1 Summary

Hypermedia presentations are documents which are not printed on paper but make use of a computer screen for their display. Fig. 1 shows an example hypermedia presentation. The example shows text, video, images and audio combined together into different presentations and includes choice points where a reader can select other presentations to view. We use the term *media items* for the pieces of text, video etc. and call the choice points *links*. When combining media items into a presentation, temporal relations among the items specify when each item should appear on the screen and for how long.

A hypermedia presentation can be generated at play-back time from an underlying document which specifies the various aspects of the presentation. To allow presentations to be played on different software systems, a model of the

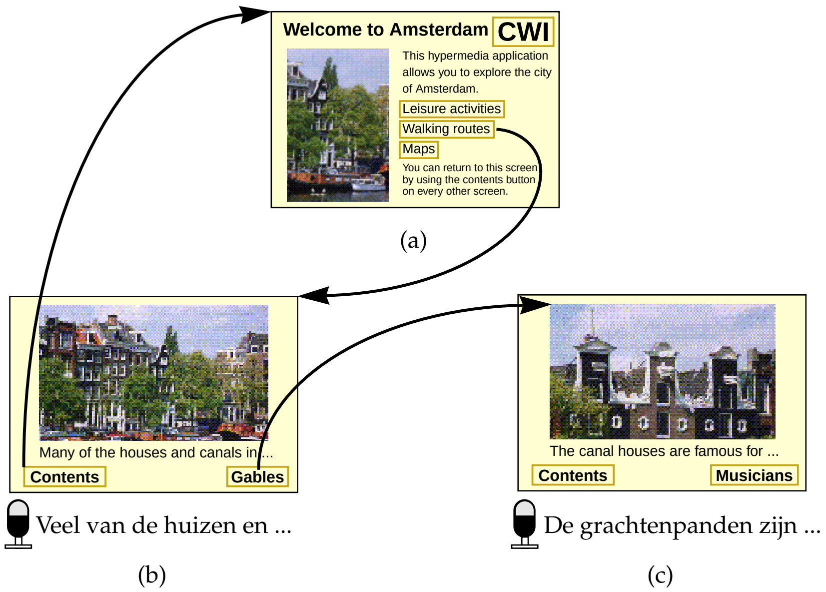

Figure 1. An example hypermedia presentation

underlying document is needed. An analogy is with word processors, where authors are able to use the system of their choice and the documents can be transferred among systems, because the underlying models are similar. Another advantage of having an explicit model is that documents can be processed for other reasons, such as creating multiple versions for different end-user platforms, or changing the visual styles of the document. It is this degree of compatibility and processability which we seek to achieve with a model of hypermedia documents.

Given a model for hypermedia documents, an authoring system can be created to support the creation, editing and deletion of the constituent parts of documents.

This thesis first states the requirements and defines a document model for hypermedia. It then analyses the user interfaces in existing authoring systems for multimedia documents and goes on to state the requirements for a complete hypermedia authoring environment. Finally, it describes the CMIFed authoring system, implemented by members of the CWI multimedia group.

1.1 Requirements for a model for hypermedia

Looking again at the example in Fig. 1 we can deduce a number of the requirements for a hypermedia document model.

Firstly, there are a number of media items. It should be possible to specify which items are included in the presentation, and which part of the item is to be displayed. The latter is useful for avoiding duplication of similar media items. The data type of each media item also needs to be known so that the playback system can interpret it.

Each media item is displayed at some position in the window and with a specific size. For example, the heading in Fig. 1(a) is at the top left of the scene and does not overlap with the CWI logo to its right. Timing information is also needed, e.g. to specify that all four items in Fig. 1(a) are to be displayed at the same time. Each media item needs an associated start time and duration.

As well as information per media item, other information involving multiple media items is needed. For example, the scene in Fig. 1(a) is a collection of the four items on the screen. Structuring information such as this must also be included as part of a document model.

The choice points must also be specifiable. These include the information about where a reader can click on the screen (for example the lightly shaded boxes in Fig. 1) and what the destination is. Links in hypermedia become quite complex since the presentation consists of multiple items, of which some can be continuous media, such as video and audio. When the reader follows a link only some of the presentation may change—for example in Fig. 1 (b) and (c) the *Contents* text item remains unchanged when following the link from the *Gables* text item.

While the emphasis of the document model is to ensure that a presentation can be reproduced on the basis of the stored document, it can also be used to support the retrieval of different parts of a document. Each media item, or part of a media item, can represent a real-world object or concept. The document model can be extended to include media independent descriptions of the information contained in the media items.

These are thus the requirements for a document model for hypermedia. In the thesis we show that existing models of hypertext (in particular the Dexter model) and multimedia (the CMIF model) are insufficient for describing all the required aspects for hypermedia. In particular, the following aspects are missing from existing models.

- The specification of which part of a presentation is affected on following a link. For example, should the whole presentation be replaced by the destination of the link, should only one part be replaced, or should the destination appear in addition to the original presentation?
- The inclusion of style information describing how the source of the link should transform into the destination of the link. For example, it is not visually appealing if the presentation stops playing, the screen goes blank and the destination only starts playing a few seconds later. A feeling of continuity is given if the presentation continues playing but fades out gradually while the new presentation fades in.
- The inclusion of media-independent descriptions for parts of media items. For example, allowing an author to associate the notion of "house" with parts of the images in Fig. 1.

We conclude that a new model for hypermedia documents is required.

1.2 A hypermedia document model

In chapter 3 we define the Amsterdam Hypermedia Model (AHM) which satisfies the requirements derived in chapter 2.

