Available online at www.sciencedirect.com
ScienceDirect
Int. J. Human-Computer Studies 67 (2009) 1048–1059
International Journal of Human-Computer Studies
www.elsevier.com/locate/ijhcs

Mobiphos: A study of user engagement with a mobile collocated–synchronous photo sharing application

Nirmal Patel[a,*], James Clawson[a], Amy Voida[b], Kent Lyons[c]

[a]Georgia Institute of Technology, 85 5th Street, Atlanta, GA, USA
[b]University of Calgary, 2500 University Dr. NW, Calgary, Alberta, Canada T2N 1N4
[c]Intel Research, 2200 Mission College Blvd. Santa Clara, CA 95054, USA

Received 19 December 2008; received in revised form 14 July 2009; accepted 7 September 2009
Available online 17 September 2009

Abstract

Photographs have always been artifacts for creating memories and engaging in storytelling activities with others. To date there has been much research in the HCI community towards sharing of both analog and digital photographs. With recent advances in network technology further research has been done with photos being shared almost immediately after capture. However, most of the research has focused on synchronous sharing with groups of distributed users and little has been done to focus on how synchronous capture and sharing could benefit a group of collocated mobile users. To help start exploration in this area we have created Mobiphos. In this article we present how synchronous capture and sharing affects how groups of mobile, collocated users engage with their environment and each other while touring a city. We also discuss the design guidelines of Mobiphos and the implications for future photoware for the mobile, collocated context.
Published by Elsevier Ltd.

Keywords: Mobile; Photo sharing; Collocated; Synchronous

1. Introduction

Photographs have always been an artifacts for remembering and storytelling. They help those who capture photographs not only remember the context but also allow them to engage others in the experiences documented in the photographs. The act of capturing photographs not only reflects the capturer's own interests but the social norms which surround the method and objects of photography (Chalfen, 1987).

Digital cameras and camera phones have enabled the number of photographers to increase while also increasing the number and variety of photographs captured. Innovations in networking technology, such as MMS, have also enabled easy sharing of photographs to many people immediately after capture thereby increasing the number of participants that can engage in the context of capture.

Traditional analog photography forces a temporal break between the capture and sharing of photographs. While digital photography allows for capture and sharing using a digital display, the sharing is still limited by the size of screen and the time and method taken to share the photograph. In this paper we discuss how our research has enabled immediate and effortless sharing for collocated groups of social users, helping them engage in the photography of the group and enjoy a shared experience.

One such example of a shared social experience may be a group of friends touring a city for the first time, a genre of photography called "Camera Recreation" (Chalfen, 1987). While touring, it is likely that all participants will want to take photos of the major landmarks. It is also likely that each person will take more directed photographs that fit their personal interests. These could be anything from random objects found during the tour to more up-close photographs of various aspects of the landmarks. Beyond photographs of landmarks and environment it is common

*Corresponding author. Fax: +1 404 894 0673.
E-mail addresses: nirmal@gatech.edu (N. Patel), jamer@cc.gatech.edu (J. Clawson), avoida@ucalgary.ca (A. Voida), kent.lyons@intel.com (K. Lyons).

1071-5819/$ - see front matter Published by Elsevier Ltd.
doi:10.1016/j.ijhcs.2009.09.002

to see photographs taken of other group members engaging in tourist activities. Most such experiences end with the friends asking for certain pictures to be e-mailed or uploaded to photo sharing websites such as Flickr.

In Clawson et al. (2008), we presented the Mobiphos application and our initial findings. Mobiphos was designed to explore the impact of highly synchronous photo capture and sharing in group photography. We conducted a quantitative and qualitative study of how Mobiphos is used in the context of a group of friends touring a city. We recruited participants and gave them a tour map, highlighting historical landmarks. The participants were then instructed on the use of Mobiphos and the tour began. Multiple researchers shadowed the participants to gather field notes as well as provide any technical assistance should it be needed. In Clawson et al. (2008) we presented the analysis of the field notes along with post-trial focus group. From these data we identified seven themes of use: (1) collective photography, (2) the situated, shared experience, (3) where the individual meets the group, (4) rhythms of use, (5) collaboration and competition, (6) gift giving, taste, and identity, and (7) spectrum of appropriation.

In this work, we add to the already presented themes of use by performing an analysis of the log data and the content of the photographs captured by the participants when on the tour. We inspect the photographs and analyze their content exploring the participants' engagement with various aspects of the experience. By using the photographs as sample points into what the participants found to be engaging, we have arrived at three high-level categories which we present in Section 6: participant engagement with touring, participant engagement with other members of the group, and participant engagement with the environment. We present the counts of the total number of photographs for each section, excluding photographs that were too unclear to categorize. We also present the design motivations for Mobiphos in regards to existing literature as well as the evolution of Mobiphos from pilot study to final version. Finally, we describe the design cycle for Mobiphos in further detail. Specifically, we look at existing work in the space as well as our own pilot testing to show how the design evolved to the current version and discuss ways in which the interface design affected the level of engagement the users had with the system.

2. Related work

Literature concerning the capture and sharing of photographs has discussed them in various technological contexts such as physical sharing in the Kodak Culture (Chalfen, 1987). Additionally, Becker (1982) showed that capture and sharing are not separate practices, but the relationships between those who produce and consume are reciprocal. The collaborative nature of these relationships allow for both parties to not only engage in the act of photography but also engage in a dialogue by using photographs to respond to or influence others.

With traditional analog photography, there is an enforced temporal break between the capture and sharing of photographs. However, once photographs are available the sharing of digital and analog photographs can be done with collocated or remote users. Analog photographs can easily be copied and distributed physically, but advances in networking technology have allowed digital photos to be shared with many people at very little expense. Digital photos can be transmitted from person to person, via email, or shared with many people with various photo distribution services such as Flickr (Miller and Edwards, 2007), blogs (Nardi et al., 2004) or other online communities (McDonald, 2007). While face-to-face sharing of photographs has been studied, with much of the research exploring discussion around the photos (Balabanovic et al., 2000; Chalfen, 1987; Crabtree et al., 2004; Frohlich et al., 2002), little has looked at how users who are engaged in the same context can make use of photographs to enhance the conversation.

