An MPEG-21 Web Peer for the consumption of Digital Items

Abstract
MPEG-21 enables content consumers to access and interoperate with a large variety of multimedia resources and their descriptions in a flexible manner. Considering the great heterogeneity that presently exists across the entire multimedia content chain and the growing importance of open standards to facilitate the interoperations across environments, applications and formats, an MPEG-21 Peer was developed to process and present complex multimedia content, represented as MPEG-21 Digital Items. The novelty of the work essentially relies on the adoption of a Web Services architecture, based on a single Digital Items processing core available for all types of terminal devices.

Keywords: multimedia consumption; MPEG-21; Digital Items; MPEG-21 Peer; Web Services.

1 Introduction

With the advent of the WWW and the proliferation in the mass market of a diversity of multimedia-enabled end-user devices, consumers increasingly expect to be able to access any kind of content anywhere and at any time regardless of the capabilities of their terminals. To this reality adds the large volumes of multimedia content available on-line, presenting a multiplicity of formats and the ever growing acceptance of the network-centric paradigm from the general public. An access to content that meets user expectations and demands must take into account all the different aspects of this heterogeneous scenario. Nowadays, one of the reasons for the popularity of networks is the information accessibility that they offer: the consumers now can reach the multimedia contents from almost anywhere. This is a challenge for the recent standards, technologies and business models, and hence the generation of new tools and applications for the multimedia content consumption is more oriented for the Web to allow remote networked.
A thorough characterization of both the content as well as of the context of usage is therefore needed to enable the selection and delivery of the content in the most adequate conditions given the current context of usage. MPEG-21 aims to provide a solution that satisfies these needs, thus augmenting the multimedia experience of the users. However, dealing with all this issues implies that the solution can itself be complex. Applications for the processing of MPEG-21 are required, including for the transparent presentation of richer information to the end user.
MPEG-21 does not specify the technologies to be used and the manner of integrating other tools when developing MPEG-21 applications. This allows great

flexibility for implementers. Consequently, some incompatibilities may occur between the tools and applications, provided by different producers, within the MPEG-21 multimedia framework. A question derived from this: how to implement an interoperable application within MPEG-21 framework? The use of Web Services concepts is a solution, as they define a distributed architecture ensuring maximum interoperability and transfer of data in an efficient manner.
This paper describes an MPEG-21 application named MPEG21 DI Browser that implements a view on the presentation and consumption of meta-information and multimedia content represented according to the MPEG-21 standard, and that aims at providing a friendly and intuitive interface that can be used to give users a richer experience. The goal is to achieve interoperability and also to enable the usage, visualization, and interaction with the multimedia contents on different types of terminals.
The remainder of this paper is structured as follows: an introduction to MPEG-21 and its relevant parts is made in the next section; then a brief state of the art survey of the MPEG-21 applications is presented; after that the implemented MPEG21 DI Browser is described and then an usage scenario for this application is illustrated; in a separate section is discussed the integration of the results in an European project, and finally the future work and some conclusions are made.

2 MPEG-21

MPEG-21 aims at describing how the various elements of the multimedia content delivery chain fit together to create a truly interoperable multimedia framework where content is conveyed as Digital Items (DI) (ISO/IEC N6388, 2004). MPEG-21 presents itself as a solution for augmenting the consumption of multimedia content. The objective is to provide a platform that will enable users to access multimedia content in a transparent way over a variety of networks (dialup, broadband, mobile, wireless, broadcast, etc.) and devices (computers, PDAs, media players, mobile phones, digital TVs, etc).
This standard is currently composed of 17 parts: 1 - Vision, Technologies and Strategy; 2 - Digital Item Declaration (DID); 3 - Digital Item Identification and Description (DII); 4 - Intellectual Property Management and Protection Components (IPMP Components); 5 - Rights Expression Language (REL); 6 - Rights Data Dictionary (RDD); 7 - Digital Item Adaptation (DIA); 8 - Reference Software; 9 - File Format (FF); 10 - Digital Item Processing (DIP); 11 - Evaluation Tools for Persistent Association; 12 - Test Bed for MPEG-21 Resource Delivery; 14 – Conformance testing; 15 - Event Reporting (ER); 16 - Binary format; 17 - Fragment Identification for MPEG Media Types; 18 – Digital Item Streaming.
For the remainder of the paper we will assume that the reader is familiar with MPEG-21. If it is not the case, the reader should refer to (ISO/IEC N6388, 2004), (Bormans, 2003) or (Burnett).

