

A Tool for Creating, Editing and Tracking Virtual SMIL Presentations

Suzanne Little¹, Jane Hunter¹

¹ DSTC Pty Ltd, Level 7, GP South, University of Queensland,
St Lucia, Queensland, Australia 4072
{slittle, jane}@dstc.edu.au

Abstract. The ability to easily find, edit and re-use content adds significant value to that content. When the content consists of complex, multimedia objects, both the difficulty of implementing such capabilities and the added-value are multiplied. The work described here is based on an archive of SMIL presentations, built and indexed using tools, developed by the authors, which enabled digitized videos of lectures to be automatically synchronized with their corresponding PowerPoint slides and recorded metadata about the lecture context, origin of the video and slide files and the temporal alignment information. As the archive grew, it became clear that the ability to edit, update or customize existing presentations by deleting, adding or replacing specific slides without re-filming the entire lecture was required. This paper describes an application that provides such functionality by enabling the easy editing, repurposing and tracking of presentations for web-based distance learning.

1 Introduction

The use of digital, synchronised, multimedia presentations in education is increasing and as the volume of presentations available grows, so does the need for tools to facilitate the re-use and management of these presentations. Presentation editing tools, which permit updates, additions and concatenations of existing presentations, whilst maintaining synchronization and updating metadata, can improve the quality of the content and reduce the time involved in production and indexing. Simple functionality such as *cut*, *copy*, *paste* and *delete* would enable educators to re-use existing presentation segments, update presentations with more recent or topical material and customize presentations for particular audiences, without having to re-film the entire lecture each time.

This work is part of an ongoing project [1] in which a set of tools were developed to automate the synchronisation of digitized lecture videos with their corresponding PowerPoint slides. These tools, described briefly in the next section, enable the cost-effective, streamlined production of SMIL presentations and associated metadata for archival and searching. Such online presentations provide an extremely useful educational resource, particularly in a distance learning environment. However, as the size of the presentation archive grew, so did the need to edit, update or customize existing presentations by deleting, adding or replacing specific slides without re-filming. The economic benefits associated with such re-use are obvious. This paper describes the *PresentationEditor* application which was developed to satisfy this need.

In developing an editor for the synchronised multimedia presentations a number of requirements were identified. Firstly the editor should maintain the synchronization of the video and PowerPoint slides, automatically, without need for user input. Secondly the metadata created for the presentation must be updated to reflect the changes made, who made them and when they were made. The system also needs to record details of the composition of new 'virtual' presentations (i.e., presentations that are composed of edited segments of previous presentations, but which have never been presented in reality) to enable accurate, fine-grained (full-text or keyword) searching across the presentations. Finally it is important that the editor is easy to use and lightweight. No knowledge of SMIL or the underlying structure of the presentation should be required in order to use the application.

In the next section we describe previous related work and how the work described here differs from and builds on top of that.

2 Previous Work/Background

DSTC's SMIL Presentation Project [1] developed a suite of tools to cost-effectively create and index synchronized PowerPoint/Video presentations for deployment either on the Internet, CD-ROM or a LAN. The *PresentationLogger* records the time that each PowerPoint slide is displayed. The *PresentationIndexer* collects the metadata about the presentation in a semi-automated fashion and creates an XML file containing a complete description of the presentation, its

contents and structure. The *PresentationBuilder/Archiver* creates the finished SMIL presentation and stores the metadata as a separate stand-alone XML file or inserts the metadata into a database. An Internet search and browse interface [2] has also been built and can search either flat XML files or utilise the database for more efficient searching. Figure 1 shows the system components and architecture. In this paper we describe the *PresentationEditor* component, at the bottom of Figure 1.

