Interactive Multimedia for Dummies

Regina Gehne and Chris Jesshope

Hull University

r.gehne dcs.hull.ac.uk & c.r.jesshope@dcs.hull.ac.uk

ABSTRACT

This paper will describe some significant enhancements to a multimedia, authoring tool that has been developed for use by lecturers and teachers, rather than multimedia professionals. The tool uses a novel paradigm that eliminates time and hence any synchronisation in the development and editing of multimedia documents. Instead it uses the notion of a strict sequence of media elements. The media is packaged and compressed and is delivered by a browser plug-in to users. The media can be streamed over low-bandwidth modems making it suitable for delivery in any environment. Feedback from a large base of users of the existing tool has led us to develop a new version of this software that adds new capabilities, while maintaining the goals of a broad user base and an easy-to-use interface. The new capabilities include the introduction of interactive media and the incorporation of a richer set of media elements. The original tool produced linear presentations which are rather like a video on playback and we have added a hyperlink capability. The design and implementation of the interactive media has been a big challenge and the constraints that we had to meet were a requirement to maintain the streaming property of our presentations, in the presence of user choice, and to maintain the ease of use of the tools. The paper will discuss the choices made in our design. It will discuss the user interfaces and the presentation of the now more complex multimedia presentation and it will demonstrate the use of the tools and give examples of the new pedagogical techniques that can be used with the new tool.

Keywords

Multimedia authoring, interactive multimedia, low-bandwidth streaming, easy-to-use user interface

INTRODUCTION

The AudioGraph toolset has been developed to enable on-line teaching by providing the equivalent of face-to-face lectures as web-based multimedia documents. The goal of this project has been two-fold. To develop the tools and to experiment and evaluate their use in on-line teaching scenarios in a conventional university environment. That goal has been achieved and there are a number of publications on both our own and others' research using this tool, for example: Segal, (1997), Pearson and Jesshope (1988) and Jesshope(1999, 2000a, 2000b, 2001). This paper focuses on further developments to these tools based on our own evaluation and also feedback that we have had from upwards of 1000 registered users of the tool.

The AudioGraph software itself is the result of some 5 years of research on three campuses, Surrey University, Hull University and Massey University, see: Jesshope and Shafarenko (1997), Jesshope Shafarenko and Slusanschi (1998), Jesshope (1999), and Gehne and Jesshope (2000). AudioGraph can be downloaded from the NZEdSoft web site (http://www.nzesoft.com/) and the tools are available free of charge to anybody wishing to use them.

The methodology of teaching that the tools supports is very similar to what has been named Just in Time Teaching (JITT), by Novak and Patterson (1998). We also believe that the use of on-line multimedia in education not only liberates the students from time and geographical location, but also addresses issues of learning style. A growing body of literature in the impact of learning styles, e.g. Montgomery. (1998) and Felder (1987,1993), as well as end-user feedback, has driven the developments described in this paper. A summary of the results from Montgomery is still highly relevant and is reproduced below.

- 67%of the students learn best actively, yet lectures are typically passive;
- 57%of the students are sensors, yet we teach them intuitively;
- 69%of the students are visual, yet lectures are primarily verbal;
- 28%of the students are global, yet we seldom focus on the ``big picture."

Multimedia, especially interactive multimedia can overcome these barriers.

THE AUDIOGRAPH TOOLS

First we will give a general introduction to the AudioGraph tools, which comprises two pieces of software, an authoring tool and a player, which is a plug-in that enhances a web browser's functionality, enabling it to play the AudioGraph presentations. The authoring tool, the AudioGraph Recorder, is used for producing the multimedia content, which is a web-ready recreation of a teaching session. Unlike some other tools, such as Tegrity (http://www.tegrity.com/), AudioGraph records the presentations in the teacher s own time and not by capturing a live class. The results are very similar but usually more polished. The AudioGraph Recorder is available on both Macintosh and Windows platforms, but the authoring tool for the latter, is still at version 1 whereas the Macintosh authoring tool has been enhanced as described in this paper.