Advances in technology have allowed users to share photos almost immediately after capture. Many researchers have explored the social impacts of this network enhanced photography (Mäkelä et al., 2000). Voida and Mynatt (2005) have explored how instant messaging has increased the distributed sharing of photos. Advances in mobile messaging systems (MMS) and the increasing affordability of wireless data transfer have increased research efforts in the area of serendipitous photo capture and the synchronous sharing of those photos with distributed people (Battarbee, 2003; Kindberg et al., 2005; Koskinen et al., 2002; House et al., 2005). While MMS is a common way for sharing images from camera phones, it is not well suited for quickly sharing photographs with a group of people nor does it directly support collocated individuals.

With the technical limitations of the MMS system and other commercial mobile messaging systems, recent research has focused on technology oriented solutions for supporting sharing of photographs in a near-synchronous manner with distributed groups. A mobile system that shared photos with groups of remote friends, determined by buddy lists, was developed by Counts and Fellheimer (2004). While MMS is not ideal for group sharing it has been used by groups to support awareness of activities when individuals are distributed. Researchers have explored how groups that are spread over multiple areas or attending large crowd events use commercial media sharing technologies to maintain context and engage in each others' surroundings while remote (Jacucci et al., 2005; Salovaara et al., 2006; Sarvas et al., 2005).

Despite the increase in technologies to enable wireless, synchronous sharing, very little research has been done to see how these technologies could be employed to capture and share photos in a collocated setting (Ashbrook et al., 2006). Kindberg et al. (2005) studied the use of camera phones to explore the variety of ways that camera phone

photos were shared. While some sharing was accomplished using MMS, the majority sharing was done using the display on the camera phone in collocated contexts:

> There was little evidence of a strong 'capture and send' culture ..., the study data showed that two thirds of the images examined were captured to share ... . The majority of image-sharing (one third of all images) took place face-to-face on the phone itself, often in the moment .... [In addition] sharing involved impromptu storytelling, passing the phone to someone else, or swapping phones with a friend (Kindberg et al., 2005).

Despite the findings of Kindberg et al. (2005), there have been few technologies designed specifically to support synchronous capture and sharing. With the exception of recent work from Salovaara et al. (2006), almost all work has focused on small portions of the mobile, collocated-synchronous capture and sharing context. For example, Kohno and Rekimoto (2005) focused on technology to automate the creation of collocated groups of users. The algorithms presented a combination of image analysis, wireless proximity measures and digital compass based orientation as a heuristic for determining the devices that belong to people in a collocated group. Kohno and Rekimoto (2005) also presented a few interesting designs based on this orientation including one that shows timelines for each person in the group around the edge of the display based on the relative position of the people in the group. While this technique makes use of the physical orientation of the users to make understanding the interface easier, the design implications found in Salovaara et al. (2006) suggest that the separation of the timelines by owner detracts from creating a true group experience. Unfortunately, there was no user evaluation to accompany these technologies, so it is not clear as to how much affect the separated timelines would have compared to the "common space" discussed in Salovaara et al. (2006).

Work done by Kun and Marsden (2007) has focused on evaluation of techniques for designing co-present photo sharing on mobile devices. Specifically, they have focused on how the storytelling activities described by Kindberg et al. (2005) can be augmented when each person in a collocated group has their own device. They choose to focus on a WYSIWIS (What-You-See-Is-What-I-See) interaction, thereby making every display in the group show the same thing. This provides some of the grounded context of a group display but without the requirement of extra hardware beyond the mobile devices already carried by the users. Multiple floor control policies were designed and implemented to determine what the role of technology would be in supporting the storytelling activity. It was found that an explicit control system, where control would have to be explicitly requested and released, was most preferred. An interesting result from this was that despite the ability to request and release the control from within the technology, it was common for participants to form an ad hoc verbal policy for requesting control. This result is supported by previous work which shows that impromptu social protocols can be effective in small groups and sheds light on the importance of working with face-to-face conversation to provide a smoother experience (Brinck and Gomez, 1992; Greenberg and Marwood, 1994).

Most recently, Salovaara et al. (2006) has focused on extending photoware to mobile devices and studying them in the field with live trials. The mGroup system allows groups of users to capture, share, and annotate photographs to create *Media Stories*. In contrast to Mobiphos, which was designed as an additional mode to stand-alone digital cameras, mGroup was designed specifically for mobile phones and therefore allows for a wider variety of functionality including the input of text annotations. While mGroup was used by participants in collocated situations, it was not uncommon for groups to split up which separated the devices. Due to the lack of face-to-face conversation in these situations, it was important for mGroup to allow users to add additional information to photographs. In the design and evaluation of Mobiphos, we have focused exclusively on participants who are both mobile and collocated throughout the trial.

With inspiration derived from these few exceptions (Kohno and Rekimoto, 2005; Kun and Marsden, 2007; Salovaara et al., 2006), we have created a system designed specifically to enable collocated, synchronous photography and studied how the engagement in the task, environment, and group are affected by the immediate availability of sharing photographs within a social group. We also explore how interface design can serve to give the group a greater engagement with each other while experiencing their collocated context.