The application presented here - MPEG21 DI Browser - has currently implemented part 2 and part 10.

2.1 Digital Items and Users

The MPEG-21 standard defines two essential concepts of the standard. These are: Digital Item (DI) and User. A DI is defined as:

> *"A Digital Item is a structured digital object with a standard representation, identification and meta-data within the MPEG-21 framework. This entity is also the fundamental unit of distribution and transaction within this framework" (ISO/IEC N6388, 2004).*

All types of information that can be combined through MPEG-21 are encompassed in Digital Items which are "the currency within MPEG-21 framework" (S.Lauf, 2005). Digital Items combine multimedia resources, other digital objects and related metadata. It also encompasses the description of the relations between all these elements. Examples of DI are: Multimedia presentations, music albums, collections of e-learning objects, TV program guides, lists of videos and their descriptions, digital libraries, etc.
Any entity that interacts with Digital Items within the MPEG-21 multimedia framework is a User (ISO/IEC N6388, 2004).
Due to the relevance for the contents of this paper, we will briefly present some essential concepts of parts 2 and 10 of the standard, respectively Digital Item Declaration (DID) and Digital Item Processing (DIP).

2.2 Digital Item Declaration

Part 2 of the standard specifies the Digital Item Declaration Language (DIDL). This language is used to generate the Digital Item Declaration (DID), typically an XML document, which describes the structure of the DI. Inside the DID, standardized MPEG-21 mechanisms are used to either convey the resources embedded directly in the DID, or to provide information on the location of the resources that make up the DI (ISO/IEC FDIS 21000-2, 2005). This part defines several elements to be used for describing the structure and composition of a DI. Bellow we briefly present some of these elements:

- *Container*: groups more *Items* and/or sub-*Containers*;
- *Item*: contains *Components* and/or sub-*Items*. *Item* represents "the lowest level of granularity transacted by Users within the MPEG-21 framework" (ISO/IEC FDIS 21000-2, 2005);
- *Component*: binds *Resources* to a set of descriptions that contain control or structural information about the resource but do not contain information describing the content within;

- *Descriptor*: associates information with the enclosing element (*Item*, *Container*, *Component*, *Choice*, *Annotation*, *Descriptor*, *Selection*, etc.).
- *Choice*: describes a set of *Selections* that can affect some conditioned elements such as: *Items*, *Descriptors*, *Component*;
- *Condition*: specifies the enclosing element as being optional and links it to the *Selection*(s) that affect its inclusion;
- *Selections*: describes a specific decision that affects one or more conditioned elements in Digital Item;
- *Annotation*: embeds a set of information about another identified element of the Digital Item without altering or adding to that element;
- *Assertion*: defines a configured state of a *Choice* and of the associated *Selections*.

The DID model enables the support for a static User - Digital Item interaction. Users are given by the possibility of limitedly configuring parts of the DID using the *Choice/Selection* mechanism (ISO/IEC FDIS 21000-2, 2005).

2.3 Digital Item Processing

Part 10 of the standard specifies the syntax and semantics of tools that may be used to process Digital Items (ISO/IEC FDIS 21000-10, 2005). On receiving a static declaration, as it is the case of the DID, Users have nothing that indicates how the information should be processed, since this is out of the scope of part 2. Digital Item Processing (DIP) intends to cover all aspects of processing a static Digital Item from the User perspective, by enabling MPEG-21 Users to specify a selection of preferred procedures to apply to the Digital Item when configuring, validating, processing and consuming it. DIP provides the means for a dynamic User – Digital Items interaction. Digital Item Processing (DIP) intends to cover all aspects of processing a static Digital Item from the User perspective, by enabling the MPEG-21 Users to specify a selection of preferred procedures to apply to the Digital Item when configuring, validating, processing and consuming it. DIP provides the meanings for a dynamic interaction User – Digital Items.

DIP includes in basic operations (Digital Item Basic Operations - DIBOs) such processing as Digital Item downloading, rights management, media resources downloading, printing or playing resources, presentation of the Digital Item or part thereof, etc. Also this MPEG-21 part defines the mechanism of using extended operations (Digital Item eXtended Operations - DIXOs) ensuring the addition of User specified functionality to a DID.

