USING MPEG-BASED TECHNOLOGIES FOR BUILDING PERSONALIZED MULTIMEDIA INTERACTIVE ENVIRONMENTS WITH CONTEXT-AWARENESS REQUIREMENTS

Development of an application for Interactive Television

João Benedito dos Santos Junior, Iran Calixto Abrão and Thelma Virgínia Rodrigues
Pontifical Catholic University of Minas Gerais – PUC Minas, Poços de Caldas – MG, Brazil

Mario Guglielmo
Multimedia Division, Telecom Italia Laboratory, Turim, Italy

Keywords: New MPEG Standards; Interactive Video and Audio; Context-Awareness

Abstract: We are using MPEG-4 technology to build applications to be used in real environments. One of these applications allows for teacher to send real-time lessons to this/her students or to record them. The Tele-Learning System under development includes: a) on the teacher side: a recording workstation with two cameras, microphone, specific MPEG-4 software; b) an IP network or an MPEG-2 TS satellite link; c) on the student side: a PC with special MPEG software, and a special board if receiving from satellite. This research focuses on the broadcast scenario where a satellite board is used in a PC. Thus, the work covers how to send the lesson even to a student that is not connected to the intranet, using a satellite link, either over IP embedded in the MPEG-2 TS or directly over MPEG-2 TS. For the security part it may be necessary to have a low-band return channel implemented, for example, through a mobile phone. The satellite environment may require the redesign of the User Interface and the retargeting of the elementary streams parameters in order to match specific requirements and features of the medium. At this point, new interaction criteria have been established from distribution of MPEG-4 media objects and MPEG-7 scene descriptions on network environments. Furthermore, context-awareness aspects are being added for providing personalization on the teaching-learning environment and MPEG-21 is being studied for applying to new multimedia requirements.

1 INTRODUCTION

Universities (and high schools) and commercial enterprises (manufacturers, service providers, ICT companies, among others) generate an increasing demand for more efficient means and supports to teaching, training and upgrading activities. In the different environments "efficient" has different meaning: from the reduction of time required by a student to get the subject of an academic course to the minimization of training costs and unproductive hours for the upgrading of a company personnel. In this context, "*e-Learning*", "Distance Learning", "Corporate Training" are few of the names used to indicate these kinds of services in several combinations with the "multimedia" attribute.

It is widely accepted that the answer to this demand can be provided by the new technologies, such as MPEG-4 for media coding and streaming and MPEG-7 for information description and indexing (MPEG4, 2000a; MPEG-4, 2000b; MPEG-7, 2000).

The common understanding is that by means of those technologies it will be possible to improve the quality, at the moment often quite poor with respect to content media, of the multimedia courses. The way in which the content is selected, merged and organized is the other key issue that imposes the creation of new professional characters: the

Benedito dos Santos Junior J., Calixto Abrão I., Virgínia Rodrigues T. and Guglielmo M. (2005).
USING MPEG-BASED TECHNOLOGIES FOR BUILDING PERSONALIZED MULTIMEDIA INTERACTIVE ENVIRONMENTS WITH CONTEXT-AWARENESS REQUIREMENTS - Development of an application for Interactive Television.
In *Proceedings of the Seventh International Conference on Enterprise Information Systems*, pages 180-183
DOI: 10.5220/0002526301800183

instruction architects. Easy "navigation" of the material and the actual possibility of accessing the relevant piece of information by "few" clicks are very important usability aspects that can determine the spreading of the service and its success or, on the contrary, its failure.

Availability and accessibility are important aspects to consider beside the costs. Distance learning platforms require suitable delivery media solutions that can be based on both telecommunication and broadcast infrastructures and the different hypotheses are in principle able to satisfy different requirements and could be a further interoperability issue (Battista et al., 1999). The management and protection systems proposed so far have not been able to protect anything for long time and new flexible and dynamically adaptable solutions are required.

