In Proc. IEEE Int. Conference on Multimedia Computing and Systems (ICMCS) '99, June 7–11, 1999, Florence, Italy.

Multimedia Document Models – Sealed Fate or Setting Out for New Shores?

Susanne Boll, Wolfgang Klas, Utz Westermann
Databases and Information Systems (DBIS),
Computer Science Department, University of Ulm, Germany
{boll, klas, westermann}@informatik.uni-ulm.de

Abstract

Existing multimedia document models like HTML, MHEG, SMIL, and HyTime lack appropriate modeling primitives to fit the needs of next generation multimedia applications which bring up requirements like reusability of multimedia content in different presentations and contexts, and adaptation to user preferences. In this paper, we motivate and present new requirements stemming from advanced multimedia applications and the resulting consequences for multimedia document models. Along these requirements, we discuss HTML, HyTime, MHEG, SMIL, and ZYX, a new model that has been developed with special focus on reusability and adaptation. The analysis and comparison of the models show the limitations of existing models, point the way to the need for new flexible multimedia document models, and throw light on the many implications on authoring systems, multimedia content management, and presentation.

1. Introduction

The initial requirements to multimedia documents were the modeling of the temporal and spatial course of a multimedia presentation and also the modeling of interaction. To offer suitable support for multimedia applications, standardization activities started and commercial tools were evolving. The development and the passing of standards took quite a long time, while in the meantime very sophisticated commercial multimedia authoring tools came to the market that support their own proprietary format and only by now start to cross the bridge to standards.

However, we think that both the standards and the commercial tools developed so far only partially offer the necessary prerequisites for multimedia document modeling as next generation multimedia applications extend the initial requirements by far: demand for *reusability* of the media including entire documents and parts of documents, modeling of *adaptation* to user specific needs and context, and *wide-spread use* in the Internet.

Why do we consider these to be the new requirements to multimedia documents? As authoring of multimedia information is a very time consuming and costly task the reuse of material is definitely of high interest simply from an economical point of view. But reuse by means of "cut and paste" obviously can not be a solution, rather distinct and fine-grained reuse of multimedia content is highly demanded. Personalization and adaptation of information systems to personal needs and personal interests become more and more important (e.g., [4]). The trend to offer a user the most suitable and narrowed down multimedia information can be seen in research prototypes from different research areas, e.g., an adaptive tutorial agent [22], adaptive textbooks on the WWW [7], personalized news paper [14], personalized delivery of news [16], etc. Personalization and adaptation in consequence call for the enhancement of multimedia content with metadata to allow for the targeted context-specific selection of multimedia content. Another new requirement to a document model is its Internet-applicability, i.e., how it can cope with the demands of the heterogeneous environment of the Internet.

Within our project "Gallery of Cardiac Surgery" [15] (Cardio-OP)[1], that aims at the development of an Internet-based and database-driven multimedia information system in the domain of cardiac surgery, we find a representative application that explicitly requires a model for multimedia material which can be extensively reused in different contexts. Based on a multimedia content repository, the system is going to serve as a common information and education base for its different types of users, physicians, medical lecturers, students, and patients, who are provided with multimedia information according to their user specific requests, their different understanding of the selected subject, their location and technical infrastructure. For example,

[1] Cardio-OP - Gallery of Cardiac Surgery - is partially funded by the German Ministry of Research and Education, grant number 08C58456. Our project partners are the University Hospital of Ulm, Dept. of Cardiac Surgery and Dept. of Cardiology, the University Hospital of Heidelberg, Dept. of Cardiac Surgery, an associated Rehabilitation Hospital, the publishers Barth-Verlag and dpunkt-Verlag, Heidelberg, FAW Ulm, and EN-TEC GmbH, St. Augustin. For details see also URL www.informatik.uni-ulm.de/dbis/Cardio-OP/

a high quality multimedia presentation, dynamically composed during a lecture at the university campus, should be available for students at home for revise although they do not need the high quality of videos or images. Therefore, either the multimedia material must be delivered to the student with a lower quality and lower bitrate or high data volume parts like video are replaced by "comparable" but less voluminous parts like a slide show. To achieve this kind of functionality, a suitable multimedia document model that allows for flexible and context-dependent reuse of a multimedia document and parts of it is needed.

