

Component-based Hypervideo Model: High-Level Operational Specification of Hypervideos

Madjid Sadallah
DTISI - CERIST
Alger, Algérie
msadallah@mail.cerist.dz

Olivier Aubert
Université de Lyon, CNRS
Université Lyon 1, LIRIS,
UMR5205, France
olivier.aubert@liris.cnrs.fr

Yannick Prié
Université de Lyon, CNRS
Université Lyon 1, LIRIS,
UMR5205, France
yannick.prie@liris.cnrs.fr

ABSTRACT

Hypervideo offers enhanced video-centric experiences. Usually defined from a hypermedia perspective, the lack of a dedicated specification hampers hypervideo domain and concepts from being broadly investigated. This article proposes a specialized hypervideo model that addresses hypervideo specificities.

Following the principles of component-based modeling and annotation-driven content abstracting, the *Component-based Hypervideo Model* (CHM) that we propose is a high level representation of hypervideos that intends to provide a general and dedicated hypervideo data model.

Considered as a video-centric interactive document, the CHM hypervideo presentation and interaction features are expressed through a high level operational specification. Our annotation-driven approach promotes a clear separation of data from video content and document visualizations. The model serves as a basis for a Web-oriented implementation that provides a declarative syntax and accompanying tools for hypervideo document design in a Web standards-compliant manner.

Categories and Subject Descriptors

H.5.1 [Multimedia Information Systems]: Video; H.5.4 [Hypertext/Hypermedia]: Architectures, Navigation

General Terms

Design, Experimentation

Keywords

Hypervideo, Annotation, component-based modeling, Time and Synchronization, Timeline Reference, CHM

1. INTRODUCTION

Video-based information becomes today an invasive social phenomenon that involves a wide and growing range

of digital consumers. Hypervideo is an attractive technology that aims to bring new interaction modalities to video-centric hypermedia documents. Currently, several definitions of hypervideo exist [3], depending on the considered point of view. While some works emphasize the hypermedia aspect of adding information to digital video so that users can activate video hyperlinks and access the added rich content, other authors highlight the storytelling aspects of dynamically creating non-linear and user-defined navigation paths into the document. In this work, we define hypervideo as being *an interactive video-centric hypermedia document that includes an audiovisual content - a set of video objects - augmented with several kinds of data in a time synchronized way; the added content supplements the audiovisual part around which the presentation is organized in space and time.* The integration of content-enriched video offers additional interaction and navigation alternatives and additional information levels.

While initial hypervideo concepts can be traced back to the late of nineties [11], the availability of appropriate hardware and software has prevented this kind of documents from being broadly investigated. The need for theoretical foundations for hypervideos motivates the present work. The proposed component-based approach is a means to conceptualize annotation-driven hypervideos structure and behavior. High level components are provided to emphasize the presentation and interaction features while lower level ones allow more complex component building.

2. HISTORY AND RELATED WORK

Following Ted Nelson's hypermedia model extension to include "branching movies" or "hyperfilms" many researchers addressed the field of interactive video. The Aspen Movie Map Project or HyperCafe [11] added interactivity and branching to videos, using specialized software. Basic forms of web-based hypervideos can be achieved for instance through Youtube annotations, but many projects like *VideoClix*¹ and *Popcorn.js*² experiment with other approaches.

2.1 Hypervideo Modeling

Hypervideos as General Hypermedia/Multimedia. Models and systems for continuous media integration in hypermedia were discussed since the Amsterdam Hypermedia Model (AHM) [7] proposal, providing mechanisms for structuring,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

DocEng'11, September 19–22, 2011, Mountain View, California, USA.
Copyright 2011 ACM 978-1-4503-0863-2/11/07 ...\$10.00.

¹<http://www.videoclix.tv/>

²<http://popcornjs.org/>

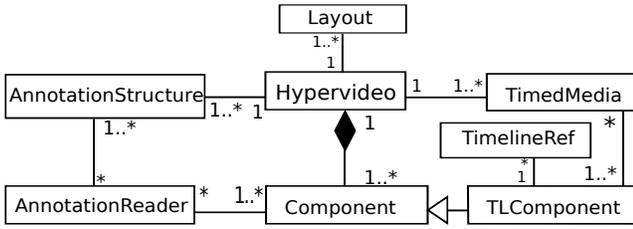


Figure 1: CHM General Overview

linking and synchronizing dynamic multimedia content. Other efforts like NCM (Nested Context Model [10]) also aimed at introducing synchronization in hypermedia.

