The MPEG-7 Multimedia Content Description Standard and the XML Schema Language. *

E. Terzi A. Vakali J. Fan M.-S. Hacid
Department of Computer Sciences, Purdue University
West Lafayette, IN 47907 - 1398
{edt,vakali,fanj,mshacid}@cs.purdue.edu

Abstract

Web accessed multimedia applications have been widely adopted under enhanced requirements for fast searching, browsing and retrieval. Efficient multimedia data content description is a neccesary and critical issue towards which several solutions have been proposed. In the present paper, we focus on the MPEG-7 standard as an efficient multimedia content description tool. MPEG-7 is a newly introduced standard from the MPEG commitee and a brief description of the standard and its components is given here. MPEG-7 Language-Schema as used towards efficient Audio/Visual content description is discussed more extensively. As decided by the MPEG-7 committee it the Description Definition Language (DDL) of the standard should be based on the XML Schema. The several extensions to this schema, necessary for satisfying the requirements of MPEG-7 are discussed, while other functionalities of the schema which seem not necessary under current specifications are also commented.

1. Introduction

Multimedia data is the main form of data being circulated over the Internet and therefore the objective of today's multimedia applications is not only to acquire, process and distribute multimedia content, but also to be able to provide relatively easy access to the available information. However, identifying the desired information, by retrieving or filtering, is becoming more and more difficult due to the large amount of information being generated at a daily basis, and to the increasing abilities of consumers that use the new technologies and their functionalities. Therefore, the need for a powerful solution for quick and efficient identification (searching, filtering, etc.) of the multimedia data of interest to the user is obvious as emphasized by the MPEG comittee five years ago at Tampere (Finland) meeting. The result was a new standard for "Multimedia Content Description" widely known as MPEG-7 ([19]). MPEG-7 does not standardize the format of multimedia data, like the previous mpeg attempts, but it provides a flexible, interoperable way of representing and describing audiovisual information regardless of storage, coding, display, transmission, medium or technology. Therefore, MPEG-7 is a format of organizing both semantical and structural metadata so that multimedia searching, filtering, browsing and summarization will be effectively facilitated within the applications that support the standard. We should notice that the need for complex multimedia data descriptions is accompanied with the need for interoperability of the metadata descriptions.

An overview of the MPEG-7 standard and its objectives as well as the state of its development and its expected impact on the digital libraries of the future is given in [12]. More detailed description of the standard with special emphasis to every one of its components is provided at the committee drafts [25]. In fact, MPEG-7 standardizes ways to define Multimedia Descriptors (Ds), Description Schemes (DSs) and the relationships between them. The descriptors correspond to the data features themselves, while the descriptor schemes refer to more abstract description entities that are perceivable by the humans. These low-level and high-level description entities as well as their relationships are represented using the Description Definition Language (DDL), also specified by the standard. The Descriptors, Descriptor Schemes and the Description Definition Language are the three of the four basic parts of the MPEG-7 standard. The fourth one refers to the system level of the applications and standardizes coded representations of descriptions to enable efficient storage and fast access.

Although MPEG-7 is still in its infancy, many fundamental research efforts on MPEG-7 Ds, DSs, and DDL have been presented in the past few years. The works of [1, 3, 4, 13, 14, 15, 16, 17, 18] introduce new Ds and DSs

*This work is supported by NSF under 9972883-EIA, 9974255-IIS, and 9983249-EIA, a grant from the state of Indiana 21th century fund, and by grants from HP, IBM, Intel, NCR, Walmart, and Telcordia.

for describing multimedia and particularly video content. Emphasis on the DDL part of the MPEG-7 standard is given on the works of [5, 6, 7, 8, 9, 10, 11]. Moreover, there are a lot of research and commercial applications that have been proposed and intended to exploit the advantages of the MPEG-7 multimedia content description scheme, and particularly its applicability for web searching and accessing of multimedia data. According to the authors of [26] a total of 25 applications were found in July 2000.