The key feature of the tools is their ease of use. They have been designed from the outset to be simple, with a clear and intuitive interface. This has meant reducing the number of concepts that the user is confronted with and of course this limits the capabilities of the tool, when compared to tools that are used by multimedia professionals. Thus AudioGraph, unlike other tools, can be learnt in an afternoon and requires little time to produce professional looking multimedia web sites. Typical preparation to presentation time vary from 2:1 to perhaps 10:1 for a complex, animated presentation(ref). This compares very favourably with professional authoring ratios, which are between 100:1 to 200:1.

AudioGraph makes use of a range of media elements, such as images, direct voice recording, vector graphics and pen annotation. These tools provide analogues of tools used in face-to-face education and hence the teacher immediately feels at home with the use of the tool. Images replace the slides used in a face-to-face class; the vector graphics and pen input the various drawing devices, such as whiteboard, blackboard, flipchart etc. and of course, spoken voice is the essence of a face-to-face presentation. Video is not supported as we believe it to be poor pedagogical value for the bandwidth required. Still or quasi-moving images can be used at the author's discretion but do increase storage and download requirements.

A key issue in the debate of enabling the lecturer with this technology, is that the experts who would otherwise produce the multimedia are often too far removed from the teaching area, as well as good teaching practice. This results in professional CDs or web sites that favour gloss, such as animation for its own sake, rather than sound educational content and the use good pedagogical techniques. If the teacher and media expert work closely together they can ameliorate this problem but this only increases the cost of the educational outputs, as there are now two people working on the production, an educator and a multimedia specialist. Again the cost of the multimedia becomes prohibitively expensive for all but the largest of audiences.

Another requirement of this project was that the media should occupy only a small space on the web server and, more importantly, be accessible to the students by modem.

AudioGraph has been used in a variety of ways and some examples of its use are illustrated below:

- the presentation of on-line material to both internal and extramural university students;
- the facilitation of on-line training of equipment and software;
- the replacement of video instructional courses;
- as a means of asynchronous technical communication in virtual organisations;
- as a means by which school children can present their study to their peers and to their teachers.
- as a means of teacher evaluation; and
- as a means of sending electronic greetings.

The AudioGraph Principle and Realisation

This tool is based on what we call the AudioGraph principle, which states that the media elements be arranged as an ordered collection and are played in strict sequence, regardless of playback timing. This is a simplifying principle that effects both authoring and playback. In authoring, there is no concept of time and synchronisation to worry about and in playback the presentation quality is independent of download speed and speed of the computer. No loss of synchronisation is ever seen, even in the presence of a slow internet connection or very old computers.

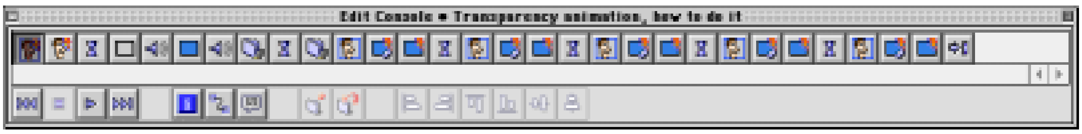


Figure 1. AudioGraph Recorder's Edit Console showing the sequence of media components - images, pauses, sounds and user drawn rectangles. Each icon corresponds to something heard of seen in the presentation.

Normally, complex programming over time is required when using a multimedia-authoring tool. There are tools that provide an exception to this, but these can not really considered to be multimedia authoring tools. For example PowerPoint allows different media elements to be placed on slides and these may be output to make web presentations. However, PowerPoint provides no functions or display for the sequencing or editing of these media components.

In the AudioGraph this is not the case, we provide a window, which controls precisely the sequence of components. Each media component is represented as an icon and the sequence of icons is displayed in a window called the *Edit console*, which is shown in figure 1. This is linked to the display window so that when an icon is selected in the control panel, everything up to and including that media element in time is displayed. Thus the *Slide window* can easily display different snap-shots of the presentation at different stages of its progress. The Edit console represents the sequence in time, from left to right. You can see the tape-recorder-style controls in the bottom left corner of this window for previewing the presentation. The slide window on the other hand, shows exactly what the learner will see in the presentation when viewed in the web browser. Thus the unique difference between the AudioGraph recorder and other presentation packages, is the ability to see and edit any slide in the presentation at any point in its delivery, between its start and end.