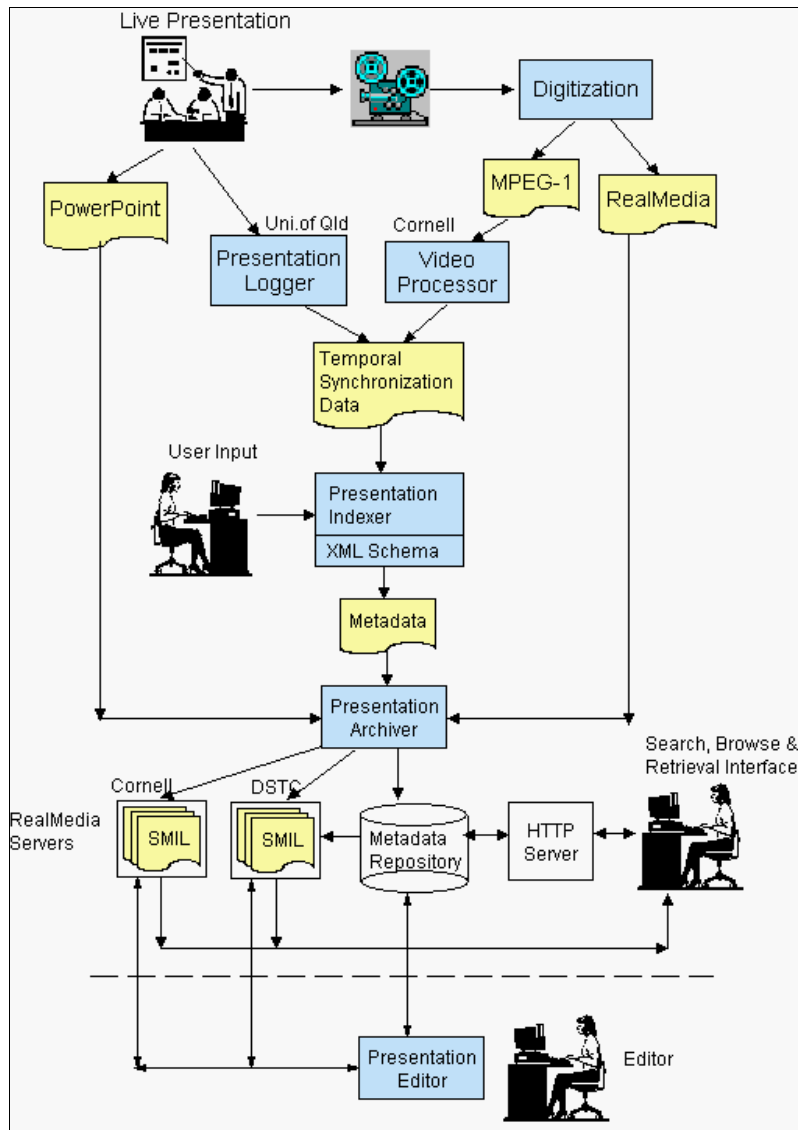


Figure 1: System Architecture

Synchronized Multimedia Integration Language (SMIL 1.0) [3] is a W3C Recommendation designed for choreographing web-based multimedia presentations which combine audio, video, text and graphics in real-time. It uses a simple XML-based mark-up language, similar to HTML, which enables an author to describe the temporal behaviour of a multimedia presentation, associate hyperlinks with media objects and describe the layout of the presentation on a screen. Since this work was started, the W3C SYMM (Synchronised Multimedia) Working Group has released the SMIL 2.0 Recommendation [4] which extends the functionality contained in SMIL 1.0 by enabling interactivity in multimedia presentations, advancing the timing model and improving the accessibility features. In the work described here we have used SMIL 1.0 to layout the presentations spatially and to temporally synchronize the video and PowerPoint slides. Possibilities for improvement exist as SMIL 2.0 is more widely deployed and more SMIL 2.0 tools/players become available.

SMIL has the following advantages to offer:

- Dynamic generation – similar to HTML this makes for easy, cheap alterations to presentations as opposed to complex and intensive digital editing.

- Platform independence – as a W3C recommendation, SMIL is not a proprietary technology and hence does not tie the implementation to particular platforms or programming languages;
- Network and client adaptability – SMIL provides a switch tag which can be used to dynamically choose the most appropriate media object to stream, depending on client display capabilities or connection speed e.g., use audio instead of video for low bandwidths;
- Ready availability of SMIL players – there are nine SMIL 1.0 players available covering a wide range of platforms [5]. The most popular of these are Apple's QuickTime 4.1, Microsoft's Internet Explorer 5.5 Browser and RealNetworks' RealPlayer 8.