In this paper[2], we identify and present the requirements for suitable multimedia document models imposed by upcoming advanced applications like Cardio-OP. We then introduce the reader to the standard multimedia document models HTML, HyTime, MHEG, SMIL, and to ZYX [1], a model that has been developed by our group with advanced applications in mind. Along the requirements, we analyze and compare these document models. We expose the limitations and drawbacks we find with the existing standards and show how these can be overcome with new document models like ZYX. Concluding, we give an idea of the implications of this new kind of document models which call for a concerted action on developing next generation authoring and management tools for advanced multimedia applications.

2. Requirements to Next Generation Multimedia Document Models

In this section, we identify requirements to multimedia document models. These can be divided into the traditional requirements *temporal model*, *spatial model*, and modeling of *interaction*, which we consider to be imperative for any multimedia document model, and the advanced requirements *reusability*, *adaptation*, and *presentation-neutral* representation, which we expect to be demanded more and more by future multimedia applications. The requirements constitute a metric along which selected multimedia document models are analyzed in Section 4.

Temporal model. A temporal model (see also [17, 24, 5, 8]) describes temporal dependencies between the media elements of a multimedia document. One can find four types of temporal models: *point-based* temporal models, *interval-based* temporal models, and *event-based* temporal models. Another way to specify temporal relations between media elements is by the use of *scripts* – programs written in a scripting language which can comprise temporal synchronization operations.

[2] An extended version of this paper can be found in [2].

Spatial Model. Three approaches of positioning the visual elements on the presentation medium can be distinguished: *absolute positioning* based on a coordinate system, *directional relations* [19], using relations like *strong-north* and *weak-north* (to specify overlapping), and *topological relations* [6] using relations like *disjoint*, *meet*, and *overlap*.

Interaction. For the modeling of user interaction, one can distinguish two basic types of interaction: *navigational interactions* providing for control of the flow of a presentation and *design interactions* manipulating the visible and audible layout of a presentation.

Reusability. Reusability of document content can be characterized along three dimensions: *granularity* of reuse, *kind* of reuse, and support for *identification* of reusable components. The granularity of reuse determines *what* can be reused. Regarding multimedia document models, we can distinguish at least three levels of granularity of reusable components: reuse of complete multimedia documents, reuse of fragments of multimedia documents like single scenes or chapters, and reuse of individual atomic media elements such as a video or audio.

For all three levels of granularity we distinguish the different kinds of reuse, i.e., *how* material can be reused for the composition of new documents. Identical reusage provides for reuse of components with all their original temporal, spatial, and interaction relationships and constraints, whereas structural reusage separates the layout from the structure of components and reuses only the structural parts.

Before we can reuse components, identification and selection of the component must take place. This calls for metadata and mechanisms for classifying, indexing, and querying components.

Adaptation. The presentation of multimedia documents, preferably, is adapted to the user's context. For this adaptation, we distinguish between *adaptation to personal interest*, which adapts the contents of a document to the user's interests, and *adaptation to technical infrastructure*, which technically adapts the infrastructure available to a user. For example, consider a professor on campus who is interested to see in-depth multimedia material on coronary artery bypass grafting, and an undergraduate student at home who needs to get only an abstraction of the same material.

Depending on *when* different "alternatives" are defined that can be exploited for adaptation, we distinguish between *static adaptation* and *dynamic adaptation*. With static adaptation the adaptable alternatives must be known and included in the document at authoring time. Whereas for dynamic adaptation, the available alternatives are determined depending on the specific context at presentation time.

Presentation-neutral Representation. In a heterogeneous network environment, it is desirable that the material of a multimedia application can be presented with minimal implementation effort. Thus, it makes perfect sense to reuse existing presentation software, e.g., HTML browsers, MHEG engines. As a consequence, the multimedia material has to be modeled in a *presentation-neutral* way, i.e., independent of the actual realization of a presentation. This is a challenging problem as it calls for automatic conversion of the multimedia document model used for the presentation-neutral description of multimedia content into the multimedia document model used for presentation of the multimedia content.

In general, two major characteristics influence the convertability between multimedia document models [21]: *multimedia functionality*, which is the expressiveness of the modeling primitives of a multimedia document model, and the *semantic level*. A document model describes a document on a high semantic level if the document's structure is specified rather than its presentation. In order to allow for automatic conversion, the description of multimedia content should take place on a high level of semantics.

3. Existing Multimedia Document Models

In this section, we briefly introduce the most relevant standards for multimedia document models: HTML, SMIL, MHEG-5, and HyTime. Then ZYX as an example for a document model aiming at overcoming the deficiencies of traditional models with regard to advanced applications is described in some more detail.