This paper aims to present an overview of the newly introduced MPEG-7 standard and the way the multimedia (audio/visual) content will be described using the MPEG-7 tools. More specifically, Section 2 has an overview of the MPEG-7 standard describing its basic components and the way they interact. Section 3 focuses on the Description Definition Language part of the MPEG-7 providing a summarization of the requirements raised towards the DDL, and classifying the proposals submitted to the MPEG-7 committee. The finally adopted XML schema language as the MPEG-7 DDL is described in further detail in Section 4. Conclusions and remarks of the MPEG-7 overview are given in Section 5.

2. MPEG-7 : A Brief Overview

2.1. MPEG-7 Components

The MPEG-7 standard aims to satisfy the need of an efficient, compact flexible and interoperable representation of multimedia content, by standardizing a framework for compact, efficient and interoperable descriptions of audio visual content. MPEG-7 is expected to improve current multimedia applications and enable new exciting ones. Some examples for such applications are distributed processing,content exchange, personalized representation and efficient storage/retrieval of multimedia content. In other words, MPEG-7 is used to describe and search multimedia data and does not deal with how the data will be encoded (in the way the other MPEG standards have). The main idea of MPEG-7 is that there are high level and low level descriptors. The low level descriptors are called Descriptors (Ds), they are directly derived for the data and they are instantiated as *Descriptor Values*. The high level descriptors are called Description Schemes (DSs) and are usually added by human classifiers. MPEG-7 is not concerned on how the descriptors are ascertained or how they are searched, but only on how they are recorded. For this part the Description Definition Language (DDL) is introduced. DDL will allow the definition of new description schemes and new descriptors tailored to the needs of specific applications. The DDL is also used for describing the inter and intra relationships between the Ds and the DSs schemes. The whole description will be encoded, based on coding schemes that satisfy compression and transmission requirements.

MPEG-7 Descriptors (Ds) are mainly designed for describing the following types of information:

- *low-level audio-visual features* :such as color, texture, motion, audio energy, and so forth; The descriptors corresponding to low-level features are extracted automatically.
- *high-level features* : such as semantic objects, events and abstract concepts, content management processes, information about the storage media, and so forth. The high-level descriptors cannot be automatically extracted, but they require human intervention.

MPEG-7 Description Schemes (DSs) are models of the multimedia objects and of the universes that they represent. Therefore, they represent the data model for the description. Additionally, they specify the types of the Descriptors that can be used within a given description, and the relationships between these Descriptors or even between other Description Schemes. As emphasized in [13], the main difference between a DS and a D is that within a given description a Descriptor can neither be extensible nor can it be reducible with respect to the number of descriptor values contained. Therefore the interrelation of Ds and DSs build a hierarchical model where DSs correspond to the upper levels of the hierarchy, while Ds are in the lower parts. This hierarchical structure indicates that DSs are composed of other DSs and Ds.

The Data Definition Language (DDL) is a core part of the MPEG-7 standard. It provides the solid descriptive foundation by which users can create their own Description Schemes and Descriptors. The DDL provides the syntactic rules to express and combine Description Schemes and Descriptors. According to the definition in the MPEG-7 Requirements Document the DDL is " a language that allows the creation of new Description Schemes and, possibly, Descriptors. It also allows the extension and modification of existing Description Schemes." The DDL is not a modeling language such as Unified Modeling Language (UML) but a schema language capable of representing the results of modeling audiovisual data (i.e. DSs and Ds).

MPEG-7 DDL should provide descriptions that satisfy the following requirements:

- The description should express spatial, temporal, structural, and conceptual relationships between the elements of a DS, and between DSs.
- It must provide a rich model for links and references between one or more descriptions and the data that it describes.
- must be platform and application independent and human- and machine-readable.