2 THE SYTEM ARCHITECTURE

In recent times, the growing demand of *e-learning* systems brought to the choice of standard multimedia technologies, to allow learning contents being compatible with many other systems and reaching flexibility and reusability.

In this context, MPEG-4 was adopted; MPEG-4 features make this standard really suitable for *e-learning* applications; the main reasons are (MPEG-4, 2000a; MPEG-4, 2000b): a) the compression ratio reached by MPEG-4 Audio and Video encoders allows to achieve good audio-video quality; b) MPEG-4 Systems provide the splitting of the audio-visual signals in elementary, synchronized objects; c) a high security level is achieved, and it is guaranteed by a key, and not by algorithms.

The system represented in **Figure 1**, is composed of several parts: a) didactical sources: audio, video (two analog or digital video sources) and teacher slides; b) one or more PC (Control Stations) that encode, compress and possibly encrypt the learning elementary media; finally there are two software modules, a "streamer", that sends on different channels (IP multicast, satellite, among others) and a recorder that locally stores, in real-time, the whole MPEG-4 scene; c) the network connection; d) a PC for each student: it receives, decrypts and presents to the user the multimedia interactive lesson.

We present now the system details for the description of software modules, divided into three categories, described separately: a) modules for the creation of the real-time lesson; b) modules for the network distribution; c) modules for the fruition of the lesson.

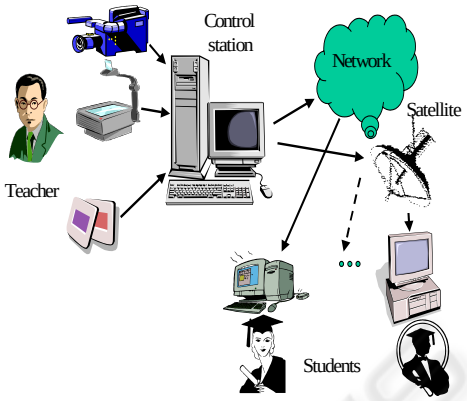

Figure 1: Diagram of the system.

The system tries to reproduce all these actions: this is reached making use of both local and remote interactivity from the teacher side and from the student side.

3 FRUITION SYSTEM

The student position is conceptually the opposite of the teacher one: the application installed on the student PC is a special software that implements all the necessary functions for the fruition of the lesson (live or recorded).

This application from the user point of view is an area in a page of a web browser, while internally is composed of many blocks, listed below: a) the Network Receiver accesses multimedia contents from the appropriate device (LAN IP, satellite, modem, among others); the output is made up of a series of interactive streams; b) the MP4 file reader accesses contents from local MP4 files; c) decoders for audio, video, JPEG, among others, to decode elementary streams; d) security modules (optional): this module decrypts the lesson received; e) compositor: it updates the student interface according to the commands received; f) student interface: the student can decide which stream of the lesson he wants to see in full-screen mode, navigate through the slides (among the slides already transmitted), have a break during the lesson or terminate it.

The scenario in which the student application works is the following: the student is provided with a multimedia PC and the special software to receive the lesson in MPEG-4 format. Through a LAN or satellite link he connects to the server that stores the pre-recorded lessons or to follow the current live lesson. After the beginning of the lesson, he can interact with the scene: thanks to the particular

description of the graphic interface implemented during the creation of the scene itself, the direction of the lesson is left on the user interface: in particular the student can select the stream he wants to see, the slide showed. He can also interact with the teacher with a chat system: when the student has a question he can send it using the chat system, and the teacher can choose the way of answering: text reply to the single student, to all the students or by microphone. The student has also the possibility to download a file that the teacher has sent to everybody (i.e. a document, a presentation, an executable file, among others). Another way of interaction is given by the questions that sometimes the teacher can ask: the students have to answer in real-time and send back the reply, so the teacher has immediately the results of the test.

4 SYSTEM EVOLUTION

The system is expected to evolve in three directions: the delivery to different kinds of networks, the adaptation to new emerging standards, and the context-awareness. In this scenario, the main aspects of each one of these directions are presented in the next sections.