HTML [20] is the standard document model we encounter on the WWW. Based on SGML, it defines a syntax to enrich text pages with structural and layout information. There are efforts to allow for the dynamic manipulation of the structure, layout, and content of a HTML document with scripting languages, a technique which is also called Dynamic HTML (DHTML). SMIL [10] is an effort by the W3C aiming at synchronized multimedia presentations on the web. A SMIL document provides synchronization of continuous media elements and constitutes an integrated presentation. The ISO MHEG-5 standard [12] provides an object-oriented data model for multimedia documents. The standard defines a hierarchy of classes featuring attributes, actions, and events. These classes constitute a framework for describing multimedia presentations. Instances of the classes are glued together using *links*, which resemble event-condition-action rules and define the behaviour of a presentation. MHEG-6 [13, 9] defines an interface between an MHEG-5 presentation engine and Java applets which can further influence the behaviour of a presentation. HyTime [11, 18] allows for the enrichment of SGML document type definitions (DTD) with well-defined multimedia semantics allowing for the structural description of multimedia content. The main primitives provided by HyTime are *architectural forms* (AF). Any SGML element type can be associated with an AF thereby inheriting its multimedia semantics and attributes.

The ZYX model [1] developed by our group describes complete or fragments of multimedia documents by the means of a tree. The nodes of the tree are called *presentation elements*. Each presentation element has got a *binding point* associated with it. Such a binding point can be bound to one *variable* of another presentation element, thus creating the edges of the tree. The presentation elements are the generic elements of the model. They can represent *atomic media elements* (e.g., videos, images and text) or *operator elements*, which combine presentation elements with certain semantics. There are operator elements that allow for temporal synchronization, definition of interaction, adaptation, and for the spatial, audible, and visible layout (the so-called *projector elements*) of the document.

It is possible to delay the process of variable binding by leaving variables unbound. This allows for the definition of *templates* which can be customized to a specific problem at a later point in time. Furthermore, a tree can be encapsulated by a *complex media element* which can then be used in other trees like any other presentation element. Unbound variables of an encapsulated tree are exported by the complex media element allowing for the encapsulation of templates. Thus, a complex media element is somehow a black box view of a ZYX tree.

4. Analysis

In this section, we analyze how the multimedia document models HTML, MHEG-5, HyTime, SMIL, and ZYX introduced in the previous section fulfil the requirements outlined in Section 2. A more detailed analysis of the document models can be found in [2].

Temporal Model. HTML does not offer constructs to specify temporal synchronization between media elements. With DHTML, the temporal course of a presentation can be programmed in a scripting language. MHEG-5 specifies the temporal course of a presentation by means of its link concept. Thus, the temporal model of MHEG-5 is event-based. HyTime provides a means to define n-dimensional coordinate spaces which can include time dimensions. Media elements can be placed into this coordinate space. Hence, the temporal model of HyTime is point-based. SMIL follows an interval-based approach to temporal synchronization. The presentation intervals of media elements can be coordinated by the use of the schedule elements *par*, which specifies parallel presentation of n intervals, and *seq*, which allows for sequential presentation. The temporal model of ZYX is closely related to the temporal model of SMIL. ZYX defines the temporal operator elements seq and par which resemble the parallel and sequential elements of SMIL. Furthermore, the $loop$ operator element and the $delay$ temporal operator elements are defined providing for loops and delays during the presentation. The ZYX temporal model must be considered as interval-based.

Spatial Model. All examined document models define the spatial layout of a presentation by absolute positioning. In HTML, control of the spatial layout can be achieved by the use of tables and frame sets. MHEG-5 provides attributes for visual media elements defining the coordinates of their presentation area. HyTime provides n-dimensional coordinate spaces including spatial dimensions. SMIL allows to define fixed rectangular regions, so-called *channels*. In a similar way, spatial layout in ZYX is defined by the use of spatial projector elements. A spatial projector element defines rectangular region of the presentation area in which the subtree below the spatial projector element is presented.