MPEG Committee has recently decided to adopt the XML Schema Language as the MPEG-7 DDL. However, because XML Schema language has not been designed specifically for audiovisual content, certain extensions have been recently proposed as necessary in order for it to meet the MPEG-7 DDL requirements. The XML Schema Language has been chosen as the most appropriate to be used at the DDL component of the MPEG-7 Standard mainly because it is a schema language rather than a grammar for the XML documents. The XML schema specification consists of the specification of the Structures and the Datatypes. The Structures provide facilities for describing the structure and constraining the contents of the XML documents, while the datatypes propose facilities for defining datatypes. Consequently, MPEG-7-specific parsers will be developed by adding validation of these additional constructs to standard XML Schema parsers. Therefore, the DDL part of MPEG-7 is considered to be XML oriented, to extend the XML Schema (as specified by [25]), to define the MPEG-7 data models and finally, to extend the MPEG-7 when it is necessary.

2.2. Conceptual View of MPEG-7

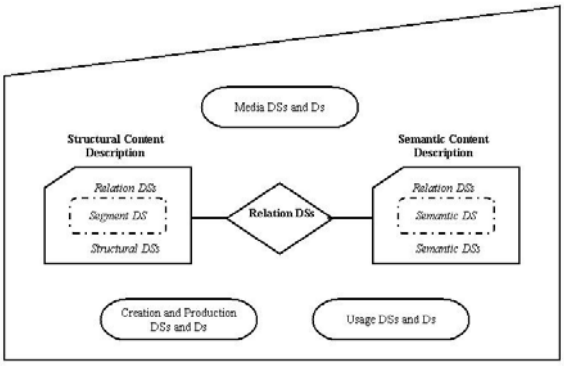

Figure 1. Description of Multimedia Content.

Intuitively, the concept behind the MPEG-7 methods, used for content description and management is to consider both physical and semantic characteristics of the video data. The content of multimedia data (and video data in particular), is adequately described if both their *structural* and *semantic* aspects are highlighted. The integration of both structure and semantic features is the most important contribution of the MPEG-7 standard. In more detail:

- The structural description of the content is based on the SegmentDS. The latter mainly describes the spatio-temporal segments of Audio Visual content and the way the structural contents are composed into hierarchical structures.
- The semantic description, deals with descriptions of the narrative worlds depicted or related to the audio-visual content, as well as objects, events, places and time in the narrative worlds.For this part of the description the entity that can be used as a basis is the Semantic DS.
- A full description of the audio visual content, requires both the structure and semantic aspects, as well as their interrelation to be defined. MPEG-7 provides the basic DSs and Ds that can be associated with the structural and the semantic descriptions of the content, and even the tools to create new DSs that can be tailored to the needs of a specific application.

Figure 1 visualizes the relationships between the components of the Multimedia Content Description based on the following viewpoints ([22, 24, 25]):

- *Creation and Production* : Typical features like title, creator, classification and purpose of the creation are included in this viewpoint. This information is usually provided by the author of the context since it cannot be automatically extracted by the content. Consequently, this viewpoint is related to high level description of the content as it has been defined in Section 1.
- *Media* : This viewpoint refers to the description of the storage media; typical features related to this description viewpoint are the storage format, the encoding of the AV content, and the elements for the identification of the media.
- *Usage* : This viewpoint is about meta-information related to the Usage of the content. More specifically, it defines the right holders, access rights, as well as publication and financial information related with the content.
- *Structural Aspects* : Description from the the structural prespective. The description is structured around segments that represent physical, spatial, temporal, or spatiotemporal components of the AV content. Each segment may be described by the signal-based features (color, texture, shape, motion, audio).
- *Conceptual Aspects*: Description of the AV content from the viewpoint of its conceptual notions.

3. Description Definition Language: An Overview

Although a lot of research effort has focused on generating Ds and DSs for video indexing, comparatively little

