mTable: Browsing Photos and Videos on a Tabletop System

Patrick Chiu, Jeffrey Huang[*], Maribeth Back, Nicholas Diakopoulos[°], John Doherty, Wolf Polak, Xiaohua Sun[†]
FX Palo Alto Laboratory, 3400 Hillview Ave., Bldg. 4, Palo Alto, CA 94304, USA
[*]Media and Design Laboratory, Ecole Polytechnique Federale de Lausanne, Switzerland
[°]School of Interactive Computing, Georgia Institute of Technology, Atlanta, GA, USA
[†]Communication & Media Dept., Clarkson University, Potsdam, New York, USA
{*lastname*@fxpal.com}

ABSTRACT
In this video demo, we present *mTable*, a multimedia tabletop system for browsing photo and video collections. We have developed a set of applications for visualizing and exploring photos, a board game for labeling photos, and a 3D cityscape metaphor for browsing videos. The system is suitable for use in a living room or office lounge, and can support multiple displays by visualizing the collections on the tabletop and showing full-size images and videos on another flat panel display in the room.

Categories and Subject Descriptors
H5.2 [**Information interfaces and presentation**]: User Interfaces - *Graphical user interfaces*.

General Terms
Design, Human Factors.

Keywords
Tabletop display, multimedia visualization, photo browsing, photo labeling, video browsing.

1. INTRODUCTION
Our goal is to support browsing photos and videos in a more natural environment, not being confined to a desktop PC with the Windows metaphor and mouse & keyboard input. Toward this purpose, we designed the *mTable* system, which is intended to be used in a living room or office lounge environment. It resembles a coffee table with a dark tinted plastic surface. When the display is not on, it looks like a table and can be used that way. When the display is turned on, a hi-res display (1080p LCD) is revealed through the tinted surface.

Inside the table is a PC, and the input controls consist of gamepads that can be operated simultaneously by multiple users. A trackball is also available for some of the applications. Other hardware devices that can potentially simplify the input are touch interfaces (e.g. [9]) and tangible interfaces (e.g. [7]).

In this work, we focus on the multimedia applications and describe a set of them for visualizing photos and thumbnails, labeling photos, and browsing videos. Some of these applications are loosely coupled; for example, the metadata from the photo labeling application can be shared by the photo visualization applications.

2. VISUALIZING PHOTOS, THUMBNAILS
Our early prototype was based on a traditional force-directed visualization method [6] for data objects (photo thumbnails) by modeling forces between the objects or "magnetic" labels, resulting in clusters of similar objects. One drawback is that an object that is similar to two groups will end up in an awkward location between the groups; this is generally a problem when the user is interpreting and organizing the data for analysis and design (see [2], [8]) or for exploring photos [10].

To improve the interactive grouping process, we developed a technique for implicit brushing and target snapping [10]. Instead of automatically moving the photos around the display as with force models, an object that is related to other groups is decorated with visual hints in the form of arrows, with the arrow pointing to its corresponding group and the color of the arrow matching the border color of the group. The user can move the photo object across the large display into a desired group by clicking on the appropriate arrow, which sends the photo (with animation) to the target group.

Furthermore, as the group memberships are updated, a background query is performed for each group based on the text metadata of the photos in each group. While the retrieval mechanism is similar to existing systems (e.g. [4]), our interface is designed to support retrieval in a highly visual manner by manipulating the photo thumbnails. For more details, see [10].

Figure 1. mTable.

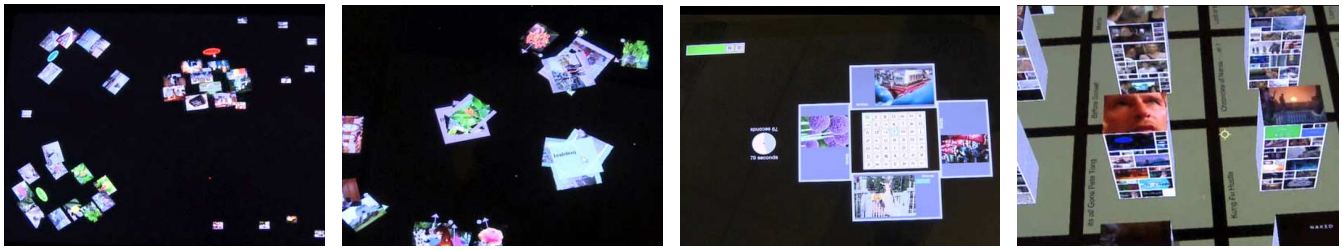

Figure 2. (a) Force-directed Layout, (b) Implicit Brushing & Target Snapping, (c) PhotoPlay, (d) MediaMetro.