Interaction. HTML provides links which allow for navigational interaction. Using DHTML, more sophisticated interactions like design interactions can be programmed by the use of scripts. MHEG-5 provides a small set of basic interaction classes for the modeling of navigational, and design interactions. HyTime provides mechanisms to define the spatial and temporal coordination of a presentation by the use of n-dimensional coordinate spaces, which require to know all spatial and temporal positions of media objects in advance. This excludes ad-hoc navigational interaction by the user. The concept of links in SMIL provides for navigational interaction. But no support is given for the specification of design interactions. The ZYX model defines two types of interaction elements, *navigational* interaction elements and *design* interaction elements. An example for a navigational interaction element is the *link* element that allows to specify hypertext structures as in SMIL or HTML. The design interaction elements are interactive versions of projector elements. For example, with respect to the typographic projector that allows to set font, size and style of a text, the *interactive typographic projector element* specifies that these values can be set interactively when the document is presented.

Reusability. HTML allows to reference whole documents and single media elements via *uniform resource locators (URL)*. However, it is not possible to reference just a fragment of an HTML document. Thus, reusability is only supported on the highest and lowest level of granularity as identified in Section 2. As there is no clear distinction between structure and layout of an HTML document, reuse can only be identical. In order to support selection and identification of documents, HTML allows for the specification of metadata by means of attribute-value pairs in the head of a document.

Considering the granularity of reuse in MHEG-5, it is important to notice that the media elements of an application are structured into *groups* which can be explicitly addressed. Hence, groups constitute the units which can be reused. As a document is a group, it is possible to reuse entire MHEG-5 documents. Likewise, scenes are MHEG-5 groups and, hence, could be reused in principal. Since scenes can refer to objects global to an MHEG-5 document which are contained in the application object, it is not possible to reuse scenes which depend on such global objects. Therefore, there is no general support for reusability at the level of document fragments. As media elements do not include the underlying raw medium data but also just can refer to this data, groups can share at least the raw media data allowing for reuse at the level of media elements. Since MHEG-5 aims less at modeling the structure of a multimedia application but at representing its final presentation form, which includes the layout, groups can only be reused identically. The identification and selection of groups to be reused is a serious problem in MHEG-5 as no metadata can be assigned to MHEG objects.

HyTime allows for reusability on all levels of granularity identified in the requirements section. Single media elements and complete documents can be referenced as entities and therefore be reused. Moreover, using *locators*, parts of a HyTime document can be referenced by name, position, or even by the use of a powerful query language. As any SGML DTD can be made HyTime-compliant, HyTime documents describe rather the structure of a document than its presentation semantics. Thus, HyTime allows for structural reuse. Moreover, because HyTime is independent of a DTD, a DTD can be provided with support for classification of (parts of) documents, e.g., by the use of attribute-value pairs. Hence, HyTime offers support for selection and identification of reusable components.

As SMIL can reference complete documents and single media elements by the use of URL, it allows for reuse at the according levels of granularity as defined in Section 2. However, SMIL does not support the reuse of fragments of documents. SMIL separates layout specifications, which have to go into the head of the document, from the structural specifications given in the body. But as both kinds of specifications are closely interrelated, SMIL provides only for identical reuse. Like HTML, SMIL allows to define meta-attributes within the head element of a document. Such meta-attributes can be used to specify metadata providing support for selection and identification.

The ZYX document model has been designed to support all levels of granularity of reuse. To support reusability of media elements, atomic media elements are provided which can be reused in any ZYX specification tree. Likewise, complex media elements which encapsulate specification trees can be reused in any other tree. As the encapsulated specification trees can smoothly range from small logical parts of a document to entire documents, ZYX supports reuse both on the level of entire documents and fine-grained document fragments. Moreover, the ability to encapsulate templates in complex media elements provides for the reuse of document templates. The ability to delay the process of variable binding allows for the separation of the presentation elements building the structure of a document and the projector elements determining its layout. This allows for structural reuse of ZYX specification trees. As complex media elements may include projector elements defining visual and audible layout this provides for identical reuse of components, as well. Concerning selection and identification of reusable elements, ZYX allows media elements, either complex or atomic, to be annotated with key-value pairs.

Adaptation. Since HTML does not offer any mechanism to specify adaptation of a document to a user's interests or to technical infrastructure, we consider only DHMTL here. DHTML allows to dynamically manipulate the structure and content of HTML documents by the use of scripts. Therefore, adaptation to user interest or technical infrastructure can be programmed with the help of a script. In a first step, such a script has to determine the user or system profile, for example by a database query. In the second step, the script has to change the structure of the HTML document according to the profile. As the author of the script must encode and thus know at authoring time all adaptation alternatives, this kind of adaptation must be considered as static.