research has been done on schemas capable of defining the structure, content and semantics of multimedia documents and of enabling validation and high-level automated checking. However, such schemas will be necessary for the development of cost-efficient, user-friendly, semi-automatic metadata generation and editing tools for video. The main goal of the schemas for multimedia content-based description is to effectively describe the metadata associated with the multimedia data. Therefore the problem of effective multimedia content description is equivalent to the problem of effective metadata representation. A summary of the proposals submitted to the MPEG-7 committee for the DDL part is given in [7], while an extensive comparison of the existing schemas for video metadata representation is given in [9]. In [5] the authors compare the functionalities of RDF and XML schema and propose a combination of the two in order to enhance the interoperability between the metadata descriptions of different applications. Finally, in [6] an approach of dynamic and flexible mapping among metadata standards is described. In order for this to be done, a metadata term "thesaurus" has been developed, which provides the additional semantic knowledge that is non-existent within declarative XML-encoded metadata descriptions. This section will focus on the schema that can be adopted for multimedia content description. More specifically, we first address the requirements for such a description language and secondly we categorize the existing and proposed schemas based on the aspect of the description on which they focus.

3.1. Requirements of the Description Language

A schema able to adequately describe multimedia content should satisfy mainly the following requirements ([7, 9]):

- *Hierarchical Structure Definitions:* The schema must be able to constrain the structure to a precise hierarchy in which complete multimedia documents reside on the top level. For example in case of video documents, the whole video documents reside on the top of the hierarchy, and these in turn contain sequences, which contain scenes, which contain shots, and shots contain frames and so on.

- *Compositional capabilities:* The Multimedia Content Description Language should allow new DSs and Descriptors Ds to be created and existing DSs to be modified or extended.

- *Unique identification:* The Description Derfinition Language should allow a unique identification of Ds and DSs. Example: An example of unique identification is a namespace, which enables the unique qualification of element names and relationships and avoids name collisions on elements which have the same name but are defined in different vocabularies/domains.

- *Data Types:* The DDL should:

  1. provide a set of primitive data types e.g. text, integer, real, date, time/time index, version etc.
  2. be able to describe composite data types such as histograms, graphs, RGB values, enumerated types etc.
  3. provide a mechanism to relate Ds to data of multiple media types of inherent structure, particularly audio, video, audio-visual presentations, the interface to textual description, and any combinations of these.
  4. allow various types of DS instantiation: full, partial, full-mandatory, partial-mandatory.

- *Spatial, Temporal and Conceptual Relations:* Spatial relationships such as neighboring objects and temporal relations such as sequential or parallel segments should be supported. Moreover, given such relationships between two classes, it should also be possible to constrain specific attribute values of these classes.

- *Relationship between description and data:* The description language should supply a powerful model for links and references between one or more descriptions and the data that it describes.

- *Platform Independence:* The Multimedia Description Scheme should be platform and application independent.

- *Availability of Existing technologies:* Representation of Multimedia content that has no compliance with the existing technologies cannot be applicable and useful since it would require a lot of effort for the corresponding technology to be developed and become mature.

- *Grammar and Parser Requirements:* As far as the parsing phase is concerned, the description language should follow an unambiguous and easily parsed grammar, while the parser should be able to validate the values of the properties, the structures, the related classes and the valuesof the properties of the related classes.

- *Human readability:* It is preferable both the schema and the description output to be human readable.

3.2. Categorizing MPEG-7 DDL proposals

According to [7], the proposals sumbitted for the MPEG-7 DDL part can be subdivided into three broad categories:

- **semantic-based approaches:** The semantics-based approaches start with a knowledge-base schema language like RDF- Resource Description Framework([20])- or OKBC-Open Knowledge Base Connectivity ([2]). For the expression of this knowledge a *syntax* should be used.

- **grammar-based approaches:** The grammar-based approaches are mainly based on XML. They actually start with the XML syntax and through richer constraint mechanisms such as data typing and inheritance they allow storing varying degrees of semantic information. In these cases some presentation related aspects can be enhanced by specifying spatial and temporal control layouts.

- **presentation-based approaches:** According to the authors knowledge concerning the proposals submitted for the DDL part of the MPEG-7 only one of them could be classified to this category and it is based on SMIL (Synchronized Multimedia Intergration Language) [21].