IMPLEMENTING INTERACTIVE MULTIMEDIA

This paper is primarily concerned with how we have implemented interactive media in the AudioGraph. Current AudioGraph lectures can be interactive, but only because of the environment that they are presented in. AudioGraph lectures comprise html pages with the AudioGraph slides embedded within the web pages as links. The AudioGraph slide itself is a strict sequence of media elements that has a start and a finish. The standard output from the tool is a single index page, with one link to each slide in the presentation. The problem with this is that to generate anything more than a linear sequence of material requires the author to also edit the html pages, which is possible but adds to the design time.

Our goal therefore was to create AudioGraph presentations with non-linear characteristics, so that with different input from the person viewing it, the output would follow different paths through the presentation. This means that there are certain choice points and that at these points there are links to different parts of the material, which will depend on the user's input. There are a number of design issues in creating a tool that would do this. A number of questions have to be asked:

1. what is it that is being linked to? Html links are to pages or to anchors within a page.
2. When and how should the links be active? Remember a multimedia presentation has a time element.
3. How can we deal with streaming in the presence of the user's choice?
4. Finally, how are these links represented and displayed in the authoring tool?

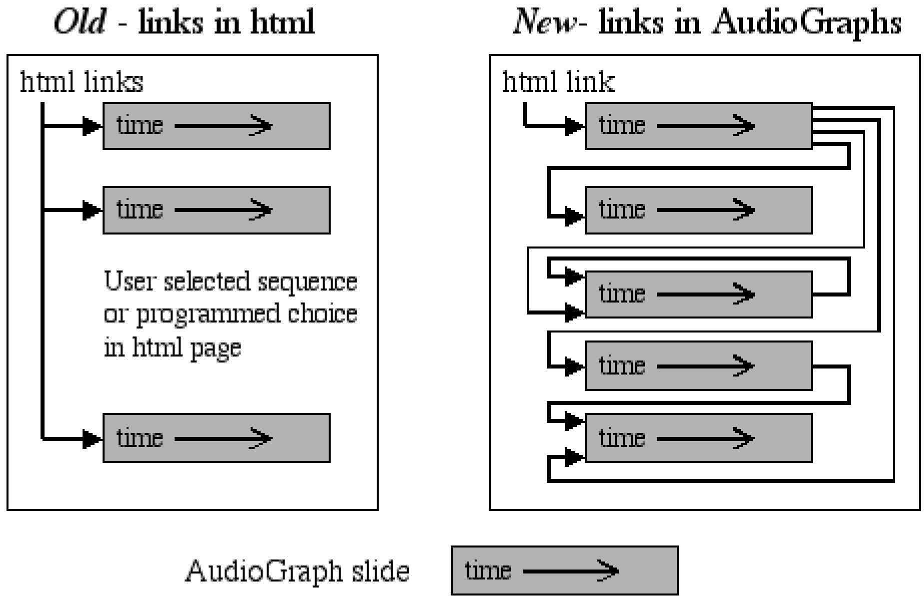
Editing and file model for links

In AudioGraph a presentation is a set of linked files, in version 1 the linking was in html. Each file is a sequential presentation, which was played by the AudioGraph plug-in. Similarly, each episode or slide in the presentation is an editable unit represented by the combination of edit console and slide window, see figure 1. On export to the web, each slide produces one .html file and one embedded .aep file of MIME type application/vnd.audiograph. It was important that we did not move too far away from our current editing model, which is simple and intuitive. Keeping this same file and editing model therefore was our first approach and we found that it provided answers to both points 1 and 3 above. That is links would be to episodes only and a users choices selects the next file to stream to the them.

In version 2 therefore we maintain the episode as the unit of presentation and an episode is still sequential. What we add are links in the presentation and those links can be to any other episode (including itself), but only to the start of those episodes. The user's choice by mouse clicking on a linked area in the presentation, will determine which of potentially many links are taken. Streaming is solved because there are no non-linear paths within a presentation, that are non-deterministically chosen. The presentation is linear and has choice points anywhere within it that exit that episode and link to another.

