Navigation in Structured Multimedia Documents using Presentation Context

Loay Sabry-Ismail — Cécile Roisin — Nabil Layaïda

Opéra Project, INRIA Rhône-Alpes,
655 avenue de l'Europe, F-38330 Montbonnot Saint-Martin.
e-mail: {Loay.Sabry,Cecile.Roisin,Nabil.Layaida}@inrialpes.fr

ABSTRACT. Navigation in multimedia documents differs from the navigation in classical documents the fact that temporal information must be handled at each instant of the presentation. In Madeus, a multimedia authoring and presentation tool developed in our project, we have defined a presentation context which holds this information. In this paper, we will discuss how presentation context is defined, calculated and used for the temporal navigation. We introduce the notion of Tic and Tac instants which correspond to the most meaningful instants in the presentation of multimedia documents, and we describe how they can be used to navigate through documents.
KEY WORDS : multimedia document, temporal navigation, authoring.

1. Introduction

Madeus is a prototype of an interactive authoring environment for editing and presenting multimedia documents, taking into consideration the four dimensions of multimedia documents: logical, spatial, temporal and hyperlinking [5]. The main goal is to provide high level editing functionalities (declarative language with ad hoc user interface for incremental document specification process through easy skips from the editing state to the presentation state and vice versa), with an efficient coherence checker (qualitative, quantitative and indeterministic), powerful static and dynamic formatting operations, and, finally, an efficient execution support for presentation.

The presentation of multimedia documents differs from the presentation of static documents mainly because of the temporal dimension to be taken into account both internally (inherent temporal properties of media objects composing the document, such as video and audio) and externally (author's specified temporal relationships between the media objects). Therefore, the state of media objects in a multimedia document is continuously changing as the presentation progresses in time. For example, an active media object in the past can be inactive in the present. Also, for a continuous media object (such as video), the currently displayed frame changes as the time elapses. The presentation is an instant based process where at each instant of the presentation we have a current instant and a set of significant relative past and future instants.

One of the main requirements expressed by authors of multimedia documents is the ability to get a global view of their documents. This need is also encountered in hypertext systems and much work has been done in this area [2][3][7][8].

Another way of improving the users' perception of the presentation of multimedia documents, consists in providing temporal navigation facilities, such as fast forwarding and rewinding in audio or video and hyperlinking. In this paper, we identify temporal navigation needs for authors/readers of multimedia documents and we show how they are realized in Madeus. The key point of our approach is to take advantage of the internal representation of documents (hypergraph) to provide relevant ways of temporal navigation.

Navigation in multimedia documents is a navigation not only in space within document contents, but in time as well. To implement this kind of navigation, some temporal information about the state of the document at different instants of its presentation must be managed. In this paper, we are interested mainly in the definition and management of the required temporal information for navigating in multimedia documents.

The rest of the paper is organized into two main parts: the first one analyses navigation requirements for multimedia documents as they are handle in our system; basically, we identify new navigation functions taking into account the temporal dimension of documents. The second describes how temporal information is managed in Madeus system in order to implement navigation facilities.

2. Navigation requirements for multimedia documents

Multimedia documents in Madeus A multimedia document is composed of media objects which are temporally and spatially related and which can be hyperlinked. The temporal organization of the document is called the *scenario*. Media objects can be continuous (like video and audio), discrete (like text, images and graphics) and indeterministic (like a button of interaction on which a user can click). The set of temporal relations used to relate objects in Madeus [4] is composed of a set of quantified Allen operators [1] together with a set of causal operators (for example, the LIST-PARMIN operator expresses the interruption of a set of objects when one of the objects has finished, as in the Menu of Fig. 1).