The DIP operations are called in methods defined in Digital Items. These methods are named Digital Item Methods (DIMs) and can be regarded as a “menu” of User interaction possibilities with the Digital Item (ISO/IEC FDIS 21000-10, 2005). When referring to DIMs, there are several roles that a User might undertake. For example, as a consumer a User can interact with a Digital Item via the execution of a DIM. As a creator, a User can author a DIM to be included in a Digital Item. As

MPEG-21 compatible hardware or software for processing a Digital Item, a User can provide an execution environment for DIMs, where the implementations of basic operations (DIBOs) should be appropriately included in the environment (ISO/IEC FDIS 21000-10, 2005).

3 MPEG-21 Applications

MPEG-21 has introduced the term Peer to characterize a "*device or application that compliantly processes a Digital Item*" (ISO/IEC N6388, 2004). Such processing includes the creation and presentation/rendering of DIs.

MPEG-21 Peers for visualizing and interacting with DIs are still scarce. In the existing implementations, solutions have been found for ensuring accessibility, flexibility and performance of the Peers on specific terminal devices (e.g., only PC, or only PDA, or mobile phones).

The research group Multimedia Lab pertaining to Ghent University (Belgium) has developed a generic MPEG-21 terminal for the processing of DIs and DIMs, implementing the functionalities of part 10 of the standard. This is an online terminal which communicates with a streaming server, through SOAP messages, for dynamically configuring the video streams referenced by DIs.

Klugenfurt University (Austria) brought a considerable contribution to the MPEG-21 DIA standardization but they also developed various MPEG-21 applications, among which two are more relevant for the purposes of this paper: the DIBuilder application for creating DIs; and the DIConsumer application for the consumption of DIs created with DIBuilder.

Wallongong University (Australia) is an example of another institution that manifests interest in implementing MPEG-21 applications. One of their most relevant projects consists of an MPEG-21 Peer for mobile devices. Their objective was to develop an MPEG-21 Peer for mobile phones for demonstrating an efficient and flexible architecture which encompasses a sufficient range of MPEG-21 technologies to realize the consumption, authoring and transmission of DIs.

Enikos is an Australian company focused exclusively on MPEG-21 technology and solutions. The two commercial products of this company are DICreator and DIBrowser. DICreator is a tool that permits to easily create, edit and view DIDs. DIBrowser is an application capable of browsing any local or remote DID, presenting its internal structure as a tree. Enikos also released a new product: the MPEG-21 DIP Desktop Peer which is a demo application implementing the DIP functionality.

Adactus is a software company (from Norway) offering solutions for delivery and presentation of adapted multimedia for mobile terminals. **mobilize** is the base product of Adactus and consists in a cross-platform system for content delivery. MPEG-21 technology was used in the **mobilize** system to enable the content providers to adapt the content to the capabilities of terminals, to report the content usage and consumption and to protect the content carried in DIs.

Among all the existing MPEG-21 Peers, none is portable to many or all types of devices. Also, there is no implemented solution to ensure the consumption of complex DIs on thin devices (such as mobile phones or PDAs).
The MPEG21 DI Browser Peer intends to solve the portability problem and proposes a modality of processing and visualizing complex DIs on majority of terminal devices.

4 MPEG21 DI BROWSER

The MPEG-21 standard does not specify how a terminal application should present the DI to Users. Hence, multiple applications, implementing different views, may be developed. The adequacy of an application is usage scenario specific. Due to current familiarization of Users with the World Wide Web and to the explosion of web application, it was decided that MPEG21 DI Browser would be web oriented. A modular and incremental design approach was adopted, supporting initially a restricted set of MPEG-21 functionality (DID and DIP) to which new ones will be added. This approach is suited to the evolving nature of the MPEG-21 standards.

4.1 Application functionality

The functionalities implemented in the current version of this MPEG-21 Web Peer are: downloading DIs stored in a remote repository; validating the respective DIs against the normative DID schema; processing DIs "by pieces" for ensuring an adequate consumption on all devices (including thin devices such PDAs and mobile phones); navigating the contents of remote DIs in a web page style; presenting the assets within DIs; processing the User preferences and selections for accessing the restricted contents from a DIs; and executing DIMs methods for dynamic interaction User-Digital Item (currently, DIP functionality works only on PC version).