Figure 2 illustrates the Replay interface for browsing and viewing SMIL presentations.

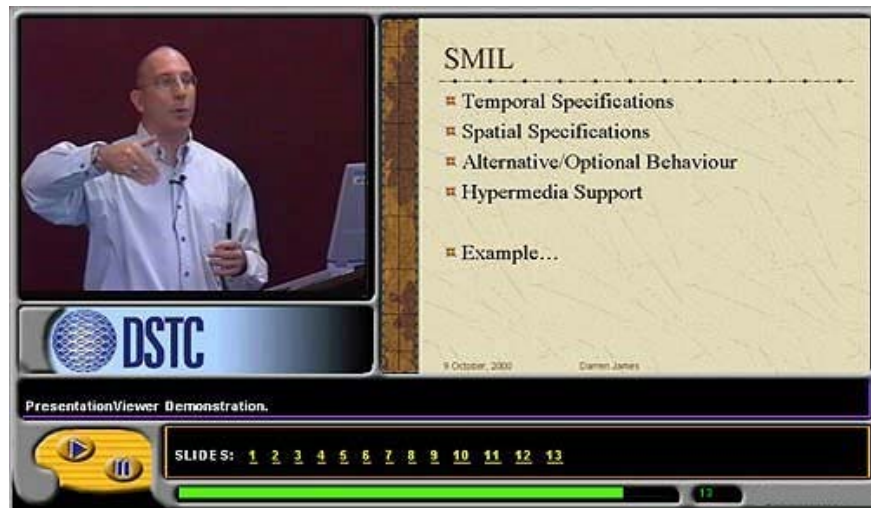


Figure 2: SMIL Presentation Replay Interface

A number of commercial and research tools are available for multimedia and SMIL editing. These include RealPresenter [6], RealSlideshow [7], Fluition [8] and GRINS [9]. RealPresenter and RealSlideshow are specifically for generating multimedia presentations while Fluition and GRINS are general SMIL editing tools. While the Real tools provide good support for generating basic multimedia presentations, they lack the metadata required for fine-grained searching and archiving of multimedia presentations or for recording changes or tracking versions. In addition, the presentation interfaces are somewhat restricted and the Slideshow tool, which has perhaps the most user-friendly interface, is limited to only providing audio with the slides. GRINS and Fluition are both very capable SMIL editors but assume a level of technical knowledge far above that which could reasonably be expected in this case. They also do not provide any support for complex and fine-grained or editing event metadata.

Hence the major goal of the work described in this paper was to develop a tool that provides a simple, intuitive, graphical user interface which supports presentation editing and a metadata model to enable both fine-grained searching and re-use tracking. The next section describes the system design and functionality.

3 System Description

The *PresentationEditor* is a Visual Basic application that exploits the temporal information and media links contained within a presentation's *metadata* (XML) file to divide the presentation into manageable self-contained segments which can be manipulated to implement easy editing. Each segment consists of a video component (video source, start and end times) and its associated metadata (pointer to corresponding PowerPoint slide, keywords, text etc.), expressed as an XML file. Using a graphical user interface, these segments can be cut, paste, reordered or moved within and between presentations to create new virtual presentations. Once the rearrangement of the presentation is complete, the editor generates the new RealMedia video and the new SMIL file which represents a synchronized presentation. The *PresentationEditor* also includes search facilities to help the user find relevant segments for insertion into the new presentation and previewing facilities for viewing draft presentations during editing.

Figure 3 illustrates a typical editing process in which the user creates a new virtual presentation by combining parts of existing presentations with new segments that have been filmed to replace outdated sections of the original presentation.