3. Mobiphos application

As initially presented in Clawson et al. (2008), Mobiphos is a collocated photo sharing application designed to be an additional mode which could be enabled for digital cameras being used among a group of friends engaged in a social activity, such as sightseeing. Mobiphos implements some standard features found in all digital cameras such as a digital viewfinder, a photo capture mechanism, and the ability to browse and view photos at full size. Additionally, all photographs captured by a group member are wirelessly transmitted to all other Mobiphos devices that are part of the group. Mobiphos behaves differently from standard digital camera software in that it allows the user to simultaneously capture photographs and view thumbnails. The combination of view-finding and thumbnail modes allows users to capture images and monitor as the photographs other group members stream to their device. In the end, all of the users will have all copies of photographs captured by the group thereby creating a group repository.

There are the multiple parts to the Mobiphos application. Nominally, it is a digital camera that allows users to press a button to capture a photograph. Thumbnail

Fig. 1. A screenshot of the Mobiphos interface with the thumbnail timeline in mid-animation. The viewfinder is in the top-right and thumbnails are along the left and the bottom of the display. The colored border on the images indicates who captured the photograph. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

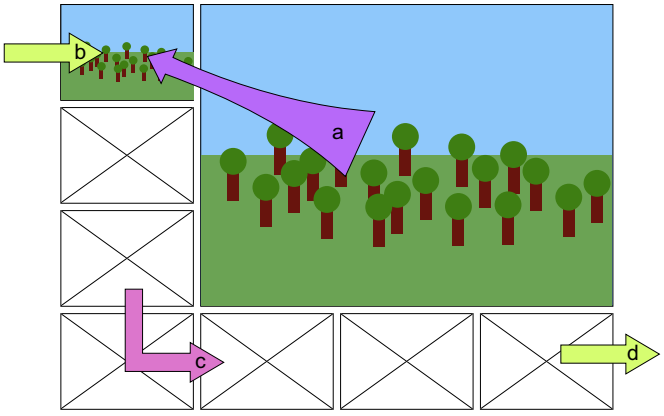

Fig. 2. The viewfinder is in the top-right of the display. When the user captures a photograph, the picture from the viewfinder animates into the top-left corner (a). When an image comes from another user, it is also placed in the top-left (b). Either of these events causes the timeline to animate, wrapping around the bottom-left corner (c) and the oldest image is moved off screen (d).

browsing is also provided so that users can review the captured images. The images are saved locally to a solid state drive and then sent over a wireless network to all other users in real-time. The user interface has been modified both to support the real-time nature of the application as well as to provide necessary information about the photographs to the users. In standard digital camera software, there are separate modes for viewfinding/capture and thumbnail review. With photos arriving on each device when they are captured by others in real-time, we chose to combine these modes in Mobiphos to better enable users to engage with both the capture and sharing aspects of the Mobiphos experience.

In our design, one screen is used to show both the viewfinder and recently captured photos (Figs. 1 and 2). The top-right $\frac{3}{4}$ of the display is used to show the viewfinder. The thumbnails are organized around the viewfinder area in an L-shape. Each thumbnail is $\frac{1}{16}$ of the total screen size allowing us to place seven thumbnails, at a time, onto the display along with the viewfinder (Fig. 2). Three thumbnails are positioned to the left of the viewfinder, one in the bottom-left corner and three below the viewfinder. The most recent thumbnails are placed in the top-left corner and wrap around the viewfinder from newest to oldest. This combination of modes allows for easy capture of photographs and reviewing of photographs arriving from other users without switching modes.

As photographs are captured, they scale down from the area of the viewfinder and move to the top-left corner of the screen (Fig. 2). Photographs captured by other users also appear in the top-left corner. In both cases, animation is used to show the user where the picture is arriving from. In the case of photographs taken by other users, the photo animates in from the left edge of the screen. As new photographs move into the timeline, the existing photos move down and out of the screen through the bottom-right corner. All captured photographs receive a colored picture frame overlay. These colors correspond to the device used to capture the photograph. We will discuss later how this design decision was instrumental in allowing the users to easily combine the face-to-face discussion about the experience with the photographs captured.

A directional keypad allows users to browse the timeline. By pressing up or left, the user can scroll the timeline backwards and look at older photographs. By pressing down or right, the user can move towards the most recent photographs. In both cases, animation is used to help the user understand how they are browsing the timeline. When scrolling the timeline towards older photographs, the timeline moves in the opposite direction of presented in

Fig. 2 from (d) to (c) to (b). While holding down a directional button, the animation will continue to gain speed. Upon release the speed fades quickly and the timeline snaps to the nearest whole photograph. Also another button allowed the user to jump directly to the front of the timeline.

As with traditional digital cameras, users can enlarge the thumbnails. In Mobiphos this is accomplished by using the touchscreen and tapping the thumbnail. When tapped, a thumbnail animates from its spot in the timeline to the area of the viewfinder. The thumbnail's location in the timeline is then shown as a gray rectangle with a white frame. Users can then use the directional buttons to move through the timeline and enlarge each photo one at a time. To send the photograph back to its spot in the timeline, the user simply taps the enlarged photograph.

To help maintain usage context among other group members, all other members are alerted when a user enlarges a photograph. This alert places a colored dot in the top right corner of each user's display. The dots represent the users who are currently viewing an enlarged photograph. By tapping on the dot users will be presented with an alternate screen that shows which photos are being inspected by all users. At this point the user can either dismiss this screen or tap on a thumbnail to enlarge the photo.

3.1. Implementation

Due to the unavailability of programmable, stand-alone digital cameras, Mobiphos was implemented on a Motorola E680i Linux based camera phone (Fig. 3). The E680i was specifically chosen because when held in a landscape orientation, the placement of buttons were close to that of a standard digital camera. There is a button on the top right of the device used as the capture button. Additionally, there is a small directional pad which has become common on digital cameras as a way to navigate a grid of thumbnails. The application was developed in Python. PyGame, a wrapper for SDL, was used to create the user interface elements, and the standard Python socket library was used for networking the phones together over WiFi. Additionally, we developed Python modules to allow for direct interaction with the camera.