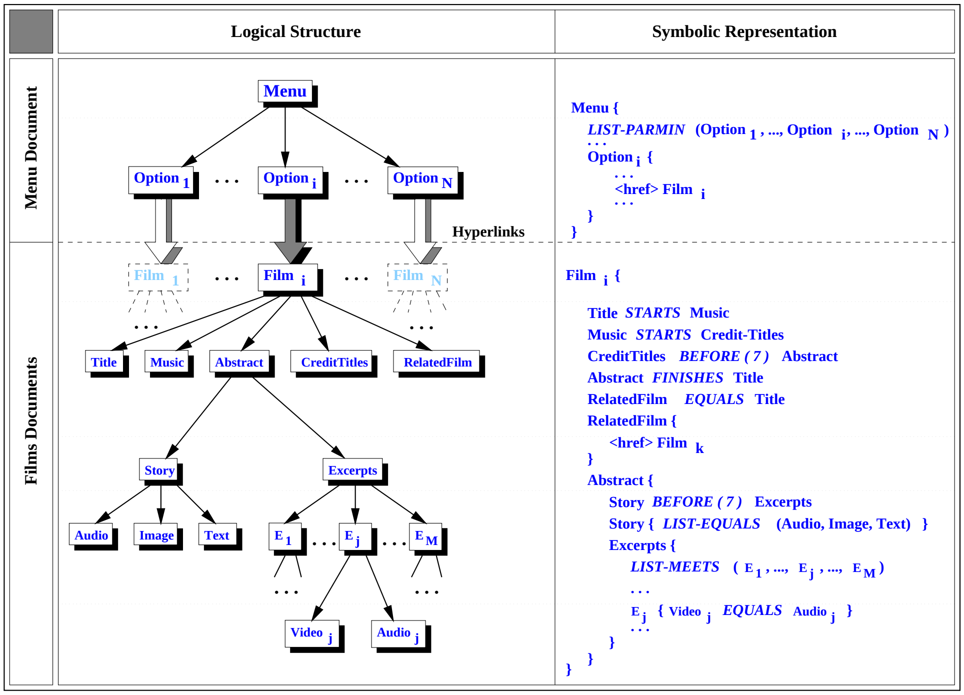

Figure 1. *Logical structure and symbolic representation of film library*

In Fig. 1, we can see an example of a multimedia document which is used hereafter as a working example: on the left side, the logical structure of the document is given, and on the right side, temporal relations are defined between the objects. This example represents a library of films excerpts where a film can be selected from a **Menu** of films **Options**. For each film in the menu, there is a corresponding scenario **Film$_\mathbf{i}$** in which the **Title**, the **Music**, the **Story**, the **RelatedFilm** button and a sequence of **Video/Audio** clips of the film are temporally related, as expressed in the symbolic representation (right side of Fig. 1). For example, the Title and the Music start at the same time (STARTS relation), the RelatedFilm button is active as long as the Title is presented (EQUALS relation). The RelatedFilm button allows the user to jump to another related film (of the same actors, for example). It is important to notice that temporal relations are defined among sibling objects; the hierarchical structure provides nesting for the definition of the temporal organization.

Navigation in Madeus Two classes of navigation can be identified:

- **Context-dependent navigation:** The navigation is defined by explicit objects declared in the document at some precise points. This includes

internal and external hyperlinks defined by the author. An internal link can be either an inclusion-link (a related scenario is included inside the current one when the link is activated) or a reference link (the execution jumps to a another temporal point when source link is activated, with the ability of coming back to the source point). In this paper, we will consider only this last class of links because these interactions require temporal management..
For instance, in the example of Fig. 1, while watching the sequence of Video/Audio clips of one film, the user can follow a hyperlink to another film related to the first film (for example, the same actress plays in it), watch some sequence of its Video/Audio clips and then come back to proceed with the first film.

- **Context-independent navigation:** It is independent of the subject presented by the document. This includes generic presentation manipulation operations like stop, resume, or fast forwarding and rewinding. The user can accelerate the speed of the scenario to get a global overview of the document. He may also want to run step by step from one significant instant to another, such as the beginning of each video/audio clip of the film excerpts.

In all these navigation operations, the system must be able to know the (temporal) state of every object. We describe in the next section the internal document representation and the management of time information in our system.