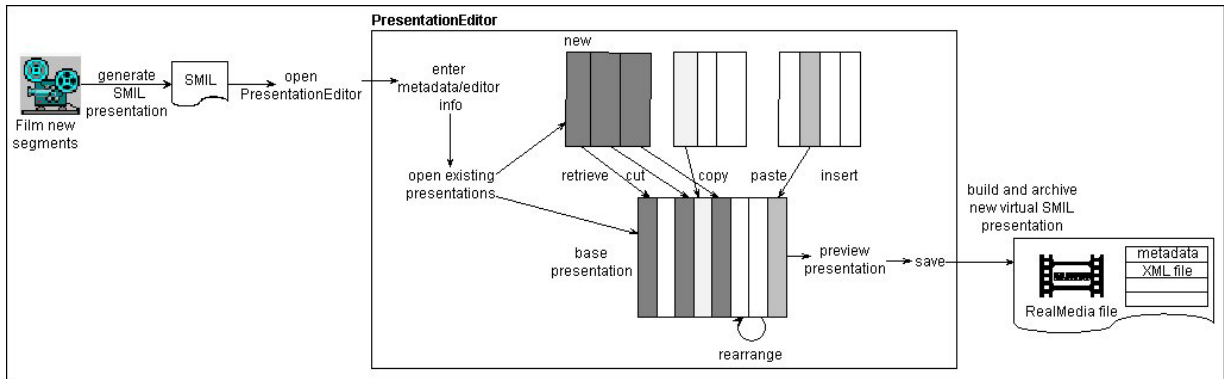


Figure 3: Steps in the creation of a virtual presentation

The PresentationEditor incorporates and integrates the following functionality:

- Search tools to find and retrieve relevant presentation segments;
- Ability to cut, copy and paste of segments both between and within presentations;
- Re-ordering of slide/segments through drag-and-drop;
- Direct updating of metadata for the presentation;
- Automatic adjustments to indexing and temporal synchronization information resulting from post-production editing of the video;
- Updating and correction to segment contents;
- Previewing of draft presentations during editing;
- Saving of updated metadata XML files;
- Re-construction of edited video files (RealMedia) and creation of new presentations (SMIL).

Figure 4 illustrates the graphical user interface to the PresentationEditor. The interface includes a video replay window, a PowerPoint slide window and an editing panel. In the example in Figure 4: extra slides have been added to the original presentation (“6: Metadata Standards” and “5: Metadata Requirements”); the original slide 6 has been deleted and the order of the slides in the original presentation have been rearranged using drag-and-drop tools in the bottom editing panel.

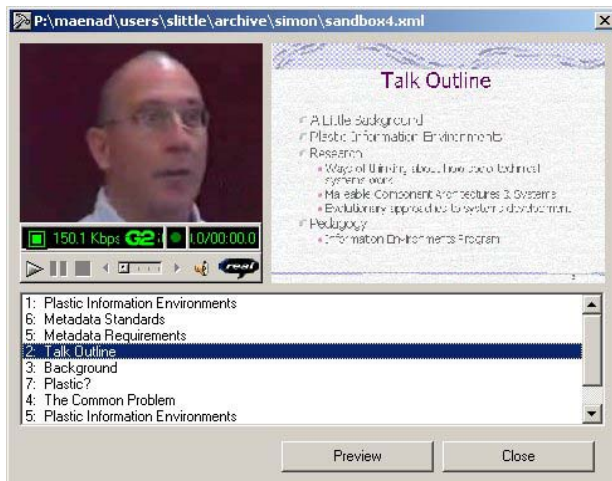


Figure 4: PresentationEditor Interface

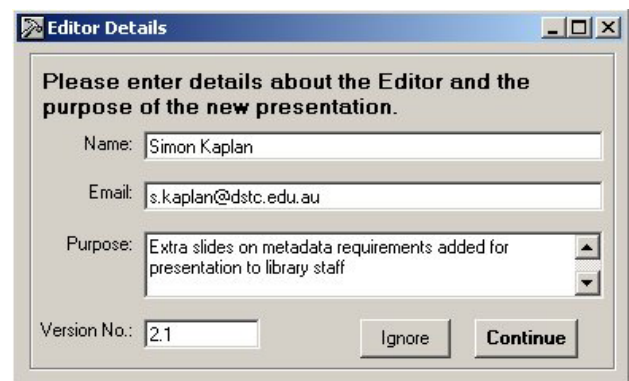


Figure 5: Editing event details form

4 Changes to the Metadata Model

A metadata model was developed for the original system, which enabled both coarse and fine-grained discovery of complete presentations and individual slide-segments respectively. The resource discovery of complete presentations was supported through metadata such as: *presentation location*, *lecturer*, *department*, *institution*, *description*. The search and retrieval of segments was possible through either full-text or keyword searches over the textual content of

each PowerPoint slide. The segment-level metadata consisted of: *source, keywords, text, start and end time, duration, related materials*. The XML Schema corresponding to the complete metadata model for the original system is available at [10].

