To appear in *Computer Standards & Interfaces*

# Integrating the Amsterdam Hypermedia Model with the Standard Reference Model for Intelligent Multimedia Presentation Systems

Lynda Hardman, Marcel Worring*, Dick C.A. Bulterman
*CWI, Amsterdam, The Netherlands*
**WINS, University of Amsterdam, The Netherlands*

The standard reference model (SRM) for intelligent multimedia presentation systems describes a framework for the automatic generation of multimedia presentations. This framework, however, lacks an explicit document model of the presentation being generated. The Amsterdam hypermedia model (AHM) describes the document features of a hypermedia presentation explicitly. We take the AHM and use it as a basis for describing in detail the stages of generating a hypermedia presentation within the SRM framework, which we summarise in a table. By doing so the responsibilities of the individual SRM layers become more apparent.

## 1. Introduction

Intelligent Multimedia Presentation Systems (IMPS) deal with the dynamic creation of multimedia presentations optimally geared towards the needs of a user. In broad terms, the created presentation should define what is presented to the user (the content), where it is presented (the spatial layout) and when it is presented (temporal layout). Given a collection of user goals and availability of resources these three aspects leave open an immense number of possible presentations. An IMPS is a reasoning system aimed at selecting the optimal one.

The Standard Reference Model (SRM) defined in [BFMR97] considers the decomposition of this process into well defined layers. In the content layer a rough outline of the content and structure of the presentation is made, in the design layer every part of the presentation is designed and, in the realization layer, the design is realized into a full specification of the presentation in a platform independent way. These are complemented with a control and a presentation layer, mainly dealing with the input and output from and to the user. A number of expert modules are defined which aid the layers in their reasoning process.

The SRM considers the three aspects of the document in parallel. In the current version of the SRM it is not made precise which part of each of the document aspects is defined in which layer. The reason for this omission is the lack of a well defined document model. The Amsterdam Hypermedia Model (AHM) [HaBu97], [HaBR94] was explicitly designed to capture all of the above aspects and hence can serve in the further specification of the model. The main goal of this contribution is to integrate the AHM with the SRM leading to a more precise model without losing expressiveness. It provides each of the layers with explicit concepts on which to base their reasoning. Consequently, it also defines for which concepts a layer should communicate with the experts defined in the SRM and when communication with the other layers is required.

The structure of this contribution is as follows. After describing an example hypermedia presentation in section 2 we present the main features of the AHM in section 3. In section 4 we show how the AHM can be integrated with the SRM. In section 5 we give a brief discussion of our own work on generating hypermedia presentations.

## 2. Example Presentation

We provide an example of a hypermedia presentation[1], Fig. 1. This consists of a number of multimedia presentations through which a reader can navigate. The initial screen is a scene composed of four media items: two text objects, an image of the modem and an animated image of the presenter. These start simultaneously and remain on display until the reader selects to follow one of the links from each of the three sentences to an explanation given by the presenter. The first explanation is composed of two subscenes, one following the other. The first subscene shows an animation of the presenter played alongside a spoken commentary. While this is playing an image of the modem is shown, which is then changed in time with the commentary to highlight the on/off switch. The presentation continues without any interaction from the reader (indicated by the synchronization arc in the figure), and the presenter explains about LED 11. Note that the two buttons allowing the reader to jump back to the initial screen or go on to the next explanation remain on the screen throughout the scene.