3. Temporal information management in Madeus

In this section we present the execution structure used in Madeus, then we introduce the notion of presentation instants called Tics and Tacs. Finally, we show how temporal information of the presentation is calculated at each needed instant.

3.1. *Execution structure*

The execution structure used in Madeus is a hierarchical structure of nested components in which leaves are basic media objects and nodes are composite objects. Each composite object is defined by an extended constraint network, a Directed Acyclic Graph (DAG), built from its enclosing components by a set of relational operators [1][5]. The resulting structure for a multimedia document is a hypergraph where the **Root** object is the highest level composite object in the hierarchy representing an abstraction for the whole document. We define also $\mathbf{S_{Root}}$, the set of basic and composite objects involved in the scenario of that document.

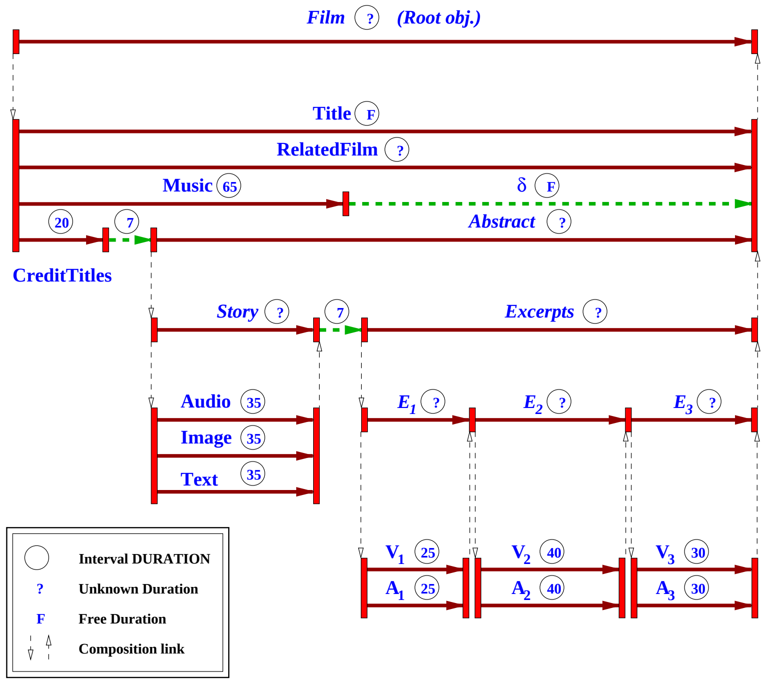

Figure 2. *Initial hypergraph of the film library*

This execution structure, detailed in [4], is illustrated in Fig. 2 for the film library example. We can see the initial hypergraph structure of the film excerpts scenario: the highest level composite object **Film** acts as the **Root** object and basic objects are the leaves of this hypergraph, such as **Music**, $\mathbf{V_1}$ and $\mathbf{A_1}$ objects. Only user specified durations are shown in Fig. 2 (such as **Music**, **Audio**, **Image**, $\mathbf{V_1}$ and $\mathbf{A_1}$). Duration of composite objects are marked to be unknown (such as Film, Abstract, Story, Excerpts, ... etc.). Some basic objects are marked to be free (as **Title** and the delay δ following **Music**), i.e. the static formatter will freely choose a value for their duration from an interval of possible values specified by the user.

Each object in Madeus is associated with a set of variables and methods for its management by the system. In the rest of this paper, we are interested in the following variables and methods (related to the temporal and structural information of objects):

- **State**: such as active, suspended and terminated.
- **Nature**: BASIC or COMPOSITE.
- **StartTime**: instant at which the object starts playing.
- **EndTime**: instant at which the object ends playing.
- **Expected duration**.