The development of the *PresentationEditor* demanded certain changes to the original metadata model. The recording of additional metadata associated with “editing events” was required to ensure effective tracking of re-use and version control. This additional metadata included:

- details of the person who performed the editing (the editor’s Name and Email);
- when and why they did this (Date.Modified, Purpose);
- the version number of the new presentation;
- the source of new segments and their associated media objects.

Figure 5 shows the form displayed during the creation process for the new virtual presentation, to gather this additional metadata. The Date.Modified field is updated automatically by the application. The user is given the option to ignore this metadata update if they feel that there have been insufficient changes to warrant the presentation being defined as a completely new version or new presentation.

In addition, certain metadata associated with the original base presentation may no longer be applicable (e.g., presentation location) and may need to be deleted or modified. Changes to the “segment” metadata were also required. A “segment” is no longer purely a video component but is now a presentation component – containing both slide and video components which require two sets of associated temporal data. The slide component consists of a pointer to a PowerPoint presentation, the slide number within that presentation, an image location and the temporal location of the segment within the presentation. This timing information will change if the segment is relocated within the presentation. The video component consists of a source file, start and end times and duration. This temporal information identifies the section of video within the source video that should be displayed with the segment and does not change if the segment is moved. These alterations support the concept of a segment object that can be independently moved within and between presentations and therefore provides critical support for effective editing.

The changes to the metadata model described here demanded alterations to the *PresentationIndexer* application and the legacy metadata files in the existing archive. Because the metadata is expressed as XML files, these modifications were relatively simple to implement. Once the *PresentationIndexer* was updated, the legacy metadata XML files were rapidly regenerated. Alternatively the metadata files could have been easily regenerated through mechanisms such as XSLT or a simple script. The XML Schema corresponding to the new metadata model is available at [11].

5 Problems and Limitations

Whilst the segmentation approach has a number of advantages in this context, it may not be suitable for all lecturing styles. The indexing and synchronization of our SMIL presentations is based on the assumption that the video can be precisely segmented according to slide changes and that within each segment, the lecturer only refers to the currently displayed slide. However some presenters may refer to slides other than the currently-displayed one or begin an introduction to the next slide whilst the previous slide is still displayed. Such situations can lead to incongruities between slides and video content. Irrelevant and distracting information may be present or valuable and important information could be misplaced.

A lack of continuity between contiguous video segments recorded at different times or at different places, is an obvious problem that may occur within “virtual” presentations and distract viewers from the content. A presenter may be wearing different clothes, or have a different hairstyle between contiguous slides. Background and lighting may vary widely within a single virtual presentation. The extent to which this variation disturbs viewers and detracts from the learning experience needs to be assessed.

Originally the presentation incorporates a single continuous video but after editing and the rearrangement of presentation segments, the video consists of a collection of separate video segments in a particular order. The segmentation and reorganisation of the video components can be time- and resource-intensive and result in decreased video quality and reduced playback performance. Consequently the *PresentationEditor* provides two alternative options for generating the virtual presentation. The first option involves dividing and then recombining the video to create a single digital video object which plays in parallel with the synchronised slide images in exactly the same manner as the original presentations. The second approach involves the creation of a SMIL file which defines the start and end times of a sequence of video sections - giving the illusion of a continuous video file playing in parallel with the slide images. Both solutions offer advantages and disadvantages. Choosing to divide and recombine the video is extremely time consuming, CPU-

intensive and restricted by a number of factors including video frame sizes. It also requires a large amount of storage space and the quality of the final video is reduced. Never-the-less, the playback quality is reasonable with no loading or seek problems and there is a single video file associated with the presentation. Using SMIL timing to artificially rearrange the video file results in greatly reduced playback quality. The player must load the video file to the requested start time at the beginning of each presentation segment which can cause disruptions in the synchronisation. However this second approach is cheaper, faster, more flexible and less reliant on external factors such as video size or quality and, since the video file itself is not changed, the video quality remains the same.

