Using mobile phone contextual information to facilitate managing image collections

Jakob Eg Larsen Maciej Luniewski
Technical University of Denmark
Department of Informatics and Mathematical Modeling
Richard Petersen Plads, Building 321
DK-2800 Kgs. Lyngby, Denmark
jel@imm.dtu.dk, maclun@gmail.com

ABSTRACT
In this paper, we describe a prototype application that utilizes the embedded sensors in advanced mobile phones to infer meaningful contextual information, with the potential to support the users in managing their personal information. Contextual information such as time, location, movement, surrounding networks, devices, people, and application data is used to semi-automatically annotate information in our current proof-of-concept prototype. The application allows the derived contextual information to be annotated as tags to available content and thereby facilitating the processes involved in personal information management. We hypothesize that information inferred from embedded mobile phone sensors can offer useful contextual information for managing personal information, including the domain of interest here, namely image collections. This has potential for individuals as well as groups managing shared image collections or other types of information.

Keywords
Personal information management, context, context-awareness, metadata, image collections, mobile phones, embedded sensors

INTRODUCTION
Personal information management can require tedious efforts by the person managing the information even in modern interactive tools aiming to support the activities involved. This involves activities such as to create, store, organize, retrieve and use information to handle daily activities. For instance organizing the information into hierarchical structures in email and file systems. The concept of tagging has become popular in Web 2.0 social networks characterized by user generated and organized content. By labeling the content with one or more tags different ways of navigating and retrieving the information is made possible. However, manually labeling content with tags requires efforts from the user. This introduces a dilemma of storing and retrieving personal information. If less effort is spent on storing the information (e.g. a piling strategy) more effort is required to retrieve the information and vice versa [7]. By choosing a personal filing strategy individuals implicitly choose a balance of effort spent on storing and retrieving information respectively. Thus a challenge for tools aiming to support personal information management could be to minimize, to the possible extent, the effort needed for both storing and retrieving information items.

We introduce a proof-of-concept application implemented for modern mobile phones, which allows the user to semi-automatically annotate information with contextual information [4] derived from the multiple embedded sensors. In the specific application prototype facilitating the management of images captured using the embedded camera has been the focus. The semi-automatically annotated tags can be used for storing, retrieving and managing collections of images. By presenting the derived contextual tags to the user the aim is to support the users in recalling and recognizing their personal information organization when later retrieving the information. The importance of such contextual information in personal information management was emphasized by Barreau [1].

Even though the focus here has been on image collections, we speculate that the principles of semi-automatic metadata annotation (tagging) can apply to any information type dealt with as part of everyday personal information management. The collective sources of sensor information provide descriptions of the multiple contexts in which the user of the device is part of on a daily basis. The rich set of sensor information available on modern mobile phones provide multiple ways of potentially inferring contexts such as location, places, people, environment, activities and movement. When tags are explicitly chosen by the user it also serves as implicit feedback to the system in order for similar contextual situations to be recognized at a later stage and thus over time improve the likelihood of suggesting useful tags.

CONTEXT GALLERY APPLICATION
We have constructed a proof-of-concept prototype application Context Gallery for navigation and semi-automatic annotation of images by means of contextual information available through a mobile phone. The application contains two parts. The first is a gallery that provides multiple interfaces for the user to navigate, browse, and retrieve images by means of the annotated tags. The second part of the application contains modules that continuously read the multiple sensors available and establish the current user's context. Each module can be switched on and off by the user.

The prototype application is developed for Nokia N95 mobile phones, which provide the following set of embedded sensors:

- GPS – for location discovery (accuracy in meters)
- GSM cellular information – for place discovery (accuracy of meters to kilometers)

- WLAN – for discovery of places
- Bluetooth – for device and people discovery
- Application data – calendar information (activities)
- Phone specific data – state, profile (state information)
- Light sensor – available when capturing images with the built-in camera.
- Accelerometer – to determine movement patterns (not included in the present prototype)
- Microphone – could be used to establish noise level and activities (not included in present prototype)