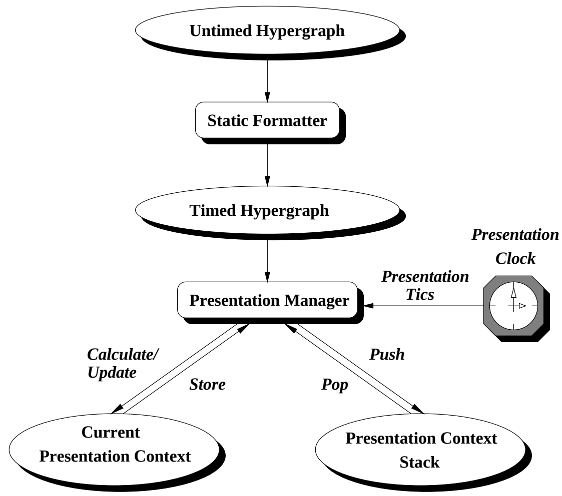

Figure 3. *Time management in Madeus*

3.2. *Architecture of Madeus*

As declared previously, one of the main goals of our system is to provide the user with a reliable editing system. Therefore, we try to verify statically (and incrementally for efficiency purpose) the coherence of each editing operation performed by the user.

In Fig. 3, we can see two of the main components (the *static formatter* and the *presentation manager*) of Madeus system and how temporal information is managed by them. Next sections are devoted to the role of these components in the time management process.

3.3. *Time management*

To handle navigation in multimedia documents, we need to define a *presentation context* (its precise definition is given in section 3.4.1) that holds the temporal information about a presentation at a certain time instant. Before we proceed, it is useful to introduce the notion of presentation instants referred to as presentation Tics and Tacs in the sequel.

Definition of presentation Tics: A Tic is a time instant in the lifetime of the presentation of a scenario at which a set of presentation actions must be carried out. So the presentation of a scenario can be represented as a sequence of Tics (as clock tics), such as the Tic at t= 100 in Fig. 4. Each two consequent Tics are separated by a *time step* that is the minimal interval needed for the presentation of the objects (for example 1/30 sec or 1/25 sec is required for playing video frames).

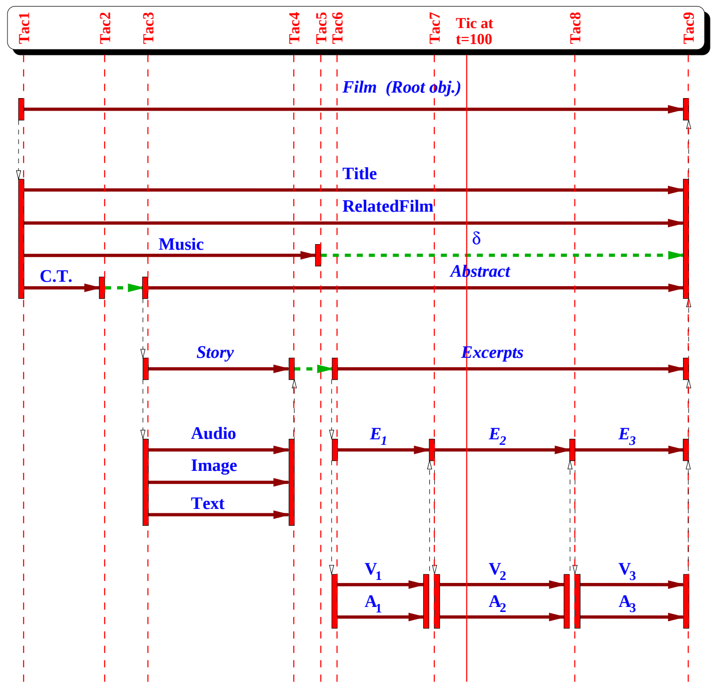

Figure 4. *Tacs of the films library scenario*

Definition of presentation Tacs: Tacs are the Tics corresponding to the time instants at which basic (and composite) objects start and finish. Tacs of the film library are presented in Fig. 4 (from Tac1 to Tac9): the hypergraph representing the document is augmented with vertical lines, each of them being set at the beginning or the end of an object. For example, at Tac1 instant, **Film** and **Music** objects start, while at Tac7, excerpt **E1** (with **V1** and **A1**) finishes and excerpt **E2** (with V2 and **A2**) starts.

