

Capturing Rich Media through Media Objects on Smartphones

Florian Güldenpfennig, Wolfgang Reitberger and Geraldine Fitzpatrick

Vienna University of Technology, Argentinierstr. 8, Vienna, Austria

{florian.gueldenpfennig, wolfgang.reitberger, geraldine.fitzpatrick}@tuwien.ac.at

ABSTRACT

New mobile phones with built-in sensors provide the opportunity to easily capture vast amounts of multimedia data. In this paper we present a file format to explore new ways of relating digital content: the *media object (mob)*. Mob was designed as a tool for serving three concerns: first, to provide the user with an object structure that allows them to capture and associate multimedia and contextual data (photo, video, text, locations and so forth) especially on mobile phones; second, to put them in a creative frame of mind when recording media; third, to learn more about peoples' common practices as well as desires around digital photography and digital support for memory. We report on an exploratory user study featuring 7 participants over 2-5 weeks and conclude with a set of future requirements.

Author Keywords

Archival, creativity, data capture, digital photography, memory, mobile phones, multimedia, research through design, smartphones

ACM Classification Keywords

H3.m. Information storage and retrieval: Miscellaneous.

INTRODUCTION

In recent years, HCI has increasingly been concerned with digital memory systems (Sellen and Whittaker, 2010) and the question of how digital technology can further support us in preserving valued moments, objects and other digital information. Advancements and developments in engineering such as immense storage capacities or ubiquitous cameras give us the unprecedented opportunity to create new forms of digitally supported memories. What can be 'remembered' can be anything that comes in an electronic form or that can be captured digitally: photographs, audio recordings, emails, location, etc. One of the challenges is around supporting people creating order over digital material and making sense of it. For instance, Whittaker et al. (2010) raised the issue about the difficulties in retrieving digital pictures of salient family events.

In this paper, we introduce a new file format *media object (mob)*, and associated applications (App; plural: Apps) *mobRecMobile* and *SimpleMobView*, for supporting the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

OZCHI'12, November 26–30, 2012, Melbourne, Victoria, Australia.

recording of digital resources for memories. It was designed to serve three particular concerns. First, it provides the user with an object structure or container that allows them to capture and associate multimedia and contextual data (photo, video, text, locations and so forth) especially on mobile phones. Second, media object attempts to put the users into a creative frame of mind when recording media. Third, we seek to learn more about peoples' common practices as well as desires around digital photography, memories and reminiscence by providing the user with our novel media object tools and seeing what they make of it. To explore this, we gave the mob applications *mobRecMobile* and *SimpleMobView* to 7 test participants for recording and replaying mob files. In this paper we describe the mob concept and report on the participants' engagement with the system, and their thoughts about digital photography and memories in conjunction with the media object. Findings from the study also point to requirements for future mob systems.

THE MEDIA OBJECT

Media object (mob; plural: mobs) is a file format that allows the bundling of a theoretically unlimited number of arbitrary files using XML for organizing the content. Along with images, a rich variety of additional media files (e.g. video, audio, text or additional media objects) and metadata (ID, date, location, access statistics, etc.) can be stored into one memory file. The underlying concept of this data structure is to extend traditional media files, e.g. images, beyond common metadata information while keep everything in one aggregated place. Technically speaking the media object is a container that can wrap any kind of information.

In this way mob is similar to other products. For example, *Evernote* is a commercially successful software platform for "remembering everything" (2012). Moreover, there are standards such as *MPEG-7* that share common aspects with mob as they also employ XML to describe multimedia material. The value in creating our own interpretation of a multimedia container is that it enables us to have an extendable structure and to customize the media object software to the users' needs and to the researchers' focus of interest. It also allows us to have open access to the collected data. Hence, the focus of this paper is not the technical innovation of the mob, but to explore the way people capture and use such objects.

The special feature of mob is that it provides a container that exists *a-priori* to data capture unlike e.g. conventional folders that are created often-times after data capture. Consequently, the users have to decide what

to fill into that container or even ‘plan’ their digital resources, namely what (multimedia) data is to be captured to make a good memory. The mobile phone then potentially offers a convenient way for creating such data. We were curious whether such a tool could engage the users in creating original digital resources for memories and what practices they would develop around this.

