TOOLS FOR THE PRODUCTION OF SMALL-FOOTPRINT, LOW-BANDWIDTH, STREAMING MULTI-MEDIA FOR DISTANCE EDUCATION

Regina Gehne and Chris Jesshope
Massey University
Private Bag 11222, Palmerston North, New Zealand

ABSTRACT

This paper describes the AudioGraph project, a collaborative project between Surrey and Massey universities in the development of multi-media authoring tools. The tools are for use by academics, rather than multimedia professionals, for the production and delivery of course-ware via the world-wide web. The goal was to place the production of the media firmly in the hands of the educators rather than the technologists. It describes the current tools and our future plans for them.

INTRODUCTION

This paper introduces tools that have been developed for the production of web-based, multimedia, for on-line teaching and training. The tools comprise an authoring program called the AudioGraph Recorder and plug-ins for web browsers to play back the recorded presentations. The goal of these tools was to enable open and lifelong access to learning, through migration of presentation to the world-wide web. This is a collaborative project between Surrey University (UK) and Massey University (NZ). The results from a previous phases of this project have been described in Jesshope and Shafarenko (1997) and Jesshope, Shafarenko, and Slusanschi (1998). This paper describes the tools and their planned further enhancement, following extensive field trials (Jesshope 2000).

The fundamental philosophy behind the design of the AudioGraph tools was always to assume that they would be used by the academics, teachers or trainers, rather than by multimedia professionals, who are already catered for by a number sophisticated and complex authoring tools. This has required us to analyse the basic requirements of on-line educators and provide an intuitive interface that allows these non-professionals to produce professional looking presentations very rapidly. Our aim was to place the production of the multimedia firmly in the hands of the educators rather than the technologists. For more background on multimedia authoring and an extensive catalogue of tools available, the reader is referred to Silgar (1999).

At the outset of this project our goals were as follows:

- to target the delivery of multimedia course-ware via the world-wide web;
- to provide simple-to-use tools capable of producing multimedia presentations with little or no experience of multimedia editing;
- to significantly reduce the industry norm of 200 hours of preparation time for every hour of multimedia presentation;
- to ensure the presentation size was small enough so that institutions could

provide a large corpus of multimedia tuition on a modestly sized server;
- to ensure that the multimedia presentations could be accessed over modest modem connection speeds (e.g. 14K) without any significant delay;
- to provide cross platform delivery of the multi-media content using a browser plug-in.

It will be shown that all of these objectives have been met and this has been a major factor in the AudioGraph being adopted by a number of universities around the world to produce on-line, multimedia course-ware. The remainder of this paper provides an overview of the tools, a detailed look at the media elements and how they contribute to the goals of small foot-print web sites and low-bandwidth down-load. Finally, the paper takes a detailed look at the automatic generation of web sites, which includes both html and Javascript code.

AN OVERVIEW OF THE TOOLS

Inevitably, in order to achieve the broad range of goals outlined above, a number of compromises have had to be made in the design of the tools. In order to put this in context, Silgar's classification of multimedia authoring tools (1999) is considered. This taxonomy identifies a number of factors: the media types available to the author, the delivery mechanism, the programming model supported and the user-interface. Using this classification, the AudioGraph system can be defined as a multimedia authoring tool that supports compressed audio, images and vector graphics as its media elements and generates presentations for web-based delivery, which are strictly sequential, using an iconic interface. One of the obvious compromises that was made, was not to support video as one of the media elements; clearly this was done in order to satisfy the requirements of a small footprint web-site and low-bandwidth streaming media. Other compromises have been made in the complexity of the programming model, which has been kept deliberately simply (strict sequence only) in order to satisfy the goals for shallow learning curve and rapid development of the multimedia presentations. Results from using the tools in extensive field trials are presented in Jesshope (2000) which demonstrates that all of these goals have been met in the tools.

The key difference between this and other multimedia authoring tools is the so called AudioGraph principle. Normally a multimedia presentation comprises a set of objects that each have a fixed time at which they appear, disappear etc. and a lot of the effort in authoring is taken up in synchronising or coordinating the elements. The AudioGraph principle on the other hand determines the presentation to be a strict sequence of media elements (and their virtual anti-element, which vanishes them). Some of the media elements may have some real time semantics, such as a fixed pause or a sound byte, while others, like the rendering of an image or vector graphic component, have no real time meaning but are rendered at the speed of the playback computer, which may be very different on different generations of computers. Whatever the playback speed, the media file structure, which follows the AudioGraph principle, maintains the strict sequence and hence the correct semantics of the presentation.

