On the road to Rich Digital Books

Rui Lopes Carlos Duarte Hugo Simões Luís Carriço
LaSIGE, Faculty of Sciences, University of Lisbon
Edifício C6, Campo Grande, 1749-016 Lisboa, Portugal
{rlopes,cad,hsimoes,lmc}@di.fc.ul.pt

Abstract

This article presents an architecture for producing rich digital books. These are books enriched with supporting media, like images and sound, to reach broader audiences and enlarge their usage possibilities. The architecture is based on automating the books' production process, increasing their availability. It also enjoys a high flexibility degree, allowing the production of books that can be read either with off-the-shelf software or specialized digital book players. One such player, capable of meeting the demands of active reading is also presented in this article.

Keywords

Rich Digital Books, Automated Book Production, Behavioral Dimensions, Adaptive Multimodal User Interfaces

1. INTRODUCTION

Before digital libraries, several barriers were in the way of everyone's right to information. Today, users can read books in the comfort of their homes, thus another level of issues becomes relevant. The blind, children, students, and the average user should have access to information. However, current technologies are unable (or become too expensive) to cope with this. Books should be enriched with new information, new media, or even rewritten. As it is impossible to have an *one size fits all* solution, different interfaces must be provided by reproduction platforms.

With this range of possibilities, manual production of rich digital books will be time consuming and error prone. Therefore, production should be as automated as possible, driving the focus of manual tasks to specialized activities (e.g. describing content's metadata). This way, book collections can be delivered to a wider audience, through diverse usage scenarios. General-purpose time based hypermedia frameworks [vO01] may produce similar results, but lack critical text-based formatting capabilities.

This article presents an architecture for rich digital books. Different reproduction platforms can be targeted by production processes, e.g. limited players or powerful players with adaptive capabilities. Such players are presented in the article, accomplishing the same goals that a set of dedicated digital reading platforms have achieved [Schilit 99], as current playback devices [DC] suffer from severe usability and accessibility flaws [Duarte 05].

2. REQUIREMENTS

Reading activities are greatly influenced by readers' goals, resulting on different commitment and attention levels. To portray the diverse kinds of reading, a categorization of reading situations was proposed in [Schilit 99]. Two dimensions – nature of engagement (passive or active reading) and breadth of the activity (single or multiple texts) – are used to categorize four conventional reading situations [Schilit 99]. While understanding the text is the goal of all situations, each one has different problems. Multiple texts entail the need to manage multiple documents and the difficulty on finding the needed information. Active reading [Adler 72] involves underlining, highlighting and annotating, either on the text or in a separate notebook, demanding an annotations management mechanism.

Digital books and digital libraries may contribute to mitigate some of these problems. Digital books offer the possibility to record, organize and search annotations entered by the reader [Schilit 98]. Digital libraries make it possible to manage book collections, and to create and explore relations between books. The possibility of sharing personal annotations within the community of members of a digital library conveys a social sense to the books [Kaplan 05].

Besides these affordances, the book's digital support opens allows for enriching its content with supporting media [Carriço 03]. This way, the book can be enriched with new multimedia content. In such scenario, the book content can also be narrated in addition or in alternative to the visual presentation, similar to Digital Talking Books [AN02].

The rich digital book in a digital library context must cope with a heterogeneous audience and a variety of situations: Researchers, students and other professionals who can benefit from the advanced navigation features, annotation support, and the integration with a large collection of related material; Pleasure readers in constrained environments can benefit from the alternative modes of presentation. An audio-centered book can be used in situations

where visual focusing is more difficult or even impossible; Both children and adults can enjoy visual and audio enhancements that enrich and complement the book's presentation; Children and other reading/writing disabled persons benefit from simultaneous narration and visual presentation when learning how to read/write; The blind can benefit from audio narration, coupled with navigation mechanisms (e.g. table of contents), and speech-based annotations; Partially sighted users will benefit from the enhancement of visual components. Therefore, books' contents (and its navigation, presentation and interaction mechanisms) should be tailored according to each situation.

