Sharing Digital Media on Collaborative Tables and Displays

Trent Apted **Judy Kay** **Mark Assad**
School of Information Technologies
University of Sydney
Sydney, Australia
{tapted,judy,massad}@it.usyd.edu.au

ABSTRACT

This paper describes the design and early experience with Cruiser – a multi-user, gestural, collaborative digital photograph sharing interface for a tabletop – and the techniques we use to share information with devices and other displays in its pervasive computing environment. The design is strongly influenced by the metaphor of physical photographs placed on the table and we have concentrated on the provision of an effective UbiComp interface that does not use a keyboard, a mouse or traditional WIMP (Windows, Icon, Mouse, Pointer) interface widgets. That is, with an emphasis on *seamlessness*. Aspects of the interface include the ability to interactively attach audio, handwriting or drawings, as well as other photographs to the "back" of an image after flipping it over; direct sharing of images with digital cameras, large displays and other UbiComp devices; and the provision of personal spaces – an area close to a user in which only they can work, which is enforced by the interface.

INTRODUCTION

We are just beginning to explore the possibilities afforded by tabletop interfaces, including the provision of some interactions that are not well supported by traditional computer displays. Face-to-face collaboration is possible, with multiple users sitting around the tabletop, each able to access computing resources. A table involves a social dimension and an environment where users can interact with a computer and each other in an informal setting. A tabletop also has the natural concept of a user's personal space – an area close to them in which they can work.

The need for technology to support natural tabletop interaction has prompted researchers to approach the *hardware* problem from a number of directions. Here, the problem is the determination of where *each* user is acting on the display. Devices such as the DiamondTouch [3] use capacitive coupling with the human body to detect multiple touches, while computer vision techniques are used in other work [7, 9] with varying results. There is also the possibility of adapting mature technologies such as SMARTBoards [5] or Mimio [10], traditionally for vertical presentation displays (e.g. whiteboards), to support multiple users on a

This work was conducted as part of the *Homes & Communities* project, a joint project between the University of Sydney and the Smart Internet CRC.

Figure 1. Cruiser in use

large horizontal surface. Furthermore, collaborative touch technology and computer vision techniques are still at an early stage and it is attractive to explore these, stylus-based alternatives especially as a *pen* is particularly natural for some interactions.

Regardless of the hardware chosen, there is a need to explore software possibilities to support natural interaction. Early work with touch screens introduced the concept of a *gesture* for manipulating virtual objects without a mouse or keyboard [6]. More recently, and closest to our work, Shen et al. [8] investigate layout techniques for multi-user browsing of photo libraries and story sharing on a circular tabletop in their Personal Digital Historian (PDH). Our work introduces novel annotation methods and attempts to provide an interface truer to the analogy of physical photographs lying a table for collaborative mark-up and social interaction.

In this paper, we will outline our techniques for interacting with the tabletop, as well as with other devices in the environment *through* the tabletop interface. These currently include a wall-sized stereoscopic display, digital still cameras, regular computing desktops and a prototype "magic mirror" distributed home messaging system we are developing. We will also discuss methods for handling input from these devices; as well as a microphone, the display framebuffer itself and multiple users in a manner that enforces ownership and privacy on the tabletop.

OVERVIEW OF Cruiser

Our work is motivated by the desire to support, for digital photographs, natural archival and mark-up activities with groups of people, and the social interaction experienced while *sharing* photographs. This was once accomplished by passing the latest pile of holiday snaps back from the developer around a coffee table, but the advent of digital

Figure 2. The projected image for 4 users

①: The blackhole (an image partially within it)

②: An image photo corner (selected by 'cyan' user)

③: A dynamically updated remote framebuffer "photo" from a laptop

④: An audio object, attached to the back of a "Keep-in-Touch" image (audio objects on people are synchronised with the magic mirror interface)

⑤: The Frame; ⑥: A flipped photo with attachments (some writing)

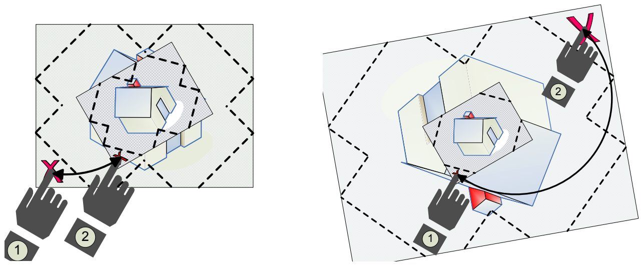
photography is changing the nature of photography and the photographic process. While this transition brings with it many possibilities in terms of copying, editing, publishing and printing, it also brings challenges in terms of presentation, viewing and sharing photographs [4], which our work is investigating.

