AUDIOVISUAL ARCHIVE WITH MPEG-7 VIDEO DESCRIPTION AND XML DATABASE

Pedro Almeida, Joaquim Arnaldo Martins, Joaquim Sousa Pinto, Helder Troca Zagalo
IEETA – Instituto Engenharia Electrónica e Telemática de Aveiro, Departamento de Electrónica e Telecomunicações, Universidade de Aveiro – Campus Universitário de Santiago, 3800-193 Aveiro

Keywords: MPEG-7, XML, NXDB, Audiovisual Archive, Multimedia, Digital Libraries

Abstract: This article presents the development of an audiovisual archive that uses the MPEG-7 standard to describe video content and a XML database to store the video descriptions. It presents the model adopted to describe the video content, the framework of the audiovisual archive information system, a video indexing tool developed to allow the creation and manipulation of XML documents with the video descriptions and an interface to visualize the videos over the Web.

1 INTRODUCTION

This article describes the work developed in the creation of an audiovisual archive that allows to index and store the content of the parliamentary video records of the Portuguese Parliament. This project appears as part of the digital library for the Portuguese Parliament, mainly associated with the system Electronic Diaries of the Portuguese Parliament (Pinto, 2001). The main objective of this project is to allow the visualization of a video of a complete session of the parliamentary debates or a small video segment of one session that corresponds to the intervention of a specific orator.

In more detail, the intention is to characterize a movie of a parliamentary session from the Portuguese Parliament, split the video in several segments and characterize them in a temporary and descriptive level. This way it is later possible to visualize segments that correspond to parliamentary interventions that contain specific characteristics.

Primarily are described the base technologies over which lays the information system, namely XML, XML Schemas, XML databases and Web Services.

It is presented the model, built with MPEG-7 elements, that allows a detailed characterization of an audiovisual content of a video from a parliamentary session of the Portuguese Parliament.

After the model it is presented the framework of the information system that has been developed, as well as its characteristics, with a special note to a video indexing tool that allows several users to index different videos from different parliamentary sessions and to the Web Viewer that makes it possible to view the videos over the web.

2 TECHNOLOGIES

2.1 XML

XML, eXtensible Markup Language, is a World Wide Web Consortium (W3C, 2002) recommendation and comes as an evolution of SGML, Standard Generalized Markup Language (ISO, 2001), a markup language. Initially, its objective was to overcome some limitations of HTML, HyperText Markup Language (W3C, 1999). XML comes as a markup language that allows relating text content with the marks by which it is delimitated.

The main difference between XML and HTML is that while in HTML all the marks that appear in a document are defined by the HTML standard in XML its possible to create marks whose syntax and semantic are specific, bringing great extensibility to this markup language.

2.2 XML Schemas

Despite the fact that an XML document presents its data delimitated by marks, nothing stops that a user interpretation is different from the one intended, not taking in regard the semantic of the marks. This

Almeida P., Arnaldo Martins J., Sousa Pinto J. and Troca Zagalo H. (2004).
AUDIOVISUAL ARCHIVE WITH MPEG-7 VIDEO DESCRIPTION AND XML DATABASE.
In *Proceedings of the Sixth International Conference on Enterprise Information Systems*, pages 536-540
DOI: 10.5220/0002614605360540

brings the need for a language that permits describing the structure of a XML document.
Initially came the DTD's (Document Type Definition) (W3C, 2000) proposed by the W3C as a way of defining a structure to the XML documents.
Later, due to some limitations of the DTD's came the XML Schemas (W3C, 2001) as a W3C recommendation.
The goal of a XML Schema is to define a way to build a XML document according to a defined structure. XML Schemas permit defining the elements and attributes of a XML document, the positions where they appear, the order of the child elements, the number of child elements, if a element may be empty or not, data types to the elements and attributes, default values to elements and attributes, etc.

2.3 XML Databases

The video descriptions are stored in a XML document with a structure as the one defined in section 3.2 and it is used a XML database to store these documents.
The DBMS (Database Management System) used is a NXDB (Native XML Database). It is called XIndice (Apache, 2003) and is based on an open-source platform developed by the Apache Foundation Software.
The use of an XML database was justified by the fact that the video descriptions were stored in XML documents, taking advantage of the functionalities associated to native NXBD's in storing and searching XML data.

