

Niels Henze
René Reiners
Xavier Righetti
Enrico Rukzio
Susanne Boll

Services surround you

Physical-virtual linkage with contextual bookmarks

© Springer-Verlag 2008

N. Henze (✉)
OFFIS Institute for Information
Technology, Oldenburg, Germany
henze@offis.de

R. Reiners
Fraunhofer Institute for Applied
Information Technology, Germany
rene.reiners@fit.fraunhofer.de

X. Righetti
VRLab, EPFL, Switzerland
xavier.righetti@epfl.ch

E. Rukzio
Computing Department, Lancaster
University, UK
rukzio@comp.lancs.ac.uk

S. Boll
Media Informatics, University of
Oldenburg, Germany
susanne.boll@informatik.uni-oldenburg.de

Abstract Our daily life is pervaded by digital information and devices, not least the common mobile phone. However, a seamless connection between our physical world, such as a movie trailer on a screen in the main station and its digital counterparts such as an online ticket service, remains difficult. In this paper, we present contextual bookmarks that enable users to capture information of interest with a mobile camera phone. Depending on the user's context, the snapshot is mapped to a digital service such as ordering tickets for a movie theater close by or a link to the upcoming movie's Web page.

Keywords Mobile interaction · Physical-virtual linkage · Physical interaction · Contextual bookmark · Content analysis

1 Introduction

Mobile devices are a part of our everyday lives as most people rely on mobile phones, PDAs, and multimedia players as personal and pervasive information devices. Ubiquitous network access has been made possible by the use of these devices, while the ever-growing amount of digital content provides information about almost everything and all kinds of digital services are available from buying a cinema ticket to renting a car. These digital content and digital services often have physical representations in the real world. Movie trailers are shown on public displays and advertisement posters promote cheap car rentals. However, a seamless connection between our

physical world and related digital services is not possible today. Mobile interaction only takes place between users, their mobile devices, and digital services.

Over the last few years, there has been an increasing interest in extending the interaction between a user and her device to an interaction that takes objects from the real world into account. Outcome of this development are first commercial prototypes, as for example by Nokia [11] and Kooaba¹. However, linking between the real and the digital world today is only possible for specific objects and services. The concept of contextual bookmarks presented in this paper goes one step further: it enables users to cre-

¹ <http://www.kooaba.com/>

ate, use, and share bookmarks of real world objects. The intention is to support a wide bandwidth of real world objects and provide its users context sensitive with digital services. The example implementation recognizes objects without the necessity of visual or electronic markers by taking contextual information into account. This introductory section will present a scenario which shows the use case of a mobile contextual bookmark application. Then, the underlying concept of contextual bookmarks will be described.

1.1 Scenario

Mary is on the way to a conference in Berlin and sits in the metro. The train is equipped with small advertising screens; suddenly, the trailer of the latest Tarantino movie pops up on the screen. Tarantino is one of Mary's favorite movie writers and she does not want to miss this movie. She points at the screen with her mobile phone and bookmarks the trailer as a movie she would like to see. During the conference's first session a talk about mobile interaction gets Mary's attention. Using her phone Mary takes a contextual bookmark of the presenter and sends it to her colleague Jim as he is interested in this topic. Jim receives the bookmark, opens the related webpage and sends an email to the presenter in order to ask him for a meeting.

Later that day in the demo session Mary spots a fascinating poster on a novel interaction technique for mobile navigation. Since the poster's authors are busy with talking to other participants, Mary uses the poster as a link to access a related video and watches it on her mobile.

It is already late afternoon, when the conference finishes and she is on the way back to her hotel. Browsing through the bookmarks Mary collected today the system automatically checks the cinemas near Mary's hotel and asks her if she wants to go to the Tarantino movie she bookmarked in the train. Mary remembers that she has an appointment with Jim tomorrow morning and refuses with a sigh. Arriving in the hotel Mary considers watching the bookmarked movie in her hotel room. The bookmark provides Mary with a link to download the trailer. Watching the trailer she advises her mobile phone to get a license for downloading the movie to the hotel's set-top box. Using her mobile Mary remotely controls the set-top box to skip the promotion trailer and watches Tarantino's great movie until she falls asleep.

1.2 Contextual Bookmarks

The system used by Mary allows her to memorize interesting information and situations she faces while on the move. It provides her with access to services which are related to physical objects. Such services can range from simple web pages, over specific services such as buying a licence for a movie to rather object-related services such

as controlling a set-top box. The basis of the idea behind the scenario is the concept of the contextual bookmark that will be described in the following.

