ARCHIVING MULTIMEDIA CONTENT DESCRIPTIONS: AN EARY ADOPTATION OF MPEG-7

Spiros Ioannou and Giorgos Akrivas

National Technical University of Athens, department of Electrical and Computer Engineering, Image, Video and Multimedia Laboratory (IVML/ECE/NTUA)

ABSTRACT

In this paper, PANORAMA, an integrated system for digitization, annotation, storage and retrieval of audiovisual data and their associated description is presented. The information model has been heavily influenced by the evolving MPEG-7 standard for description of multimedia content. XML schema has been used for defining the data structures and XML for representation of content – related information. A hierarchical and object – oriented XML database has been utilized for storage and retrieval. Subsystems for multimedia archiving, annotation and retrieval have also been constructed.

1. INTRODUCTION

The digital revolution of the past decade has resulted in great amounts of digital audiovisual information being created. The archives that are constructed need novel technologies for representing the associated knowledge about the audiovisual data, and adaptation of existing technologies in the fields of archiving, storage and retrieval. Because of the absence of automatic algorithms, human intervention is critical in many modules of the current systems.

We present, in this paper, an integrated system for managing audiovisual information that uses new technologies (MPEG-7 and XML) for representation of audiovisual descriptions.

2. INFORMATION MODEL

The status of MPEG-7, as it was available in late 1999, had provided a sufficient information model, as well as DDL, a language for creating new data types. Using DDL, MPEG-7 had defined a set of Descriptors for describing features of audiovisual content, and Description Schemes for organizing hierarchical description information. However, specific needs of the PANORAMA users, as well as the incompleteness dew to the early stage of the standard, had forced as to modify most of the DSs and create a number of new ones.

The most basic Description Schemes used by PANORAMA are described below.

A description of a whole Audiovisual document is contained in the AudiovisualDS. An Audiovisual DS contains information about physical storage (MediaInfoDS), textual information (MetaInfoDS), structural information (SyntacticDS) and summarization information (SummarizationDS).

Because of the complexity of the physical storage of some of the material owned by the PANORAMA users (especially in old cinema films), we had to store information about physical mediums in a separate entity, MediumDS, unlike MPEG-7. Specializations of this type have been CassetteDS, FilmDS and PhotoDS, as well as the two digital types, DigitalVideoDS and DigitalImageDS. Each of the mediums can contain one or more MediumSegmentDS, which has its own frame numbering. Copies between pairs of mediums are stored in MediaRelationDS.

Information about where audiovisual data is recorded is stored in MediaInfoDS. Copies of the audiovisual data in various mediums are stored in separated MediaProfileDS, which contains pointers to MediumDS. It can also contain frame numbers to specific MediumSegmentDS.

Information about segmentation of the information into separate units, such as shots, is stored in SyntacticDS. Unlike MPEG-7, which uses an unlimited number of levels, PANORAMA uses a four-level hierarchy of themes, shot groups, shots and keyframes. Each of the units can contain its own MetaInfo and MediaInfo.

Summarization information is contained in SummarizationDS. Like MPEG-7, a sequential and a hierarchical summarization are supported. The former contains pointers to a selection of shots and keyframe, while the latter consists of a subtree of the whole segmentation tree.

Annotation information desired by the users of the system includes strings (StringValueDS), integer numbers (IntegerValueDS), Dates (DateValueDS), frame numbers (TimeValueDS), boolean values (BooleanValueDS) and enumerations (DynamicListValueDS). Each of these value contains a reference to each name (which is contained in a set of names) and one or more values. These values are stored in a MetaInfoDS.

3. HANDLING OF XML DOCUMENTS

A basic design principle of PANORAMA is that data exchange among subsystems is in the form of valid XML documents. This created the need for a language capable of defining valid XML documents, a system for converting XML documents into C++ classes suitable for internal processing and a database capable of storing and retrieving XML documents.

Because of some difficulties of using DDL to define generic (non-MPEG-7) XML documents, we decided to transcript the MPEG-7 definitions from DDL to the then-emerging XML

Schema language, which is more expressive for describing hierarchy and inheritance[1].

Conversion to C++ classes had been necessary because of the difficulties in using the existing interfaces for handling XML documents in memory, such as the DOM and SAX interfaces. The module that was constructed creates a class for each complex type and a vector for each sequence of elements. Moreover, XML references are converted to C pointers.