The *PresentationEditor* operates primarily through the manipulation of flat XML files, either SMIL files or metadata XML files. The application does not support the direct editing of digital media objects such as the video files or Power-Point slides. While this functionality may be useful, it can also be complex and resource intensive and maintaining synchronization would be problematic. One of the advantages of the SMIL Presentation Tools is that they rely solely on manipulation of textual descriptions and thus avoid expensive and complex digital video editing.

6 Conclusions and Future Work

6.1 Conclusions

The *PresentationEditor* tool described here has the potential to increase the value of multimedia presentation content and enable more efficient and cost-effective production of distance learning resources by maximizing re-use, providing more effective metadata and facilitating ease of management. It provides educators with greater flexibility in the development, maintenance and management of a digital multimedia presentation archive - allowing them to improve, update and customize digital multimedia presentations for specific audiences and to provide a richer, more flexible means of learning, with a minimum of time, effort or cost.

6.2 Future Work

Our immediate objective is to deploy this application within a university department and to carry out user trials to obtain feedback from lecturers/content creators and students on the user interface and the application's usefulness as a tool for building and re-using effective online learning resources.

Future improvements in multimedia players and streaming technologies are expected to result in improvements to the SMIL presentation development tools, the replay interface and the *PresentationEditor*. For example, RealNetworks have recently released the RealONE player [12], which includes the ability to display related hyperlinks and references. Upgrading the system to SMIL 2.0, which provides greater interactivity and embedded metadata [13], should also result in further changes and improvements.

Finally, an area of research requiring more work is that of customized dynamic SMIL generation. In a manner similar to that of third generation HTML web pages [14], it may be possible to combine semantically-related mixed media objects to dynamically generate knowledge-enhanced SMIL presentations, which have been customized for particular users or a given set of constraints.

Acknowledgements

The authors wish to acknowledge the use of lecture material belonging to the University of Queensland and Cornell University and the generosity and cooperation of the presenters whose lecture content is contained within the archive: Bill Arms, Simon Kaplan, Susan Hamilton, Shazia Shadiq and Darren James. The authors also wish to acknowledge that this work was carried out within the Cooperative Research Centre for Research Data Networks established under the Australian Government's Cooperative Research Centre (CRC) Program and acknowledge the support of CITEC and the Distributed Systems Technology CRC under which the work described in this paper is administered.

References

1. Hunter J., Little S., Building and Indexing a Distributed Multimedia Presentation Archive using SMIL. European Digital Library Conference 2001, <http://archive.dstc.edu.au/RDU/staff/jane-hunter/ECDL01/ECDL01.html>
2. SMIL Online Lecture and Presentation Archive, <http://sunspot.dstc.edu.au:8888/smil/>
3. Synchronized Multimedia Integration Language (SMIL 1.0) Specification, W3C Recommendation, 15 June 1998, <http://www.w3.org/TR/REC-smil/>

4. Synchronized Multimedia Integration Language (SMIL 2.0) Specification, W3C Proposed Recommendation, 05 June 2001, <http://www.w3.org/TR/smil20/>
5. SMIL Players, <http://www.w3.org/AudioVideo/#SMIL>
6. RealPresenter, <http://www.realnetworks.com/products/presenter/index.html>
7. RealSlideshow, <http://www.realnetworks.com/products/slideshow/index.html>
8. Flution, <http://www.confluenttechnologies.com/>
9. GRINS, <http://www.oratrix.com/GriNS/index.html>
10. XML Schema for Original Presentation Archive, <http://archive.dstc.edu.au/RDU/staff/jane-hunter/ECDL01/ECDL01.html#AppendixA>
11. XML Schema for Editable Presentations, <http://sunspot.dstc.edu.au:8888/smil/edmedia/schema.xml>
12. RealONE player, <http://www.realnetworks.com/solutions/ecosystem/realone.html?src=rnhmfs>
13. The SMIL 2.0 metadata element, <http://www.w3.org/TR/smil20/metadata.html#edef-metadata>
14. Jacco van Ossenbruggen, Towards Second and Third Generation Web-based Multimedia, WWW10 2001