The E680i camera is capable of capturing images at 640×480 pixels in landscape mode while the screen is oriented naturally for phone use in portrait mode. We intended to use both the camera and screen in landscape mode. To maintain the orientation and aspect ratio of the image on screen, we reduced the resolution of our images to 480×360 (360×480 in camera coordinates). This image was then scaled down as needed for the viewfinder and thumbnails.

Our first version of the application kept all images at full resolution loaded in memory. This implementation worked well when testing the application on a desktop computer. However, the program become sluggish or unresponsive when run on the phone. The photos were saved to a file as a JPEG, but their in-memory counterparts were represented as bitmaps. The E680i has 32MB of total RAM, and holding full-size images in memory along with the rest of the phone software and application exhausted the RAM after loading approximately 60 images. To remedy this situation, we implemented a dynamic loading system which only kept smaller versions of the pictures in memory and loaded full-size versions from the filesystem as necessary.

Fig. 3. The Motorola E680i running Mobiphos.

During development, we tested the possibility of using Bluetooth to transmit files between devices. Unfortunately, the delay between the time when a photograph was captured to when it was transmitted to a single device (approximately 3 s) was too great for the type of interaction we wanted to support. This delay was compounded when increasing the number of people in the groups. During pilot testing, we found that it was common for users to verbally cue others to look at a photograph; unfortunately, due to the speed and connection delay issues of Bluetooth, the other users would have to wait so long that they eventually lost interest and did not look at the display when the photo arrived. Instead, our final implementation of Mobiphos shares images between users through a WiFi connection. Mobiphos will find all other instances of itself running on other devices which are connected to the same network. As each image is captured by a user, it is saved locally and sent to all of the other devices in the group. When Mobiphos cannot send a picture to a device, it queues the picture and sends it the next time the unavailable device appears on the network. After switching to WiFi, we noticed that the interactions between the users and the system became more synchronous.

Unfortunately, the E680i does not have built-in WiFi. Instead, we used SDIO WiFi cards to create a network that allowed the rapid transmission of images to all phones. The system can run in ad hoc mode; however, for the data presented here that capability was not yet functional, and instead one researcher carried a battery powered access point for the purposes of the experiment. Participants were not actively made aware of the router in the bag.

To facilitate data collection for our study, we added a logging component to the software. The application recorded its start and end times, any navigation of the thumbnail timeline, the scaling of images, the capture of images, and the sending and receiving of images over the network. All entries were time-stamped in milliseconds and were tagged with a unique, anonymous participant identifier.

4. Method

The goal of our study was to explore how real-time sharing can affect the way in which small social groups interact and capture photographs. We used a combination of qualitative and quantitative techniques to study groups of participants while they engaged in a self-guided walking tour of a city. Mobiphos was designed for groups of friends, so all of our recruited participants were from preexisting social groups. Due to constraints under which the evaluation was to be conducted, our participants were chosen from the local population. Additionally, identical tour routes were used by each group so that the researchers would have the ability to compare the activities at various stages in the tour across groups with some grounding provided by the context of the tour. While the participants were not strictly tourists, many of the photographs they captured were characteristic of tourist photography, with a strong focus on not just the landmarks shown on the tour map but also on unmarked landmarks that were found during the tour (Chalfen, 1987).

4.1. Procedure

The first step in the study was to have each participant complete a survey regarding their current use of digital cameras, camera phones, and photograph sharing services. We also presented them with questions regarding their most recent experience with sharing photos on their camera in a face-to-face setting as well as their most recent experience with having others share photos with them in a face-to-face setting. The researchers then showed participants how to use all of the features of the Mobiphos application and made sure that all participants knew how to use the features properly. The researchers provided the participants with a tour map which described five historical landmarks within one square mile of the research building. This map was created to look similar to a standard tourist map. The participants were then given a chance to ask any further questions, after which the tour began and no further instructions were given on how or when to use Mobiphos.

For each run, three researchers accompanied the participants. Two researchers captured field notes regarding system use. These notes covered use of technology as well as participant behavior. Due to the fluid, and in many cases subtle, interaction, it was crucial to have at least two researchers taking field notes at all times. In many cases, the groups would physically separate for short periods of time making it impossible for one researcher to properly observe all participants. The third researcher carried a battery-powered wireless router and stayed at the center of the group to maintain connectivity. This researcher also dealt with any troubleshooting required with the Mobiphos devices. The walking tour portion of the trial was approximately 60–90 minutes in duration.

Upon completion of the walking tour, the participants engaged in a photo sorting task designed to determine the usefulness of automated sharing in comparison to traditional sharing practices. The photograph sorting task had three steps. First, participants were asked to look at all of the photographs captured by the group and pick their favorite photographs. Next, participants were shown just their favorite photographs and asked to choose which photographs they had captured themselves. In the last step, participants were asked to look at the photographs they had captured and decide which photographs would be shared through traditional means if Mobiphos did not implement automatic sharing.

After completing the sorting tasks, the participants engaged in a focus group with the researchers. A set of common, open-ended questions were asked to each group. These questions prompted participants to recall their existing practices for sharing photographs in collocated situations and compare those with the experience of using Mobiphos during the walking tour. When a response required further elaboration, researchers asked follow-up questions to understand deeper themes of usage. Each focus group also provided feedback on the general usability and usefulness of the system.

4.2. Participants

Eight groups of participants were recruited for our study. Each group was composed of two to four people who were part of an existing social group. Due to the difficulty in recruiting tourists, all of the participants were either students at the local academic institution or colleagues of the researchers. The first four groups participated in a pilot study with an earlier version of Mobiphos. The pilot study helped inform the final feature set of Mobiphos. The details of the changes made between the pilot study and the final version of Mobiphos are described in Section 7.