2.4 Web Services

In a conceptual level Web Services (W3C, 2002) are services offered via the Web (Armstrong, 2003).

The main objective of using Web Services in the information system of the audiovisual archive is to create an abstraction level that allows establishing inter-application communications in a transparent way, ensuring that the system has the best modularity as possible. This kind of approach allows, in the future, the use of other DBMS's without the need to rebuild or recompile the code that builds the information system.

3 MPEG-7

The MPEG-7 standard permits the description of various types of multimedia information. One of the objectives of this standard is to permit efficient characterization of audiovisual material.

This standard does not cover the area of automatic extraction of descriptors neither specifies a search engine that can use the descriptors, permitting that software factories build their own tools raising this way the competition and functionality of the available tools.

The MPEG-7 standard uses XML and XML Schemas as a descriptive language, permitting this way high extensibility and easiness of use. This fact also allows a high interoperability, creating independence of the standard from a specific software platform or software vendor. (Martinez, 2002)

3.1 MPEG-7 Elements

The MPEG-7 standard is composed of three elements that permit creating descriptions of audiovisual content: (Martinez, 2002)
1. Descriptors (D) – Representations of characteristics, define the syntax and the semantic of each representation to each characteristic.
2. Description Schemes (DS) - Specifies the structure and semantic of the relations between components. These components can be either Descriptors or Description Schemes.
3. Description Definition Language (DDL) – Permits the creation of new Description Schemes and Descriptors and the extension or modification of existing Description Schemes.
MPEG-7 consists of seven parts (Martinez, 2002). The Multimedia Description Schemes part was used in the creation of the model presented further ahead.

3.2 MPEG-7 model

Figure 1 presents the model of description built with MPEG-7 elements and shows the Description Schemes that where used to describe the video content of a parliamentary session.

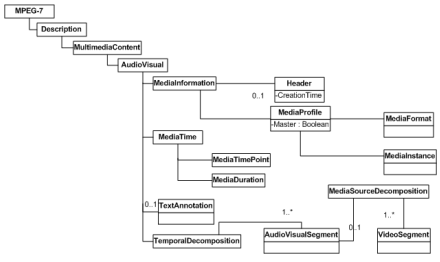

Figure 1: MPEG-7 description model

The first element in the model is the MPEG-7 element. This element indicates that the content of the XML file is a MPEG-7 description. After this element appears the Description element followed by a MultimediaContent element, which indicates the type of content that is going to be described. The fallowing element is the AudioVisual element. This element represents the total audiovisual content, in this particular case a complete video of a parliamentary session of the Portuguese Parliament. The MediaInformation element contains information about the video codification and the location of the audiovisual content and the MediaTime element contains information about the duration of the complete video. The TemporalDecomposition element indicates that there is a temporal decomposition of the audiovisual content. From this element derives one or more AudioVisualSegment elements that represent each segment of the audiovisual content described. Each segment contains the necessary information for its correct characterization and identification. Associated with the audiovisual content may exist a TextAnnotation element that permits adding textual information that characterizes the audiovisual content, namely textual notes and keywords. Finally the MediaSourceDecomposition and VideoSegment element permit the characterization of sub-segments of a video segment, increasing the granularity of the audiovisual archive system.
A more detailed explanation of the model can be obtained in a previous article (Almeida, 2003).

4 AUDIOVISUAL ARCHIVE INFORMATION SYSTEM FRAMEWORK

Figure 2 presents the audiovisual archive information system framework. This framework is based in the classic model of three layers: data layer, logic layer and presentation layer.

The data layer is composed of three components that store information. The first repository is a video collection with the debates from the Portuguese Parliament. The second is a relational database that contains information about the interventions of orators from the parliament. The third component is a XML database that stores the video descriptions.

The logic layer is composed of a group of technologies that have been used in order to permit the construction of a distributed information system for the audiovisual archive, based on the client-server model.

Finally, the presentation layer presents the video indexing tool and the web viewer, being this interfaces available to interact with the audiovisual archive.