These issues have an impact not only on the reproduction platform, but also directly on the book's characteristics. To be able to meet all the proposed goals, the book's format must be presentable in some reproduction platform, and enriching materials must be selected according to the book's content, allowing book authors and publishers to easily create different editions. Also, through production automation based on a normalized digital library, the capabilities offered by reproduction platforms will be met.

To fit all these aspects in a rich digital book creation process, production architectures must be modular, easing both book creation and maintenance tasks, as follows: provide modular and composable content processing configuration; define a rich content format supporting multimedia composition; allow the addition of content to a central repository; support addition of new material to the content; provide a clear separation between the book's content and its UI; and define a reusable specification of UIs, enforcing coherence amongst different usage scenarios.

Having flexibility in the book production process raises issues regarding production time users. As such, three user profiles must be supported: top level user, power user, and developers. The first is typically a user with little technical expertise (e.g. digital library maintenance staff), whose tasks relate to specify and annotate content, and control profile-based batch production. The second profile relates to those who have high knowledge over digital publishing and, therefore, should have full control of book production processing steps, to create production profiles (geared towards publishers' specificities). Lastly, developers are specialized in creating digital publishing components.

In order to support the different rich digital books scenarios, a multimedia repository must be available both at production and reproduction times supported by the following: continuous identification, classification, and organization of multimedia items; converting, and structuring data into a normalized format; establish relationships between media items, based on different criteria (e.g. semantic or composition based); lastly, support online query and retrieval of multimedia items, to be integrated into manual and semi-automatic content production tools.

3. RICH DIGITAL BOOKS PRODUCTION

A flexible architecture for automated production of rich digital books was defined, based on the requirements previously gathered (see Figure 1). An input is fed to production (by top level users), where it is transformed, augmented, and/or simplified according to a production profile (e.g. based on publisher's requests). At the end, a tailored rich digital book is delivered. As documents are specified in (or transformed into) XML languages in both ends of the architecture, the document flow was implemented with XML technologies (created by domain expert developers).

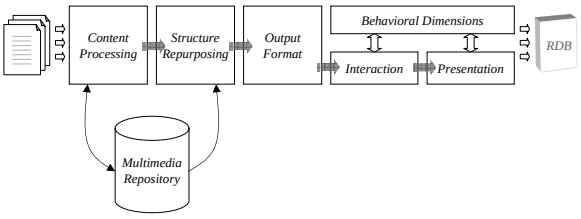

Figure 1. The production architecture.

3.1. Content Processing

The increase on the production and use of rich contents requires an effective, efficient, and reliable multimedia content management. To handle these issues, the processing architecture's first concern deals with different tasks centered on book content processing. As several input formats may be fed, a normalization task is performed initially.

Afterwards, content reasoning tasks should be performed. These can be classified as manual, semi-automatic, or fully automated, depending on the content's richness and complexity (semantic analysis *vs.* automated syntactic analysis). These tasks are supported by a multimedia repository through its rich modeling capabilities and content based multimedia storage, indexing and classification, to facilitate on retrieval and integration facilities (see Figure 2).

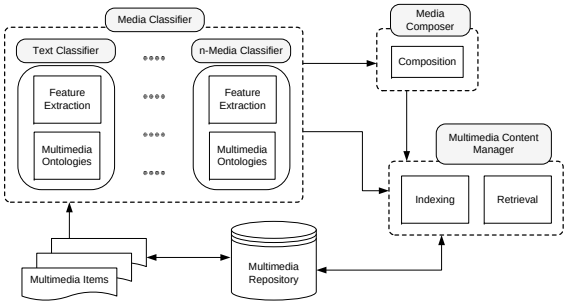

Figure 2. Multimedia items classification.

3.2. Structure Repurposing

The second concern relates to content structure repurposing. Initially, different tasks provide powerful content reasoning features, enforcing content reuse mechanisms and multimedia repository enrichment with the current state of the rich digital book's content. Afterwards, structure extraction tasks can be applied to the normalized content. A simple example is the extraction of a table of contents as independent structured content modules. Lastly, some

control over content structures can be performed, regarding their complexity (e.g. if targeted to low-end devices).

