MediaMetro: Browsing Multimedia Document Collections with a 3D City Metaphor

Patrick Chiu, Andreas Girgensohn, Surapong Lertsithichai[*], Wolf Polak, Frank Shipman[†]

FX Palo Alto Laboratory	[*]Silpakorn University	[†]Texas A&M University
3400 Hillview Ave., Bldg. 4	Faculty of Architecture	Department of Computer Science
Palo Alto, CA 94304, USA	Bangkok, Thailand	College Station, TX 77843, USA
{lastname}@fxpal.com	surapong@post.harvard.edu	shipman@cs.tamu.edu

ABSTRACT

The MediaMetro application provides an interactive 3D visualization of multimedia document collections using a city metaphor. The directories are mapped to city layouts using algorithms similar to treemaps. Each multimedia document is represented by a building and visual summaries of the different constituent media types are rendered onto the sides of the building. From videos, Manga storyboards with keyframe images are created and shown on the façade; from slides and text, thumbnail images are produced and subsampled for display on the building sides. The images resemble windows on a building and can be selected for media playback. To support more facile navigation between high overviews and low detail views, a novel swooping technique was developed that combines altitude and tilt changes with zeroing in on a target.

Categories and Subject Descriptors

H.5.1 [**Information Interfaces and Presentation (e.g., HCI)**]: Multimedia Information Systems.

General Terms

Design, Human Factors.

Keywords

Multimedia applications, 3D user interfaces, city metaphor, visualization, navigation, video, multimedia document.

1. INTRODUCTION

Multimedia documents are becoming more widely available with rapid advances in technologies for content creation and distribution. Activities and events are often captured on more than one type of media. For example, in a meeting or seminar, video along with slides and text notes may be captured (e.g. [2]). Information and concepts are also increasingly being communicated through more than a single medium; examples include educational CD-ROMs and news services in which text articles are accompanied by images and video clips.

For browsing a collection of multimedia documents, a good user interface metaphor can help the user understand the structure and relationships of the individual media elements, and make

navigation more effective through familiar way-finding conventions. Furthermore, a representation that visually exposes the contents of its various media elements allows the user to examine documents quickly. For navigation, it is important to enable users to move seamlessly from seeing an overview of the whole collection to examining individual objects.

In this demo, we present a solution based on an interactive 3D visualization using a city metaphor. As an intuitive way to represent information, the city metaphor has been explored in previous works (e.g. see [3], [7]). The basic idea is that information elements are mapped to visual elements of a city such as buildings, districts, paths, etc. [6]. Applying the city metaphor to multimedia document collections, we have designed and built an application called *MediaMetro*. See Fig. 1.

With the current graphics performance of desktop PCs, an interactive 3D visualization can be generated in real-time that shows visual summaries for a large number of the multimedia documents. In contrast, earlier 3D file browsing systems such as SGI FSN [4] only rendered documents as abstract geometric objects. Moreover, the swooping navigation technique in MediaMetro takes advantage of the visual richness by moving seamlessly from a high altitude overview of the collection's structure to a low altitude detail view of a document's contents.

The MediaMetro prototype was implemented in Java3D. We used a desktop PC (2.6 GHz Pentium 4 with 1.0 GB of RAM) with an NVIDIA GeForce FX 5600 graphics card with 256 MB of VRAM, running Windows XP. A large amount of memory (about 0.5 GB) is required to render the 207 buildings in the dataset shown in Fig. 1, using fairly high resolution textures varying from 128 to 256 pixels in width. With this setup, the performance was sufficient for smooth and responsive real-time user interaction.

2. VISUALIZATION & INTERACTION

The directories of a collection are laid out in the MediaMetro city landscape using rectangle packing algorithms similar to quantum treemaps [1] so that the directory tree structure is preserved. Our grid layout algorithm generates recursively from the bottom-up, resulting in a relatively square representation of a tree where each branch of the tree also approximates a square.

Each multimedia document is represented by a 3D rectangular building with different constituent media types shown on different sides of a building. In the multimedia collection at our lab, video is the primary media type supplemented with optional slides and text. Each building's roof displays a single representative keyframe, its façade shows a storyboard or Video Manga [9];

these are generated via content analysis algorithms. Visual summaries with thumbnails of the slides and text appear on the building sides. These images resemble windows on a real building. Each image is selectable for viewing in a pop-up player for the corresponding media.

XML metadata in the format of HTML Image Maps [5] are generated for the roof and visual summaries. The image map specifies the keyframe or thumbnail hotspots and playback points into a video, image sequence, or pages of text.

The navigation metaphor for MediaMetro is flying around in a helicopter. Navigation is performed by clicking the buttons around the application window (Fig. 1), or more compellingly by using a joystick with the application running full screen. As the user moves around the 3D landscape she sees an overview at a high altitude where the rooftop keyframes and the groupings of the buildings are apparent, and she sees a detail view at a low altitude where the finer grained storyboards and thumbnails on the building sides are visible.

