Issues in Temporal Representation of Multimedia Documents

Nabil Layaïda

OPERA project, INRIA Rhône–Alpes.
655 avenue de l'Europe, 38330 Montbonnot, France.
e–mail: {Nabil.Layaida}@inrialpes.fr

1 Introduction

The recent advances in multimedia systems, together with the advent of high speed networks, paved the way to a new generation of applications. In particular, authoring environments have found in multimedia the means of increasing the richness of information contained in electronic documents.

One of the goals of the Opera team is designing an authoring environment for multimedia documents, called MADEUS, which meets the following requirements:

- a high level of **expressiveness** for both spatial and temporal dimensions;
- a **user–friendly interface** allowing highly interactive design process, scriptless and structured–based editing and automatic production of spatial and temporal layout;
- **the portability and cross–platform interchange** of multimedia documents;

To achieve this research goal we first focused on finding a "good" representation of time in multimedia documents. This representation is required to capture the temporal dimension of media objects like video, audio, etc. and is also used to temporally organize objects with respect to each other. Time representation and management is clearly the main difference between multimedia authoring environment and the mostly static editing/viewing ones found in HTML–based tools.

We present in this position paper a set of relevant issues that may help in the design of multimedia document representation. These are based on our experience in designing and implementing an authoring and browsing tool called MADEUS. The document representation used is very similar to HTML documents in that users can create their own presentations even in line mode. This is because the document

source integrate both elements descriptive markup for audio, video, motion style and temporal relations between these elements. We think that adding temporal behaviour to HTML documents must follow the same principles that brought it to success : the simplicity of the format which allows a rapid construction of information and easy reuse.

2 Representation of multimedia information

In order to measure the quality of a given time representation used to model the temporal organization of multimedia documents and to evaluate comparatively those used in different document systems [2][3][12], we need to define some criteria. In the following section, we present key points that affect significantly the quality of the temporal representation of multimedia documents. Each of these issues is discussed in order to arise the problems facing the design of time based documents.

Expressiveness of the synchronization language

The time representation must capture the different temporal characteristics of basic objects in a homogeneous way. Media objects may be static like text, still images, etc., or dynamic like video, audio, interaction, etc. A video or an audio are active when they are played. An interaction, like a button, is active when it is enabled so that it is possible to interact with it through the user interface. Dynamic objects have an associated duration where two cases must be distinguished according to their temporal behaviour:

- **Controllable** objects: a range of possible durations is known at the authoring stage and it is possible to choose any value within this range during the presentation stage (audio, video, ...).;
- **Uncontrollable** objects: their durations are only known at the end of their presentation (interaction objects, uncontrollable program executions like applets).

The time representation must also capture the different ways to temporally organize this kind of objects [16][13]. Firstly, it should be possible to distinguish four temporal instants when dealing with dynamic objects: when they are mapped (unmapped) at the screen (referred in the sequel as spatio−temporal instants) and when they start (finish) playing (referred in the sequel as temporal instants). Static objects are only affected by the spatio−temporal instants. Secondly, operators which express the temporal organization of the document must support logical time: relative positioning of

objects (as expressed by the thirteen temporal relations introduced by Allen [1]); quantitative time: numerical constraints placed on object's duration and causal time: how instants causes other presentation action to happen (the end instant of one object causes the termination of another).

Finally, the time representation must be able to deal with temporal hyperlinks which gives the document readers the ability to navigate freely either inside the document or outside it. Hyperlinking elements like scenes, audio and video breaks the linearity of planned scenarios and allows semantical exploration of the document content in the same manner as HTML documents.

Declarative description and hierarchic symbolic representation

Multimedia document authors do not necessarily have any particular skills in computer programming, so the editing phase must be very intuitive and as close as possible to the familiar editing style of classical documents. This goal is obviously hard to obtain when using scripting languages where the document architecturing scheme (manipulation of the document entities through a set of associated attributes) disappears in favour of the programming one. Defining global attributes on composite objects like color, font, position, motion, etc. and applying inheritance rules on these attributes reduce the burden on document specification, increase the reuse and hide low−level presentation details. Existing document models based on declarative style [6] and descriptive mark−up of its content like HTML are more and more widely used. We think that keeping these best approach in order to disseminate time−based multimedia documents. Moreover, it is also suitable for direct manipulation authoring interfaces.

The symbolic representation of time in multimedia documents must support structural composition[10]. This gives the author the ability to organize the document in nested components in order to divide its content in logical parts[14]. It helps him also in focusing on an isolated part of the document during each step of the authoring process (sub−scenes of a multimedia scenario).

Efficient and incremental manipulation of the document

The manipulation of the temporal structure of a multimedia document in an authoring environment must be supported by efficient algorithms and data structures in order to allow interactive design. Moreover, given the progressive nature of the editing process and the accesses across the network

to the document content [17], the temporal organization of a document must be suitable for an incremental building process.

