

Flexible Views: Annotating and finding context-tagged mobile content

Frank Bentley and Crysta Metcalf

Applications Research Center

Motorola Labs

1295 E. Algonquin Rd

Schaumburg, IL 60196 USA

[f.bentley, crysta.metcalf]@motorola.com

ABSTRACT

We present the results of three studies on the use of contextual metadata tags to find personal photographs. Specifically we address the use of time, event, and location tags to allow users to find media from their collections. We found that users need to be provided with flexible ways to browse their content and that they remember quite different attributes of each piece of content. Finally, we discuss the need for flexible photo search to support photo sharing and photo talk, two key uses of photographs.

Keywords

Metadata, location, tagging, mobile search, photographs

INTRODUCTION

Mobile devices provide many opportunities and challenges for capturing, finding and sharing media. Many of today's mobile devices make it simple to capture media but many do not provide simple ways to browse for that captured media. On a mobile device with large amounts of storage, but with limited input and navigation functionality, it becomes increasingly important to develop tools that allow users to easily manage and share their digital photographs.

In the summer of 2003, we conducted a series of lab concept tests investigating the use of contextual cues such as time, event, and location to help users easily find photos on a limited interaction mobile device. These studies helped inform the design of later iterations of our application as shown in [3]. All of this work was a follow-on to several ethnographic-style studies into photo use as described in [1].

PREVIOUS WORK

Many researchers have studied the use of contextual tags to automatically annotate media and aid in image search. The most notable commercial example is the ZoneTag [11] system by Yahoo! Research Berkeley. Zone Tag allows users to automatically apply location information to photos

(City/State/Zip Code) as well as apply other contextual tags from community annotations with minimal effort; however, they do not currently provide a way to browse photos using these tags from a mobile device.

Other work has explored using hierarchical location representations to help organize photos [8] and has developed meaningful abstractions for semantic location information [5].

Chalfen [2] and Walker and Moulten [10] both found that an important way that people conceptualize their photographs is in terms of "event." Graham et al [4] developed an algorithm for clustering photos into short burst-like clusters and episode-based clusters spanning at most several hours. Users could browse these clusters using an absolute date-based hierarchy or through a calendar interface. They found that the cluster-based approach was on average 33% faster than browsing a flat structure in ACDSec.

Microsoft [9] and Kodak [7] have also explored the use of event based clusters to find photographs. The Microsoft team found that with their algorithm, 79% of the images returned in a time cluster were semantically meaningful for the two users in their trial. Kodak similarly found that that they could achieve a recall rate of 80% and a precision rate of 79%.

EVENT AND LOCATION STUDIES

Following our work on photo sharing [1], we began to develop interfaces to facilitate simple browsing of photos for reminiscing and sharing. We conducted two initial studies focused on specific metadata attributes. The first was focused on the ability to automatically determine event groupings for a set of photos and to browse these groupings on a small-screen device. Five participants shared their digital photo collections with us. Participants were asked to manually sort selected subsets of their photos into "event" groupings. We compared the manually sorted groupings against our own event clusters which were created using a time based algorithm based on EXIF date/time tags extracted from each photo. We then had participants talk about the groups of photos that the algorithm created while listening for occasions where the

City Name	Event	Photos		
Map	Event	Photos		
Map	Location	Event	Photos	
Year	Month	Event	Photos	
Year	Month	Location	Event	Photos
Season	Location	Event	Photos	

Figure 1: Interface paths for the combined application. Users could progress across any of the rows of the table in order to browse for photographs from their collection.