Description of the .mob Software

We instantiated this mob concept in two software components: *mobRecMobile* and *SimpleMobView*. *MobRecMobile* (displayed in Figure 1) was developed for mobile phones operated with Android. It enables the user to capture data and to store it into media object files. These mob files are saved to SD card or to internal phone memory and can be transferred to a desktop computer in the same way as other conventional file types (via USB cable), e.g. *JPEG* images. Besides creating and storing media objects the user can edit and view existing mob files. Editing mob files comprises primarily the following actions:

- Capture a new image / video / audio note / text note / geo location (GPS) and add it to the current mob file
- List all files contained in the current mob file
- Open and display a specific file (e.g. an image or text note) contained in the current media object

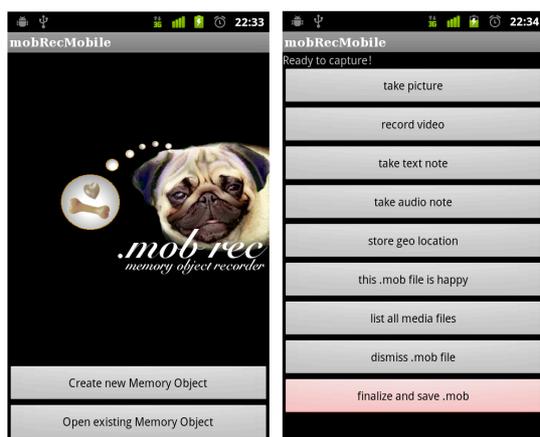


Figure 1. Start screen (left) and options menu (right) of the *mobRecMobile* App.

SimpleMobView is desktop computer software that allows the user to view media objects. It displays a list of all files that are contained in a media object. The user can select and display entries from that list. Videos and images can be viewed simultaneously. Audio notes can be replayed in the background while interacting with images, text notes or geo locations.

Both *mobRecMobile* and *SimpleMobView* make use of metadata (e.g. for usage statistics), which is also contained in the media objects. Since both mob applications employ this feature for internal and technical reasons only, we do not give a detailed description of mob metadata here.

USER STUDY

To explore what people make of the mob concept, we recruited 7 participants in their 20s and 30s from a variety of backgrounds. Every subject either owned or was provided with an up-to-date smart phone. The *mobRecMobile* App was installed on each device and the participants were asked to use it for 2 weeks. However, three participants (P1, P2, P3) with own smart phones kept using our App for 3 more weeks. Since the study is exploratory and no comparisons between participants are made, we also included this additional material into our data pool. They were given no further instructions besides the invitation to capture anything with the App they liked to record. After the test period each participant reported their experiences in a semi-structured interview (lasting between 50 and 80 minutes), while also viewing mob files on the *SimpleMobView* App. The interviews were audio taped, transcribed and then analyzed qualitatively looking for common themes. We also did a thematic analysis of the content of their mobs.

P	Duration (weeks)	n mobs created	n files in mobs
1	5	15	226
2	5	8	71
3	5	22	77
4	2	7	26
5	2	6	36
6	2	3	10
7	2	5	36

Table 1. Participants (P) in the exploratory study.

RESULTS

The participants created media objects on many different occasions. In total 482 media files (images, video, audio, text, GPS) were captured and were distributed over 64 media objects (see Table 1). These mobs differ strongly in content and in their properties. Also, the creation time-span varied from short (minutes) to longer periods of time (where mobs grew successively over multiple days). Here we provide a list with example events or occasions that have been captured into media objects:

- P1: impressions of a short trip (see Figure 2)
- P1: documentation of childhood memories (various cherished physical mementos)
- P2: long winter walks
- P2: Christmas Day and a New Year’s party
- P3: sound installation at a museum
- P3: night at the opera
- P4: indoor climbing wall visit (see Figure 3).
- P5: preparation of a special tea ceremony
- P5: beverage bottles and the sounds they make when opening them

- P6: snowfall in the city
- P6: breakfast with parents-in-law
- P7: documentation of an electronics tinkering project

The vast majority of media files contained within the media objects are images. However, all except 17 mobs are comprised of different file types (e.g. images combined with audio etc.). The most common media object consists of a set of images accompanied by a video clip or text note. Figure 2 illustrates such a typical mob file. It displays two pictures that have been enriched by P1 with text and audio notes, because this was “a most appropriate way to capture atmospheres” (P1). Two more exemplar media objects are shown in Figure 3 and 4. Both contain images as well as movies. P3 states that images are “for capturing moods and colors”, whereas movie clips are made for “things that happen”. Accordingly, she tried to capture her first impression of a fun fair with an image and made a video clip later on for showing and commenting on the children’s play in the straw (Figure 3). P4 provided orientation of an indoor climbing wall by means of a video pan shot. Further still images went into the same mob, for instance a detail picture of a climbing grip (see Figure 4).