The multimedia content description can be based purely on one of these three approaches, or can start based on one of them and try to be enhanced so that the advantages of the others to be met. Another classification of the MPEG-7 DDL proposals can be based on the target applications which will host the descriptions n this case the following categories can be discriminated:

- **Search and Retrieval Applications:** The semantic-based proposals are mainly applicable to this kind of applications since both search and retrieval would use knowledge representation and Artificial Intelligence Concepts.

- **Automated Content Description Applications:** The proposed DDLs that were focused on the structural aspect of the language and on the validation process are the main components of this category. These languages' objective is the automated generation of the validaded content descriptions for large scale applications.

- **Multimedia Navigation Applications:** This category consists of the hypertext-focused proposals. These proposals were mainly interested in using links within and between descriptions and content as a means of navigating through multimedia documents.

It is clear, that the choice of a DDL involves determining the optimum balance between simplicity and usability (on one hand) and the extensibility needed to satisfy the MPEG-7 requirements (on the other hand). Although a pure XML approach (as proposed by several organizations) is readily available, simple and cost-effective, it fails to satisfy certain fundamental DDL requirements such as data typing or co-ordinate systems. Conversely, certain other proposals offer complete extensibility but at the price of high complexity in DS design and parsing.

4 The MPEG-7 DDL

XML schema language with some changes and additions that will make it compliant with the MPEG-7 requirements has finally been adopted as the most appropriate schema for the MPEG-7 DDL. The W3C XML Schema specifications [23] are currently at the Candidate Recommendation stage. The MPEG-7 community is expected to provide implementation feedback and continue to liaise with the W3C and the XML Schema Working Group to ensure that MPEG-7 requirements are satisfied as far as possible. Therefore, MPEG-7 DDL and XML Schema are strongly related to each other and MPEG-7 DDL is expected to reflect any changes made to XML Schema. The main logical normative components to which the DDL can be broken to are *the XML Schema structural components, the XML datatype language components* and *the MPEG-7 specific extensions*.

4.1. XML Schema Structural Components and Datatypes

According to the XML Schema Specification, an XML Schema consists of a set of Schema components that can be divided into three groups, *the primary, the secondary* and the *helper components*. The *primary components* consist of Type Definitions and Component Declarations. Type definitions define internal schema components that can be used in other schema components such as declarations. There are mainly two kinds of type definition components:

- simple types: that are simple data types that cannot have child elements or attributes

- complex types: which may carry attributes and have children elements or be derived from other simple or complex types.

As far as the declarations are concerned, there are mainly two types of declarations that can be encountered within the XML Schema specifications:

- Element Declarations: An element declaration specifies a type definition for an element either by reference or explicitly, and may provide occurrence (minOccurs and maxOccurs attributes) and default information (default attribute).

- Attribute Declarations: Attribute Declarations associate a name with constraints on the presence and value of the attribute by referring to a simple datatype.

A distinction between declarations and definitions seems to be necessary at this point. Declarations are things that will be used in an XML instance document such as element or attribute declarations. The definitions on the other hand, are things that can be just used in the schema document. A definition in fact creates a new type that can be a simple, a complex type or attribute group and model group definitions.

The Attribute Group, the Identity Constraint and the Model Group Definitions as well as the Notation Declarations consitute the *secondary components* of the XML Schema. The Attributte Group definitions provide a mechanism to include a group of attributes by name within a complexType definition. Finally, the *helper components* contain the Annotations, Attribute declarations, Model groups, Particles and Wildcards. Details on each one of these components and subcomponents of the XML Schema are provided at [25]

As far as the datatyping facilities of the MPEG-7 DDL are concerned, it must be mentioned that they are based on the corresponding XML Schema facilities ([23]). There are though certain extensions to satisfy MPEG-7 specific requirements. The datatyping facilities provided within the XML Schema:Datatypes specification can be discriminated into two main categories: the *build-in Primitive Datatypes* and the *built-in Derived Datatypes*. The second type of datatypes has been derived from the primitive datatypes and are provided by the XML Schema:Datatypes. Two important data types that should be mentioned separately is the List and the Union Datatypes. A list datatype must be derived from an atomic datatype. This yields a list datatype that can contain whitespace separated lists of values of the base type. The Union types on the other hand enable element or attribute values to be one or more instances of one type drawn from the union of multiple atomic and list types.