3.3. Output Format

The third concern in the production architecture relates to output format conversion. This concern defines tasks for the conversion of normalized content structures. As different scenarios must be taken into account, different formats are supported (e.g. HTML+TIME). Richer formats allow the integration of interaction and presentation capabilities, whereas more limited platforms require simple content formatting. Afterwards, different tasks can be applied to the current processed book content state, integrating user custom constructs (such as skeleton structures for bookmarking and annotating), if applicable within the chosen format.

This concern provides the minimum output to be played on a rich digital book reproduction platform, as some platforms are rich enough to provide flexible interaction and presentation capabilities. Consequently, the specificities for these two concerns are optionally applied later on.

3.4. Interaction and Presentation

After the playback platform is chosen, and the book's content is transformed into a specific output format, behavioral dimensions are introduced in the production architecture. These dimensions define how a reproduction platform should handle interaction and presentation issues over a book's content. Corresponding concerns are introduced in production time to handle playback platforms' limitations.

The first concern relates to interaction. If the output format allows its configuration, tasks implement interaction mechanisms for specific inputs (e.g., mouse, speech). Interaction limitations are also handled (e.g. limiting speech recognition vocabularies in crowded environments).

The other concern defines how a book is presented. To ease configurability and keep UI coherence amongst different output formats, the architecture defines presentation profiles (sets of presentation rules), applied to the book's content. By combining different rules, profiles share presentation features (enforcing UI coherence). Also, by selecting richer presentation profiles, stricter behavioral dimensions are fed to the reproduction platform.

4. REPRODUCING RICH DIGITAL BOOKS

The production process flexibility outputs digital books tailored to a reader's desired format. This means that books can be available for immediate presentation on several devices or platforms, or that the production process output can be fed into a dedicated book reproduction artifact that will deal with the interaction and presentation aspects.

In the first case the production outputs ready-to-read books on commonly available presentation platforms. The output formats currently available in our production platform include print formats and other ones combining text and audio, like SMIL (which can be viewed with RealPlayer), or HTML+TIME (using Internet Explorer). Other formats, targeted at more specific user populations, may also be generated (e.g. a Braille version for the blind).

However, no book player meets all the requirements elicited. These requirements, namely the possibility to present the same content with different media and be operated by visually impaired users, demand for multimodal interaction handling. The several personalization capabilities that must be available, and the need to manage possibly simultaneous media streams, also suggest the introduction of adaptive capabilities. To meet these requirements, a player with adaptive multimodal capabilities was developed, based on the FAME framework [Duarte 06].

For rich digital books reproduction four components were identified: main book content, table of contents, annotations and supporting media. A set of dimensions is identified in each component, describing how they relate to each others during the book's reproduction, and also determine the way each component is presented and how the user can interact with it. For the mentioned components, the identified dimensions allow the adaptation of the following aspects of presentation and interaction: modalities used to present content and receive instructions; synchronization granularity unit between audio and visual presentation; audio narration speed; text that is presented marked based on attached annotations or media items; reading path followed based on reader-made annotations; component visibility; intrusion and awareness raising features of the player whenever there are supporting contents or annotations to be presented; narration behavior whenever supporting contents or annotations are presented; and Table of Contents and annotations window presentation.

The developed player (Figure 3) improves the reader's situational awareness: a marker highlights the excerpt being narrated; the current section or chapter number is highlighted in the Table of Contents; and, the current reading position is afforded by the main content scrollbar. The current section or chapter number can be narrated at anytime, providing the information for the blind. To perceive the book's structure, the user can scan the Table of Contents at anytime without dismissing the narration.

Figure 3. Rich digital book player.

Advanced annotation support offers the reader the possibility to create and manage bookmarks and notes. Raising the reader awareness to annotations during the book's presentation is achieved in two distinct ways: the show/hide annotations button flashes every time the narration reaches

a point in the text that has been annotated and the annotations window is not visible. If the annotations window is displayed, the text that has been annotated is highlighted in the main content window. The awareness to supporting media is raised similarly. Audio alerts are also provided to signal the presence of an annotation or an image.