Navigation requires some manipulation actions on the presentation context, such as:

- **Updating:** The current presentation context of a playing document must be frequently *updated* during the lifetime of the presentation. Updates of presentation context are done at every Tic instant of the presentation.
- **Storing:** When activating a hyperlink, the presentation context at the activation instant must be *stored*. This stored context will be used for resuming the original presentation instant when returning back from the destination link. We use a stack to hold the stored presentation contexts, so that the last stored context is the first to be resumed.
- **Calculating:** Each time we need to jump across the scenario towards a target instant in the past or the future, the presentation context at this target instant (target presentation Tic) must be *calculated*.

Object	Offset
Title	**100**
RelatedFilm	**100**
δ	**35**
V_2	**6**
A_2	**6**

Figure 5. *Temporal cut information for the film library scenario at Tic instant 100*

The method used for calculating the presentation context at a certain instant is presented in section 3.4.

3.4. *Temporal cut*

3.4.1. *Definition*

A temporal cut can be considered as a snapshot of the presentation of the multimedia document at a specific instant. This instant can be a past, present or future instant relative to the current instant of presentation. If we consider a temporal cut at a time instant t for example, we are interested in knowing which objects are concurrently *active* at that precise instant of time and what point in their presentation lifetimes they have reached (an offset value is calculated as the difference between time instant t and the object's start time) as shown in Fig. 5. This temporal cut information is what we have called until now the current *presentation context.*

So, the temporal cut of scenario $\mathbf{S_{Root}}$ at instant t for a document identified by its **Root** object is defined as the set of ordered pairs of all *active* basic objects together with their associated offset values. It is defined by the function $\mathbf{k}_t$ as follows:

$\mathbf{k_t}$(Root) = {(Object, Offset (Object)) / (Object $\in$ $\mathbf{S_{Root}}$)
and (Nature (Object) = BASIC)
and (StartTime (Object) $\leq$ t $<$ EndTime (Object))
and (Offset (Object) = t - StartTime (Object)) }

For example in Fig. 4, we can see that the Tic instant 100 crosses the scenario in five *active* basic objects (Title, RelatedFilm, delay δ, V2 and A2). The temporal cut information, corresponding to this instant, is shown in the table in Fig. 5. In this table we have an entry for each *active* object with its corresponding offset time.

3.4.2. *Calculation of k_t*

Both the static formatter and the presentation manager take part in calculating temporal cut information. For evaluating k_t, we need:

- The structural information of the objects, as given by the hypergraph.
- The temporal information about dates of start and end for each object (*StartTime* and *EndTime* functions); this information is calculated by the static formatter, as described in section 3.4.3.
- The object state information, i.e. if it is *active* or not, and its nature: *basic* or *composite*.

Based on this information, elapsed period of presentation of objects at the cut instant can be calculated dynamically by the presentation manager, as described in section 3.4.4.

3.4.3. *Static calculation of dates*

Calculation of dates of instants is a part of the document static formatting phase, in which the Madeus system computes (for a consistent scenario) values for:

1. the **durations** of *basic* objects,
2. the **durations** of *composite* objects and
3. the **dates** of the *start* and *end* instants of the objects.

Duration of basic objects The duration values assigned to free *basic* media objects, are either calculated values (from author specified ranges of duration) for discrete or continuous objects or arbitrary values for indeterministic objects (such as user interaction buttons). These duration values are calculated taking into account that each object must have a duration value as close as possible to its preferred one [6].

Dates and duration of objects Initially, the starting instant of the **Root** object of the hypergraph is assigned a date value equals to zero, and dates of all the starting/ending instants in the hypergraph are calculated relative to this instant.