Figure 2. “It starts! I am on my way. The train was 5 min late, but I have a good seat. Brrrrr it is freezing outside.” – Text note by P1 referring to the left image. The right image is extracted from the same media object. In an audio note P1 explains, that it was the first night at the hotel shown in the picture.

Through the analysis of the collected media objects and interview data we could identify prevalent usage patterns and observations regarding mob that occurred in the course of the study. These include:

- Most participants reported having mild to severe problems managing their conventionally recorded digital media (mostly photos) in terms of ordering or relocating data. They also see a potential for media objects to bring more order to their data collections due to the fact that the mob container is created a-priori and find the concept useful for retrieval. 5 participants stated that they wanted to continue using media objects. 2 participants appreciated the mob concept, but stated that mob needs some extra features before they could imagine using it in a real life setting (see Future Work).
- A great number of media objects were created by the participants for personal reminiscence and reflection. P3 for example talked about mobs as “virtual memory albums” (P3). Some mobs have been captured for sharing experiences with others and story telling. A

couple of mobs represent strange or funny things and have been made for pure amusement.

- Participants reported capturing media objects in addition to, rather than replacing, regular photography. It was considered as a creative act and *mobRecMobile* was reported to inspire creativity, because “the line-up of tools is always visible” (P7).
- The participants appreciated the fact, that media objects could be “unzipped” by common file archiving software in order to ensure future data access.



Figure 3. Picture of a fun fair (left). Still from a movie displaying children’s play in the straw (right).

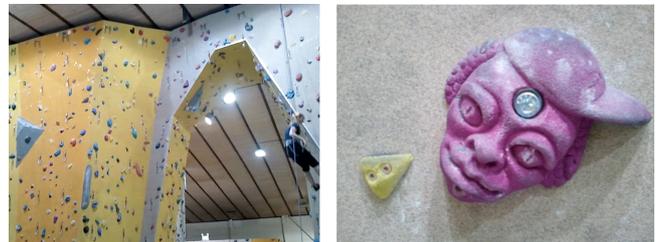


Figure 4. Still from a movie showing and commenting on a climbing wall (left). Detail picture of a climbing grip (right).

RELATED WORK AND DISCUSSION

We go on here to reflect on these findings against related work. The fact that most participants said they struggled at times to manage their conventional digital image collections is in line with the literature (Churchill and Ubois, 2008; Whittaker et al., 2010). Overall the participants see a potential for mob to structure digital memories because media objects constitute containers or folders that are created deliberately at the beginning of data capture. Hence, being created and sorted in the course of data capture, media object takes away the burden of sorting materials after the event. These mobs were also quite deliberately constructed, which leads us to hypothesize that mob files can support sense making: due to the fact that (multimedia) data is aggregated in one container the user might be provided with a valuable source for establishing context and for the reconstruction of memories.

Recent literature in HCI questions so-called life logging technologies, pointing out that memories are not recorded but reconstructed (Van den Hoven and Eggen, 2008). A more deliberate capturing of data, and alternatives to the storage and retrieval model have been proposed that put broader emphasis on functions of human memory such as reminiscence and reflection (Sellen and Whittaker, 2010; Peesapati et al., 2010). We suggest that the media object concept can eventually support these aspects of memory, given that recording and replaying mob files can provide a frame for a deliberate creation of memory resources for

later reminiscence and sense making. This assumption finds some support by the fact that the participants' creativity was encouraged by *mobRecMobile*. P4 stated in conjunction with Figure 4, that he "would never have captured the climbing grip [without mob], which actually made a nice memory". Hence, media object might continue to engage the user in capturing (multimedia) data in original ways due to the fact that "a media object is like an empty album that is inviting me to fill it [...] and showing me all the tools and possibilities at the same time" (P1). In an exploratory study by Petrelli et al. (2009) participants were motivated to put much effort into the creation of mementos. Other research to support the value of creative engagement with mobile phones and picture taking can be found in (Ljungblad et al., 2004).