We use pen-based gestures for manipulating digital photographs on a tabletop interface, and leverage modern dedicated computer graphics hardware to efficiently simulate a physical tabletop with photographs placed on it. These may be annotated, manipulated and shared using a pen. Hardware is not the focus of this work, so we have adapted the Mimio system, an off-the-shelf whiteboard capture tool, to support asynchronous multi-user pen-based interaction. However, the application is not bound to this hardware. In fact an earlier version [1] used a DiamondTouch for input. The C++ and OpenGL implementation is tested on Windows, Mac OSX and Linux to ease future adaptation.

Cruiser supports use by up to four users at the tabletop through the configuration of a static layout. Figure 1 shows the layout for two users, with the two participants sitting opposite each other. Each participant has a coloured triangular "personal space" in front of them. Figure 2 shows the user view for three or four users. A photo selected by a particular user has a border and photo-corner indicated ② and also shown on the images in Figure 3. Other objects, such as the *Black Hole* ①, remote framebuffer photo ③, audio objects and Keep-in-Touch images ④, the Frame ⑤ and writing ⑥ are described below.

General Interface Concepts

After ***selecting*** a photo by physically touching it with a pen tip, it can be ***moved*** by sliding the tip over the tabletop. We attempt to maintain the physical-virtual coupling throughout this and other actions, by keeping the contact point on the photo appear "stuck" to the pen tip. This coupling is broken in special circumstances – the centre of the photo is not allowed to move off screen, nor is it allowed to enter another user's *personal space*.



(a) A reduction (b) An enlargement

Figure 3. The *rotate/resize* gesture

The centre of the image remains fixed while the image rotates and resizes to keep the "✘" under a user's finger

To ***rotate and resize*** a photo, a user must first touch it in a *photo corner*. The photo is rotated and resized concurrently, with the centre of the photo fixed and the contact point remaining under the pen as it moves in two dimensions. For example, Figure 3(a) shows the contact point moving from position 1, inwards and in an anti-clockwise direction to end at position 2, giving the superimposed version of the photo; reduced in size and rotated 45°. Figure 3(b) shows an enlargement and 170° rotation.

To ***flip*** a photo, a user selects a triangle along one of the photo's edges and drags it across to the opposite edge. Again, we wish to keep the original point under the pen tip, so this occurs gradually and in simulated 3D. If other photos are beneath it, the user will see the flipping photo pass *through* those images, as its plane rotates on a horizontal axis through the centre of the photo, parallel to the edge whose triangle was chosen. Once flipped, the user sees a partially transparent white surface and they may attach other objects on the table to that image, which will be automatically laid out on the flipside. This is done simply by moving the object over a flipped image and releasing.

A photo may also be ***flicked***, during a *move* action. This occurs when a photo (or other movable object) is released whilst it is moving with a velocity greater than a threshold – currently 300 pixels per second. The object is given momentum and subjected to a kinetic frictional force that causes it to decelerate. This action becomes relevant when the table is large and a user might not be able to reach across the full width to pass a photograph to another user, but it can also be a trigger for other actions and events. The photo is stopped if any user selects (*"grabs"*) it, and it is not allowed to fall off the edge of the table, nor move into the personal space of a user other than that of the user who flicked it.

In the current implementation, users may send a photograph from the tabletop to a wall-sized display, where it is shown covering the wall, for wider presentation and analysis. This occurs at the edge of the table closest to the display, so the user simply *flicks* the image towards the display itself. In our lab, the display is a "3D DataWall" that allows the display of stereoscopic photographs, and a three-dimensional animation of the photograph moving from the edge of the table to the wall. The latter is a design towards seamlessness, but the need for 3D goggles to see the effect (and their unsuitability for interaction with the current tabletop) means that it is usually wasted. Flicking to the opposite edge of the table currently inserts the image into a slideshow on the magic mirror interfaces, which behave like a dynamic picture frame when not in use.

Special Actions

A user may ***copy*** a photo by an implicit action that is triggered when they try to move an unowned photo into their personal space. A copy owned by that user is created once the original would have been moved more than a threshold distance into the personal space and the copy immediately becomes the user's active selection, which they are still moving.

To ***draw*** or ***write***, a user simply uses the pen on the background. That is, they do not select a photo, but drag the pen over the black, public area or their own personal space. This mode switch is implicit, so we take steps to avoid accidental drawings.