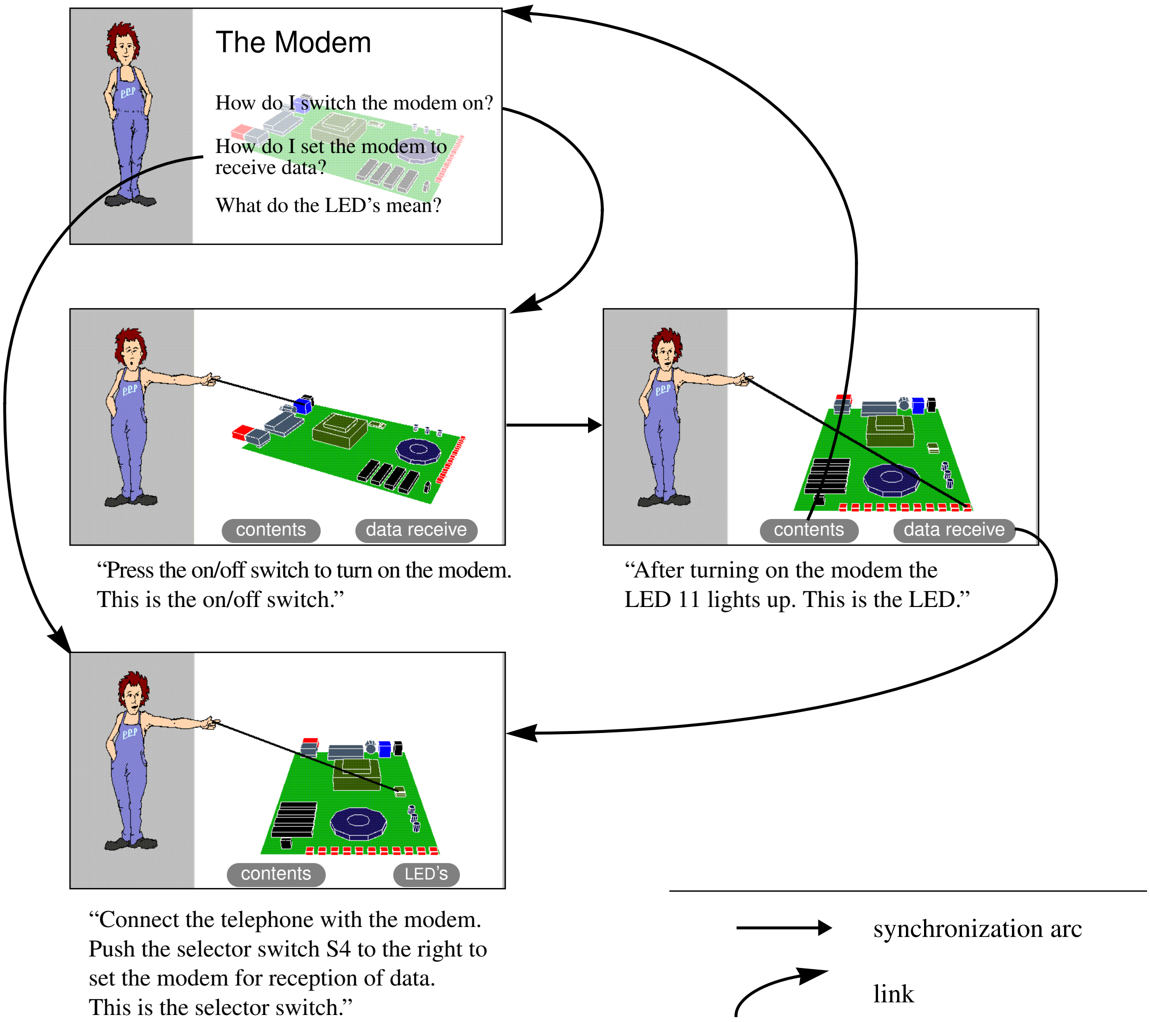


**Figure 1.** An example hypermedia presentation

1. This example is based on a generated PPP presentation explaining how a modem works. `http://www.dfki.uni-sb.de/imedia/java/applets/persona/modem.html`

## 3. Amsterdam Hypermedia Model

The purpose of the Amsterdam Hypermedia Model (AHM) is to capture multimedia and hypertext aspects in a unified document description. The AHM is a specification of the information needed for storing all aspects of a hypermedia document in a platform independent way, so that it can be presented on a number of platforms while still reflecting the author's original intentions. The model associates properties, such as duration and size, with a media item and encapsulates these in an object, termed an atomic component. The atomic component, rather than the media item itself, is used as the basis for creating presentations. Other associated properties are semantic labels and presentation specifications. The AHM is described in detail in [HaBu97]. We give a summary here for completeness.