4.2. MPEG-7-specific Extensions of the XML Schema

The following features have been added to the XML Schema Language specification in order to satisfy specific MPEG-7 requirements :

- **Array and Matrix Datatypes**
  The notion of matrix and arrays is necessary for multimedia context description and therefore mechanisms able to check whether the size of an array or a matrix is conformant either to a pre-defined value or to an attribute at the time of instantiation must be supplied by the corresponding DDL. Under this prespective, an extension to the XML schema specifications is necessary. The extension will be based on the use of list datatypes in order to provide support for such mechanisms.

  More specifically, based on the list datatype, two methods are provided for specifying sizes of (1D) arrays and multi-dimensional matrices. A new facet $mpeg7:dimension$, is provided to enable the specification of the dimensions of arrays or matrices. When $mpeg7:dimension$ is used with positive integers, the size is fixed and the parser will generate the array or matrix from a list of values, accordingly. The $mpeg7:dimension$ facet can also take a value "unbounded". In such a case, a derived simpleType can redefine these unbounded dimensions with a fixed value. Furthermore, an instantiated matrix, whose definition contains unbounded dimensions, should be used in conjunction with the $mpeg7:dim$ attributes so that any ambiguity occuring when using unbounded dimensions to be removed. For 1D arrays, the $mpeg7:dimension$ facet provides the same functionality as the existing $length$ facet. If neither the $mpeg7:dimension$ nor the $length$ facet are specified, the default is a list of arbitrary length. The length will be determined by the number of elements in the list. Preferably only one facet ($length$ or $mpeg7:dimension$) is specified. If both are specified then they must be consistent otherwise the parser will flag an error. The first three definitions of Table 1, show the definitions of matrix data types with unbounded and bounded dimensions for a two-dimensional matrix.

  In case of unbounded dimensions, the value of $mpeg7:dimension$ should be specified within the description using the reserved $mpeg7:dim$ attribute. The $mpeg7:dim$ attribute is also applicable when support of parameterized array sizes is required. In this case this attribute is used to specify the dimensio of a complex list type at the instantiation time. The $mpeg7:dim$ attribute must itself be defined in the mpeg7 namespace as a lilst of integers. Indicative XML Schema Language code that defines the "dim" attribute and shows the way this can be used for the definition of an array is given in the last four definitions of Table 1.

- Built-in datatypes for *basicTimePoint* and *basicDuration*. The *basicTimePoint* datatype specifies a time point according to the Gregorian dates, day time and the Time Zone (TZ). The format is based on the ISO 8601 norm. To reduce conversion problems only a subset of the ISO 8601 formats is used. This is also the case for the *basicDuration* datatype that specifies the duration of a time period according to days and time of day. Notice that fractions of a second are specified according to the basicTimePoint datatype. The syntax definition of the *basicTimePoint* and *basicDuration* is shown in Table 2.

The adequate and efficient description of multimedia data requires effective represention of audiovisual data. The

2 Dimensional Matrix - Unbounded Dimensions

```
<IntegerMatrix2D>
<simpleType name="IntegerMatrix2D">
 <list itemType="integer">
  <annotation><appinfo>
   <mpeg7:dimension value="unbounded unbounded"/>
  </appinfo></annotation>
 </list>
</simpleType>
```

2 Dimensional Matrix-Bounded Dimensions

```
<simpleType name="IntegerMatrix2x3">
 <restriction base="mpeg7:IntegerMatrix2D">
  <annotation><appinfo>
   <mpeg7:dimension value="2 3"/>
  </appinfo></annotation>
 </restriction>
</simpleType>
```

Element Declaration

```
<element name="IntegerMatrix2x3"
     type="mpeg7:IntegerMatrix2x3"/>
<IntegerMatrix2x3>
  8 9 4
  6 7 8
</IntegerMatrix2x3>
```