The main elements of the model are the atomic, composite and link components and the channel. These incorporate other sub-elements in order to provide the required expressiveness for the complete model.

Atomic component

An atomic component collects together all the properties that can be associated with a single media item, including a reference to the media item. The properties are: duration, spatial information, style, and media-independent descriptions. An atomic component also allows the specification of parts of a media item, termed anchors, which can be used as the start and end-points of links. An anchor requires a data-dependent specification of part of the media item and can also have associated properties.

Composite component
A composite component allows the grouping of a number of other components into an element that can be treated in the same way as an atomic component. There are two types of composite component: temporal and atemporal. A temporal composite allows the specification of temporal relations among its children, thus forming a continuous presentation. An atemporal composite allows the grouping of presentations which have no pre-specified temporal relations among one another. This allows the creation of independent scenes through which a reader is able to navigate. A temporal composite includes synchronization arcs which record the temporal relations. An atemporal composite includes activation information recording which of the child presentations should be made active when the parent is activated, e.g. by following a link to it.

Link component
A link component specifies the source and destination components for following a link in hypermedia. The component consists of a number of specifiers, each of which can act as a source and/or destination of a link. Each specifier has a reference to an anchor which may be the place the reader clicks, or may be highlighted when at the destination of a link. The specifier also denotes which parts of the presentation are associated with the anchor, e.g. the atomic component containing the anchor or a composite component representing the complete scene.

Channel
The model also includes a channel element. This brings together spatial, style and data format information in a form that can be re-used by multiple atomic components.

By deriving the model from an example presentation we demonstrate that the parts of the model are necessary for describing a hypermedia presentation. They must also be shown to be sufficient, i.e. that the documents from a broad selection of systems conform to the model. We show that the AHM is able to describe the models implicit in a selection of existing hypertext, multimedia and hypermedia systems. We therefore conclude that the AHM is a comprehensive, yet not overly complex, model for hypermedia documents.

1.3 Authoring paradigms

Different approaches to authoring multimedia already exist. We analyse a selection of existing systems and categorise the authoring approaches into a number of paradigms which represent the underlying models presented in the user interface.

We analyse the paradigms to determine their suitability for different parts of the authoring task, namely creating narrative structure, temporal information, spatial layout and links among individual presentations. Structure based systems are more suited to editing the structure of a presentation, and, where the structure reflects the temporal structure, are also useful for editing the presentation's timing. Timeline based systems are more suited to showing the timing throughout a presentation and the timing relationships among parts of a presentation. Flowchart and script based systems are best at specifying more generalised interaction, where the flowchart has a better user interface. None of the paradigms is particularly suitable for editing layout or for creating links, although the structure-based paradigm allows the different parts of the link to be specified.

We conclude that each paradigm is most suited to a particular editing task, that no single paradigm is sufficient for covering all editing aspects of a hypermedia presentation, and that several interfaces within a unified environment are required.

1.4 Requirements for authoring hypermedia

Given the AHM, we specify the functional requirements for an authoring system supporting the document model. To maintain an overview of the authoring environment it is divided into four layers: data, resource, component and document.

The data layer contains the media items themselves, thus shielding the other layers from data format dependencies.

The resource layer contains the resources used for the different aspects of the document, for example a data format resource required for interpreting the data, style information for fonts, media-independent descriptions dependent on an application domain, and layout information. The importance of including the resources as a separate layer is that each can be replaced by a similar resource while leaving the document structure itself unchanged. This allows multiple presentations to be generated from the same underlying document structure, e.g., layout can be tailored to specific output environments.

The component layer stores the document components and is supported in an authoring environment by allowing the editing of individual components. For example, a picture becomes part of a presentation when it is specified where, when and for how long it should appear on the screen.

Temporal and spatial layout play a particularly important role in hypermedia presentations, and these aspects have to be coordinated among multiple elements. The document layer allows the editing of these aspects and communicates the information to the component and resource layers. As well as stating the requirements for the document layer, illustrations of potential user interfaces are given for aspects such as editing temporal and spatial information. In particular, timeline illustrations are given for showing temporal constraints, changes in tempo, and navigating the presentation timeline.

1.5 A hypermedia authoring environment—CMIFed

Having specified the requirements for an authoring system, we describe CMIFed which was implemented for creating hypermedia documents. The importance of CMIFed from the perspective of this thesis is to show that the majority of the stated authoring requirements can indeed be implemented. For the requirements that were not implemented we are able to state what the reason was. This may have been unacceptable implementation effort, or that with hindsight they were not actually needed. CMIFed also illustrates ways of visualizing some of the aspects of the requirements, in particular temporal and spatial relationships among components.

1.6 Applications of work

The work described in this thesis has contributed to and benefited from two major international collaborations. Our model heavily influenced the development of the Synchronized Multimedia Integration Language (SMIL). This was developed by the World Wide Web Consortium working group on Synchronized Multimedia, which includes members from the CWI multimedia group. The SMIL language is being developed as a vendor-neutral multimedia document description language. Browsers are currently being developed which will enable SMIL documents to be played over the Web.

Work on the authoring system has continued as part of the Chameleon project, ESPRIT-IV Project 20597. The project benefited in two ways. Firstly it had immediate access to a hypermedia authoring environment, CMIFed, whose functionality, broadly speaking, satisfied the requirements derived in this thesis. Secondly, one of the goals of the Chameleon project is to allow the creation of a single source document which can then be (semi-)automatically adapted for playback on a range of end-user platforms. The CMIFed document format is based on the AHM, which allows it to be translated with relative ease to other formats. Target formats currently being implemented in the Chameleon project are MHEG-5 and SMIL.