The four main components of the AHM are:

- the atomic component which hides data dependencies of a media item from the rest of environment and associates presentation information and semantic attributes with the object;
- the composite component which groups atomic components and smaller presentations into larger presentations;
- the link which allows the expression of relationships among presentations;
- the channel which pre-specifies resources and allows styles to be applied to multiple media items.

In addition to these main components are a number of other objects in the model:

- media items which are the data objects which form the basis of the presentation;
- synchronization arcs which allow the expression of temporal constraints among parts of a presentation;
- anchors which allow links to be attached to parts of media items.

We describe these objects within the main four components.

### 3.1. Atomic component

An atomic component consists of content, presentation specification, attributes and anchors. This is summarised in Table 1.

| | | |
|---|---|---|
| Content | Media item reference | e.g. URL |
| | Media-dependent reference | part of text, image, video, etc. |
| Presentation Specification | Duration | derived from content,<br>specified by author,<br>or determined via surrounding structure |
| | Layout | size derived from content,<br>or specified by author,<br>size and position determined via channel |
| | Style | media item style,<br>anchor style,<br>transition effect |
| Attributes | Semantic description | related to external knowledge representation |
| Anchors | Anchor ID | for referring to anchor |
| | Value | media-dependent reference to content |
| | Attributes | semantic description, as for component |
| | Presentation specification | anchor style |

TABLE 1. AHM Atomic Component

- *Content*
  The content of an atomic component is a reference to a complete or partial media item.
- *Presentation specification*
  The presentation specification includes temporal layout, spatial layout and style information. *Temporal information* for an atomic component is its duration. This can be derived from the content for continuous media such as video or audio, can be assigned by the author, or may be derived from the composition structure.
  *Spatial layout* is assigned through a channel, which has its own size and position, plus the size and position of the content with respect to the channel. The content may have its own intrinsic size, for example for images or video.
  *Style* information is of three different types: media item, anchor and transition effect. Media item style is medium-dependent and includes, for example, font for text and background colour for all visual media types. Anchor style determines how the anchors are visualized to the user, e.g. pop-up when the mouse cursor is over the anchor. Transition effect specifies a special effect when the content starts or ends its display, for example, "fade-in" or "wipe-left".
- *Attributes*
  Semantic attributes can be associated with the component. These allow selections from an external knowledge representation to be associated with individual parts of a presentation.
- *Anchors*
  Anchors are identified objects specifying part of a media item and can be used as the basis for creating links and synchronization arcs among components. An anchor consists of an anchor identifier, an anchor value, attributes and a presentation specification. The anchor value is a media-dependent description of part of the content of the component. The attributes have exactly the same function as those for the complete component and should be selectable from the same knowledge representation. The presentation specification is anchor style only.

## 3.2. Composite component

A composite component consists of children, presentation specification, attributes and anchors, Table 2. There are two types of composition: temporal and atemporal. The latter is a collection of multimedia presentations, for example the three presentations shown in Fig. 1. The former is a grouping of media items belonging to the same presentation, for example the complete subscene about the on/off switch. We describe the temporal composite component.

| | | |
|---|---|---|
| Children | Component references | atomic or composite components |
| Presentation Specification | Synchronization arcs | timing between descendant components |
| | Duration | determined by children and synchronization arcs |
| | Styles | media item style,<br>anchor style,<br>transition effect |
| Attributes | Semantic description | related to external knowledge description |
| Anchors | Anchor ID | for referring to anchor |
| | Value | component/anchor reference(s) |
| | Attributes | semantic description, as for component |
| | Presentation Specification | anchor style |

TABLE 2. AHM Composite Component

- *Children*
  The children of a composite component are references to the atomic or composite components contained by the composite.
- *Presentation specification*
  *Temporal information* for a composite component is the duration of the children along with their temporal relationships, stored as synchronization arcs.
  *Spatial layout* is assigned via the channels corresponding to the descendant atomic components.
  *Style* information is media item style, anchor style and transition effect.
- *Attributes*
  Semantic attributes can be associated with the component.
- *Anchors*
  Anchors are identified objects specifying part of a component. The anchor value in a composite is a reference to an anchor in a descendant component.