In fact a linked presentation with AudioGraph links is still a linear presentation and is shown in the plug-ins progress bar as such. It is more accurately the union of a set of non-deterministically chosen episodes, committed by the user's choice. It is also dynamic, in that the user can go back and re-commit the choice and produce a completely different linear sequence. The is only wasted bandwidth when a choice point is made early in an episode and the user exits before of it is all downloaded. This can be avoided however, by splitting the episode at the choice point. The linking strategy for both version 1 and version 2 links is illustrated schematically in figure 2. Notice that in the new version we have a much richer potential for linking and moreover it can all be completed within a single tool, the AudioGraph recorder. Notice that a link may be to itself, and may be taken automatically, which allows for loops in presentations. A model of an AudioGraph presentation is therefore a directed graph (which may contain

Figure 2. Link model in version 1 and 2 of the AudioGraph Recorder.



loops) of AudioGraph episodes, which are the nodes in the graph and choice is made in each episode by selecting one of the arcs that starts at that node.

Authoring model for links

By design, therefore the is no significant change in the authoring model, as each episode is still sequential and is represented by the combination of Edit console and Slide window. To answer questions 2 and 4 above however, we need to consider more carefully what a link is. To that end we define a link to be an attribute of any graphical object, or collection of objects that have been grouped together. The activation area of the link is the drawn area of the object or the union of all of the drawn areas of all objects in a group. The default link tool creates a transparent rectangle and allows the author to add a link to it; this has no visual representation on the screen but simply defines an area, which activates the link. The link itself can be either another AudioGraph episode within the same presentation, or a standard URL, which could be an external AudioGraph episode or anything that can be represented in an html page. Links may be opened in the same window, in which case the episodes are concatenated together, or can be opened in a new window, rather like a pop-up window. This provides for a number of different pedagogical styles of authoring. Pop-up presentation can be used, for example to give more detail on a topic and links in the same window can be used to provide alternate pathways based on user preference or ability, as testing may be an element of the choice.

An html page which has no sequencer, without scripting and hence links are active as soon as the page is displayed. AudioGraph presentations however, have a default sequencer built into the plug-in. The issue of when and how links are activated must therefore be considered. We decided that links would only be activated when the presentations is stopped. The presentation can be stopped by clicking anywhere on the screen or the start stop button. When stopped, a link's active area is shown a change of cursor (it becomes a finger pointer) and mouse clicking on it will continue the presentation with the episode it links to. Mouse clicking in a non-active area will restart the presentation; at the beginning, if the presentation has already reached its end. We have also introduced a stop tool, whose action on playback is to stop the presentation and hence activate any links that may have been placed at that point in the presentation.

DETAILS OF AUDIOGRAPH RECORDER VERSION 2

Links display at the Lecture level

The AudioGraph recorder has a lecture window, which provides a view of the complete presentation. Two views are available in version 1, a text view and a thumbnail view, the former giving a textual view of what will be presented on the index page, complete with episode durations and the latter giving a graphical index to the different slides. In version 2, a "Links" view has been added to the Lecture Window. This currently just displays the links between episodes as a matrix of directed arrows, although our intention has been to provide link destination editing from this view as well. This will be added later. The links view of the Lecture window is shown in figure 3, it is now a resizable window.