Devices with additional sensors will simply enable more contextual information to be derived. Most of the raw sensor data is not directly usable for the end-user of the application. For instance, the coordinates acquired by the GPS sensor or the GSM network information has little meaning to most people. Therefore, a set of online services[1] are used in order to infer and translate the sensor data into human readable information, that is, meaningful descriptions and labels. The mentioned GPS and GSM information [10] is translated into place descriptions (city, region, country, etc). Weather information (temperature, humidity, clouded) for the particular region is also acquired. Bluetooth is used to discover hand held devices and through user provided feedback translated into people present. Finally, time is an obvious parameter [5], which is also translated into descriptions including season, month, weekdays, weekend, morning, afternoon, and light condition (dusk, dawn) etc. Time is also used to group images into events, dividing by levels of the time span in between the pictures taken. This is supported by calendar events (if present). Images taken within a time span is classified as belonging to an event and associated to a corresponding collection.

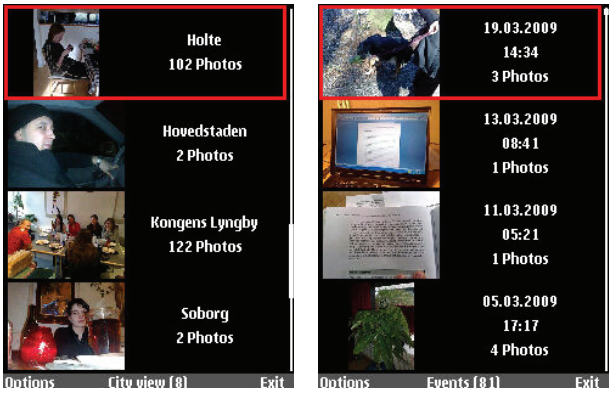

Figure 1. Screenshots from the prototype mobile application with "view by place" and "view by event"

Beyond these system derived contextual labels, the user has the option to add his/her own textual labels simply by manual input of one or more tags to be annotated to the information item (image) or collection. As mentioned above, such input can serve as feedback to the system in terms of not only annotating the particular information item, but also annotating the corresponding context in which the information is used.

Based on the multiple annotations of the images the information stored in the application repository is indexed in order to allow retrieval to be performed quickly and to establish different views of the content. Views are organized as lists of photo collections, in which images are presented in grids of a chosen size, as shown in Figure 1. The current prototype allows the user to view the information by time (days, weeks, and months), places, events, and similar images based on similar context or other user-defined search criteria. Searching by means of the annotated tags is possible by means of an interface where one or more tags can be indicated. Figure 1 contains screenshots from the prototype mobile application showing "view by place" and "view by event".

Initial experiments with the proof-of-concept prototype application have been carried out. Three test participants used the application in the greater Copenhagen area in Denmark for four weeks in order to gather contextual information. It was found that the information that could be obtained from the sensors could be translated into meaningful and useful context descriptions for annotating the images allowing multiple ways of navigating and retrieving the images. Furthermore, the prototype application was performance tested with 1700+ existing images annotated with contextual information (temporal annotation and some with location), which showed the capabilities with a high quantity of images.

DISCUSSION AND CONCLUSIONS

Related image browsing applications exist [2,3,6,9,11,12,13], where the emphasis has been on photo management and less on the general aspects of and implications for personal information management as well as facilitating information management by means of derived contextual information. Our focus is personal information management facilitated by establishing user contexts by means of multiple mobile phone sensors and enhancing this by also including the personal and manual annotations provided by the user. This provides useful contextual feedback to the system. Thus when a similar context is detected later the manually provided tags in that context could be used along with the tags derived from the embedded sensors. In the present work we have focused on contextual information about the environment inferred by means of the embedded sensors. However, additional information could be acquired either by analysis of the content itself [9], (e.g. face recognition) or from descriptions derived from tags generated in social networks. In future versions the derived context information could be shared among several users, thus enabling them to potentially benefit from contextual labels added by other users for instance in a social network. ZoneTag is an example of such an application that utilizes social networking in the image annotation process [8]. Users can share tags for specific locations (Cell ID or GPS based), that is, user contributed tags are propagated to other users in the social network, in terms of suggestions in similar locations. Context Gallery utilizes a wider range of embedded sensors and the spatial context is translated automatically (both GPS and Cell ID) with use of web services, which results in decrement of user's involvement in storing images. Contrary ZoneTag, the inference process primarily takes place on the mobile phone in Context Gallery. As a result, in situations where the context is already known (the second picture

[1] http://geonames.org/, http://opencellid.org/, http://wigle.net/, http://world-gazetteer.com/, and http://maps.google.com

taken in the same location), the location description is retrieved from the local storage without connections to web services.