### 3.3. Link component

A link allows the expression of a relationship within a collection of multimedia presentations. This is most often expressed to the end-user as a navigation action. A link consists of a list of specifiers, a presentation specification, and attributes, Table 3.

- *Specifiers*
  A specifier stores the information for the (possibly multiple) ends of the link and is itself composed of a number of parts: an anchor, a direction, a context and a presentation specification.
  The *anchor* is a list of component and anchor identifiers of components which are associated with the end of the link.
  The *direction* specifies the direction of the semantic relationship among the link-end components. This can be interpreted as a traversal direction when required.
  The *context*[2] is the collection of objects associated with the end of the link and is given by a component reference.
  The *presentation specification* for the specifier is only anchor style.
- *Presentation specification*
  The presentation specification includes temporal layout and style information.
  *Temporal information* for a link component is the duration of the transition from the source to the destination context.
  *Style* information is transition effect.

| | | |
|---|---|---|
| Specifiers | Anchor | component/anchor reference(s) |
| | Direction | FROM/TO/BIDIRECTIONAL |
| | Context | component reference |
| | Presentation Specification | anchor style |
| Presentation Specification | Duration | transition duration |
| | Transition | transition effect |
| Attributes | Semantic description | related to external knowledge description |

TABLE 3. AHM Link Component

2. This should not be confused with the context expert of the SRM. Link context is a statically specified part of the document structure which defines the scope of the link ends. We will use the full expression link context to avoid confusion.

- *Attributes*
  Semantic attributes can be associated with the component. These allow selections from an external knowledge representation to be associated with the relationship expressed using the link.

Given that a link may have multiple specifiers, each with its own context, a link is able to record the semantics among a number of source presentations and a number of destination presentations.

## 3.4. Channel

A channel is an abstract, rather than a physical, resource. A channel serves two basic purposes: to support multimedia styles (high-level assignment of properties) and to separate out runtime resource issues. A channel, Table 4, requires a media type along with a size and position with respect to some spatial coordinate axes, for example a window. Other properties can be assigned as defaults for the atomic components which are assigned to the channel.

### Multimedia styles

A channel can be used to record default atomic object properties. This use of the channel can be thought of as a generalised multimedia style sheet, i.e., not only specifying presentation styles, but also by stating defaults of the other properties. Every atomic component needs values for each one of these properties. For presentations with only a few components, these can be assigned per component. However, for presentations using multiple components this information needs to be repeated for each one. The use of a multimedia style thus saves repetitive assignment. Values assigned to the channel can be overridden in the atomic component. The more properties that require to be specified per atomic component then the more convenient it becomes to “hide” them in a channel object.

### Runtime resource issues

- *Playing media item:*
  In its most basic form, all a channel does at runtime is play a media item—it takes a chunk of data, interprets it according to a data format and displays it on the screen. It also applies the properties assigned to the media item by default from the channel, and takes into account overrides from the atomic component. A collection of active channels is (almost) equivalent to a document player — the only other information needed is the scheduling, since this requires information from multiple channel streams.
- *Allocation of resources:*
  Because the channels provide a description of the resources used by a presentation, the player software is able to make decisions about the presentation, such as how to distribute the audio streams over the available audio device(s), or whether the presentation can be played at all on the available hardware.

| | | |
|---|---|---|
| Media type | | TEXT/IMAGE/VIDEO/AUDIO etc. |
| Content | *Media item reference* | *default* |
| | *Media-dependent reference* | *default* |
| Presentation Specification | Layout | size and position defined with respect to window |
| | *Duration* | *default* |
| | *Style* | *default media item style,*<br>*default anchor style,*<br>*default transition effect* |
| Attributes | *Semantic description* | *related to external knowledge description* |

TABLE 4. AHM Channel

## 4. Relationship with the SRM

In the previous section we considered how a hypermedia presentation can be modelled using atomic components describing the content, composites and links defining the temporal layout, and channels for specifying the spatial layout. In this section we integrate the AHM and SRM by specifying for each layer of the SRM which parts of the AHM document description are the basis of reasoning. The results of this section are summarized in Table 5.