The hierarchy and inheritance that is employed by XML Schema creates the need for construction of a new database that extends the currently used relational model, which supports neither of the features. The module that was constructed creates a table of a relational database for each complex type and each of its elements. Association between complex types, inherited complex types and elements is performed via foreign keys. The database supports the COM interface, which makes it accessible by both Windows – based and web – based applications.

4. THE ARCHIVING SUBSYSTEM

The archiving subsystem enables the administrator of an audiovisual archive to enter, modify and delete physical mediums (in digital or in analog form). The production of copies is automated by issuing copy orders, where the necessary data (the mediums involved and the time codes) are taken from the database and written into an XML file, which is used by the machine performing the copy. Alternatively, copy can be done manually; in this case, the user enters the information to the system a posteriori.

Using the archiving subsystem, the administrator can also issue annotation orders, to accept annotation results and to cancel pending annotation orders.

Finally, standard database operations (creation, destruction, backup and restore) are supported, as well as directly giving commands (in XML) to the XML database.

5. THE ANNOTATION SUBSYSTEM

An annotation order is an order to an expert (the annotator) to create one or more content descriptions (to be stored in AudioVisualDS) out of one or more physical mediums. Under an annotation order, the archiving system has created an XML file containing information about the physical mediums that have to be annotated, as well as audiovisual descriptions created in previous annotation orders.

The XML file, and an MPEG version of the medium to be annotated (if available), is loaded to the annotation subsystem, which is an off-line application. Using the annotation subsystem, the annotator can view the MPEG files and create a number of programs (separate audiovisual descriptions to be stored into individual AudioVisualDS). For each program, the annotator can create its structural information (to be stored into SyntacticDS), graphically shown in a hierarchical tree. Highlighted items of the tree will consist the hierarchical summarization, while another node of the tree represents the sequential summarization. For each node of the tree, its position in the MPEG file (MediaInfoDS) and its characterization (MetaInfoDS) can be entered. Out of the given position, positions in all the copies of the MPEG file (relative media) are computed. Annotation is presented in the form of a table with the names and the values of the fields; the latter can be entered and altered by the annotator.

Once the XML description is ready, it is returned to the archiving subsystem, which updates the database.

6. OTHER SUBSYSTEMS

A web – based user interface was constructed. The interface supports user authentication, queries (both text – based and content – based) and presentation of results. The hierarchical nature of the description is mapped in a hierarchy of interlinked web pages. Streaming and downloading is supported for viewing MPEG files.

A digitization system was utilized to convert video cassettes into MPEG video. The digitization subsystem is capable of receiving XML information regarding timecodes of the cassettes to be digitized. Cinema film was converted into TV signal via telecine.

Protection of intellectual property in digitized images and video is achieved through a watermarking application that was constructed for PANORAMA. The application embodies a watermark directly into the DCT information of the MPEG and JPEG data.

A set of algorithms for automatic extraction of visual features (color, and texture information, moving objects, camera motion, faces and regions with text) was constructed in order to support content – based queries. Moreover, the shots are automatically extracted in order to save the annotator time. These algorithms are used by the annotation subsystem. The XML database handles visual features separately, in order to achieve improved performance for feature – based queries.

8. EXPLOITATION AND FUTURE WORK

PANORAMA has been used by two audiovisual archives, the audiovisual archive of the Greek National Television and the National Film Registry. The annotation subsystem has enabled each archive to define its own annotation scheme (in terms of dynamic lists).

Extension of audiovisual descriptions towards higher level (conceptual) descriptions is the goal of the EU IST FAETHON project, to which the authors participate.

9. REFERENCES

[1] ISO/IEC JTC 1/SC 29/WG 11/N3966, "Text of 15938-5 FCD Information Technology – Multimedia Content Description Interface – Part 5 Multimedia Description Schemes", March 2001, Singapore.

[2] W3C, XML Schema Recommendation, May 2001, http://www.w3.org/XML/Schema

[3] ISO/IEC JTC1/SC29/WG11 N4032, "Introduction to MPEG-7", March 2001, Singapure.

[1] In March 2000, MPEG-7 also abandoned DDL in favor of XML Schema