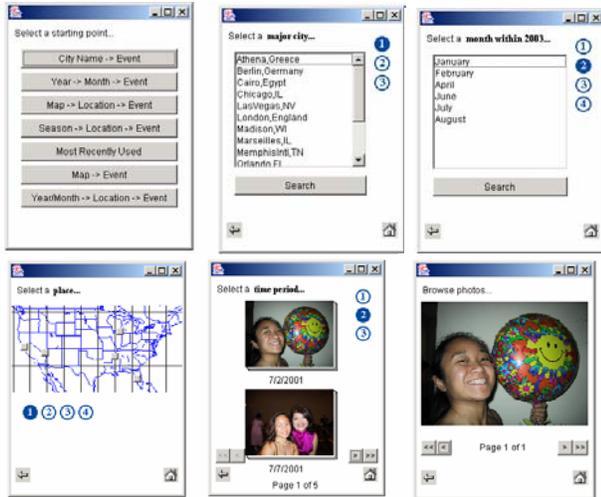


Figure 2: Screenshots from the application

descriptions covered multiple events or where successive groups of photos were from a single event. We found that events are a meaningful concept for organizing photos, but that time alone was not enough to accurately segment events. As participants often mentioned a change in location as a reason for a new event to begin, we then turned to location interfaces.

The second rapid study focused on location groupings as a way to browse photo collections. Six new participants shared their photographs with us over the course of two interviews. In the first, participants gave us a CD of their digital photos and described various information about their pictures to us (location, people, event, etc.). We recorded the locations where the photos were taken and later loaded absolute location metadata (latitude and longitude values) into our application. We provided three ways to browse for photos: a zoomable map, a city list, and a thumbnail based hierarchal browser showing a representative picture from each location.

We found that the map interface worked well on the level of the United States, but when users zoomed into a metro area they often were unable to disambiguate the different locations and often chose the wrong location grouping (7 of 28 searches).

In each of the studies we saw our users trying to use both time and location as memory aids for finding their photos, even though each interface focused only on a specific metadata attribute in isolation. Following these insights, we set out to make an interface that allowed users to combine time, location, and event information to search their photos.

INITIAL COMBINED PROTOTYPE

We created an application on top of the Media Assistant architecture [3] that would allow users to browse for their photos using a number of contextual cues. This interface (see Figures 1 and 2) provided a number of fixed paths for users to select from to browse for photographs from their collections. The screens were sized to run on a mobile device with a 240 x 320 screen resolution.

TIME AND LOCATION STUDY

We recruited six participants from an internal database of volunteers. Participants were selected that had a large number of digital photographs as in the first two studies. Two interviews were conducted. At the first interview, we received a CD of the participant's photographs (average 1727 photos) and had them describe aspects of the photos (location, people, date, etc.) to the best of their memory. We then tagged the photographs with location information and used the times from the EXIF tags to supply date tags as well as to make event groupings.

For each of the six interface paths shown in Figure 1, we chose what we thought would be an easy, medium, and hard photo to find based on the initial interview. For example, if participants hesitated in remembering location, or needed to look through a group of photos in order to determine location, a photo was selected from this group as a "medium" or "hard" target photo for interfaces that began with location. (Subsequent analysis showed that our selection process for selecting photos accurately predicted success rates for actually finding the photos).

Two weeks after the first interview, we returned to the participants and asked them to find each of the selected photographs in the specific interfaces that we had chosen. At the start of the second interview, we demonstrated the interfaces and how to perform the technique of thinking aloud using the photo collection of one of the interviewers. Then, participants were able to browse through the interface for three minutes to discover how various screens work and where some of their photos might be stored. During this time, they were also practicing the task of thinking aloud.

After the three minutes were up, or they felt comfortable to continue, we presented them with each of 18 selected photos and asked them to find the given photo in a particular interface. Interface order was balanced across participants. At the end of the interview, a final set of twelve pictures was presented to the participants and they were allowed to choose from the six interface paths to find each photo. They were also free to return to the starting

screen and retry the search using other interfaces if they wanted to do so.

OBSERVATIONS

Our major finding was that a predefined search path is not the most appropriate way for a user to find his/her own personal content. In a variety of circumstances, users remember different pieces of context about a picture. For example holidays like birthdays are often remembered by month as in TL3; “This was Emily’s birthday but I don’t remember which birthday, maybe if I could pick by month... you wouldn’t necessarily know the year.” TL4 also expressed such a desire; “This picture was for my dad’s birthday and I believe it was in October so that gives me the month. And I think it was around 2001 or 2002.” However for other photos, there are examples where the same participants remembered different information (e.g. locations and seasons, seasons and years, seasons without years, sub-locations and months, etc.). TL4: “This is in my boyfriend’s brother’s house in TN. I can’t remember exactly when we went but that was in Memphis.”