### 4.1. Control layer

This layer selects a goal from a set of goals which have still to be met. It influences the message to be expressed by a presentation, but does not make any decisions directly affecting the instantiation of parts of a particular presentation.

### 4.2. Content layer

The content layer takes a single goal passed on by the control layer and refines it to a list of subgoals. These are in turn transformed into a set of communicative acts and relations among these acts. This is carried out by four separate, but communicating processes—goal refinement, content selection, media allocation and ordering. Goal refinement is a task which has to be carried out before any parts of a presentation are created, so we do not discuss it further here.

The content selection process communicates with the application expert, and decides on the semantic content of the presentation. This can be reflected in the document structure using AHM attributes. These can be associated with a component or anchor to express the role of a particular part of the presentation, e.g. recording the knowledge structure used to guide the generation of the presentation in the first instance. The attributes should be taken from the knowledge representation used by the application expert. As an illustration, the picture of the modem at the bottom of Fig. 1 may have the associated attribute “modem showing switch S4”. The system is then able to re-use the document later for further processing, for example, the user retraces a few steps and then formulates a different goal, and the system is not obliged to regenerate everything. A further advantage is that the description of the object can be used either to find an existing media item in a database or to generate it.

| AHM / SRM | Atomic | Composite | Link | Channel |
|---|---|---|---|---|
| Control | - | - | - | - |
| Content | semantic attributes,<br>media type,<br>anchor ID’s,<br>anchor attributes | children,<br>semantic attributes,<br>anchor attributes | links among composites:<br>anchors, direction, link context;<br>semantic attributes | channel media type |
| Design | style specifications | temporal constraints within composite,<br>styles,<br>anchor ID’s,<br>anchor styles | link anchor style,<br>transition temporal constraints,<br>transition effect | spatial constraints among channels |
| Realization | media item,<br>anchor values,<br>size,<br>duration | synchronization arcs,<br>duration,<br>anchor values | transition duration | position, size |
| Presentation | - | - | - | - |

TABLE 5. The SRM layers in which the parts of the AHM are generated

The media allocation component assigns a particular medium to a communicative act, producing a media communicative act. In terms of the AHM, a decision has been made as to which media type will be used. This determines, not the selection of a particular channel, but the restriction of potential channels to those of the appropriate media type. Note that while the media items have not been finalised in the content layer, this has no influence on the ability to create structures in the AHM. Since the atomic component and not the media item is the basis of a presentation, compositions and links can be created independently of the final media items.

The result of the ordering process is a, not necessarily linear, ordering of the media communicative acts. Hyperlinks, composition, timing relations and spatial layout are all different ways of expressing ordering relations among media items. When a communicative act is to be expressed using a number of media items these methods can be traded-off against one another. For example, rather than display many items together, links can be made among smaller groups of the items, or the sequential playing of items may be a suitable alternative for laying out items next to each other at the same time, [WeWi96]. These different ways of "ordering" the material require to be captured within the content layer, but may not be finalised until the design or realization layer. We thus restrict the content layer to specifying a structure of composites and links, and leave the decisions on temporal and spatial layout to other layers.

The composition structure defines a hypermedia presentation consisting of a collection of multimedia presentations with relations among them, for example the different explanations of the modem in Fig. 1. The collection of presentations is an atemporal composite, containing a number of children each of which is a temporal composite. One communicative act can be instantiated as an atomic or composite component, and a relation can be instantiated as a link. Although a link requires the specification of anchors, these are dependent on the media items which have not yet been determined. However, on the basis of semantic attributes, anchors can be assigned to components. These are specified using only anchor ID's which can later be instantiated with the appropriate anchor values within the media items. The direction and link context can already be specified, since these are independent of the final media items.

