by DIA specify information that would help Users form an optimal adaptation strategy.

In particular, it's possible to specify the relationship between the constraints, feasible adaptation operations satisfying the constraints, and associated utilities (qualities). The utility values can be in the form of either subjective or objective measures. These values are derived based on different adaptation operations and the operations' parameters. The DIA standard defines a wide variety of possible constraints, adaptation operations, and utility metrics to enable expression of a rather rich set of relationships.

Metadata adaptability

The metadata adaptation tool specifies hint information that can be used to reduce the complexity of adapting metadata (XML) instances. The specified hints have been designed for two specific classes of applications. The first is filtering and scaling of relatively large description instances, and the second is integrating two or more description instances.

Given a large description instance that should be sent and processed on a mobile device, filtering and/or scaling the metadata could help to overcome processing power, memory, and bandwidth limitations. To support this operation, the

DIA standard specifies attributes such as the size of the metadata, total number of elements, and depth and number of occurrences of certain elements. To integrate description instances—for example, to combine one description that includes low-level features of each digital video shot with a second description containing semantic descriptions for specific digital video segments—it's useful to know whether the descriptors are invariant to the adaptation and what the average value of certain descriptors are.

Session mobility

The Digital Item Declaration (DID) is the second part of MPEG-21. It specifies a uniform and flexible abstraction and interoperable schema for declaring the structure and makeup of Digital Items. As a basic functionality, we can declare a Digital Item by specifying its resources, metadata, and their interrelationships. Further, we can configure the DID with Choices and Selections tools that are part of the declaration. The instantiation of Choices and Selections in the DID is the *configuration-state* of a Digital Item.

In DIA, *session mobility* is the transfer of configuration-state information, pertaining to the consumption of a Digital Item on one device, to a second device. This enables the Digital Item to be consumed on the second device in an adapted way. During this transfer, application-state information, which pertains to information specific to the application currently rendering the Digital Item, can also be transferred.

Concluding remarks

In summary, DIA offers a rich set of tools to assist with the adaptation of Digital Items. It offers standardized tools for the description of usage environments, tools to create high-level descriptions of the bitstream syntax to achieve format-independent adaptation, tools that assist in making tradeoffs between feasible adaptation operations and constraints, tools that enable low-complexity adaptation of metadata, and tools for session mobility.

Moving forward, the MPEG-21 committee is considering amendments to the specification—for instance, tools that provide further assistance with modality conversion and tools that relate more specifically to the adaptation of audio and graphics media. Furthermore, also being actively considered is how to express the rights that a User has to perform adaptation and how this expression fits into a system that governs those rights.

Further reading

For more information about MPEG-21, including overviews, FAQ sheets, and working documents, visit http://www.chiariglione.org/mpeg. Complete MPEG-21 DIA schemas and description examples are available at http://www.midgardmedia.net. An overview of the goals and achievements of MPEG-21 is also available.[2] **MM**

References

1. A. Vetro, C. Christopoulos, and T. Ebrahimi, eds., Special Issue on Universal Multimedia Access, *IEEE Signal Processing Magazine*, vol. 20, no. 2, Mar. 2003.
2. I. Burnett et al., "MPEG-21: Goals and Achievements," *IEEE MultiMedia,* vol. 10, no. 6, Oct.–Dec. 2003, pp. 60-70.
3. MPEG MDS Group, MPEG-21 Multimedia Framework, Part 7: Digital Item Adaptation (Final Committee Draft), ISO/MPEG N5845, July 2003; http://www.chiariglione.org/mpeg/working_documents/mpeg-21/dia/dia_fcd.zip.

Readers may contact Anthony Vetro at Mitsubishi Electric Research Labs, 201 Broadway, Cambridge, MA, 02138, USA; avetro@merl.com.

Contact Standards editor John R. Smith at jsmith@us.ibm.com.