List of Positive Integers for Dim attribute

```
<simpleType name= "listOfPositiveIntegerForDim">
 <list itemType="positiveInteger"/>
</simpleType>
```

Dim Attribute Definition

```
<attribute name="dim">
 <simpleType>
  <restriction base="listOfPositiveIntegereForDim">
   <minLength value="1"/>
  </restriction>
 </simpleType>
</attribute>
```

Defining a matrix based on Dim Attribute

```
<complexType name="NDimIntegerArray">
 <simpleContent>
  <extension base="listOfInteger">
   <attribute ref="mpeg7:dim"/>
  </extension>
 </simpleContent>
</complexType>

<element name="IntegerMatrix"
    type="NDimIntegerArray" />
```

Using the Integer Matrix DataType

```
<IntegerMatrix mpeg7:dim="2 4">
  1 2 3
  5 6 7
</IntegerMatrix>
```

Table 1. Defining Matrices

Definition the basicTimePoint Datatype

```
<simpleType name="basicTimePointType">
 <restriction base="string">
  ="\-?P(\d+D)?(T(\d+H)?(\d+M)?(\d+S)?(\d+N)?
      (\d2f)?)?(\d+F)?((\-|\+)\d2:\d2Z)?"/>
 </restriction>
</simpleType>
```

Definition the basicTimeDuration Datatype

```
<simpleType name="basicDurationType">
 <restriction base="string" >
  <pattern value="\-?P(\d+D)?(T(\d+H)? (\d+M)?(\d+S)?
      (\d+N)?(\d2f)?)?(\d+F)?((\-|\+)\d2:\d2Z)?"/>
 </restriction>
</simpleType>
```

Table 2. Basic time Time Point and Duration Data Type

representation of visual information is made with the use of matrices, while audio information is stored in vectors and therefore MPEG-7 DDL should provide support for the corresponding data types. Moreover, multimedia data are mainly characterized by their timing and synchronization constraints, which should be explicitely represented. This requirements make obvious the introduction of the time-related datatypes described before. Finally, it should be pointed out that there are XML Schema features that are currently not necessary for satisfying the MPEG-7 DDL requirements. However, there is no guarantee that these XML Schema features will not be necessary at some point in the future. The most characteristic example of such a feature is the Notation Declarations.

5. Concluding Remarks

This paper focuses on the notion of Multimedia Content Description and it is based on the MPEG-7 specifications towards these goal. More specifically it summarizes the MPEG-7 objectives and its basic components towards these. Special Emphasis is given on the Description Definition Language part. The requirements of the language that should be adopted are specified while the proposals submitted to the MPEG-7 commitee are categorized based on their main goals and applications to whithin which they can be used. The DDL proposed in the final Comittee Draft submitted this March (2001), based on the XML Schema, is described in further detail. The basic components of the XML Schema Language and the extensions proposed are described in further detail while some indicative examples for the presented cases are provided.

References

[1] A.B. Benitez, S.Paek, S. Chang, A. Puri, Q. Huang, J.R. Smith, C. Li L.D. Bergman and C.N. Judice: "Object - Based Multimedia content description schemes and applications for MPEG-7", *Signal Processing: Image Communication* , Vol. 16, pp. 235-269, 2000

[2] R.J. Douglass, DiamondBack Systems Inc. for DARPA Consortium: "Knowledge Representation Language for MPEG-7 DDL", P124, MPEG-7 AHG Test and Evaluation Meeting, Lancaster, Feb 1999

[3] C. Dorai and R.M. Bolle: "On the Evolution of Videotext Description Scheme and Its Validation Experiments for MPEG-7",*Proceedings of ACM Multimedia Conference*, 2000.

[4] Q. Huang, A. Puri, A. B. Benitez, S. Paek and S.-F Chang: "Proposal for MPEG-7 intergration description scheme for multimedia content", *Proposal to ISO/IEC JTCI/SC29/WG11 MPEG99/P477*, Lancaster, UL, February 1999.

[5] J. Hunter, C. Lagoze: "Combining RDF and XML Schemas to Enhance Interoperability Between Metadata Application Profiles", *accepted at WWW10*, HongKong, May 2001.