AHM inspired the definition of SMIL (Synchronized Multimedia Integration Language) which is an ideal candidate for developing general purpose hypervideos. However, hypervideo support has not received much attention [12]. Moreover, SMIL is a huge standard and the specification itself contains no formal model definition [8]. Its concepts are all-purpose and do not precisely describe the hypervideo specific features, even though some proposals [10] have been made to address this issue.

Hypervideo as Annotation-based Multimedia. Video not being intrinsically structured, annotations are needed to add semantic description and content enrichment. They provide the foundation for enriched experience of video content, for instance by breaking the linearity of video [9] or enriching it [4]. We consider here an annotation as a piece of data associated to a spatio-temporal fragment of a video [1].

VideoAnnEx³ and MediaDiver [9] allow to annotate videos and generate video centric presentations. Hyper-Hitchcock [5] and Advene [1] allow viewers to not only interact with annotated videos but also to create their own annotations and to share them.

The Need for Dedicated Models. Informal models are generally used by existing systems to describe hypervideos. However, the implied representations are mainly technically driven and do not proceed from the definition of hypervideo data models besides the conventional hypermedia/multimedia ones. For instance, hypervideo raises many important rhetorical and aesthetic issues [11, 3] and stresses common hypermedia authoring and reading concerns that risk to overstrain the cognitive capacities of the user and put him under time pressure during navigation [12], provoking a user disorientation [3]. Moreover, hypermedia modeling of such documents results in an inefficiency and a weakness in expressing the logical hypervideo patterns [6] since the underlying concepts are too general to grasp and characterize hypervideo details. Consequently, hypervideo specificities need a dedicated document model able to consider these issues.

3. THE CHM PROPOSAL

3.1 General Overview

The proposed *Component-based Hypervideo Model* (CHM) is based on the presentation of annotations through nested low- or high-level components.

³<http://www.research.ibm.com/VideoAnnEx/>

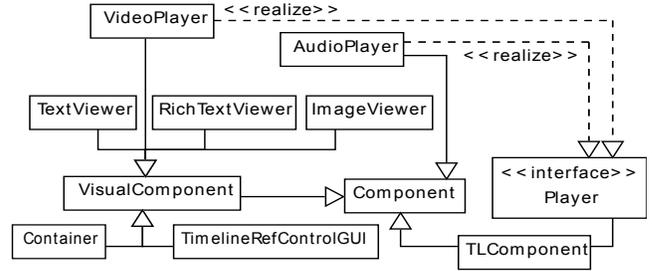


Figure 2: CHM Plain Components

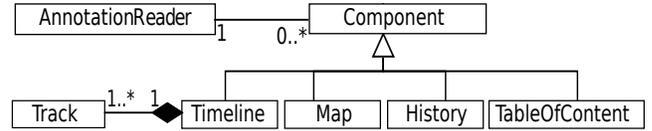


Figure 3: CHM High Level Components

Figure 1 presents a general overview of the model. A hypervideo is formed by a set of low and high level components, building blocks that represent the formal information and composition units. A hypervideo references at least one main audiovisual document accessed through the **TimedMedia** element that addresses at least a temporalized stream, audio or video. A **TimedMedia** played through a player component conveys a timing reference to the document, expressed by the *TimeLine Reference* (TLR) that synchronizes the rendering **TLComponents** related to the played **TimedMedia** element. Many players (therefore, many TLRs) may be present within the same document, defining different hypervideo sub-documents.

The CHM *annotation* concept expresses the association of any data or resource to a video fragment for enrichment, linking or visualization purposes. An annotation is defined as any information associated to a spatio-temporal video *fragment*. The annotation data may be generated by third party software such as Advene [1] or IBM’s VideoAnnEx annotation tool.

CHM Plain Components. Figure 2 presents the basic data components that form a hypervideo. Specific synchronized *display components* offer interactive interfaces for rendering temporalized data, provided as annotations through *AnnotationReader* components.

Document content viewers such as **TextViewer**, **RichTextViewer** and **ImageViewer** allow the display of textual and graphic content, such as a synchronized *Wikipedia* content. The **TimelineRefControlGUI** element defines a graphic user interface for controlling and interacting with TLRs.