The remaining four groups participated in the study procedure described above. The participants were comprised of nine males and four females with an average age of 26 years. All of the participants owned technology used for digital photography, and all engaged in some form of technology-enabled photo sharing. Of the thirteen participants, eleven owned digital cameras and ten participants owned camera phones, at the time of the study. Twelve participants had used a camera phone to take photographs, and six had used a camera phone to send photographs to another camera phone user. All of the participants had

engaged with online services for photo sharing such as Flickr or Facebook, while eleven used online photo sharing sites to share photographs with friends and family.

4.3. Data analysis

Different techniques were used to analyze the data gathered. Statistical methods were used to evaluate the data gathered from the post-tour photo sorting task. The field notes gathered during the tour and those gathered during the focus group were examined with an inductive qualitative analysis. In our previous work, we analyzed the qualitative data along with the quantitative data to determine the best way to understand the effect of the Mobiphos application (Clawson et al., 2008). In this paper, we present the analysis of the log data and photographs captured by the participants to explore the uses of Mobiphos. Additionally, we explore the changes in design from the pilot to final version of Mobiphos in relation to the observed participant behavior.

5. Results

5.1. Interaction with mobiphos

The four groups and thirteen participants took a total of 479 photographs during the walking tour for an average of 36.8 images (SD = 14.1) per participant. On average each group captured 120 photographs. Participants chose to use the enlarge feature an average of 38.6 photographs (SD = 25.3). Of the photographs which were enlarged 10.6 (SD = 8.3) on average were taken by the participant and 28.0 (SD = 16.7) on average were captured by other group members.

5.2. Photo sorting results

The results of the photo sorting task were analyzed to better understand how many of the photographs shared by Mobiphos would not have been shared if the sharing had occurred manually instead of automatically. In the first step of the photo sorting task, participants were asked to choose their favorite photos, from all of the group members' photographs. We found that 22.49% of the photographs chosen were taken by other participants. We then asked the participants to look at only photographs they had captured and choose the ones they would share with their group if they were sharing them in their traditional manner, such as with Flickr or through email. The results of this task showed that of the photos marked as favorites in step one that were not captured by the participant, 60.4% would not have been shared by the person who captured that photograph. If Mobiphos did not automatically share photographs, each participant would not have received 13.58% of the photographs they marked as their favorites through the traditional sharing manner. The users might never have seen, much less received copies of, these photographs unless the photographs were shared at the moment of capture. When asked about the mismatch between the photographs wanted by fellow participants and the ones they marked for sharing, many reasons were given for not choosing certain photographs. Our initial thought was that the photographs were embarrassing in some way and so were actively not chosen for sharing. However, it was found that many of the photographs which were wanted but not shared held meaning for the participant who wanted them but were inconsequential to the participant who had captured the photograph.

5.3. Content of photographs

The types of photographs taken using the Mobiphos system included the iconic landmark photograph characteristic of tourist photography described in Chalfen (1987) as well as the social, playful, mundane, and serendipitous photography characteristic of camera phone use (Battarbee, 2003; Kindberg et al., 2005; Koskinen et al., 2002; House et al., 2005). In addition, participants crafted new types of photographs that were fostered by the particular affordances of collocated–synchronous photography (Clawson et al., 2008). Below we present an analysis of all of the photographs captured. Our analysis is inspired by the types of photographs described in previous work regarding both tourist and camera phone photography. By analyzing the photo content, we arrived at three high level categories: photographs of the tourist landmarks, photographs of people, and photographs of the environment.

Our first category includes characteristic tourist photography as described in Chalfen (1987) with a focus on the landmarks which were along the tour. Each group captured photographs of all of the landmarks described on the tour map. Landmarks were photographed from wide-angle views, as participants neared the building and close-ups to capture more detail such as stain glass windows. Additionally, photographs of signs or plaques showing details of the landmarks were also captured. Table 1 provides a numerical breakdown of the 127 (26.5% of total) photographs captured of landmarks shown on the tour map.

A large number of the photos captured featured the group of participants and accompanying researchers. In this category we include all photographs captured of group

Table 1
A numerical breakdown of all landmark photographs.

Photographs of landmarks	
Wide-angle	41
Close-ups	38
Landmark signs	29
Entire buildings	19
Total	127

Table 2
A numerical breakdown of all photographs of people.

Photographs of people	
Capturing the experience	98
Photographs of non-group members	20
Feet	9
Posing	7
Photographs of self	3
Total	137

Table 3
A numerical breakdown of all photographs of the environment.

Photographs of environment	
Buildings	59
Street	30
Non-landmark signs	28
Nature	25
Close-ups	20
Vehicles	8
Total	170

members using Mobiphos, engaging with the landmarks, or the research team observing the groups. In all there were 137 photographs of participants, research team members, or outside people. The majority of the photographs captured of other people were participants using the Mobiphos application. Twenty of the photographs were of the research team or non-group members. Table 2 provides a numerical breakdown of the 137 (28.6% of total) photographs captured of people.

The majority of all photographs captured were of other parts of the environment. In this category we include all buildings that were not on the tour map, wide shots of streets, shots of nature, plaques or signs not related to landmarks, close-up shots of objects on the street or vehicles. These photographs account for 170 (35.5% of total) of photographs captured. Table 3 provides a breakdown of all photographs categorized as photographs of the environment.

6. Discussion

By using the photographs as sample points into what the participants found to be engaging, we arrived at three high-level categories: participant engagement with the task, engagement with other members of the group, and engagement with the environment. Finally, we explore the evolution of the Mobiphos design from pilot to current version and discuss the motivation for each aspect of the Mobiphos design in relation to observations made during pilot testing and prior work in the area of mobile, collocated photoware.

6.1. Engagement with touring

Each group captured photographs of all of the landmarks described on the tour map. In many cases group members would arrive at a landmark as a cohesive group and then separate in order to capture various aspects of the landmark before rejoining again. It was not uncommon to see multiple pictures of the landmarks taken from a variety of angles. It was common to have one participant move far away from the landmark to take a single wide-angle shot that captured the entire landmark while others would explore the landmark in greater detail, taking close-up shots of various aspects of the landmark such as stain glass windows or lighting fixtures. Photographs of landmarks were also captured from distances in between viewing the entire landmark and minute details. These photographs were generally captured from across the street as a group approached a landmark.