4.2 Web Services based architecture

A client-server architecture was adopted and Web Services concepts were applied for increasing the interoperability and portability to heterogeneous terminal devices. This architecture delegates on the server the responsibility of DIs processing, leaving the client with the content presentation. The separation of the graphical interface from the business logic facilitates the addition of other GUIs implemented in different programming languages and at the same time it allows the flexibility of re-using the processing core in other application involved in the consumption of MPEG-21 DIs.
MPEG21 DI Browser is composed by two main sub-systems: the client - WDI Browser and the server - IDI Browser. The application functionality is based on the mentioned server application which has total control over navigation and the User's interactions with DIs.

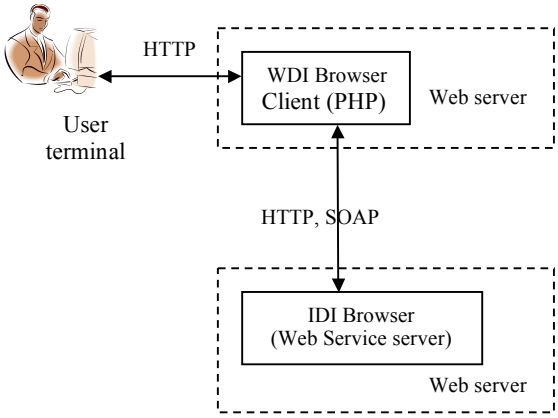

Figure 1 MPEG21 DI Browser Peer

The messages exchanged between client and server sub-systems are based on normative DID representations (which are XML) containing the DIDL element that is requested to be visualized at User terminal at a determined moment. The end-Users can access this MPEG-21 Peer using a common web browser (e.g. IE, IE Mobile, Firefox Mozilla, Netscape, Opera, etc.).

Web Services Server

The IDI Browser server application is the central unit of the implemented MPEG-21 Web Peer. Its main role is to provide the MPEG-21 specific processing and to control the navigation and all the interactions with DIs. The IDI Browser is structured on modules (of Java classes), where the top module is responsible for generating the Web Service operations. This modularity allows easy integration with other external modules adding new functionality that may not be specific for MPEG-21.

On the server side a caching mechanism for the elements previously processed and visualized was implemented. The caching and management of the previous states of browsing in DIs assure that end-Users can navigate backward in any moment. Also, the IDI Browser stores and processes the User's preferences regarding the access/visualization of the conditioned elements in the DIs.

The technology used to develop the IDI Browser Web Services was Sun Java Web Service Developer Pack 2.0 (JWSDP 2.0).

Web Client

The client – WDI Browser – is a PHP application, running on an Apache 2 web server, and generates the HTML pages corresponding to DIDL base elements. It has a modular structure which permits the creation of new web interfaces for other

types of terminal devices. Currently, an XSLT Processor module, implemented on the client sub-system (WDI Browser), provides the graphical web interfaces for PC and PDA devices.
DIP functionality was added as an applet application to the client side and is based on the MPEG-21 DIP reference software (ISO/IEC FCD 21000-8, 2007).

4.3 Object Model

During the processing of DIs it is necessary to manipulate the MPEG-21 information in a manner that ensures easy access to its elements and an efficient processing, contributing to the overall performance of the MPEG21 DI Browser. Therefore it is useful to create a corresponding object representation for mapping the XML elements of the MPEG-21 parts. The object model maps the elements defined in parts of the standard and also the relationships between them, into a set of related classes, implemented in an object oriented language (Java). The mapping of the MPEG-21 elements into object instantiations and vice-versa is made with a specific tool (Apache XMLBeans available at http://xmlbeans.apache.org/) that needs of a set of rules – normatively defined in XML schema files – in order to reflect the relationships between these elements.
However, the original XML content of the DID will be stored in the object model and will be used as an alternative in some particular situations (e.g., looking for some specific attribute in a tag element that is not defined in MPEG-21 and as a consequence does not have any representation in the object model).
The Object Model module is rather more specific to MPEG-21 processing than to MPEG21 DI Browser; the created Object Model can also be used / integrated in other MPEG-21 applications.
The current Object Model includes the implementations for the following MPEG-21 parts: 2 -Digital Item Declaration, 3 - Digital Item Identification and 10 - Digital Item Processing.