CHM High Level Components. For hypervideo document design, we propose an extensible set of high level components shown in figure 3, built upon the plain ones. The *History component* presents a visual representation of the visited nodes, to alleviate user disorientation. The *Document maps and ToCs* (tables of contents) display structuring annotations through interactive graphical or hierarchical layouts. The **Timeline** component is a graphical interactive

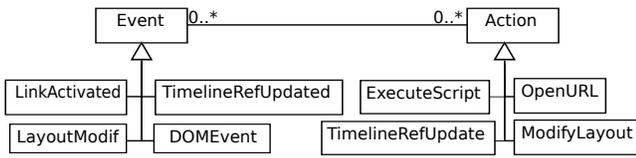


Figure 4: Hypervideo events and actions

representation of annotations through hypervideo time. It displays elements along a timed axis on different tracks.

CHM Spatial Model. Derived from the SMIL spatial model, the CHM spatial model is intended to accommodate the implementation platform specificities. **VisualComponents** are placed within **SpatialRegions**, whose global placement is defined in the root **Layout** element.

CHM Temporal Model. The document temporal specification is achieved through a timeline-based model. The explicit time scale of document components is defined by a *Timeline Reference* (TLR), a virtual time reference attached to a video playback component or to the global document. Time-based components are activated/deactivated when reaching specific timecodes provided or computed by reference to the TLR. The access and control of a TLR is performed through its “state”, “position” and “duration” attributes. Any update of the TLR position or state affects all the related components playback.

CHM Event-based Model. In addition to standard hypermedia links, CHM hypervideo links can be defined in space and time and are unidirectional. There is no separate link-like in AHM - within the model; SMIL and HTML also do not use separate link components.

Differently from AHM, CHM does not rely only on a link-based model. The dynamic behavior of a CHM hypervideo is represented by an event-based mechanism, expressed by the **Event** and **Action** elements shown in figure 4. Many actions can be associated to an event, among which: (1) the **OpenURL** action allows the display of a target URL, (2) the **ModifyLayout** action allows to modify the content and placement of elements, (3) the **UpdateTimelineRef** action specifies a state or position change of the timeline reference and (4) the **ExecuteScript** action allows the execution of a user specified set of operations.

3.2 Example

The proposed model is intended to both conceptually analyze common hypervideo documents, and help the design of new ones. We present here a CHM hypervideo document, represented in figure 5 - annotated using Advene⁴ and designed through our model implementation⁵ -, to give deep insight into the Nosferatu movie⁶. The hypervideo demonstrates some high level components such as data readers, enrichment content viewers, video players, hotspots, timelines, maps and tables of contents. Figure 6 is the abstract representation of the most important components. The example is simplified by defining one spatial region and a unique time-

⁴<http://www.advene.org/>

⁵<http://www.advene.org/chm/API/>

⁶<http://www.archive.org/details/nosferatu>

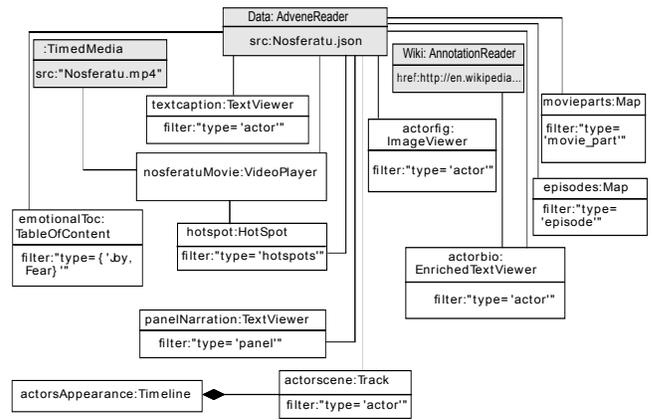


Figure 6: Nosferatu Hypervideo Modeling

line reference, both not represented in the figure. Each class instance is provided with a set of attributes and we present in the figure only the filter used to query the data readers.

4. ARCHITECTURE IMPLEMENTATION

4.1 Rationale

Different programming languages and for various architectures can be used to implement CHM. In order to demonstrate some practical uses of the model, we have developed a publicly accessible Web-based prototype. It relies on a declarative syntax and the development of a set of JavaScript libraries for the model implementation in a Web-standards compliant manner. For rendering the audiovisual content, we make use of the latest HTML5 standard proposal that introduces native browser support for video. The CHM spatial model is supported by the HTML layout model via cascading stylesheets. As there is still no established standard way to add temporal behavior to Web documents, we have used a javascript implementation of SMIL Timesheets [2].