One participant commented that the process of splitting up and quickly capturing multiple angles of the landmarks was like "distributed work". While split up, the participants were able to engage with each landmark in a more in-depth way than if each participant was expected to take all of the photographs by themselves. The immediate sharing enabled by Mobiphos allowed participants to be more engaged with each other as they discussed which aspects of the landmark they had chosen to capture. While the participants did split up to take photographs of the landmarks, they rarely moved outside of verbal communication range. In the few instances where this did happen, it was common for only one of the group members to leave and return after a short interval of time (<1 minute). Additionally, while this group member was outside verbal range, the WiFi signal was strong enough to allow for photographs to be transmitted back to the waiting group members, thus allowing the rest of the group to monitor what the separated member was capturing.

6.2. Engagement with the people

The majority of photographs captured of people showed group members interacting with the Mobiphos application. "Meta-photography", where one participant captured a photograph of another participant capturing another photograph, was quite common. Other users were captured scrolling through the timeline to look at pictures that were captured earlier. In many cases the use of the browsing functionality suggested a transition period where the users were no longer engaged with the touring but interested in what the other participants had captured. The data gathered from the application logs and photographs suggest that these browsing events occurred primarily when there were no landmarks in the immediate vicinity. Other photographs involved the group walking, consulting the map, and looking for the next landmark. There were few of these photographs captured, but this is not surprising because consulting the map was usually a group

activity, at which point there was a disengagement from the system and engagement with the task and the map.

Only two photos were captured of participants posed in front of landmarks, which is interesting in light of the large number and high detail of photographs capturing landmarks. There are numerous possible reasons that might explain the lack of this type of photography. Posing in front of a landmark is generally done to have proof that one visited the landmark. In our case, the tourists were native to the area and as such would be able to revisit the landmark at any time. Additionally, it is possible that the combination of the unique capabilities of Mobiphos and the lack of true novelty in the landmarks encouraged more exploration of the system capabilities, which overrode the traditional practice of capturing photographs of the tourist in front of the landmark. While there was limited interest in having photographs of oneself in front of the landmarks, there was interest in capturing the landmarks themselves. We believe these group capture activities led to a stronger bonding of the group over the photographs, enough so that participants commented on how they were adding to "their collection" and not taking "my photographs".

There were also photographs in which the subject of the photograph was posed without a landmark. Of these photos many are captures of participants from multiple angles. In one case, a participant held their arm in such a way as to be able to take a photo of his hand. He then asked his group members to take pictures at various positions along his arm. In the focus group it was found that this was an attempt to use the L-shaped thumbnail area of the interface to create a panorama image of the participant's arm. In other situations, photographs were taken of people, not for preservation, but to add to the conversation. For example, there were eight photographs where participants captured their own or others' feet. From reviewing the field notes and asking about the pictures, it was revealed that most of these pictures were taken by accident due to a participant's finger being over, and accidentally triggering, the capture button while walking with the camera by their side. While these photos were accidental, in one instance, the capturing of feet prompted a discussion during the trial concerning the abundance of "feet pictures", and a response was made by taking a picture of a participant's backside. Unfortunately, the photograph was not seen in the moment and was later a cause of embarrassment for all parties involved.

6.3. Engagement with the environment

When looking at the photographs of non-landmark buildings, we see three primary buildings captured. Two of these buildings are major corporate headquarters, iconic in their own right, near our institute. The other building is the home of the research lab where the tour started and ended. All of the building photographs are found at the beginning of the tour before the first landmark was encountered. The photographs of the research buildings are believed to be photographs taken as a means of internalizing the capabilities and limitations of the system. While the corporate buildings were not displayed on the tour map, as it only had historical landmarks, they were considered to be landmarks by the participants.

This category included many photos that we have labeled "found objects". In many instances participants would capture these objects for themselves, but the object would then become a point of engagement for the group. At one point a user knelt to capture a photo. The other participants immediately began monitoring their Mobiphos devices to wait for the new photo. The photo captured was of a manhole cover which prompted a discussion of what it was and why it was captured. Here we see how a photograph captured because of one participant's engagement with the environment translated into a photograph that was engaged by the entire group. In focus group interviews, the participant who captured the photograph revealed that she would not have shared the photo if the automatic sharing was not available. The lack of sharing would not have been out of a need to keep the photograph private, but because the participant felt that the rest of the group members would have no interest in seeing the photograph. In another case, a participant chose to capture an object embedded in the street while the other participants had walked past. Upon realizing that they had left someone behind, they came back and watched as they received a photograph of a fork embedded in the concrete. This led to the participant who had captured the photograph to make a joke about "coming to a fork in the road". Again, we have an example of how a "found object" that engaged one participant became a point of discussion for the entire group. In a converse situation, a group encountered a dead bird in the street and one participant expressed an interest in capturing a photograph. Before she captured the photograph, there was a discussion among the three participants concerning the decision to actually take the photo. In this case we have an instance of engagement with the group affecting further engagement with the environment.

7. Design motivations

As stated previously, four groups participated in a pilot version of the study and four groups in the study presented here. The purpose of this pilot study was to validate aspects of the Mobiphos design. In this section, we discuss how feedback from the pilot study affected our understanding of the participants engagement with the system and the implications of the feedback on the design of Mobiphos. By analyzing the interface changes made and the effects of these changes, along with the influence of prior work, we elucidate the design decisions behind the current state of Mobiphos.