The editing interface

The main editing interface that implements the AudioGraph principle is a pair of windows: the slide window, which shows the presentation as it will be at any point in the presentation sequence and the edit console, which represents the actual sequence of media elements in iconic form. The user can select any icon within the edit console and the slide window displays a static presentation as it would appear when that media element has been displayed. Figure 1 gives a demonstration of this duality in the interface.

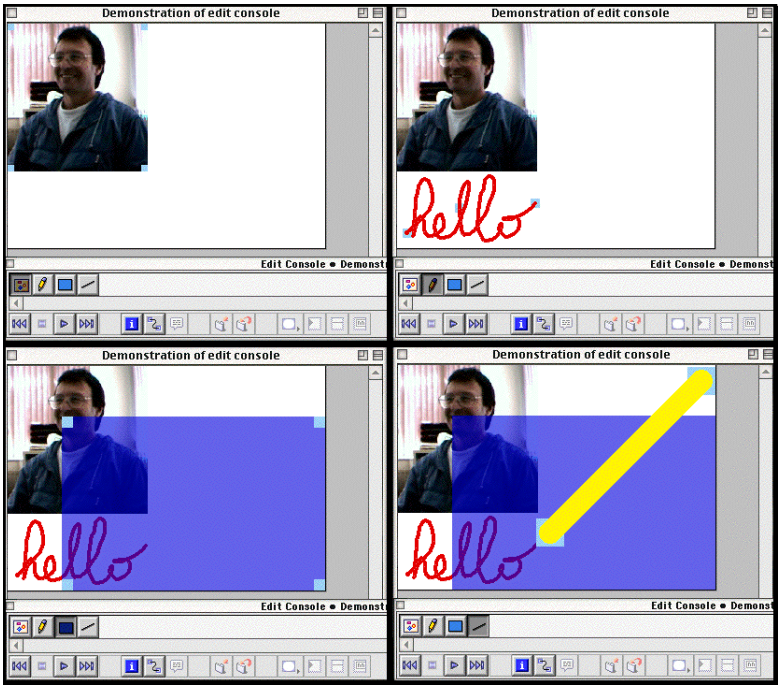

Figure 1 A composition of screen shots showing the edit console and slide window which illustrate the AudioGraph sequence and in which the correspondence between the iconic and actual representation of each media element is shown.

In this diagram notice the iconic representation of the four media elements in the edit console, namely an image, a freehand vector graphic element (the handwriting), a rectangle (notice the transparency) and a straight line. Elements can be selected in either window but if selected in one, they also appear selected in the other. Notice that only those elements up to and including the icon selected in the edit console are displayed in the slide window, where the selected elements(s) have end markers. Any selected element can be picked up and moved in either window. Thus, the edit console shows the presentation as a sequence of media elements, which can be moved around in the temporal axis that it represents and the slide window always shows a corresponding snapshot of the spatial representation of the presentation.

Notice also that the edit console has a number of controls, which are for previewing the dynamic presentation. These navigate, play and stop playback. The other controls are used for editing and grouping together the media elements.

The tools

Figure 2 shows the tool bar, which is used to select an appropriate tool to generate the media elements of the presentation.

Fig. 2

From the top left there are various vector graphic tools, namely freehand line - used for handwriting, straight line, rectangle - open and filled, ellipses and finally, arcs. The next two tools are a highlighter and an eraser, then comes the image placement tool and the link to url tool, then the scroll and the text tools and finally the sound byte and pause tools. The remainder of the window provides some tools for editing the attributes of either the tools themselves or the media elements that have already been placed. If elements are selected, then it is these that are edited, if not, then the attributes of the tool are edited. In this window, there are tools for editing colour and line thickness, which are the most commonly used attributes. There are other attributes and these can edited from a separate window which can be displayed at any time for that purpose.

Most tools are just selected and used, by drawing in the slide window. The more complex tools, like sound and image placement etc., bring up their own dialogue. Detailed documentation of the tools can be can be found at the NZEdSoft web site, from which the tools may also be down-loaded (NZEdSoft 1999).

THE MEDIA ELEMENTS

Images

The normal basis for a presentation using the AudioGraph tools is a set of presentation graphics, such as might be produced by a presentation package like PowerPoint. The slides are input to the AudioGraph as images, using a container file (Scrapbook file). PowerPoint directly generates scrapbook files of images as one of its standard output formats and there are utilities such Print-2-pict (NZEdSoft 1999), which do this as a printer driver, for when an application does not support the scrapbook format. On reading a scrapbook file, the AudioGraph will generate a number of slides to be annotated with the media elements, each of which will already contain the image of that slide. This is not the only way that a presentation can be constructed and Jesshope (2000) outlines and gives examples of a number of different development techniques. Images can be imported from a screen capture, or a cut or copy operation from another application, via the clipboard. Moreover, any format of image supported by QuickTime, can be read directly from a file.