MHEG-5 defines classes for variables whose contents can be tested. Variables can be used to choose between different branches of a presentation. Thus, a profile defining user interest and technical infrastructure could be modeled using variables. However, the problem is how such a profile is set. MHEG-5 allows to set variables only from within a document. User-specific adaptation would require to make the determination of the profile a part of the MHEG-5 document. In MHEG-6, the MHEG engine could call a Java applet which retrieves the actual values for a given profile and then sets the variables of the document. So, with the use of MHEG-6, adaptation of a presentation to user interest or technical infrastructure is possible. Since all adaptation alternatives must be specified within a document at authoring time, this is static adaptation.

Since HyTime can be used with any concrete DTD, it is always possible to define specific attributes with elements of a DTD that characterize (parts of) documents or media elements in terms of user interest or technical properties like bandwidth needed, resolution or frame rate required. It is also possible to check for values of such element attributes by using the Query-Locator. But the results of such queries checking attribute values are fully determined by the concrete document content and cannot be modified by external parameters like those in a user or system profile. Hence, it is not possible to adapt a HyTime document according to external parameters like a profile.

SMIL offers the *switch* element to model alternative presentation variants. Using this element, different adaptation alternatives can be specified inside the document at authoring time. Thus, the switch element allows for static adaptation. The selection of the alternatives is guided by simple predicates which include parameters set outside the SMIL document. These parameters are predefined by the standard and describe mainly technical features like the available bandwidth. This allows to adapt a SMIL document to technical infrastructure.

The ZYX model offers operator elements to support adaptation to a user's profile by means of *switch elements* and *query elements*. Comparable to SMIL, a switch element allows to specify different presentation alternatives for a part of the document, thus allowing for static adaptation. In contrast to SMIL, the scope of a switch statement is not limited to predefined parameters but allows for arbitrary variables and predicates. Thus, ZYX allows for adaptation to user interest and system structure. The switch element is useful if all adaptation alternatives are known to the author of a document. In order to allow for dynamic adaptation, the *query* element is provided. This element is a placeholder for a media element which is described by the means of a query. The query is represented by a set of key-value pairs. When the document is selected for presentation the query element is evaluated and replaced by the complex or atomic media element best matching the set of key-value pairs with regard to the user profile.

Presentation-neutral Representation. As HTML does not clearly separate the layout of a document from its structure, the semantic level of a HTML document description is not as high as HyTime though it is comparable to SMIL. However, HTML offers only extremely limited multimedia functionality (even simple temporal synchronization is not possible). To offer more multimedia functionality, DHTML must be employed. Since DHMTL scripts must imperatively implement multimedia functionality, their use extremely reduces the semantic level of a document description. Thus, neither HTML nor DHMTL are well suited for presentation-neutral representation of multimedia document content.

MHEG-5 primarily aims at a detailed and platform independent description of a presentation, i.e., the layout, of a document. To achieve this goal, the standard provides MHEG-5 with rich multimedia functionality. However, the description of the structure of an MHEG-5 document is very poor. Hence, the level of semantic modeling is very low. Thus, MHEG-5 cannot support presentation-neutral representation of multimedia documents.

Since HyTime mainly specifies the structure and semantics of a multimedia document, it is very well-suited for presentation-neutral representation of multimedia documents. HyTime offers specification of document content at a high semantic level though it lacks multimedia functionality (especially in the area of interaction).

In contrast to HyTime documents, a SMIL document describes in more detail the presentation of the document but less detailed the structure of the document. However, SMIL offers more multimedia functionality than HyTime. Compared to MHEG-5, the description of SMIL documents takes place on a higher level of semantics though lacking the comprehensive multimedia functionality of MHEG-5. Hence, SMIL ranks between HyTime and MHEG-5 with respect to its support for presentation-neutral representation.

As it is possible in ZYX to separate structure and layout of a document due to the ability to delay the process of variable binding and to encapsulate templates in complex media elements, the semantic level of a document description is quite high and thus the model is suited for presentation-neutral representation of multimedia document content. The amount of multimedia functionality offered by ZYX exceeds SMIL but ranks below MHEG-5.