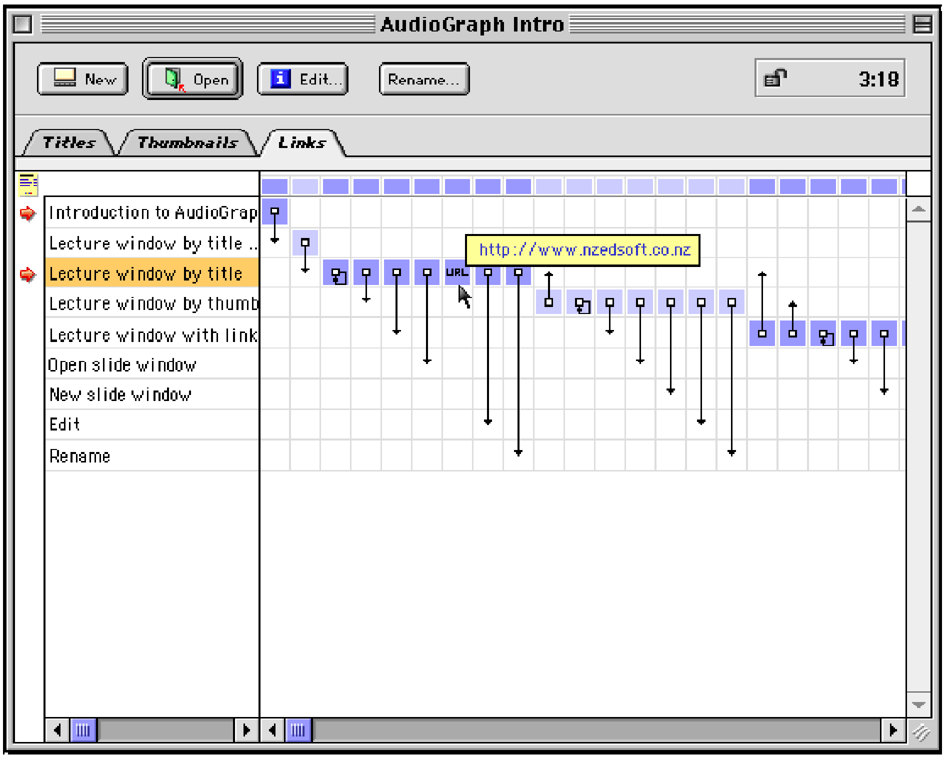


Figure 3. The links view of the lecture window, showing both URL and internal links between episodes

The colour coding of the cells is to help seeing which link belongs to which slide. Slide titles and links list can be independently scrolled horizontally. If the slide links to an URL, the URL gets displayed when the cursor is over the URL field, as illustrated.

Finally the arrows in the index column to the left of the slide titles indicate which slides should be placed on the HTML output page. In this case there will be just two entries, as the other episodes are all reachable by AudioGraph links.

Text Tool

In version 1, text could only be created by copying and pasting from another application. Version 2 enhances the media AudioGraph produces by providing text component. It is a hybrid between a rectangle and an image. It has the visual attributes of a rectangle, i.e. colour, line width, transparency, filled/unfilled etc., but gets drawn and erased in the background or image layer, whereas rectangles are drawn and erased in the foreground or annotation layer. Additionally, the text itself has its own attributes like size, font style, colour and justification. Text font, size, style, colour and justification within its rectangle (left, centred, right) can be edited in the Visual Attributes Editing Dialogue.

Text gets exported as a PNG image for display by the plug-in. This solution is chosen to mitigate any problems with a user's installed fonts. As a bitmap, the text is displayed and aligned exactly as it is in the Recorder, regardless of whether the fonts used for creating the text are present on the viewer's machine or not. The only other way to ensure this alignment would be to export fonts, which is expensive in file size even with scalable fonts. PNG compression does a good job of optimising the image files as we support the full standard, which supports bit depths, from 1 to 32 bits, with transparency in both indexed and alpha channel images.

On the Macintosh, there is an option in the Appearance Control Panel to turn anti-aliasing on for fonts bigger than a specified size. This gives the letters in a font a smoother appearance by gradually blending the text colour with the background colour. For big fonts this makes a real difference to the text. We have a similar option in the Visual Attributes Editing Dialogue now, called "Smooth text drawing", which turns anti-aliasing on before drawing, no matter if this is turned on in the Appearance Control Panel. The catch is that this results in bigger exported text, because the PNG image will need to store more transparency information. A non anti-aliased text component requires at most 2 bits per pixel, this represents black, white, the text colour and background colour. An anti-aliased text component would have a requirement for perhaps 16 levels of transparency in the colour index table for the background colour, to create the smooth effect on the jagged edges, requiring at least 5 bits to represent each pixel and more than doubling the file size.

To get a feeling for file sizes, a typical bullet point on a slide requires about 2.5 KBytes for non aliased text and 6.5Kbytes for anti-aliased text, which is just seconds of download time on a modem. For comparison, a minute of speech requires 100Kbyte.

Both images and text can be persistent across slide boundaries during a link and again this feature is implemented in order to minimise download time. A large detailed image may require up to 100KBytes ,This can also be switched on in the Visual Attributes Editing Dialogue.

RECORDING SOUND

Sound quality

The sound recording interface and algorithms have been changed. Version 1 of AudioGraph used only GSM compressed sound for exported presentations and only PCM encoded sound in the lecture document. GSM requires 13 Kbps or about 100Kbytes per minute of recorded sound. PCM requires much more (depending on the sampling rate). For example at 44KHz, a mono recording requires 2.6 Mbytes per minute of recorded sound. This feature had two negative effects:

- File sizes were very large for lecture documents
- Exporting presentations was very slow because every sound component had to be converted from PCM to GSM, every time a file was exported.