The hypergraph is visited starting from the **Root** start instant in a depth-first order, with respect to the global temporal order of starting/ending instants of objects. Each time we meet the starting instant of a composite object, we go down one level in the graph hierarchy to calculate the date of its ending instant. The date of the starting instant of the first child in the subgraph is inherited from the starting date of its parent (as shown by the descending arrows in Fig. 6). For example, the starting instant of **Story** (27) is inherited from the starting instant of its parent (**Abstract**). When we reach the lowest levels of the graph hierarchy (i.e. those which contain only basic intervals), we calculate the dates for starting and ending instants at these levels by summing

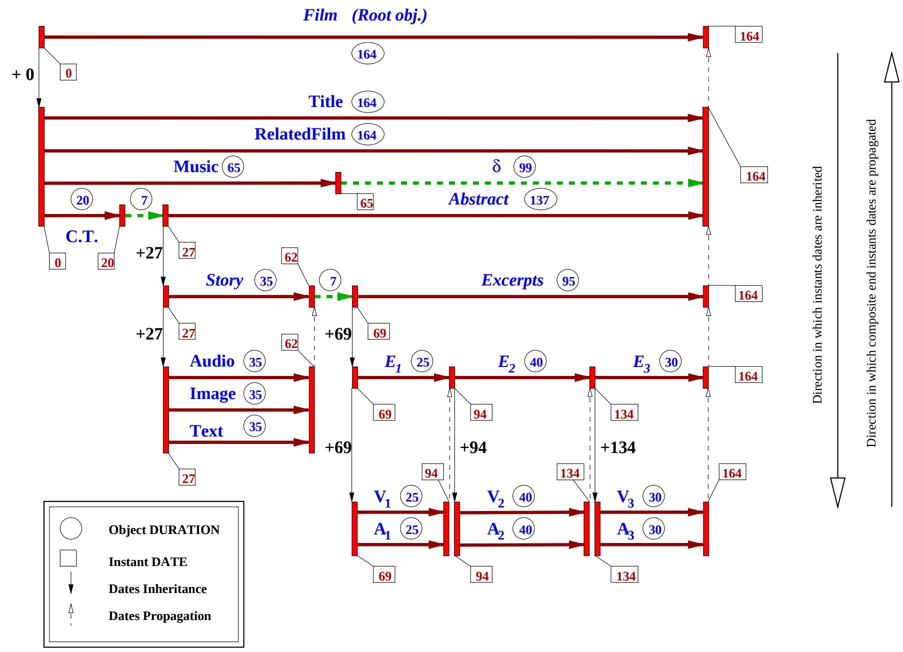

Figure 6. *Timed hypergraph*

the duration of sequential objects. We can then return one level upwards in the hierarchy towards the ending instant of the enclosing object for assigning the duration and the date of the ending instant of this composite object.

For instance in Fig. 6, objects **Abstract**, **Story**, **Excerpts**, **E1**, **E2** and **E3** are composite objects and they obtain their duration values (137, 35, 95, 25, 40 and 30, respectively) from the subgraphs in the hierarchical levels just below them, and they are equal to the difference between the calculated dates of their *end* and *start* instants.

Now the graph with its information (calculated dates for starting/ending instants) is ready to be used dynamically by the presentation manager for calculating presentation contexts.

3.4.4. *Dynamic manipulation of the presentation context*

When Madeus system needs to calculate a presentation context at an instant in the past or the future during the presentation, the presentation manager (PM) calculates dynamically the required context. The PM starts from the higher level graph in the hierarchy and gets the set of objects which satisfy the condition stated in section 3.4.1, i.e. PM applies the $\mathbf{k}_t$ function. The $\mathbf{k}_t$ function is recursively applied for each active composite object until the lowest levels are reached (basic objects) in order to cover the whole hypergraph.

The PM also performs the operations stated in section 3.3: updating the current presentation context at each Tic instant, storing the context when a link is activated and managing stack operations (pushing and popping of presentation contexts).

4. Temporal cut for navigating in multimedia documents

Navigation in a multimedia document using Tacs can have several forms depending on the choice of which Tac points to be used as time markers. Three classes of navigation using Tacs have been stated:

Step by step navigation The user can ask to go from one instant to the nearest Tac instant (whether in the past or in the future); in this case, the user can rapidly navigate to-and-fro without the need for waiting for the document to advance with its normal play speed to reach the desired instant. In doing so, navigation will be controlled by the set of Tacs calculated and ordered automatically by Madeus system.