3. LABELING PHOTOS

Having good text labels or tags for the photos is important for creating a visualization that reflects the properties and structure of the photo collection. Since it is difficult to automatically label photos by image analysis, other methods have been developed; in particular, game applications can be successfully leveraged for labeling images like the Web based ESP game [12].

For a tabletop system with collocated players, we have designed and built PhotoPlay [5], a computer board game for labeling photos. It resembles popular traditional word games by having a grid of letters (e.g. Scrabble, Boggle), plus four rotating photos randomly selected from a collection specified by the players. The players simultaneously use gamepad controllers to select letters to form words and tag the photos. To make it more fun and challenging, some letters of the alphabet are missing, and a letter selected by a player is timed out (unusable for 20 seconds).

At the end of each round, the players judge whether the labels are valid by voting using the gamepad buttons. The validity is debated with discussion among the players to promote social interaction. Through a formative design evaluation, playing with personal photos resulted in more specific labels such as named entities than when playing with randomly selected online photos [5].

4. BROWSING VIDEOS

For browsing videos in a multi-display environment that has a tabletop display and a vertically mounted flat panel display across the room, we extended the MediaMetro [3] application, enabling it to send video playback commands over the network to a video player application. In this configuration, the user browses a video collection on the tabletop, and watches a selected video on the flat panel display across the room.

The MediaMetro application visualizes the video collection using a 3D city metaphor. The directory tree of the collection is mapped to a "city grid" layout by algorithms similar to treemaps [1], and each video is represented by a "building" rendered as a rectangular box with keyframes from the video textured onto the faces of the building resembling window panes. The textures with video keyframes are constructed with the Manga [11] technique.

The user navigates the 3D cityscape visualization with a gamepad. MediaMetro has novel navigation features for flying between a bird's eye overview to a ground-level detail view [3]. A keyframe positioned under the crosshair can be activated by pressing the "fire" button, which results in sending a video playback command to the flat panel display across the room. The video plays at the time point corresponding to the keyframe.

5. CONCLUSION

In summary, we presented mTable and showed how a set of applications suitable for tabletop systems can be used to visualize, label, and browse multimedia collections in a living room or office lounge environment.

6. REFERENCES

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

[2] Beyer, B., Holtzblatt, K. *Contextual Design: Defining Customer-Centered Systems*, Morgan Kaufmann (1997).

[3] Chiu, P., Girgensohn, A., Lertsithichai, S., Polak, W., Shipman, F. MediaMetro: Browsing multimedia document collections with a 3D city metaphor. *Proc. of ACM Multimedia '05, Demo*, pp. 213-214.

[4] Cutting, D., Karger, D., Pedersen, J., Tukey, J. Scatter/Gather: a cluster-based approach to browsing large document collections. *Proc. of SIGIR '92*, pp. 318-329.

[5] Diakopoulos, N., Chiu, P. PhotoPlay: A collocated collaborative photo tagging game on a horizontal display. *UIST '07 Adjunct Proceedings*, pp. 53-54.

[6] Fruchterman, T. M. J., Reingold, E. M.. Graph drawing by force-directed placement. *Software: Practice and Experience*, 21(11): 1129-1164 (1991).

[7] Ishii, H., Ullmer B., Tangible Bits: Towards seamless interfaces between people, bits and atoms. *Proceedings of CHI '97*, pp. 234-241.

[8] Scupin, R. The KJ Method: A technique for analyzing data derived from Japanese ethnology. *Human Organization*, 56(2): 233-237 (1997).

[9] Shen, C., Everitt, K., Ryall, K. UbiTable: Impromptu face-to-face collaboration on horizontal interactive surfaces. *Proceedings of Ubicomp'03*, pp. 281-288.

[10] Sun, X., Chiu, P., Huang, J., Back, M., Polak, W. Implicit brushing and target snapping: data exploration and sense-making on large displays. *Proc. of AVI '06*, pp. 258-261.

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

[12] von Ahn, L., Dabbish, L. Labeling images with a computer game. *Proceedings of CHI '04*, pp. 319-326.