This has been changed in version 2 and it is now possible to set the export quality and choose speech quality (GSM), music quality (Ogg) or linear samples (PCM). This setting will determine how the sound is stored in both the lecture document and the exported presentation. In the first release of version 2, exported presentations will only support speech quality sound but the lecture document will support linear samples and speech quality.

Sound qualities supported or to be supported

PCM quality

PCM stands for pulse code modulation. The sound is sampled at a specified rate, say 44 thousand times a second, which is CD quality. What is stored then is one sample for each sampling period. The samples are normally 2 bytes each. This is called PCM. PCM is currently supported in the lecture document and will soon be supported in the exported presentations.

Music quality

There are a number of music quality compression algorithms, perhaps the most popular is MP3. These compression techniques use the human hearing characteristics to remove redundant information from the data. Although not yet implemented we intend to implement an Ogg-Vorbis compression scheme (http://www.ogg-vorbis.com/). This supports a variable-rate, music-quality compression of sound. At high stream rates 64 to128Kbps there is little qualitative difference between this and CD-quality PCM, and yet the data required is 5-10 less. It can be downloaded on a high-speed modem connection. Ogg is not yet supported.

GSM

GSM is a sound quality compression scheme. It is based on a model of the human vocal tract and compresses speech well but not music. It uses only a limited sampling rate 13KHz and requires 50 times less data than CD quality PCM. It can be downloaded on a low-speed modem connection. GSM is now supported in both lecture documents and exported presentations.

Conversion between sound qualities

The sound parameters, such as quality, sampling rate, etc. are set in the Preference dialogue, and these apply to all sounds recorded until the parameters are changed. It is also possible to edit the quality of an individual sound annotation in the Edit Console, by double-clicking a sound. This editing will determine the sound quality for any subsequent saving of that annotation. It will also allow the annotation to be re-recorded with the new parameters.

Without re-recording, a change from a low to a higher quality is ignored when saving, because there is not point in using more data to store a lower quality sound annotation. However a change from a high quality to a lower quality will take effect on saving, even without re-recording the sound. This will reduce the sound's quality and also reduce all file sizes. The key point is that compression reduces the data required to store sound and also reduces the quality. Once that quality is lost, it can not be recovered by re-encoding the data, only by re-recording the clip.

Volume activated detection of sound

Volume activated detection (VAD) of sound has also been introduced in version 2. What this means is that it is now possible to select the recording tool and to compose your speech with both natural and inadvertent pauses and to have only the active speech recorded as sound and the pauses encoded as pause annotations of a given duration. Before, this would have all been encoded as one sound annotation, including the silence, which would require between 100K bytes to 2.5 Mbytes per minute to encode. A pause annotation requires just a few bytes for whatever length of pause. VAD works with both speech quality sound and linear samples.

On-the-fly compression

The recording tool can now compress sound as it is being recorded. This feature uses the time you spend speaking to convert the PCM samples to speech quality sound if speech quality is selected. Because of this, there is no conversion when the lecture document is saved or when the document is exported to a presentation.

VAD sensitivity adjustment

A great deal of empirical evaluation effort has been put into simplifying the VAD parameters. On initialisation the VAD recorder measures the ambient noise in the environment, it uses this to decide when to start recording sound or when to express the silence as a pause. The sensitivity of silence detection can be adjusted. The VAD recorder has a number of parameters that determine how it works, these include triggering threshold, pre and post buffers to gain continuity before and after triggering and these have all determine the quality and sensitivity of detection. These have been linked to the single slider control labeled "VAD sensitivity". The sensitivity adjustment procedure therefore, is to start with the setting to the right, with high sensitivity, and to reduce the sensitivity until all of the sound you want to record is captured. Once set it is very stable for the same equipment.

CONCLUSIONS

We present in this paper, extensions to an existing tool that has already been proven to be effective at providing on-line teaching content. The enhancements implemented have been suggested by users of the tool but their implementation has been designed so as not to lose the original ease-of-use of the AudioGraph. One new concept has been introduced, which is one that most users will have familiarity with, that is the hyperlink. Hyperlinks have been introduced quite naturally into the Recorder tool, by allowing any visual object to have a link attached to it. That visual object's drawn region then becomes the activation area for the link. Links are only active on playback, when the presentation is in the stop state and links can only be to the start of AudioGraph episodes or slides. In this way presentations with an arbitrary non-linear flow can be streamed file by file from the server with the user only downloading those components of a presentation that they select by their choices in taking links. Thus we retain the second key feature of AudioGHraph on-line teaching, low-bandwidth downloads.