Summary. Summarizing, we can say that none of the examined standards for multimedia document models HTML, MHEG-5, HyTime, and SMIL offers sufficient support for all requirements arising from advanced multimedia applications. HTML can hardly be characterized as a multimedia document model because it lacks support for even the most basic multimedia requirement, a temporal model. Though HTML can become a quite powerful multimedia document model by the extension to DHTML, it still lacks support for reuse at all levels of granularity and suffers from a low semantic level of content description which leaves DHTML unsuitable for presentation-neutral description of multimedia content. This is also the case with MHEG-5. Although MHEG-5 offers a high multimedia functionality, it mainly describes the presentation and not the structure of a multimedia document and, therefore, cannot be employed for presentation-neutral modeling of multimedia document content. Furthermore, reuse at the level of fragments is severely hampered due to the unflexible scene-based document structure. Powerful support for reuse is the strength of HyTime. Moreover, HyTime describes document content at a very high semantical level and, thus, is perfectly suited for presentation-neutral modeling of document content. However, the lacking capability of interaction modeling and modeling of adaption is a serious drawback. In contrast to HyTime, SMIL offers the modeling of static adaptation to technical infrastructure and navigational interaction. Furthermore, the semantic level of a SMIL document description ranks between MHEG-5 and HyTime and, hence, is quite well suited for the presentation neutral description of multimedia document content. However, reuse at the level of fragments is not possible as is the modeling of design interactions.

Since ZYX has been designed with the fulfilment of the advanced requirements in mind, it offers reuse on all three levels of granularity, static and dynamic adaptation to user specific needs, a quite high semantic level of document description, and presentation-neutral representation of multimedia content. Regarding the traditional requirements, sufficient multimedia functionality has been provided to allow for interesting multimedia presentations including design interactions.

5. Conclusion and Future Work

Driven by our advanced multimedia information system application Cardio-OP, we first have identified an advanced set of new requirements for multimedia document models: *reusability of multimedia content*, *adaptation of multimedia content to user needs and interests*, and *presentation-neutral description* of the structure and content of multimedia documents. These requirements complement the more traditional requirements for multimedia document models, i.e., temporal, spatial, and interaction modeling, well known so far. We then have presented an analysis of the relevant standard formats and models, i.e., HTML, SMIL, MHEG-5, HyTime, including the ZYX model which has been designed to meet the advanced requirements. We have presented the capabilities and identified the limitations of these models. The shortcomings of standards call for a new initiative for next generation multimedia document models. As illustrated by ZYX [1] it is very well possible to push the limits of existing approaches and to meet the new requirements.

We would like to point out that the implications of our analysis and of approaches trying to resolve the shortcomings of existing models are significant: There arises an urgent need for appropriate authoring tools that support fine-grained reuse of multimedia content, adaptability of content to user needs and individual interest, and, as a direct consequence, the presentation-neutral representation of material, e.g., in a database. When developing the multimedia content repository of Cardio-OP based on ZYX we made this painful experience. Our group has already developed a DataBlade module for the object-relational database system Informix Dynamic Server / Universal Data Option capable of managing ZYX documents and fragments [3]. We currently develop an authoring tool and a presentation engine for ZYX, since presentation-neutral representation of multimedia content as well as adaptation support directly impacts the design of authoring and presentation tools.

References

[1] S. Boll and W. Klas. ZYX — A Semantic Model for Multimedia Documents and Presentations. In *Proc. of the 8th IFIP Conference on Data Semantics (DS-8): "Semantic Issues in Multimedia Systems"*. Kluwer Academic Publishers, Rotorua, New Zealand, January 1999.

[2] S. Boll, W. Klas, and U. Westermann. A Comparison of Multimedia Document Models Concerning Advanced Requirements. Technical Report - Ulmer Informatik-Berichte Nr.99-01, University of Ulm, Germany, February 1999. http://www.informatik.uni-ulm.de/dbis/Cardio-OP/publications/TR99-01.ps.gz.

[3] S. Boll, W. Klas, and U. Westermann. Exploiting ORDBMS Technology to Implement the ZYX Data Model for Multimedia Documents and Presentations. In *Proc. of Datenbanksysteme in Büro, Technik und Wissenschaft (BTW99), GI-Fachtagung*, Freiburg, Germany, March 1999. Springer.

[4] D. C. A. Bulterman. User-centered Abstractions for Adaptive Hypermedia Presentations. In *Proc. of the 6th ACM Multimedia Conference*, Bristol, UK, September 1998.

[5] A. Duda and C. Keramane. Structured temporal composition of multimedia data. In *Proc. IEEE International Workshop on Multimedia- Database-Management Systems*, Blue Mountain Lake, August 1995.