[6] J. Hunter: "MetaNet - A Metadata Term Thesaurus to Enable Semantic Interoperability Between Metadata Domains", *to be published Journal of Online Digital Information*, March 2001

[7] J. Hunter and F. Nack: "An Overview of the MPEG-7 DDL Proposals", *Signal Processing: Image Communication Journal, Special Issue on MPEG-7*, Vol 16 pp 271-293, 2000.

[8] J.Hunter: "A Proposal for a New MPEG-7 Description Definition Language (DDL) Grammar", *49th MPEG Meeting*, Melbourne, Australia, October 1999.

[9] J. Hunter and L. Armstrong, "A Comparison of Schemas for Video Metadata Representation",*WWW8*, Toronto, May 10-14, 1999

[10] J. Hunter: "An "Improved" Proposal for an MPEG-7 DDL", *47th MPEG Meeting*, 15-19 March, 1999, Seoul

[11] J. Hunter: "A Proposal for an MPEG-7 Description Definition Language", *MPEG-7 AHG Test and Evaluation Meeting*, 15-19 Feb 1999, Lancaster

[12] J. Hunter: "MPEG-7 Behind the Scenes",*D-Lib Magazine* Vol.5, No.9, Semptember 1999.

[13] J.-R. Ohm, F. Bunjamin, W. Liebsch, B. Makai, K. Muller, A. Smollic and D. Zier: "A set of visual feature descriptors and their combination in a low-level description scheme", *Signal Processing: Image Communication* 16(2000), pp. 157-179. pp. 211-234.

[14] S. Paek, A. B. Benitez and S-F. Chang : "Self-Describing Schemes for Interoperable MPEG-7 Multimedia Content Descriptions", *Electronic Imaging:Visual Communications and Image Processing*, IST/SPIE'99, San Jose, CA, Jan 1999.

[15] S. Paek, A. B. Benitez, S-F. Chang, C.S. Li, J.R. Smith, L.D. Bergman, A. Puri, C. Swain and J. Ostermann: "Proposal for MPEG-7 image description scheme", *Proposal to ISO/IEC JTCI/SC29/WG11 MPEG99/P480*, Lancaster, UL, February 1999.

[16] S. Paek, A. B. Benitez S-F. Chang, A. Eleftheriadis, A. Puri, Q. Huang C.-S. Li, J.R. Smith and L.D. Bergman: "Proposal to MPEG-7 video description scheme",*Proposal to ISO/IEC JTCI/SC29/WG11 MPEG99/P481*, Lancaster, UL, February 1999.

[17] E. Rehm :"Representing Internet Streaming Media Metadata using MPEG-7 Multimedia Description Schemes", *ACM SIGMM Proceedings*, 2000.

[18] P. Salembier, R. Qian, N. O' conor, P. Correia, I. Sezan and P. van Beek: "Description Schemes for video programs, users and devices", *Signal Processing: Image Communication* 16(2000) pp. 211-234.

[19] MPEG Home Page. http://www.cselt.it/mpeg/

[20] "Resource Description Framework (RDF) Schema Specification", WD-rdf-schema-19981030, W3C Working Draft, October 1998. http://www.w3.org/TR/WD-rdf-schema.

[21] "Synchronized Multimedia Integration Language", WD-smil-0202, W'3C Working Draft, 2 February 1998. http://www.w3c.org/TR/WD-smil.

[22] "MPEG-7, Daring to Describe Multimedia Content", 2001.

[23] The W3C Recommendation Track http://www.w3.org/Consortium/Process/Process-19991111/tr.html

[24] ISO/IEC JTC1/SC29/WG11 N 3752, October 2000.

[25] ISO/IEC JTC1/SC29/WG11 N4002, N3966, N4006, N4001, MPEG 01/ March 2001, Singapore

[26] ISO/IEC JTC1/SC29/WG11 N 3546, July 2000 and ISO/IEC JTC1/SC29/WG11 N 3545, July 2000