4.4 Development considerations

Some development considerations were made with the purpose of providing the same processing core for displaying the contents of DIs on all types of terminal devices. The DIs are processed and visualized in a progressive manner. The processing of DIs consists in preparing them for browsing/consumption. The navigation in DIs is made from the top DIDL element to the bottom nodes. At each step of the navigation, only a single element is processed and displayed to the end-User; the respective element is a *Container* or an *Item*. There is no limit regarding the depth of the hierarchic level corresponding to a sub-*Item* or sub-*Container* element. This allows the consumption of large and complex DIs on thin-devices.
For optimization reasons, the elements are simplified in order to maintain only the relevant information for this version of the MPEG-21 Peer. For example, during the processing, the elements that do not meet some conditions will be discarded; the

sub-elements will be filtered and only their first textual descriptions and the corresponding identifiers will be kept; the comments are not relevant for the common Users and may also be removed; and some additional information specific for the application implementation may be included for management purpose of the correspondence between the base elements (*Items* / *Containers*). This filtering of the elements does not impose any restrictions and the DIs are still in conformance with the standard.
The presentation is also made in a "step by step" style. To each element (*Item*/*Container*) processed at a certain moment is associated a web page. The hierarchical relationships between the elements are defined by hyperlinks. This web philosophy was easily adopted due to the similarity between the tree structure of the DID representation and the manner of navigation in websites.

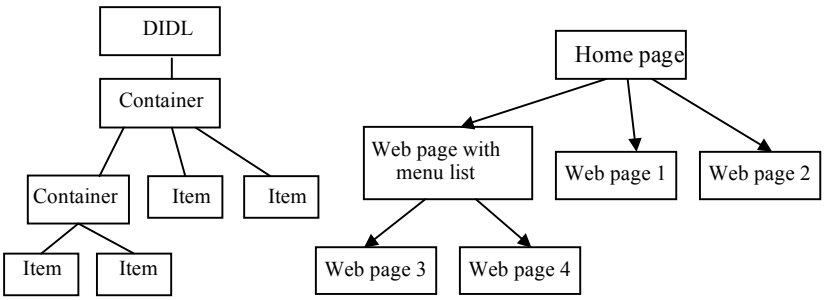

Figure 2 Similarity between a DID representation (on the left) and a website (on the right)

A base element of type *Item* may contain information that is not made available to the User unless he performs some selections. During navigation, when *Choice* elements exist in *Item*, the User is asked to express his preferences related to some parts of the content. In order to allow the Users to visualize the restricted information, an intermediary web page is necessary for making those selections. For example, a selection may be set for choosing the appropriate video file format to be played on the User's terminal device.

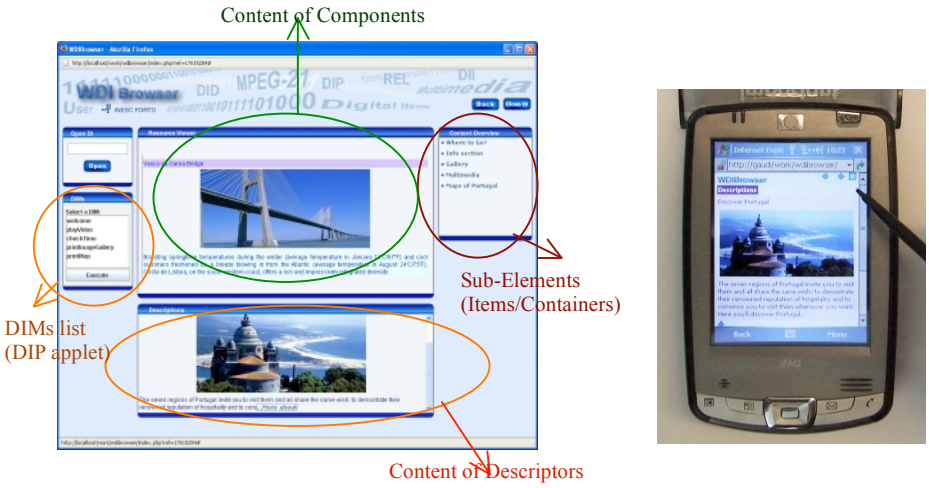

Figure 3 WDI Browser layouts (PC - on the left, and PDA - on the right)

5 Use scenario for MPEG21 DI Browser

How can the MPEG21 DI Browser be used? To answer this question a usage scenario will be presented next.