Further findings from research on personal photography and camera phone use apply to our data set as well. Van House describes personal photography as "multiple overlapping technologies: of memory; relationships; self-representations and self-expression" (2011; p. 126). Kindberg et al. (2005) identified social, personal, affective and functional purposes for the use of mobile phone cameras. We can find some evidence for all of this in our data. Furthermore, we hypothesize that the media object will be able to support all of these categories due to mob's properties described above. Here, the fact that mob combines arbitrary files certainly plays an important role.

A rather well known technology, which also can take multiple scans or shots of a memento, is the *Family Archive* (Kirk et al., 2010). There are various other research efforts that try to extend regular picture taking, e.g. *Audiophotography* (Frohlich and Tallyn, 1999). However, in contrast to the mentioned technologies, media object at its essence is a general concept or file format specification and is designed to be instantiated into various mob devices, so is a more portable solution.

CONCLUSION AND FUTURE WORK

Encouraged by overall positive feedback and the fact that three participants voluntarily kept using the applications, we plan to conduct further research into the media object to improve its implementation and user experience. At the



Figure 5: Dedicated mob player on a bookshelf. The device displays random mob files at the touch of a button.

same time we seek to better understand contemporary amateur photography and reminiscence through the engagement with the media object. This study revealed a number of requirements and exemplary features that should be considered for the further development of mob recorder and viewer software:

- Enable more creativity and provide tools for play: e.g. support sketching with a stylus, feed music into the mobs, save images of web pages
- Facilitate sharing of mobs and story telling through mobs by means of a web service that displays media objects in the browser
- Implement dedicated mob player devices. One prototyped instance is shown in Figure 5. Enable serendipitous exploration of one's mob collection
- Consider linking mobs to physical mementos by means of tags (e.g. barcodes, RFID) or optical characteristics (pattern recognition)

As the memory object is designed to be easily extendable we plan to adapt the current implementations of *mobRecMobile* and *SimpleMobView* in order to meet the above listed criteria and hence make mob our research vehicle for looking more into peoples' everyday data capture and desire for preserving (digital) memories.

REFERENCES

- Churchill, E., and Ubois, J. Ps and Qs: Designing for digital archives. *Interactions* 15, 2 (2008), 10–13.
- Evernote. <http://www.evernote.com/>. Last access: August 13th, 2012.
- Frohlich, D., and Tallyn, E. Audiophotography: practice and prospects. *Ext. Abstracts CHI 1999*, ACM Press (1999), 296–297.
- Kindberg, T., Spasojevic, M., Fleck, R., and Sellen, A. The ubiquitous camera: An in-depth study of camera phone use. *IEEE Pervas Comp* 4, 2 (2005), 42–50.
- Kirk, D.S., Izadi, S., Sellen, A., Taylor, S., Banks, R., and Hilliges, O. Opening up the family archive. In *Proc CSCW 2010*, ACM Press (2010), 261–270.
- Ljungblad, S., Hakansson, M., Gaye, L., and Holmquist, L.E. Context photography: modifying the digital camera into a new creative tool. *Ext. Abstracts CHI 2004*, ACM Press (2004), 1191–1194.
- Peesapati, S.T., Schwanda, V., Schultz, J., Lepage, M., Jeong, S., and Cosley, D. Pensieve: supporting everyday reminiscence. In *Proc CHI 2010*, ACM Press (2010), 2027–2036.
- Petrelli, D., van den Hoven, E., and Whittaker, S. Making history: intentional capture of future memories. In *Proc CHI 2009*, ACM Press (2009), 1723–1732.
- Sellen, A.J., and Whittaker, S. Beyond total capture: a constructive critique of lifelogging. *Communication ACM* 53, 5 (2010), 70–77.
- Van den Hoven, E., and Eggen, B. Informing augmented memory system design through autobiographical memory theory. *Personal Ubiquitous Computing* 12, 6 (2008), 433–443.
- Van House, N. Personal photography, digital technologies and the uses of the visual. *Visual Studies* 26, 2 (2011), 125–134.
- Whittaker, S., Bergman, O., and Clough, P. Easy on that trigger dad: a study of long term family photo retrieval. *Personal Ubiquitous Computing* 14, 1 (2010), 31–43