In summary, in the content layer the parts of the AHM that need to be generated are: the atemporal structure including the children of each composite; the atomic components; the links among components including anchor ID's, direction and link context; labelling of each part with semantic attributes; and the media type of the atomic components and thus the channel.

### 4.3. Design Layer

The design layer allows the processes of media selection and layout design to carry on in parallel, imposing no ordering requirements for which of these should be finalised first. The design layer is split into two communicating processes: media design and layout design.

The media design component makes decisions on the "look and feel" of the presentation. It can communicate with the design expert and make decisions on styles for atomic and composite components, anchors, and transitions. These are the media item styles, anchor styles and special effects described in section 3.

Some style aspects do not at first sight appear to fit into the AHM, for example, PPP Persona points to a part of the internal workings of a modem with a stick, Fig. 1. However, this can be described in AHM terms as follows. The pointing stick is an anchor style for highlighting a particular part of the graphic. Another way of visualizing the same information would be to have the graphic part flash when the relevant part of the commentary is spoken. Indeed, the Persona himself may be considered as "only" visualizing a pointing convention with which we are familiar from the physical world.

The layout design component is responsible for the spatial and temporal arrangement of the media items in the presentation plus the transition durations. In this layer the constraints used for determining the spatial and temporal layout need to be generated. These specifications should be internally consistent before being handed on to the realization layer, which may, none-the-less, be unable to satisfy them requiring the generation of a new set of constraints. The generation of the spatial layout design should occur along those lines described in [Graf97].

In summary, the parts of the AHM which are determined in the design layer are the media item, anchor and transition styles, the temporal and spatial layout constraints for each temporal composite, plus the temporal constraints for transitions.

### 4.4. Realization Layer

In this layer the final decision is taken on the media items to be played in the presentation and their corresponding spatial and temporal layout. Again, the layer is split into two communicating processes, paralleling those in the design layer—media realization and layout realization.

The media realization component ensures that appropriate media items are chosen. These may already exist, or may be generated specifically for the task. In either case, in AHM terms the atomic component content becomes firmly specified, along with the associated channel and position within the channel. The content of an atomic component is specified either as a reference to an existing media item or as a complete description in an appropriate specification language. Once a media item has been determined, the anchor values corresponding to the anchors can also be determined. For example, in the content layer for generating the presentation in Fig. 1 it is determined that the on/off switch is to be highlighted within the picture of the modem. In the realization layer the picture of the modem is instantiated and the switch can be described in terms of the image. The anchor then relates the semantic description of the switch, stored as an attribute, with the media-dependent description of the visualization of the switch in the image, the anchor value.

The layout realization component calculates the final temporal and spatial layout of each temporal composite based on the constraints specified in the design layer. The duration of an atomic component may be derived from the content for continuous media, or from the composition structure where, for example, static and continuous media are played alongside each other. The duration of a composite component can be calculated on the basis of its children plus any other temporal constraints specified in the design layer. Synchronization arcs in the AHM can be used to specify starting times of components in relation to one another. Where links define possible paths among separate multimedia presentations, there is also a duration perceived by the user when the link is followed. This is the link transition duration, which is also finalised in this layer.

The layout of the presentation is calculated by a layout realization process, such as that described in [Graf97]. By specifying the spatial layout of the atomic components the channels are also determined. This results in the selection of existing applicable channels. The exact layout of an object is determined by the channel through which it is displayed plus the position and scaling of the object relative to the channel.

In summary, in the realization layer the final media item is specified along with corresponding anchor values; size, position, duration and start time of all components, including link transitions, are calculated; channel selection, plus the size and position of each channel is finalised.