The main problem in the direct use of *e-learning* is not technical, and is due to the scarce penetration of PC and related culture in these countries. In fact it is a reality to rent bandwidth in a satellite link and to send multimedia lessons over it using DVB-MPE (Multi Protocol Encapsulation). This way IP packets directly coming from the streaming engine are encapsulated as they are onto MPEG-2 Transport Packets (MPEG-2, 1994).

Interactivity can be further enhanced if the terminal is equipped with large storage capacity. This functionality can be implemented through standard languages, and this takes us to another big area of extensibility: MPEG-21 and standard Learning Objects (MPEG-21, 2001).

In parallel, different bodies are moving towards the definition of standardized languages, most having in common XML (eXtensible Markup Language) (Connoly, 1997) associated with different Schemas or Document Type Definitions (they both describe the syntactic structure of the document). This common choice and the intrinsic extensibility of the language facilitate the integration of the work of different bodies.

4.1 Applying Requirements of Context Awareness

Context can be viewed as being any information that can be used for characterizing the status of an entity in one specific case. One entity can be one person, one place or one object relevant for any type of interaction between user and application, including user and application itself (Dey & Abowd, 2000; Pascoe et al., 1999). Therefore, one system is context-aware when it uses the context for providing relevant information and/or services to the user, and this relevance feature depends on user's tasks (Abowd, 1999; Schmidt, 2000).

These parameters can be obtained based on five arguments: a) *where* is the user (WHERE); b) *who* is the user (WHO); c) *how* the user works (HOW); d) *when* it can do what (WHEN); e) *what* the user is doing (WHAT). Context-awareness aspects can be relevant when associated to environment, in which the application is inserted, providing information for adaptability. The words HOW, WHERE, WHAT, WHO and WHEN denote semantics that can be very complex. In this point, for example, the semantic HOW is referring to "*what resources are being used*" expresses, implicitly, "*how the user is using these resources*".

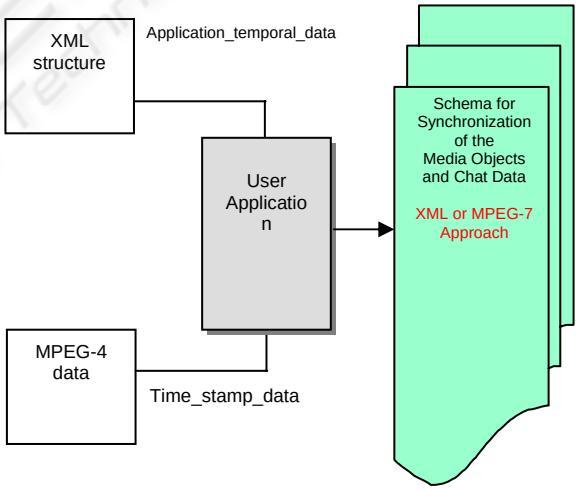

Figure 2: Basic model for integration of the new interaction criteria.

At the level of implementation, a field <TIMEDATE> of the HEADERSESSIONCHAT *xml_struct* can be also divided into sub fields, allowing more granularity in the synchronization processes with the MPEG-4 elementary streams. For example, a suggestion is the specification of the following fields: <day>, <month>, <year>, <hour>, <minute> and <second>. This additional contribution is a proposal for content description and

personalization of the teaching-learning environment, using the further resources of the technologies like MPEG-4 and MPEG-7.

As a concrete example, the XML structure, generated by a chat application on the teacher side, can be used for synchronizing the MPEG-4 objects with the questions that have been formulated by the students during the lesson. Thus, a publisher process can be used for presenting statistics about the lesson, considering the temporal relations described by chat time parameters and time stamps of the MPEG-4 objects, as illustrated in **Figure 2**.