[6] M. J. Egenhofer and R. Franzosa. Point-Set Topological Spatial Relations. *Int. Journal of Geographic Information Systems*, 5(2), March 1991.

[7] J. Eklund, P. Brusilovsky, and E. Schwarz. Adaptive Textbooks on the WWW. In H. Ashman, P. Thistewaite, R. Debreceny, and A. Ellis, editors, *Proc. of AUSWEB97, Queensland, Australia*, pages 186—192. Southern Cross University Press, July 1997.

[8] N. Hirzalla, B. Falchuk, and A. Karmouch. A Temporal Model for Interactive Multimedia Scenarios. *IEEE Multimedia*, 2(3):24–31, Fall 1995.

[9] P. Hofmann. MHEG-5 and MHEG-6: Multimedia Standards for Minimal Resource Systems. Technical Report, Technical University of Berlin, April 1996.

[10] P. Hoschka, S. Bugaj, D. Bulterman, et al. *Synchronized Multimedia Integration Language – W3C Working Draft 2-February-98*. W3C, URL: http://www.w3.org/TR/1998/WD-smil-0202, Februar 1998.

[11] ISO/IEC. *Information Technology - Hypermedia/Time-based Structuring Language (HyTime)*, 1992. ISO/IEC IS 10744.

[12] ISO/IEC JTC1/SC29/WG12. *Information Technology – Coding of Multimedia and Hypermedia Information – Part 5: Support for Base-Level Interactive Applications, ISO/IEC IS 13522-5*. ISO/IEC, 1995.

[13] ISO/IEC JTC1/SC29/WG12. *Information Technology – Coding of Multimedia and Hypermedia Information – Part 6: Support for Enhanced Interactive Applications, ISO/IEC IS 13522-6*. ISO/IEC, 1996.

[14] T. Kamba, K. Bharat, and M. C. Albers. The Krakatoa Chronicle - An Interactive, Personalized Newspaper on the Web. 1993.

[15] W. Klas, C. Greiner, and R. Friedl. Cardio-OP — Gallery of Cardiac Surgery. In *Proc. of IEEE International Conference on Multimedia Computing and Systems (ICMCS'99)*, Florence, Italy, June 1999.

[16] W. Klippgen, T. D. C. Little, G. Ahanger, and D. Venkatesh. The Use of Metadata for the Rendering of Personalized Video Delivery. In *[23]*, New York, 1998. McGraw-Hill.

[17] T. D. C. Little and A. Ghafoor. Interval-Based Conceptual Models for Time-Dependent Multimedia Data. *IEEE Transactions on Knowledge and Data Engineering*, 5(4), August 1993.

[18] S. R. Newcomb, N. A. Kipp, and V. T. Newcomb. "HyTime" – The Hypermedia/Time-Based Document Structuring Language. *CACM*, 34(11), November 1991.

[19] D. Papadias, Y. Theodoridis, T. Sellis, and M. J. Egenhofer. Topological Relations in the World of Minimum Bounding Rectangles: A Study with R-Trees. In *Proc. of the ACM SIGMOD Conference on Management of Data*, San Jose, May 1995.

[20] D. Raggett, A. Le Hors, and I. Jacobs. *HTML 4.0 Specification – W3C Recommendation, revised on 24-April-1998*. W3C, URL: http://www.w3.org/TR/1998/REC-html40-19980424, April 1998.

[21] L. Rutledge, J. van Ossenbruggen, and D. C. A. Bulterman. A Framework for Generating Adaptable Hypermedia Documents. In *Proc. ACM Multimedia Conference*, Seattle, November 1997.

[22] V. Schöch, M. Specht, and G. Weber. ADI — An Empirical Evaluation of a Tutorial Agent. In T. Ottmann and I. Tomek, editors, *Proc. of the ED-Media and ED-TELECOM 1998, Freiburg, Germany*. Association for the Advancement of Computing in Education, June 1998. URL http://apsymac33.uni-trier.de:8080/ADI.html.

[23] A. Sheth and W. Klas. *Multimedia Data Management - Using Metadata to Integrate and Apply Digital Media*. McGraw-Hill, New York, 1998.

[24] T. Wahl and K. Rothermel. Representing Time in Multimedia Systems. In *Proc. IEEE International Conference on Multimedia Computing and Systems*, pages 538–543, Boston, MA, May 1994.