The paper also describes some other features introduced into the AudioGraph, including volume activated detection of sound, which records only active speech and encodes silence much more efficiently as parameterised pauses. One of the challenges in this implementation was defining a user interface and set-up procedures to make the use of this sophisticated user programmable for environmental and technical parameters, such as microphone sensitivity etc. This has been achieved with a simple environment test and a slider control, which has been set up to change the three major parameters based on an empirical evaluation of a range of systems.

ACKNOWLEDGMENTS

We gratefully acknowledge the support for this work under the TILE project funded under New Zealand's NERF funding scheme (http://www-tile.massey.ac.nz).

REFERENCES

Felder, R.M. and L.K. Silverman (1988). ``Learning and Teaching Styles in Engineering Education," Engineering Education, 78 (7), 674-681, April 1988.

Felder, R.M., K.D. Forrest, L. Baker-Ward, E.J. Dietz, and P.H. Mohr (1993). ``A Longitudinal Study of Engineering Student Performance and Retention: I. Success and Failure in the Introductory Course." Journal of Engineering Education, pp. 15-21, Jan. 1993.

Jesshope C. R. and Shafarenko A. (1997) Web Based Teaching: a minimalist approach, *Proc. Second Australasian Conference on Computer Science Education*, ISBN: 0-89791-958-0, pp16-23, (Association for Computing machinery Inc.).

Jesshope, C. R. Shafarenko A. and Slusanschi H. (1998) Low-bandwidth multimedia tools for web-based lecture publishing, *IEE Engineering Science and Educational Journal*, **7 (4)**, pp148-154., also published in: *IEE Computing and Control Engineering Journal,* **9 (4),** pp156-162 and on-line at: C. R. Jesshope, A. Shafarenko and H. Slusanschi (1998) Low-bandwidth multimedia tools for web-based lecture publishing, IEE Computing Forum, http://forum.iee.org.uk/forum/library/, September 1989.

Jesshope C. R. (1999) Web-based Teaching - Tools and Experience, *Australian Computer Science Communications*, **21,** (1), pp27-38, ISBN 981-4021-54-7, Proc Australasian Computer Science Conference, ACSC99, Auckland, Jan 1999, (Springer).

Jesshope C. R. (2000a) The use of streaming multi-media in microelectronic education, *Microelectronics Education,* Kluwer Academic (London), ISBN 0 7923 6456 2, pp45-48.

Jesshope C. R. (2000b) The use of multi-media in internal and extramural teaching, *Proc Lifelong Learning Conference*, Central University of Queensland (Brisbane, Australia), ISBN 187 6674 06 7, pp257-262.

Gehne R. and Jesshope C. R. (2000) Tools for the production of small-footprint, low-bandwidth, streaming multi-media for distance education, *Proc Lifelong Learning Conference*, Central University of Queensland (Brisbane, Australia), ISBN 187 6674 06 7, pp240-244.

Jesshope C. R. (2001) Cost-Effective Multimedia in On-line Teaching , in Educational Technology & Society 4 (3) 2001ISSN 1436-4522 (http://ifets.ieee.org/periodical/vol_3_2001/jesshope.html)

Montgomery, S. (1998) Addressing Diverse Learning Styles Through the Use of Multimedia, (1998). Available: http://www.vpaa.uillinois.edu/tid/resources/montgomery.html

Pearson M. and Jesshope C. R. (1988) Multi-campus teaching using computer networks, *Proc. of the Third Australasian Conference on Computer Science Education* pp 106 - 111, July 1998 (Association for Computing machinery Inc.).

Segal, J. (1997). An evaluation of a teaching package constructed using a web-based lecture recorder. ALT-J, 5 (3), 32-42.

Novak, G. & Patterson, E. (1998). Just-In-Time Teaching: Active Learner Pedagogy with the WWW. Paper presented at the IASTED International Conference on Computers and Advanced Technology in Education, 27-30 May 1998, Cancun, Mexico,