Our initial prototype and experiments has shown that it is possible to infer meaningful contextual information from embedded mobile phone sensors, such as, GPS, Bluetooth, cellular and WiFi network information, and other low-level sensors along application information, such as the calendar. The contextual information can potentially be used to facilitate management of personal information such as in the information domain of interest here, namely image collections. Contextual information has the potential to provide multiple ways of searching and retrieving the information at need, depending on the recall by the user. However, further experiments are needed in order to establish to what extent this approach apply for other types of information in order to support the activities involved in personal information management. In particular usability tests are needed in order to establish the usefulness in recall based on use of the semi-automatic annotations provided.

ACKNOWLEDGMENTS

This research is supported by the Danish Strategic Research Councils through the project "Measurement Systems for Ethical Capital in the Experience Economy". The authors would like to thank the participants which took part in the initial experiments with the prototype application and Forum Nokia for the equipment used in the experiments.

REFERENCES

[1] Barreau, D. Context as a factor in personal information management system. Journal of the American Society for Information Science, 1995. Vol. 46(5): 327-339.

[2] Davis, M., House, N., Towle, J., King, S., Ahern, S., Burgener, C., Perkel, D., Finn, M.,Viswanathan, V., Rothenberg, M. MMM2: Mobile Media Metadata for Media Sharing. CHI '05 extended abstracts on Human factors in computing systems. New York, USA : 13th Annual ACM International Conference on Multimedia , ACM Press, 2005. pp. 1335-1338.

[3] Davis, M. and Sarvas, R. Mobile Media Metadata for Mobile Imaging. 2004 IEEE International Conference on Multimedia and Expo (ICME). 2004, Vol. 3, pp. 1707-1710.

[4] Dey, A.K., Abowd G.D. Towards a Better Understanding of Context and Context-Awareness. Lecture Notes in Computer Science. 1999, Vol. 1707, pp. 304-308

[5] Graham, A., Garcia-Molina, H., Paepcke, A., T., Winograd. Time as Essence for Photo Browsing Through Personal Digital Libraries. Proceedings of the Second ACM/IEEE-CS Joint Conference on Digital Libraries. 2002, pp. 326-335.

[6] Kim, S.K., Lee, J.W., Lee, R., Hwang, E.H., Chung, M.G. User-Friendly Personal Photo Browsing for Mobile Devices. ETRI Journal, 2008. Vol. 30(3): 432-440.

[7] Larsen, J.E.: NEXUS – A unified approach to personal information management in interactive systems, Ph.D. thesis, Technical University of Denmark, 2005.

[8] Naaman, M., and Nair, R. ZoneTag's Collaborative Tag Suggestions: What is This Person Doing in My Phone? IEEE Multimedia 2008, 15(3):34-40.

[9] O'Hare, O., Lee, H., Cooray, S., Gurrin, S., Jones, G.J.F., Malobabic, J., O'Connor, N.E., A.F., Smeaton, Uscilowski, B. MediAssist: Using Content-Based Analysis and Context to Manage Personal Photo Collections. Springer Berlin / Heidelberg. 2006, pp. 529-532.

[10] Trevisani, E. and Vitaletti, A. (2004). Cell-ID location technique, limits and benefits: an experimental study. Sixth IEEE Workshop on Mobile Computing Systems and Applications, WMCSA 2004. pp. 51–60.

[11] Tvarozek, M. and Bielikova, M. Personalized Faceted Navigation for Multimedia Collections. Washington, DC, USA : Proceedings of the Second International Workshop on Semantic Media Adaptation and Personalization, 2007. pp. 104-109.

[12] Wilhelm, A., Takhteyev, Y., Sarvas, R., Van House, N., and Davis, M. Photo Annotation on a Camera Phone. Vienna, Austria : ACM Press, Conference on Human Factors in Computing Systems, 2004. pp. 1403-1406.

[13] Yee, K. P., Swearingen, K., Li, K., Hearst, M. Faceted Metadata for Image Search and Browsing. Ft. Lauderdale, Florida, USA: Conference on Human Factors in Computing Systems, 2003. pp. 401-408.