To ***attach audio*** to a photo, a user *dwells* on that photo. They are prompted to record their audio and repeat the dwell when they are finished speaking. A circular audio object is created that is automatically attached to the back of that image and numbered. To replay the audio, a user *dwells* on the audio object (after *flipping* the image if necessary). This enables a user to attach one or more pieces of audio to a photo, for example as an annotation or a message.

Special Objects and Places

Personal space objects are coloured triangular elements drawn on the display, which demarcate the exclusive area for each user. For example, once the centre point of a photo is contained within a user's personal space, no other user may select that photo. This enforces a range of social conventions surrounding what constitutes acceptable behaviour for items in public space versus one's personal or private space.

The ***Black Hole*** is a semi-translucent photo with swirls. It can be manipulated as a typical photo (moved, rotated, resized, etc.) but is never obscured by other photos and cannot be copied or moved into a personal space – it is a shared resource. The Black Hole has a "sphere of influence" that affects the size of other photos near or in it. Figure 4 shows a user moving a photo into the Black Hole. Moving from the right to left, the photo gets smaller as it moves closer to the centre until it is hidden. This occurs as the *touch point* (indicated by a "✘"), rather than the photo *centre* is moved closer to the centre of the Black Hole. This gives the photo an appearance of being gradually "sucked" into the Black Hole.

The Black Hole is our take on the trash can. As the Black Hole is moved, photos in its sphere of influence move with it. Photos may be retrieved by either reducing the size of the Black Hole (allowing partially hidden photos on the fringes to be retrieved) or by dragging the Black Hole into your personal space. In the latter case, the Black Hole itself does not move into your personal space, but photos that

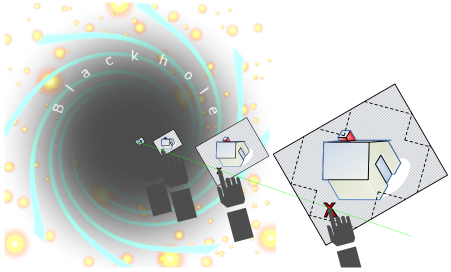


Figure 4. Putting an image into the Black Hole

are unowned or owned by you[1] will "fall out", into your personal space. *Dwelling* on the Black Hole causes photos already within it to be permanently deleted, and all other movable photos to slowly move towards the Black Hole. This can be used to clear the table and release resources when many high-quality images are causing the interface to lose responsiveness.

The ***Frame*** is a transparent blue image. Like the Black Hole, it can be manipulated as a typical photo and is never obscured. However, a user *is* allowed to move it into their own personal space, but if a different user selects it, the Frame will immediately move to the closest point not within that personal space. The Frame is an object used to *capture* areas of the display, similar to capturing a photograph with a real camera, where each Frame acts as a viewfinder. To activate the Frame's audible "shutter" action a user holds their pen tip on it and *dwells*. Users see a flash and pixels beneath the Frame are *grabbed* (from the frame buffer) and loaded in as a new photo. This allows photos to be grouped as a collage or *cropped*. For example, one might crop an image of a person from one photo and attach that to another. By capturing writing or a drawing, the user can produce a flippable image, which can then have other items attached.

The ***Remote Framebuffer*** image is photo object that can be manipulated just like any other, but is updated dynamically from a networked computer's screen contents using the VNC protocol. For example, a laptop running a VNC server can be brought to the table and have its display duplicated (as a photo) in order to display slides from a presentation or discuss a document (which may be copied, rotated, resized, etc. by each user). In future we wish to be able to interact with the remote computer *via* the tabletop by sending mouse cursor events, but we are yet to design a seamless way to distinguish this interaction from moving the photo object itself. We also wish to give further meaning to the attachment of objects to the framebuffer photo. For example, attaching writing or audio might send a command to the remote computer and attaching a photo might simply store a copy. However, this would require software running on the laptop beyond the (generic) VNC server software.

Other Events

New photos appear as a result of a number of events: when a digital camera or removable storage device is plugged into

[1]Thus a user may "delete" a photo they own by moving it into the Black Hole and other users may not remove it

the table, as the result of a Frame capture, or when *sent* an image from a networked device. In each case, the photo slides in from a corner of the surface, increasing in size as it moves into a spiral around the centre of the table, in order to draw attention. Once in the centre, it is also shown above other images until an object is selected.

Processing photographic-quality raster images directly from a compressed format[2] into a texture is resource-intensive and we wish the interface to remain responsive during this process. Thus, texture mipmap generation and other tasks not initiated by the user are executed in background threads, with only a negligible pause as we load the processed textures into video memory.