We have seen in this section a set of requirements which serves as basis and guidelines for inserting time based information for multimedia documents. These highlights are given to open a discussion of how multimedia objects can be integrated in the document representation together with the temporal synchronization information. The main question is to know if the document behaviour is to be represented by the means of new formats or better to keep it behind programming languages like Java and ScriptX ?. The limit in now drawn in HTML by the OBJECT tag that opens the room for including any type of objects these documents.

3 Conclusion

We have discussed in this paper some important issues for temporal representation and synchronization of multimedia documents. We have designed an experimental tool that manages temporal information internally as an extended temporal constraint network [11]. The document representation used in MADEUS is mainly based on descriptive markup of media items and Allen's like temporal operators. We think that it is powerful enough and suitable for both multimedia authoring and presentation.

Among the characteristics of the MADEUS prototype, we gave a particular attention to the following aspects:

- the expressiveness of the temporal representation which captures the qualitative, quantitative, causal and indeterministic aspects of multimedia objects and presentations.
- a layered application architecture which allows the production of temporal layout, temporal navigation through hyperlinks and run–time monitoring of the presentation by a dynamic temporal formatter.

From the implementation experience of MADEUS, we came to the conclusion that the applying the same techniques to HTML documents may help in disseminating time–based document. Even if the limitations due to network bandwidth remain the main problem facing the wide spread of such documents, the design of new document formats including multimedia objects and time–based information is still lacking.

This work is currently undertaken as part of the Opera project. The goal of this project is to develop an editorial environment for the construction, manipulation and storage of complex multimedia documents.

Bibliography

[1] ALLEN J.F., Maintaining Knowledge about Temporal Intervals, *Communications of the ACM*, Vol. 26(No. 11), pp. 832–843, november 1983.

[2] BUCHANAN M.C. and ZELLWEGER P.T., Automatic Temporal Layout Mechanisms, *Proceedings of the First ACM International Conference on Multimedia*, pp. 341–350, Anaheim, California, august 1993.

[3] BULTERMAN D.C.A., VAN ROSSUM G. and VAN LIERE R., A Structure for Transportable, Dynamic Multimedia Documents, *USENIX Conference*, pp. 137–1559, Nashville, TN, june 1991.

[4] DECHTER R., MEIRI I. and PEARL J., Temporal Constraint networks, *Artificial Intelligence*, 49, pp. 61–95, 1991.

[5] DIRECTOR 4.0, User's Guide, Macromedia.

[6] GOLDFARB C. F., *The SGML Handbook*, Oxford University Press, Oxford, 1990.

[7] ISO/IEC 10743, *HyTime Information Technology Hypermedia/Time–based Structuring Language (HyTime)*, november 1992.

[8] ISO/IEC 14478–1, *Information Processing Systems –– Computer Graphics and Image Processing –– Presentation Environments for multimedia Objects (PREMO)*, (1), may 1996.

[9] ISO/IEC 13552–1, *Information Tehnology Coded Representation of Multimedia and Hypermedia Information Object (MHEG)* , (1), october 1994.

[10] KHALFALLAH H. and KARMOUCH A., An architecture and a data model for integrated multimedia documents and presentational applications, *Multimedia Systems*, pp. 238–250, Vol. 3, 1995.

[11] LAYAÏDA N. and SABRY–ISMAIL L., Maintaining Temporal Consistency of Multimedia Documents Using Constraint Networks, *Multimedia Computing and Networking 1996, M. Freeman, P. Jardetzky, H. M. Vin, ed., pp. 124–135, SPIE 2667*, San–José, USA, february 1996.

[12] LITTLE T.D.C. and GHAFOOR A, Interval–Based Conceptual Models for Time–Dependent Multimedia Data, *IEEE Transactions on Knowledge and Data Engineering (Special Issue: Multimedia Information Systems)*, Vol. 5(4), pp. 551–563, august 1993.

[13] PEREZ–LUQUE and LITTLE T.D.C., A Temporal Reference Framework for Multimedia Synchronization, *Proceedings of Sync'95*, may 19, Virginia, 1995.

[14] QUINT V. and VATTON I., Grif: An Interactive System for Structured Document Manipulation, *Text Processing and Document Manipulation, Proceedings of the International Conference*, pp. 200–213, Cambridge, 1986.

[15] VILAIN M. and KAUTZ H. A., Constraint Propagation Algorithms for Temporal Reasoning, *Proceedings of AAAI*, pp. 377–382, Philadelphia, august 1986.

[16] WAHL T. and ROTHERMEL K., Representing Time in multimedia, *Proceedings of the International Conference on Multimedia Computing and Systems*, pp. 538–543, Boston, 1994.

[17] WEITZMAN L. and WITTENBURG K., Automatic Presentation of Multimedia Documents Using Relational Grammars, *Proceedings of the Second ACM International Conference on Multimedia*, pp. 443–451, San Francisco, California, october 1994.

Issues in Temporal Representation of Multimedia Documents

1 Introduction 1
2 Representation of multimedia information 2
3 Conclusion 4
Bibliography 5