### 4.5. Presentation Display Layer

The SRM separates playing of a presentation from its specification in a similar manner to the AHM, and the AHM itself does not go any further with modelling the runtime behaviour of a presentation. We have carried out work on displaying a presentation [RJMB93] conforming to the AHM, where player software interprets the information supplied in a source document, calculates an appropriate run-time schedule, based on the timing relationships finalised in the realization layer, and does the actual display of the constituent media items. A system conforming to the SRM may similarly generate a document which could be played by a separate player, such as that described in [RJMB93].

## 5. Discussion

We have introduced the AHM as an explicit document model into the SRM, and have given our interpretation of the parts of the model that should be derived in each layer. The three layers that generate parts of the model are the content, design and realization layers. The content layer should establish the overall connections among different presentations, the design layer should specify the constraints among the components making up a presentation, and the realization layer should ensure that these constraints are resolved to a complete platform independent description of the presentation. The four main components of the AHM are used as follows. Atomic components allow a presentation to be built up without any initial explicit reference to the media items. Composite components allow the grouping of components, along with temporal relationships, into a single presentation. Links allow the creation of relationships among presentations. Channels allow similar styles and layouts to be re-used, rather than defining new positions and sizes for each individual atomic component. A similar notion of consistency is embodied through design styles in [WeWi96].

Our own design work on automatically creating hypermedia documents, described in [WoBH96], takes a slightly different approach to that of the SRM. We consider that a large supply of media items already exists, and that these can be searched through in order to find representations of the concepts to be included in the presentation. The author (who may also be the end-user) can select the topic of interest and have the system make a selection of relevant media items. These can be collected together in appropriate groupings and displayed using predefined templates of channels. This takes a simpler approach to the rhetorics of the presentation through having the author take the steps carried out in the control layer and goal refinement in the content layer. This requires less analysis work on the goals that have to be met. Our approach does not demand the creation of media items but the selection of appropriate ones from an existing database. This in turn requires semantic annotation of items in order to be able to retrieve them. We are carrying out work parsing video segments, and objects within the segments, to facilitate the annotation process. Other work is being carried out on a prototype for assembling items into a presentation for display on the World Wide Web.

## Acknowledgements

The authors would like to thank Thomas Rist for his suggestions for improvements to this paper.

## References

[BFMR97] M. Bordegoni, G. Faconti, M.T. Maybury, T. Rist, S. Ruggieri, P. Trahanias and M. Wilson (1997). A Standard Reference Model for Intelligent Multimedia Presentation Systems. In this issue.

[Graf97] W. Graf (1997). Intelligent Multimedia Layout: A Reference Architecture. In this issue.

[HaBu97] L. Hardman, D.C.A. Bulterman (1997). Document Model Issues for Hypermedia. In “The Handbook of Multimedia Information Management”, edited by W.I. Grosky, R. Jain and R. Mehrotra. Prentice Hall.

[HaBu95] L. Hardman and D.C.A. Bulterman (1995). Towards the Generation of Hypermedia Structure. First International Workshop on Intelligence and Multimodality in Multimedia Interfaces, Edinburgh, UK, July.

[HaBR94] L. Hardman, D.C.A. Bulterman and G. van Rossum (1994). The Amsterdam Hypermedia Model: Adding Time and Context to the Dexter Model. *Communications of the ACM*, 37 (2), Feb, 50 - 62.

[RJMB93] G. van Rossum, J. Jansen, K S. Mullender and D. C. A. Bulterman (1993). CMIFed: A Presentation Environment for Portable Hypermedia Documents. In Proceedings: *ACM Multimedia ’93*, Anaheim CA, Aug, 183 - 188.

[WeWi96] L. Weitzman and K. Wittenburg (1996). Grammar-Based Articulation for Multimedia Document Design. *Multimedia Systems*, 4: 99-111.

[WoBH96] M. Worring, C. van den Berg and L. Hardman (1996). System Design for Structured Hypermedia Generation. Visual Information Systems ’96, Melbourne, Feb, 254 - 261.