Keep-in-Touch images

The "Keep-in-Touch" (KIT) images are photos of people that can be manipulated like the others, but are synchronised to a system of audio messaging appliances. These are designed to facilitate message communication within a geographically distributed family. One appliance is the "Magic Mirror", which hides the display behind a half-silvered mirror. The user interacts with the device using gestures over light sensors mounted in the frame. Another device is a touch screen. The messages are sent to members of the family by selecting the recipient's photo and recording a message. Once a message is recorded it is distributed to all the KIT appliances in the network.

Cruiser integrates with this system by displaying the family photos as images on the table surface. The user is able to flip these images, and record a new message for the family member. When a message is recorded, it is copied to all the other display appliances. If a message is recorded on any of the other appliances, it will appear on the back of the respective image on the Cruiser table. This allows seamless multi-way interaction between a number of appliances.

DISCUSSION AND FUTURE WORK

Our implementation is robust enough that it has been used for numerous demonstrations both in and out of the laboratory and was a 'video highlight' at the Australian CeBIT exhibition, 2005 [2]. This gave us an opportunity to gain vast amounts of informal feedback over the course of the three-day exhibition. Generally, user reaction was very positive and all visitors who chose to interact were able to use the interface effectively in less than a minute. In most cases, a user's own exploration of the interface revealed the functionality of the Black Hole, as well as moving, rotating, resizing, flipping and the activation of a "pie menu" that we have been trialling. However, the Frame was rarely discovered until the user was told about it and, while writing was simple to begin, the *dwell* requirement and the need for a single stroke was problematic.

Surprisingly, no visitors expressed an immediate desire for a touch interface, despite SMARTBoard exhibiting vertical touch displays only metres away. In discussion, the attractiveness of a touch interface was recognised but most expressed the desire for a keyboard (or pen) for textual input. The potential for multiple-finger and whole-hand gestures did not get raised, but many users expressed a desire for *video* in place of the still images, which is currently broaching the limits of what consumer computer hardware can provide. It should be noted that as CeBIT is a *technology* exhibition most visitors were computer savvy and familiar with cursor-based manipulation – in [1] we present a comprehensive study with elderly, non-expert participants and a touch interface (where the two-finger gesture we supported proved to be very difficult for the users).

[2]we support jpg, png, gif, bmp, tif, pcx, tga and others, of any size

CONCLUSION

We have presented the interaction features of an interface prototype to facilitate manipulation, sharing and annotation of high-quality digital photographs on a tabletop called Cruiser. While robust and computationally efficient, development is continuing as we improve the interface based on initial feedback before a thorough empirical evaluation is conducted. Early observations indicate clear interest in this kind of interface, based on informal feedback from visitors to our demonstration at a technology exhibition.

REFERENCES

1. Apted T., Kay J., and Quigley A. A study of elder users in a face-to-face collaborative multi-touch digital photograph sharing scenario. Tech. Rep. TR576, School of IT, University of Sydney (2005).
2. CeBIT Australia. `http://www.cebit.com.au`. Website (2005). Verified 2005-06-16.
3. Dietz P. and Leigh D. DiamondTouch: A multi-user touch technology. In J. Marks and E. Mynatt, eds., *Proc. of* 14^{th} *annual ACM symposium on User Interface Software and Technology*, pp. 219–226. ACM Press, Orlando, Florida, USA (2001). ISBN 1-58113-438-X.
4. Howard D.M. *Sharing Digital Photos: The Future of Memories*. Microsoft Press, 1^{st} edn. (2003).
5. Martin D.A. Interactive display system (SMARTBoard). US Patent 5,448,263, Smart Technologies Inc., Calgary, California, USA (1991).
6. Minsky M.R. Manipulating simulated objects with real-world gestures using a force and position sensitive screen. In *SIGGRAPH Comput. Graph.*, 18(3):195–203 (1984). ISSN 0097-8930.
7. Nakanishi Y., Sato Y., and Koike H. EnhancedDesk and EnhancedWall: Augmented desk and wall interfaces with real-time tracking of user's motion. In *Proc. of UbiComp'02 Workshop on Collaborations with Interactive Walls and Tables*, pp. 27–30 (2002).
8. Shen C., Lesh N.B., and Vernier F. Personal digital historian: Story sharing around the table. In *ACM Interactions*, 10(2):15–22 (2003). ISSN 1072-5520.
9. Song L. and Takatsuka M. Real-time 3D finger pointing for an augmented desk. In *Proc. of* 6^{th} *Australasian User Interface Conference, CRPIT 40* (2005).
10. Virtual Ink Inc. `http://www.mimio.com/.` , Mimio Website (2005).