The player supports active reading of books, by allowing the reader to annotate the text and categorize the annotations. From these categories, several behaviors can be devised for further readings of the same book: reading of the annotated material; reading material of specific categories; and associate different reading speeds to different categories. This allows user creation and reading of trails that may constitute sub-stories, argumentation paths, etc.

In order to minimize user interventions, the player can adapt its behavior regarding presentation of annotations and other supporting material. If alerting the user to the presence of images or annotations is usually met with positive feedback, and the user decides to view the image or annotation, then the player will update its behavior to showing the image or annotation, instead of alerting to its presence. But, if the user chooses to ignore the alerts, the player will then update its behavior to stop alerting to the presence of images and annotations.

5. CONCLUSIONS

To increase the availability of digital books in general and DTBs in particular, it is essential that the production process moves from the manually assisted process that it is today, to a mostly automatic one. This article presented a rich digital book production architecture to move us closer to such a vision. With the automation of the production process, the architecture enables goals such as providing the same *brand* for a whole collection offered by a digital library, or preparing special editions of books targeted for specific audiences. To allow for the production of such books, the architecture is defined by mechanisms to normalize content, perform structure repurposing and define output formats. The preparation of special editions is made possible by the close integration with a multimedia repository, which provides contents to enrich and complement the book. During the production process, contents are added to the repository for future uses.

The described architecture proved its flexibility by producing books that can be read in readily available tools, or targeted at specialized digital book players. In the first case, output formats encompass print books, presentations combining text with narrations, and even Braille. Specialized players can offer features like advanced navigation, or annotation management and sharing that allow active reading practices, and increase the possible spectrum of readers and reading situations. One such player, developed to meet these goals, was also presented in this article.

6. ACKNOWLEDGEMENTS

This work is funded by Fundação para a Ciência e Tecnologia, through grant POSI/EIA/61042/2004, within the context of project RiCoBA (Rich Content Books for All).

References

[Adler 72] Mortimer J. Adler and Charles Van Doren. *How to Read a Book*. Simon and Schuster, New York, 1972.

[AN02] ANSI/NISO Specifications for the Digital Talking Book, 2002. Available at http://www.niso.org/standards/resources/Z39-86-2002.html.

[Carriço 03] Luís Carriço, Nuno Guimarães, Carlos Duarte, Teresa Chambel, and Hugo Simões. Spoken books: Multimodal interaction and information repurposing. In *Procs. of HCII'2003, International Conf. on Human-Computer Interaction*, pages 680–684, Crete, Greece, 2003.

[DC] Daisy consortium playback tools. Retrieved January 18, 2006 from http://www.daisy.org/tools/playback.asp.

[Duarte 05] Carlos Duarte and Luís Carriço. Users and usage driven adaptation of digital talking books. In *HCII '05: Procs. of the 11th International Conf. on Human-Computer Interaction*, Las Vegas, Nevada, USA, 2005.

[Duarte 06] Carlos Duarte and Luís Carriço. A conceptual framework for developing adaptive multimodal applications. In *IUI '06: Procs. of the 11th International Conf. on Intelligent User Interfaces*, Sydney, Australia, 2006. ACM Press.

[Kaplan 05] Nancy Kaplan and Yoram Chisik. In the company of readers: the digital library book as "practiced place". In *JCDL '05: Procs. of the 5th ACM/IEEE-CS Joint Conf. on Digital libraries*, pages 235–243, Denver, CO, USA, 2005. ACM Press.

[Schilit 98] Bill N. Schilit, Gene Golovchinsky, and Morgan N. Price. Beyond paper: supporting active reading with free form digital ink annotations. In *CHI '98: Procs. of the SIGCHI Conf. on Human factors in computing systems*, pages 249–256, Los Angeles, California, United States, 1998. ACM Press.

[Schilit 99] Bill N. Schilit and Morgan N. Price *et al.* As we may read: The reading appliance revolution. *Computer*, 32(1):65–73, 1999.

[vO01] Jacco van Ossenbruggen, Joost Geurts, Frank Cornelissen, Lloyd Rutledge, and Lynda Hardman. Towards second and third generation web-based multimedia. *The 10th International World Wide Web Conf.*, pages 479–488, May 1-5 2001.