Structural navigation Structural navigation is defined as the navigation between objects located at the same hierarchical level of the hypergraph. Each level of the hypergraph has its corresponding set of Tacs that can be used for structural navigation. For example in Fig. 4, the set of Tacs Tac6, Tac7 and Tac8 is used in the structural navigation in the hierarchical level containing the excerpt objects $\mathbf{E_1}$, $\mathbf{E_2}$ and $\mathbf{E_3}$.

User defined navigation An author can select any group of Tac points to be used as functional Tacs for navigation. This set of Tacs can be, for example, these corresponding to the start instants of the **Story** object of each film in the film library.

5. Conclusion

In this paper, we set out the notion of presentation context for multimedia documents, holding the temporal information of the presentation at a certain presentation instant. We have introduced how to identify the most meaningful presentation instants, named Tics and Tacs. Then we have shown how to compute presentation contexts at Tics and Tacs instants to be used for temporal navigation purposes. These ideas are implemented in Madeus and they have shown satisfactory results with discrete objects; now implementation is done for continuous and indeterministic objects.

The documents handled by Madeus can be heavily dynamic and interactive. A scenario can be seen as a specification having numerous execution instances thanks to indeterministic objects (such as user interaction buttons to go to next film). It is necessary to fix values to these indeterministic durations when implementing context-independent navigation operations (minimal, maximal or mean values). Therefore, a step by step navigation will produce a global view of a specific execution instance of the document. It could be interesting to provide the user with tuned navigation facilities, allowing him to get several execution instances.

6. Bibliography

[ALL 83] ALLEN J., "Maintaining Knowledge about Temporal Intervals", *Communications of the ACM,* vol. vol.26, num. no.11, pp. 832-843, November 1983.

[BIE 94] BIEBER M. and WAN J., "Backtracking in a Multiple-window Hypertext Environment", *Proceedings of the ACM European Conference on Hypermedia Technology ECHT94,* pp. 158-166, ACM Press, Edinburgh, September 1994.

[HAR 93] HARDMAN L., BULTERMAN D. and ROSSUM G., "Links in Hypermedia: the Requirement for Context", *Hypertext '93,* pp. 183-191, November 1993.

[JOU 97] JOURDAN M., LAYAIDA N. and SABRY-ISMAIL L., "Time representation and management in MADEUS: an authoring environment for multimedia documents", *Multimedia Computing and Networking 1997,* vol. Martin Freeman, Paul Jardetzky, Harrick M. Vin, Editors, Proc. SPIE 3020, , pp. 68-79, California, USA, February 1997.

[LAY 96] LAYAIDA N. and SABRY-ISMAIL L., "Maintaining Temporal Consistency of Multimedia Documents Using Constraint Networks", *Multimedia Computing and Networking 1996,* vol. Martin Freeman, Paul Jardetzky, Harrick M. Vin, Editors, Proc. SPIE 2667, , pp. 124-135, California, USA, January 1996.

[LAY 97] LAYAIDA N, *Madeus : un système d'édition et de présentation de documents structurés multimédia,* Ph.D. Thesis, Université Joseph Fourier, Grenoble, France, June 1997 (to appear).

[MAS 94] MASUDA Y., ISHITOBI Y. and UEDA M., "Frame-Axis Model fro Automatic Information Organizing and Spatial Navigation", *Proceedings of the ACM European Conference on Hypermedia Technology ECHT94,* pp. 146-157, ACM Press, Edinburgh, September 1994.

[MUK 94] MUKHERJEA S., FOLEY J. and HUDSON S., "Interactive Clustering for Navigating in Hypermedia Systems", *Proceedings of the ACM European Conference on Hypermedia Technology ECHT94,* pp. 136-145, ACM Press, Edinburgh, September 1994.

[YU 97] YU J. and XIANG Y., "Hypermedia Presentation and Authoring System", *Sixth International World Wide Web Conference,* Hyper Proceedings, URL: http://www6.nttlabs.com/HyperNews/get/PAPER91.html, California, USA, April 1997.