Therefore we see the need for a flexible interface that provides users with the ability to select the attributes that they currently remember to move to the right results. These attributes should be able to be chosen sequentially while a user navigates the search interface as we discovered several occasions of users starting over with new search paths as they remembered more context of a particular photo.

Another important finding is that users often do not remember the contextual metadata around an event accurately. Several times, users were convinced that a given photograph was from a particular year and later found it under a different year. Interfaces for photo search should allow for easy modification of search criteria or simple ways to see “nearby” media that does not exactly match the search criteria.

FUTURE WORK

Our work and the work of others have shown us the importance of photographs for reminiscing and storytelling purposes in people’s lives. We are interested in investigating the types of images that people share with others in a mobile environment and the reasons for sharing. We would like to investigate how contextual metadata tags and a metadata-based browsing experience can help people find images that they would enjoy seeing and remembering as well as find images to share with others. To that end, we are exploring the use of contextual community tagging both to annotate images for easy retrieval as well as to identify images that are most relevant to view or share at a given time. Our current work explores using ties between music listening and photo capture to create richer interfaces for viewing personal photos as well as photos from members of one’s social network.

ACKNOWLEDGMENTS

We thank Emilee Rader for her help with these studies.

REFERENCES

1. Bentley, F., Metcalf, C., and Harboe, G. 2006. Personal vs. commercial content: the similarities between consumer use of photos and music. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, April 22 - 27, 2006). CHI '06. ACM Press, New York, NY, 667-676.
2. Chalfen, Richard (1987). *Snapshot Versions of Life*. Bowling Green State University. Popular Press.
3. Gandhi, B.; Martinez, A.; Bentley, F., "Intelligent multimedia content management on mobile devices," *2004 IEEE International Conference on Multimedia and Expo, 2004. ICME '04*, vol.3, pp. 1703- 1706 Vol.3, 27-30 June 2004.
4. A. Graham, H. Garcia-Molina, A. Paepcke, and T. Winograd. Time as essence for photo browsing through personal digital libraries. In *Second ACM/IEEE-CS Joint Conf. on Digital Libraries*, 2002.
5. Gurrin, C., Jones, G. J., Lee, H., O'Hare, N., Smeaton, A. F., and Murphy, N. 2005. Mobile access to personal digital photograph archives. In *Proceedings of the 7th international Conference on Human Computer interaction with Mobile Devices & Services* (Salzburg, Austria, September 19 - 22, 2005). *MobileHCI '05*, vol. 111. ACM Press, New York, NY, 311-314.
6. A. Kuchinsky, C. Pering, M. L. Creech, D. Freeze, B. Serra, and J. Gwizdka. Foto²le: a consumer multimedia organization and retrieval system. In *Conference on Human Factors in Computing Systems CHI'99*, pages 496-503, 1999.
7. Loui A.C. & Savakis A.E. (2000). Automatic Image Event Segmentation and Quality Screening for Albuming Applications. Paper presented at *IEEE International Conference on Multimedia and Expo*. New York, New York.
8. Naaman, M., Song, Y. J., Paepcke, A., and Garcia-Molina, H. 2004. Automatic organization for digital photographs with geographic coordinates. In *Proceedings of the 4th ACM/IEEE-CS Joint Conference on Digital Libraries* (Tuscon, AZ, USA, June 07 - 11, 2004). *JCDL '04*. ACM Press, New York, NY, 53-62.
9. J. C. Platt. AutoAlbum: Clustering digital photographs using probabilistic model merging. In *Pmc. IEEE Workshop on Content-Based Access of Image and Video Libraries*, pages 9~100, 2000.
10. Walker, A.L. and Moulton, R.K. (1989). Photo Albums: Images of Time and Reflections of Self. *Qualitative Sociology*, 12 (2): 155-182.
11. <http://zonetag.research.yahoo.com/zonetag>