Time-based Hypermedia

Highly interactive multimedia structured documents with specific characteristics:

- Different media sources (e.g. video, audio, images, text)
- Anchors and Hyperlinks
- Powerful navigation capabilities
- Multi modal interaction capabilities
- Inherent timeline
- Synchronized presentation of content

Examples:

- SMIL documents
- XHTML+SMIL documents
- DTB documents

With real size documents (e.g. an entire multimedia book), lower quality arises in design and authoring processes from:

- High volume of synchronizing tasks between media items
- Producing different content versions for several playback platforms
- Meet publishing requirements
- Meet end user necessities and accessibility issues

Scaling up towards full multimedia collections (e.g. interactive encyclopedias or digital libraries), these problems are aggravated.

Requirements

Time-based hypermedia contents are inherently complex, therefore incapable of being produced manually, requiring:

- Automation of production processes
- Ease production tweaking aspects
- Deliver coherent content navigation, user interface and interaction capabilities

Production frameworks should focus on:

- Identify and classify multimedia units, to be reused in different authoring processes
- Provide clear modularity of contents (e.g. separate a book's content from its margin notes)
- Enable automatic, semi-automatic and manual content analysis techniques
- Ease prototype feature testing, as complex and novel navigation, interface and interaction mechanisms may appear

Source material integration must be handled according to:

- Clear separation and modularity of contents (including logical and semantic structures)
- Promotion of coherence amongst different usage settings of a content
- Identification of equivalent representations of a same content unit in different media sources
- Promotion of simple and reusable user interface specifications

All these requirements must be accomplished in tune with the different necessities of production time users:

- Top level users (whom may have little technology expertise) specify which content is going to be processed, and control collections batch processing
- Configuration managers must be able to specify which processing tasks are used to meet specific production requirements
- Developers create processing tasks and integrates them into the production framework

Abstract

We present DiTaBBu, Digital Talking Books Builder, a framework for automatic production of time-based hypermedia, focusing on the Digital Talking Books domain. As the production of recorded-speech-based Digital Talking Books is currently performed manually, with authoring tools, it becomes difficult to deliver large book collections to the public. Also, creating specific versions of books to each user profile becomes unfeasible through manual production, as different types of users have different types of needs (e.g., visually impaired people vs. children). With DiTaBBu, we enable the delivery of content in several playback platforms, targeted to specific user needs, featuring powerful navigation capabilities on the content. Also, with its flexible extension mechanisms, we can use DiTaBBu as a testbed for prototyping future capabilities.

DTBs

Digital Talking Books (DTBs) bring to the digital world analog talking books, one of the few ways to disseminate books to visually impaired people. With digitalization based on adding navigation to audio and text, ordinary tasks a common user is used to perform (such as searching a table of contents) are made available. DTBs can be also a base for creating rich DTBs, where multimedia contents and interaction capabilities are incorporated. Both DTBs and rich DTBs may be typically characterized by:

- Target users (visually impaired, partially sighted, children, students)
- Content features (audio, text, structure, navigation control, indices, complementary multimedia information)
- Interaction capabilities (bookmarks, highlights and annotations)
- Rich user interfaces (presenting multimedia contents synchronously, miscellaneous input devices)
- Diverse playback platforms (portable devices, PCs, Web browsers, SMIL players, DTB players)
The Framework

Raw content can be supplied in different formats, such as XHTML, DTB, plain HTML or even speech alignment generated content.

Structure Repurpose normalizes document structures into a Dexter-derived format (XHTML, XLink, timestamps and sync arcs), extracts new structures from content, and controls content units granularity.

Output Format transforms content structures into a specific reproduction format, such as XHTML+SMIL or plain SMIL, integrating different navigation structures and multimedia content accordingly.

Interaction integrates input devices into the content, such as mouse or voice, in the output format chosen previously. Even when the format does not provide mechanisms for a specific device, the content and its structures still can be optimized for it.

Presentation provides presentation profiles. These specify patterns for dimensioning, colouring, sync or sound. The chosen profile is then applied in the specified output format. Afterwards, playback platform tweaking is performed on the content.

Rich DTB is the result delivered by DiTaBBu, taylored to the different constraints specified by its configuration.

Implementing

DiTaBBu has been defined as an application of an XML pipeline processing architecture, APP. With this approach, separation of concerns is enforced in the definition of complex XML applications (thus meeting the requirements for DiTaBBu).

Each processing task in DiTaBBu has been defined as an APP component (written in XSLT stylesheets). Each DiTaBBu specification was mapped into an APP configuration.

Future Work

» More input formats support (such as PDF, Timed Text, etc.)
» Deliver content in other output formats (e.g. SVG, Braille, static text formats)
» Implementation of transcclusion mechanisms based on multimedia databases, document and fragment repositories, and ontologies
» Feature extraction capabilities for semantic indexing of multimedia items
» Creation of more presentation profiles that meet publishers' requirements
» Integration of time-based hypermedia design patterns in the automation process
» Conceptualize, design and implement a hybrid framework that meet requirements of taylored documents and adaptability based playback platforms

On the Web

DiTaBBu homepage, with some information and software download links can be found at:
» http://hcim.di.fc.ul.pt/postgrad/ditabbu/

The RiCoBA (Rich Content Books for All) project homepage can be found at:
» http://hcim.di.fc.ul.pt/ricoba/

The HCIM (Human Computer Interaction and Multimedia) Group homepage can be found at:
» http://hcim.di.fc.ul.pt/

LaSIGE (Large Scale Informatic Systems Laboratory) research unit, which integrates the HCIM Group, can be found at:
» http://lasige.di.fc.ul.pt/

The DAISY Consortium is responsible for DTB standardization and dissemination. DTB related specifications and resources can be found at:
» http://www.daisy.org/

Experiments

A SMIL document for handhelds, syncing voice, text and a table of contents.

An XHTML+SMIL document for PCs, syncing voice, text, table of contents and side notes, with integrated mouse interaction and sync guidance.