Imagine that John Smith, working at "Discover, Travel and Live - Tourism Agency", is producing a new presentation about Portugal in order to publish it on the agency's website to promote tourism in this country. While traveling in various regions of Portugal, he has collected many media resources: photos, videos, traditional music, and information about folklore, the Portuguese people way of live, about their culture, and general information useful for tourists. John is using an authoring application for MPEG-21 DIs to create his presentation. He combines all the media resources and information in one Digital Item and then uploads it to a specific web (remote) repository, reporting its location to his agency but also to his family and friends that are curious about Portugal.

The agency's website has its own client application (e.g., developed with .ASP) for the IDI Browser Web Services. The agency's customers can use this client application of the agency's website to access the respective Digital Item and to view the presentation about Portugal.

At the same time, John's family and friends may access the WDI Browser for opening and navigating through the Digital Item and for enriching their knowledge about Portugal.

The exemplified usage scenario demonstrates that the MPEG21 DI Browser Peer may be useful for the Users that want to consume DIs, no matter if only the IDI Browser server or the whole system functionality (client and server) are used.

6 Integration with ENTHRONE

ENTHRONE, *End-to-End QoS through Integrated Management of Content, Networks and Terminals* is an Integrated IST Project in EU Framework Program 6. It aims to provide an integrated management solution for transparent access to multimedia resources while ensuring QoS and efficient usage of respective resources and their descriptions, covering "the entire audio-visual service distribution chain, including protected content handling, distribution across networks and reception at User terminals" (http://www.ist-enthrone.org).
Among its objectives, ENTHRONE proposes the development of an MPEG-21 terminal architecture that includes a wide range of heterogeneous devices with different capabilities. The ENTHRONE terminal will be able to deal with the interactions and processing of DIs needed for the efficient managements of the terminal QoS.
Currently, efforts have been initiated for integrating the results of the presented MPEG-21 Web Peer to the terminal component of the ENTHRONE project.

7 Future work

Future work on the described MPEG-21 Web Peer is related with: improvement of the DIP implementation for being available to work with many types of terminal devices, increase its usability by adding new web interfaces (e.g. for mobile phones, set-top boxes), develop other MPEG-21 features; and with other aspects such as: browsing the local DIs stored on User's device, searching for contents and DIs in (web) repositories, managing and controlling the sessions of Users. The search for DIs and parts thereof will take into account the User preferences, the terminal capabilities and networks characteristics, in conformance with the requirements defined in ENTHRONE.

8 Conclusions

The MPEG21 DI Browser has the capacity of browsing simple and large MPEG-21 DIs, ensuring a fast navigation between the DIDL elements (*Items* or *Containers*). This application can be used for various domains of activity in which the use of multimedia content is essential.
The distributed architecture based on Web Services permits to maintain the same processing core for all types of terminal devices and this increases the system's portability.
Due to the modularized structures of its client and server components and to the Web Services architecture, this MPEG-21 Web Peer is flexible and interoperable, allowing easy integration of other external modules or tools.

9 References

ISO/IEC JTC1/SC29/WG11/N6388, Information Technology – Multimedia Framework (MPEG-21) – Part 1: *Vision, Technologies and Strategy*, Munich, March 2004.

S. Lauf, I. Burnett, MM'05, *Implementation of a Mobile MPEG-21 Peer*, Singapore, November 2005.

ISO/IEC FDIS 21000-2:2005(E), Information Technology – Multimedia Framework (MPEG-21) – Part 2: *Digital Item Declaration*, Hong Kong, January 2005.

ISO/IEC FDIS 21000-10:2005(E), Information Technology – Multimedia Framework (MPEG-21) – Part 10: *Digital Item Processing*, Busan, April 2005.

ISO/IEC FCD 21000-8, Information Technology – Multimedia Framework (MPEG-21) – Part 8: *Reference Software (Second Edition)*, Marrakech, January 2007.

Bormans, J., Gelissen, J. and Perkis, A., *"MPEG-21: the 21st century multimedia framework"*, IEEE Signal Processing Mag., vol. 20, no. 2, pp 53-62, Mar. 2003

Burnett, I., Pereira, F., Walle, R. and Koenen, R., *"The MPEG-21 Book"*, John Wiley and Sons, 2006