In the earliest stages of Mobiphos design, Mobiphos had a unified view of the digital viewfinder and photograph thumbnails. Additionally, all thumbnails, regardless of the

person who captured the photograph, were interleaved into a single timeline which wraps around the viewfinder area. Focus groups with users in both the pilot and final study reinforce the usefulness of this design decision over an interface that is more traditional, such as a separate thumbnail and viewfinder mode. First, by combining the two views, participants were able to capture photos or view enlarged photos while monitoring new photos as they arrived on the camera. Combining all of the photographs into a single timeline also helped increase the sense of the group collective, leading many participants to say that they felt they were taking photographs for "our collection", referring to the group as a whole. The timeline is an example of the "common space" discussed in Salovaara et al. (2006).

Our next design decision stems from the importance of allowing the user to intertwine interactions with the group and the device. Specifically, we found that despite using WiFi, there was still a small delay (<1 s) between when a photograph was captured and when it appeared on all other devices. In many cases, a participant would capture a photograph and the rest of the group would look at their cameras and wonder if the photograph had appeared yet. This led to repeated glancing at the display which broke eye-contact with the other group members, leading to an interruption of the conversation. In many situations, users became frustrated and simply stopped looking for new photos and only captured photographs. The pilot version of Mobiphos did not provide an explicit cue when a new photograph arrived. Based on results found in Oulasvirta et al. (2005) and the issues experienced with our pilot users, the final version of Mobiphos was made to vibrate when receiving a photograph from another user. This external cueing allowed a participant to know when to look at the screen while engaging in the face-to-face interaction with the group, thereby not interrupting the conversation to monitor the device. An additional benefit of the vibration cue was that the other group members could also hear the vibration and knew why the person holding the device was looking at the device. The vibration cue allowed group members to make and understand decisions about when to disengage from conversation to interact with Mobiphos.

Our second finding focuses on the need to augment the contextual knowledge sufficiently such that interactions can remain fluid. Many times, a participant would capture a photo and make a comment to let the other participants know that they may find the photograph interesting. While the vibration cue described above solves the problem of knowing when to look, it did not always solve the problem of which photograph was being discussed. Specifically, when multiple photographs were captured in close succession by multiple participants, extra effort was required on the part of the capturer to point out the photograph they wanted the other group members to focus on. Fundamentally, this was a problem of disambiguation; we had originally explored adding participant names to the corner of each thumbnail, but quickly found that the font size required to make the name legible resulted in a text box which could easily cover over 30% of the photo for a five-letter name. To address this issue, we added colored frames around each photograph to identify the photograph by the person who captured the photograph. Before the colored frames were added, it was common for participants to state that they did not know which photo was being discussed so would disregard the comment. When it was very important to show the photo, participants would revert to traditional sharing practices of showing other participants their screen. In the post-tour focus groups conducted after using the final version of Mobiphos, many participants commented on how it was easy to wait for verbal cues such as "that's a good one ..." or "check that out ..." and then simply monitor the device and watch a photo with the colored frame of the speaker to appear. The mapping of color to capturer became a part of the contextual knowledge of the Mobiphos experience and along with the vibration allowed users to quickly identify the photograph they wished to look at based on external physical/verbal cueing. While the users were not particularly interested in claiming ownership over the photographs, the additional piece of information was enough support these in-the-moment comments and enabled the users to engage more with the photographs as they were captured.

The last design issue we would like to discuss is one that appears when the interactions with Mobiphos begin to occur in the context of "reminiscing talk" as described by Frohlich et al. (2002). While a combination of the physical/verbal cues and the explicit notifications from the system were sufficient for organizing the focus of participants when capturing and sharing immediately, when the context was lost, it became difficult for participants to organize their displays around a specific photograph. In particular, when the tour was winding down and participants were returning to the research lab, participants were likely to scroll through the timeline and enlarge photographs they found interesting. Without the context of having just captured the photograph it was difficult for one participant to relate to others how to find this photograph within the timeline on their own device. At this point, most defaulted to the traditional practice of sharing a single screen. Inspired by existing work done by Kun and Marsden (2007), we decided to implement the focus window described earlier. Because of the flexibility that the ad hoc floor control policy showed in Kun et al., we decided to explore that concept a little further. Our interaction method worked by allowing every participant to know what photographs were being enlarged by the rest of the group. However, unlike Kun et al., our interaction did not enforce the WYSIWIS interface on the displays of the rest of the group, it simply provided each participant with the option of enlarging the same photograph as the rest of the group. While WYSIWIS worked well for Kun et al., it is unlikely to work in the context of a system devised for simultaneous capture and sharing. When asked about this feature in the post-tour focus group, many

participants speculated on how the usefulness of this feature is not during the tour but instead likely afterwards when participants are back at their hotel or at a pub, where the context is lost and a more direct pointer to the photo is needed to get everyone on the same page.

8. Future work

This study, along with Clawson et al. (2008), is an initial foray into examining how a mobile system to support capturing and synchronously sharing photographs can affect the group photography experience. With further iterations to Mobiphos, we would like to explore how our application works with more standard camera hardware. How does transitioning from camera phone quality images to higher quality digital cameras affect the capture style of the user? We would also like to explore the use of the cellular mobile phone network to accomplish wireless sharing of photos. This would enable us to explore mixed situations with groups of distributed users (as in Jacucci et al., 2005; Salovaara et al., 2006; Sarvas et al., 2005), with groups collocated users as with Mobiphos, as well as the transitions between. We would be able to better explore the role of face-to-face engagement, or the lack of it, in the application design. Exploring mobile collocated–synchronous mediated experiences more generally with other mobile media such as video or text could provide a different perspective on our findings.

Finally, we would be very interested in deploying our application in more realistic settings. It would be very interesting to see how larger groups would utilize Mobiphos and to see how it scales to support more users. Likewise, it would be interesting to deploy the application in more complex social settings and for longer periods of time to explore the impact of the application's real-time sharing capabilities. For example, a longitudinal deployment in which we send the system overseas with a group of summer study-abroad students and monitor their use of the system for a three-month period could reveal totally new rhythms and patterns of use as the social dynamics of the group change over time.