When generating web presentations the AudioGraph supports only the PNG image format. This is a relatively new and open standard for web images (Roelofs, 2000), which combines all of the advantages of both the GIF and the JPG web formats. PNG stands for Portable Network Graphics. Like JPEG this format supports up to a 32-bit representation of an image pixel; GIF on the other hand supports only an 8-bit pixel or up to 256 colours. Moreover, like GIF (but unlike

JPEG) the PNG format supports loss-less compression, which means that the image that is regenerated is exactly the same as the image that was compressed and transmitted over the network. Another advantage of the PNG format is that it supports an alpha channel, which provides for variable transparency of the image on a pixel-by-pixel basis. In fact the PNG format is a very flexible and well designed standard that uses no more bits than is required to represent the colours in the image and has a very effective compression algorithm that includes filtering to optimise the compression. It is one contribution to the small-foot-print AudioGraph web sites.

It has already been shown, in figure 1, that the AudioGraph supports transparent vector graphics. When this is combined with the pixel-by-pixel transparency of the non-vector, PNG graphics it does not take much to imagine that very sophisticated results can be achieved in a presentation.

Sounds and pauses

As video has not been supported as a media element in the AudioGraph, the most data-intensive media element is sound. Fortunately there are very efficient compression algorithms for transmitting sound over a network. This is especially true if, as is expected, the sound input is going to speech. The AudioGraph uses the GSM compression technique (Scourias 1999), which is optimised for speech. An uncompressed speech stream of 16-bits accuracy and sampled at 8KHz would require 128 Kbits/sec of bandwidth for transmission over the network. When compressed using GSM, the 16-bit/sample speech would require only 13.2 Kbits/sec thus providing a compression ration of 128/13.2 = 9.7 or approximately 10 to 1.

Further savings are made by not transmitting a continuous stream but dividing the stream into sound bytes, which can be timed if necessary by adding pauses of arbitrary duration. In future versions we are planning to use a continuous voice activated recorder that breaks down the continuous stream into small sound bytes and timed pauses using a voice activated detector or VAD.

In the current version, sound can be recorded directly, using the sound tool and a microphone, producing sound bytes of up to 1 minute in duration. In addition sound of any sample rate and precision supported by QuickTime can be cut or copied from other applications and pasted into the AudioGraph using the clipboard. Compression takes place, as with images, on generating a web site.

WEB SITES

The most recent work undertaken to improve the AudioGraph tools has been in the generation of web sites from a presentation. It is now possible to build more sophisticated web sites, without any web editing at all. The basic structure of the site has not changed and it still comprises an index of links for each of the slides within a "lecture". Figure 3 shows the lecture interface, which lists all slides within a presentation. It allows the user to specify a meaningful name for each slide, lists the duration of each slide and the total duration of the lecture.

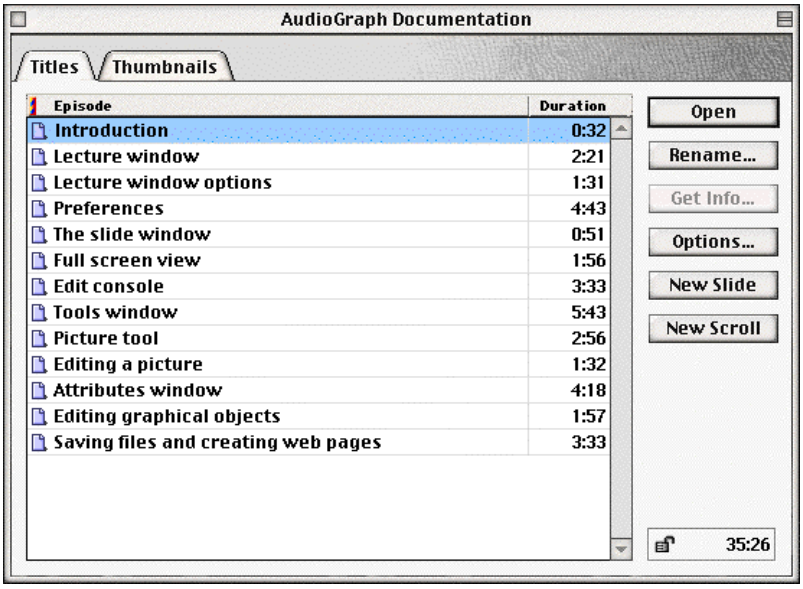

Figure 3. The Lecture window, showing the characteristic of each slide in a presentation, including duration of slides and total presentation.

This structure forms the basis of the web site that will be generated for this presentation. Within the Options dialogue, which allows various attributes to be set for an individual slide, such as size, background colour etc., there is a new option which allows the user to specify whether a slide is displayed in the same window as the index or whether it is to be opened within a new window.