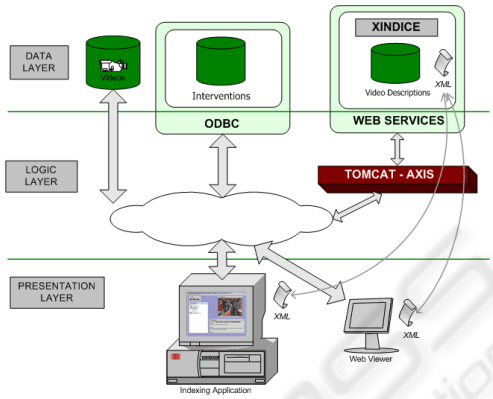

Figure 2: Audiovisual Archive information system framework

4.1 Data layer

4.1.1 Videos

The parliamentary videos are stored in a video server and organized according to a hierarchic structure to allow the use of an automatic method of recovery. The videos names can be obtained by the expression S[ns]L[nl]SL[nsl]N[nsp] , where ns , nl , nsl and nsp correspond to the number of the series, legislature, legislative session and parliamentary session. For example, in the case of a video from session number 2, 8.th legislature, 1.st legislative session, 1.st series the name of the video will be S1L8SL1N2.

4.1.2 Interventions database

The interventions database is stored in a legacy system. This database has information about the interventions of orators in each session of the Portuguese Parliament. From this database it is possible to obtain information about the name of the speaker, the summary and the pages where the intervention is written in the paper Diaries of the Portuguese Parliament.

4.1.3 Video description database

The database with the video description is a native XML database. This database is where the indexed video descriptions are stored. For each indexed video there is a record in the database, represented by a XML file that contains all the information

necessary to decompose and characterize a video of a parliamentary session.

4.2 Logic layer

This layer guaranties independence between the data layer and the presentation layer.
In the connection to the relational database with the interventions information's it is used the familiar technology of ODBC (Microsoft, 2003).
In the case of the XML database with the video descriptions it was created a Web Service, xmldbws, to allow the communication with the presentation layer.
To implement the Web Service it was used AXIS (Apache, 2003 A) with the TOMCAT (Apache, 2003 B) HTTP server.
AXIS is a SOAP (W3C, 2003) implementation of the W3C.
The Web Service was used to ensure that the manipulation of the records of the XML database is done independently of the XLM DBMS. It has a series of methods that allow manipulating XML documents in the XML database.

4.3 Presentation layer

The presentation layer is where the applications that permit interaction with the audiovisual archive system are located.

4.3.1 Video Indexing Application

With the use of this application it is possible to create, alter and eliminate video descriptions of a video collection being indexed.

The application is an MDI (Multiple Document Interface) composed by four internal windows, each one with a specific functionality.

Figure 3: Video Indexing Application Interface

Figure 3 presents the video indexing application Interface.

The application was developed in JAVA and some JAVA packages were used to permit a quicker and more efficient development. The JMF (Java Media Framework) (Sun, 2003) package was used in the creation of the internal window that presents the video.
Another important package used was the JAXB (Java API for XML Binding) (Sun, 2003) package. With this package it was possible to compile an XML Schema with the model of the XML document and was created a group of JAVA classes. These classes were later used in the Video Indexing Application to allow an easy manipulation of the XML documents.
The information presented in the *Intervenções* window is used as a guide during the indexing process. It indicates the name of the orators, the scenes that have been indexed and the scenes that are not yet indexed. This helps the technician's job of the indexing the video.
The *Anotações* window is where the user adds temporal and textual information to a video segment. The information inserted in this window is stored in a MPEG-7 compliant XML record in the XML database.

4.3.2 Web viewer

The web viewer was developed using Microsoft .NET (Microsoft, 2003) programming environment. The main objective of developing the web viewer in .NET was to test the interoperability between programs built in different platforms. Figure 4 presents the interface of this part of the system.

Figure 4: Web Viewer interface

This viewer consists of an aspx developed with C# and basically is composed by a tree view object with a media player object.
The information presented in the tree view is obtained from the intervention database and the video descriptions XML database. To create the tree view it was implemented a Web Service Client in the .NET platform that connects to the Web Service Server implemented in JAVA.
Figure 8 presents the communication architecture of the Web Viewer interface.