9. Conclusions

Mobiphos has allowed for the blending of photograph capture and sharing in a near-synchronous, collocated and highly situated way. The concerns with how to share and with whom have been removed from the thoughts of the users to allow them to concentrate on the group experience. While the technologies used are standard, they have come together to affect and alter the group experience in ways that the group engaged with the environment and with each other. Mobiphos has also shown that with increasing advances in mobile technology, there is a potential to design applications that not only add to the external experience but also increase social bonds among friends. With Mobiphos, users have been able to blend both the individual and social aspects involved when taking photographs in a small, collocated social group to create an engaging user experience.

Acknowledgment

This work was supported in part by funding from Samsung.

References

Ashbrook, D., Lyons, K., Clawson, J., 2006. Capturing experiences anytime anywhere. IEEE Pervasive Computing 5 (2), 6–7.

Balabanovic, M., Chu, L.L., Wolff, G.J., 2000. Storytelling with digital photographs. In: CHI '00: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM Press, New York, NY, USA, pp. 564–571.

Battarbee, K., 2003. Co-experience: the social user experience. In: CHI '03: CHI '03 Extended Abstracts on Human Factors in Computing Systems. ACM Press, New York, NY, USA, pp. 730–731.

Becker, H.S., 1982. Art Worlds. University of California Press, Los Angeles, CA.

Brinck, T., Gomez, L.M., 1992. A collaborative medium for the support of conversational props. In: CSCW '92: Proceedings of the 1992 ACM Conference on Computer-supported Cooperative Work. ACM, New York, NY, USA, pp. 171–178.

Chalfen, R., 1987. Snapshot Versions of Life. Bowling Green State University Popular Press, Bowling Green, OH.

Clawson, J., Voida, A., Patel, N., Lyons, K., 2008. Mobiphos: a collocated–synchronous mobile photo sharing application. In: MobileHCI '08: Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services. ACM, New York, NY, USA, pp. 187–195.

Counts, S., Fellheimer, E., 2004. Supporting social presence through lightweight photo sharing on and off the desktop. In: CHI '04: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM Press, New York, NY, USA, pp. 599–606.

Crabtree, A., Rodden, T., Mariani, J., 2004. Collaborating around collections: informing the continued development of photoware. In: CSCW '04: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, pp. 396–405.

Frohlich, D., Kuchinsky, A., Pering, C., Don, A., Ariss, S., 2002. Requirements for photoware. In: CSCW '02: Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work, pp. 166–175.

Greenberg, S., Marwood, D., 1994. Real time groupware as a distributed system: concurrency control and its effect on the interface. In: CSCW '94: Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work. ACM, New York, NY, USA, pp. 207–217.

House, N.V., Davis, M., Ames, M., Finn, M., Viswanathan, V., 2005. The uses of personal networked digital imaging: an empirical study of cameraphone photos and sharing. In: CHI '05: CHI '05 Extended Abstracts on Human Factors in Computing Systems, pp. 1853–1856.

Jacucci, G., Oulasvirta, A., Salovaara, A., Sarvas, R., 2005. Supporting the shared experience of spectators through mobile group media. In: GROUP '05: Proceedings of the 2005 International ACM SIGGROUP Conference on Supporting Group Work. ACM Press, New York, NY, USA, pp. 207–216.

Kindberg, T., Spasojevic, M., Fleck, R., Sellen, A., 2005. The ubiquitous camera: an in-depth study of camera phone use. IEEE Pervasive Computing 04 (2), 42–50.

Kohno, M., Rekimoto, J., 2005. Searching common experience: a social communication tool based on mobile ad-hoc networking. In: MobileHCI '05: Proceedings of the Seventh International Conference

on Human Computer Interaction with Mobile Devices and Services, pp. 15–22.

Koskinen, I., Kurvinen, E., Lehtonen, T., 2002. Mobile Image. IT Press.

Kun, L.M.A., Marsden, G., 2007. Co-present photo sharing on mobile devices. In: MobileHCI '07: Proceedings of the Ninth Conference on Human–Computer Interaction with Mobile Devices and Services, pp. 73–80.

Mäkelä, A., Giller, V., Tscheligi, M., Sefelin, R., 2000. Joking, storytelling, artsharing, expressing affection: a field trial of how children and their social network communicate with digital images in leisure time. In: CHI '00: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 548–555.

McDonald, D., 2007. Visual conversation styles in web communities. In: Proceedings of the 40th Hawaii International Conference on System Sciences.

Miller, A.D., Edwards, W.K., 2007. Give and take: a study of consumer photo-sharing culture and practice. In: CHI '07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 347–356.

Nardi, B.A., Schiano, D.J., Gumbrecht, M., 2004. Blogging as social activity, or, would you let 900 million people read your diary? In: CSCW '04: Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work, pp. 222–231.

Oulasvirta, A., Tamminen, S., Roto, V., Kuorelahti, J., 2005. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile hci. In: CHI '05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, New York, NY, USA, pp. 919–928.

Salovaara, A., Jacucci, G., Oulasvirta, A., Saari, T., Kanerva, P., Kurvinen, E., Tiitta, S., 2006. Collective creation and sense-making of mobile media. In: CHI '06: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM Press, New York, NY, USA, pp. 1211–1220.

Sarvas, R., Oulasvirta, A., Jacucci, G., 2005. Building social discourse around mobile photos: a systemic perspective. In: MobileHCI '05: Proceedings of the Seventh International Conference on Human Computer Interaction with Mobile Devices and Services. ACM Press, New York, NY, USA, pp. 31–38.

Voida, A., Mynatt, E.D., 2005. Six themes of the communicative appropriation of photographic images. In: CHI '05: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 171–180.