To help users navigate more easily and rapidly between high altitude overviews and ground level detail views, we developed a novel *swooping* technique. By using a slider or simply clicking at a point on the ground, an animated path with good views is traversed—looking more vertically downward when high up and looking more horizontally when down low—by automatically combining altitude and tilt changes with zeroing in on a target. This provides a more flexible navigation technique than games such as Railroad Tycoon® 3, in which the camera angle is fixed for a given altitude. Swooping is also more suitable to the browsing task than speed-coupled flying [8], which is more like airplane than helicopter flight as it cannot hold still in mid-air, making it is difficult to examine and interact with information structures in motion.

3. CONCLUSION

We present the MediaMetro system for browsing multimedia document collections with a city metaphor. It employs a 3D visualization with multi-level map layouts of the collection's directory tree, and multi-faceted buildings displaying visual summaries of a document's constituent media. MediaMetro's swooping navigation techniques smoothly transition between overviews to detail views of the visual summaries rendered on the buildings in the cityscape.
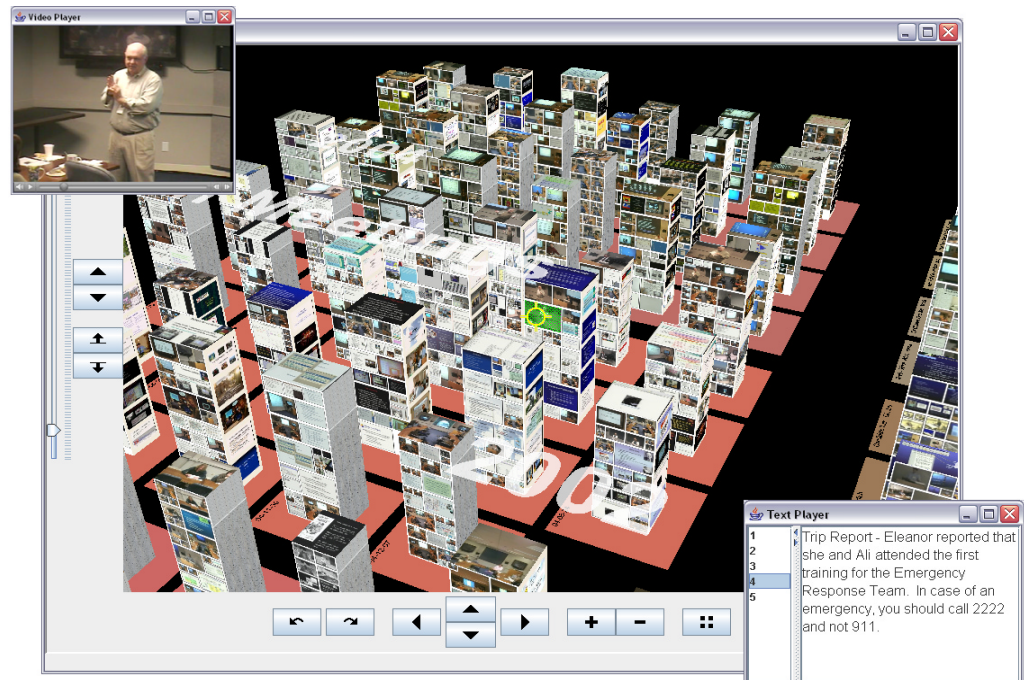

Figure 1: MediaMetro screen shot.

4. REFERENCES

[1] Bederson, B. Quantum treemaps and bubblemaps for a zoomable image browser. *Proc. UIST '01*, pp. 71-80.

[2] Chiu, P., Boreczky, J., Girgensohn, A., Kimber, D. LiteMinutes: An Internet-based system for multimedia meeting minutes. *Proc. Tenth Intl. World Wide Web Conference (2001)*, pp. 140-149.

[3] Dieberger, A. A city metaphor to support navigation in complex information spaces. *J. of Visual Languages and Computing*, 9(6): 597-622 (1998).

[4] FSN file system navigator, (1992). Silicon Graphics, Inc. http://www.sgi.com/fun/freeware/ 3d_navigator.html.

[5] HTML 4.01 Specification. W3C (1999). http://www.w3.org/TR/html401

[6] Lynch, K. *The Image of the City*. MIT Press (1960).

[7] Sparacino, F., et al. Browsing 3-D spaces with 3-D vision: body-driven navigation through the Internet city. *Proc. 3DPVT '02*, pp. 224-231.

[8] Tan, D., Robertson, G., Czerwinski, M. Exploring 3D navigation: combining speed-coupled flying with orbiting. *Proc. CHI '01*, pp. 418-425.

[9] Uchihashi, S., Foote, J., Girgensohn, A., Boreczky, J. Video Manga: Generating semantically meaningful video summaries. *Proc. ACM Multimedia '99*, pp. 383-392.