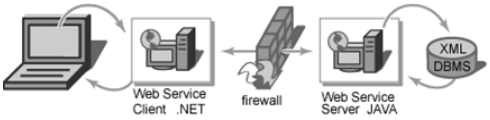

Figure 8: Web Viewer communication architecture [Source: adapted from MSDN]

The Web Viewer is represented by the Web Service Client .NET and the XML DBMS represents the videos descriptions XML database. When Web Services are used, normally, there is no need to configure the firewall. This fact is represented by the arrow that transverses the firewall.
This example shows that interoperability between applications of different platforms can be obtained using Web Services.
With this kind of approach the client only connects to the XML database once to obtain the video description. As long as the user doesn't change to another video, all the processing to obtain information to other scenes in the same video is done on the client side.

5 CONCLUSIONS AND FUTURE WORK

Building an information system that permits to describe video content is not a trivial task. It's necessary to study carefully the characteristics needed to describe the content or else it may become an unpractical system.
The audiovisual archive presented in this work is a particular example for a need of the Portuguese Parliament, but with little modifications it can be used to create a more generic system. The essential part of the work presented is the framework itself and the modularity and scalability of the system.
The MPEG-7 standard has answered completely to the needs of the system in terms of the video description. There are a vast number of descriptors in the standard that permit to describe video content in a very complete manner.
The Web Services in the logic layer permitted to create a very important abstraction level between the data layer and the presentation layer. This kind of approach permits having a high modularity in the information system of the audiovisual archive, allowing to have different technologies to support different components of the information system.
In the near future it is needed to study the behaviour of the XML DBMS in terms of search performance.

REFERENCES

Pinto, Joaquim Sousa, et. al., February 2001, "*Portuguese Parliamentary Records Digital Library*" , In Ahmed K. Elmagarmid , William J. McIver Jr, "*The Ongoing March Toward Digital Government*", Computer, Vol. 34, N.º 2, p. 38, IEEE Computer Society.

W3C, October 2002, "*Extensible Markup Language (XML) 1.1*" , http://www.w3.org/TR/xml11/ .

ISO, August 2001, "*Standard Generalized Markup Language (SGML)*", ISO 8879:1986 .

W3C, December 1999, "*HTML 4.01 Specification*", http://www.w3.org/TR/html4.

W3C, January 2000, " *Datatypes for DTDs (DT4DTD) 1.0*", http://www.w3.org/TR/dt4dtd.

W3C, May 2001, "*XML Schema Part 0: Primer*", http://www.w3.org/TR/xmlschema-0/.

Apache, March 2003, "*Apache XIndice*", http://xml.apache.org/xindice/.

W3C, November 2002, "*Web Services Architecture Requirements*", http://www.w3.org/TR/wsa-reqs .

Armstrong, Eric. et al , February 2003, " The Java Web Services Tutorial ", Sun Microsystems Press.

Martinez, José M. , July 2002, "*MPEG-7 Overview (version 8.0)*", ISO/IEC.

Almeida, Pedro et al . , January 2003, "*Descrição de vídeo com Multimedia Content Description Interface (MPEG-7)*", ISSN : 1645-0493 , Vol. 3 , N. 8 .

DSTC, March 2003, "*XMLdbGUI - Download*", http://titanium.dstc.edu.au/xml/xmldbgui/download.shml .

Microsoft, June 2003, "*ODBC - Overview*", http://msdn.microsoft.com/library/default.asp?url=/library/en-us/odbc/htm/odbc01pr.asp.

Apache, January 2003 A, "*Apache Axis*", http://ws.apache.org/axis/ .

Apache, January 2003 B, "*The Jakarta Site - Apache Tomcat*", http://jakarta.apache.org/tomcat/.

W3C, June 2003, "*SOAP Version 1.2 Part 0: Primer*", http://www.w3.org/TR/soap12-part0/.

Sun Microsystems, June 2003, "*Java Media Framework API*", http://java.sun.com/products/java-media/jmf.

Sun Microsystems, March 2003, "*Java Architecture for XML Binding (JAXB)*", http://java.sun.com/xml/jaxb.

Microsoft, June 2003 , "*Product Information for Visual Studio .NET 2003* ", http://msdn.microsoft.com/vstudio/productinfo/default.aspx.