Furthermore, MPEG-7 scene descriptions can be used for adding new personalization levels to the teaching-learning environment (Nack & Lindsay, 1999; Santos Jr. et al., 2001). The integration of MPEG-7 information with some DBMS (*DataBase Management System*), through the XML scheme, can provide new and interesting controls for both student and teacher interfaces.

5 CONCLUSIONS

This work has presented how the new standards of the MPEG family can be used of systematized form on the development of the educational applications, discussing the *e-learning* scenario developed in TILAB currently.

MPEG-4 is a standard that builds on the success of MPEG-1 and MPEG-2, two standards that have changed the audio-visual landscape. The best experts, provided by all industries with a stake in multimedia, have developed MPEG-4. It is a powerful standard, rich in functionality, encompassing other successful standards. Furthermore, MPEG-7 provides essential features for scene description, allowing the development of several personalized schemes for user interaction; MPEG-21 is being investigated for building new representation schemes for modeling of multimedia objects as universal digital items, mainly in terms of the learning-objects.

In the educational context, this paper presents a contribution for content description and personalization of the teaching-learning environment, using the further resources of the technologies MPEG-4 and MPEG-7, contributing to the increase of the personalization level in teaching-learning environments.

REFERENCES

(MPEG-4, 2000a): ISO/MPEG MPEG-4: MPEG-4 requirements, MPEG-4 applications and MPEG-4 development process. 2000.
http://mpeg.telecomitalialab.com/working_documents.htm.

(Battista et al., 1999): Battista, S. et al.: MPEG-4: A Multimedia Standard for the Third Millennium, part 1, IEEE MultiMedia 6(4) 74–83. 1999.

(MPEG-4, 2000b): Information Technology – Coding of Audio-visual Objects. MPEG-4 Systems. 2000.

(MPEG-7, 2001): ISO/MPEG N3445 MPEG-7: Multimedia Description Schemes XM (v.7.0). 2001. http://mpeg.telecomitalialab.com/working_documents.htm.

(Nack & Lindsay, 1999): Nack, F. and Lindsay, A. T.: Everything You Wanted to Know About MPEG-7: Part 1, IEEE Multimedia, pp.65 – 77, 1999.

(MPEG-2, 1994): Information Technology – Generic Coding of Moving Pictures and Associated Audio. MPEG-2 Systems. 1994.

(MPEG-21, 2001): ISO/MPEG - MPEG-21: Requirements on Digital Item Usage Environment Description. http://mpeg.telecomitalialab.com/working_documents.htm.

(Abowd, 1999): Abowd, G.D.: Software Engineering Issues for Ubiquitous Computing; International Conference on Software Engineering (ICSE'99), pp.75-84. 1999.

(Dey & Abowd, 2000): Dey, A.K. and Abowd G.D.: Towards a Better Understanding of Context and Context-Awareness; CHI 2000 - Workshop on The What, Who, Where, When, and How of Context-Awareness. 2000.

(Pascoe et al., 1999): Pascoe, J. et al.: Issues in Developing Context-Aware Computing. In Gellersen, H. (ed.): Handheld and Ubiquitous Computing. Lecture Notes in Computer Science, Vol. 1707. Springer-Verlag, Berlin Heindelberg New York, pp.208-221. 1999.

(Schmidt, 2000): Schmidt, A.: Implicit Human Computer Interaction Through Context. Personal Technologies Volume 4(2&3) June, pp.191-199. 2000.

(Santos Jr. et al., 2001): Santos Jr., J. B. dos; Goularte, R.; Faria, G. B.; Moreira, E. S.: Modeling of User Interaction in Context-Aware Interactive Television Application on Distributed Environments. Proceedings 1st Workshop on Personalization in Future TV – 8th International Conference on User Modeling. Sonthofen, Germany. July-2001. http://icmc.sc.usp.br/~joao/futuretv.

(Connoly, 1997): Connoly, D.: XML Principles, Tools and Techniques. World Wide Web Journal. O'Reilly, V.2, Issue 4, 1997.