4.2 Web-based Hypervideo Syntax

Based on CHM, a syntax for authoring hypervideos is proposed as an extension above the HTML language. Since HTML does not fully support namespace declarations, we chose to use CHM namespaced attributes associated to standard HTML elements. A javascript-based transformation interprets these attributes and dynamically generates standard HTML5 code. Complex hypervideos can therefore be authored as standard Web documents, styled with CSS, and controlled by scripts. Common Web content is written in standard HTML while hypervideo components are expressed through the CHM attribute-based syntax.

Interested readers can see the code that generates the hypervideo example of figure 5, written in the proposed syntax, by looking at the source of the <http://www.advene.org/chm/> document.

4.3 Current State of the Implementation

A first version of the proposed language and tools has been developed, and further developments are underway. We aim to enhance the language support by providing the language DOM API rather than libraries that perform content transformation. The authoring process may be eased by a graphic



Figure 5: Nosferatu Hypervideo Example

user interface (GUI), planned in future development stages. We also want to rely on cognitive theories to propose more meaningful components. The interaction possibilities can be enriched by more advanced features like offering additional mechanisms to end-users.

5. CONCLUSION

In this paper, we have proposed the *Component-based Hypervideo Model* (CHM), an annotation-based hypervideo document abstraction with a focus on a high level architectural representation. It tries to provide a hypervideo model, building upon conventional hypermedia models. A component-based approach is proposed to apprehend hypervideos and to define their behaviors. The model components allow to use explicit elements to analyse or design hypervideos, providing a high level framework while still offering lower level tools for hypervideo authoring. The model promotes a clear separation of data from the video content, through the notion of annotations. A declarative syntax for defining Web-based hypervideo documents has been proposed and a first implementation has allowed us to validate the principles. We are working on extending and refining it.

Acknowledgements. This work has been partially funded by the French FUI (Fonds Unique Interministériel) - CineCast project and the ANR (Agence Nationale de la Recherche) - Ithaca project. It is supported by the Algerian Research Center on Scientific and Technical Information (CERIST).

6. REFERENCES

- [1] O. Aubert and Y. Prié. Advene: active reading through hypervideo. In *ACM Hypertext*, 2005.
- [2] F. Cazenave, V. Quint, and C. Roisin. Timesheets.js: When SMIL meets HTML5 and CSS3. In *to appear in DocEng2011*, 2011.
- [3] T. Chambel, C. Zahn, and M. Finke. Hypervideo design and support for contextualized learning. In *IEEE Advanced Learning Technologies*, 2004.
- [4] R. Fagá, V. G. Motti, R. G. Cattelan, C. A. C. Teixeira, and M. da Graça Campos Pimentel. A social approach to authoring media annotations. In *ACM DocEng 2010*, 2010.
- [5] A. Girgensohn, L. Wilcox, F. Shipman, and S. Bly. Designing affordances for the navigation of detail-on-demand hypervideo. In *Proceedings of ACM Advanced Visual Interfaces*, May 2004.
- [6] R. I. Hammoud. Introduction to interactive video. In *Interactive Video*, pages 3–24. Springer-Verlag, 2006.
- [7] L. Hardman, D. C. A. Bulterman, and G. van Rossum. The Amsterdam Hypermedia Model: adding time and context to the Dexter model. *Commun. ACM*, 37:50–62, February 1994.
- [8] J. Jansen, P. Cesar, and D. C. Bulterman. A model for editing operations on active temporal multimedia documents. In *Proc. of DocEng 2010*, 2010.
- [9] G. Miller, S. Fels, A. A. Hajri, M. Ilich, Z. Foley-Fisher, M. Fernandez, and D. Jang. Mediadiver: Viewing and annotating multi-view video. In *CHI-Interactivity Session*, May 2011.
- [10] L. M. Rodrigues, M. J. Antonacci, R. F. Rodrigues, D. C. Muchaluat-Saade, and L. F. G. Soares. Improving SMIL with NCM facilities. *Multimedia Tools Appl.*, 16:29–54, January 2002.
- [11] N. Sawhney, D. Balcom, and I. E. Smith. Hypercafe: Narrative and aesthetic properties of hypervideo. In *UK Conf. on Hypertext*, Bethesda, US, 1996.
- [12] C. A. Tiellet, A. G. Pereira, E. B. Reategui, J. V. Lima, and T. Chambel. Design and evaluation of a hypervideo environment to support veterinary surgery learning. In *Proc. of ACM Hypertext*, 2010.