A contextual bookmark is a context sensitive link created from a physical object, event or person. It comprises a snapshot of a user's context for a specific point in time. Context information about the user are her location, the time, and may also cover information about preferences and intentions. Besides context, a bookmark contains information about the bookmarked object, event or person. This meta-information includes a name and description but also links to digital representations of the physical object, digital services, and background information. As described in the scenario persons, movies shown on public displays, and posters can serve as hooks to create contextual bookmark. However, the concept is not limited to these entities. Further examples are exhibits in museums, cooking recipes, and places. A contextual bookmark created from Tarantino's movie *Death Proof* would point to the movie itself (not to the movie's poster). It consists of the user's context at the time the bookmark was created and meta-information. The context information could comprise the user's location, a photo of the poster, the time, and the user's preferences. Meta-information would contain the type of the object, its name, and a thumbnail. Linked services could comprise the trailer of the movie, the movie in the iTunes Store or online cinema tickets booking.

An important aspect of contextual bookmarks is that context is not only important to recognize the user's intention when she creates a bookmark. The context is also necessary to provide the user with digital services she needs. That means that a bookmark of Tarantino's movie provides the user with a low resolution version of the movie's trailer if the user is in an environment with poor network coverage and the next cinema is far away. In a city like Berlin, on the other hand, the system could provide a cinema-booking services and the high resolution DRM encrypted movie itself.

The first prototype of a system implementing our concept of Contextual Bookmarks will be described in the following sections. The prototype focuses on specific aspects of the concept. However, our approach, instead of providing a solution for a single scenario, aims to support a flexible solution that enables the integration of various sensor data, media types, means of data access, additional service, and matching algorithms for different scenarios.

The remainder of this paper is organized as follows: Sect. 2 reviews the related work in the field while Sect. 3 introduces the architecture of the contextual bookmark system. The single components realizing the contextual bookmarks are elaborated in Sect. 4. We show the flow of information and interaction from capturing to using and sharing a bookmark in Sect. 5. Finally, Sect. 6 presents the lessons learned during our experiments and concludes with an outlook to future work.

2 Related work

Mobile devices are more and more used for interactions with objects in the real world. This can be distinguished in direct interactions such as touching or pointing and indirect interactions such as using the mobile phone as a remote control [3, 13]. The research presented in this paper focuses on pointing-based interactions as they allow the interaction with objects from a distance. Such interactions are seen as very intuitive as they correspond to our everyday interactions. We often point to objects with our index finger in order to show the other person something of interest or to support a statement.

Fitzmaurice was one of the first who described the vision of using a mobile device for pointing-based interactions with remote objects. He described, for instance, an application in which the user was able to point to certain locations on a map in order to get additional information [5]. Several approaches to implement such interactions are used or have been proposed in the last few years. There are, in principle, marker-based and image recognition-based approaches. Furthermore, there exist approaches in which a laser pointer attached to a mobile phone was used to interact with an object [16] or in which the infrared interface was used to establish a link [2]. The first approach has the advantage that the user gets a direct visual feedback because the user sees the dot of the laser pointer on the object. Unfortunately, there aren't any mobile phones with a built-in laser pointer available. The second approach has the disadvantage that some mobile phone vendors stopped the integration of IrDA interfaces as they were not accepted or used by users.

The NaviCam prototype, which was presented in 1995, was one of the first systems in which a mobile device with an attached camera was used for interactions with markers attached to objects [12]. Using this system it was, for instance, possible to get additional information about pictures, to interact with a calendar or to open an interactive door. Nowadays many mobile phones have a built-in camera and applications for the interpretation of visual markers are preinstalled or can be downloaded on demand. In Japan, for instance, QR codes can be found in newspapers, advertisements and posters as corresponding applications are installed in almost every camera phone [7]. Another popular example is Semapedia that uses Semacodes attached to places, sights or objects that provide links the Wikipedia articles about them [14]. Furthermore, Nokia ships their new Nokia N series phones with a preinstalled marker interpreter that can be used to extract the URL of markers shown next to an advertisement. The most important advantages of these markers are that they are easily recognizable and the extraction of the encoded information can be done on the mobile phone itself without any involvement of a server. Unfortunately visual markers affect the visual appearance of the tagged object. In addition, the user has to know how to interact with the marker.

It is a problem for some users to focus the marker in such a way that the whole marker is seen by the camera and that the resolution is sufficient for the interpretation of the marker.

The advantage of marker-less image recognition approaches is that no marker has to be attached to the picture that disturbs its appearance. Fritz et al. presented a system in which a PDA with attached camera was pointed onto sights in order to get more information about them [8]. The PhoneGuide is a similar system but focuses primarily on the interaction with exhibits in a museum in order to get more information about them [6]. Most of these systems are implemented using a client-server approach in which an image is sent to a server that then extracts certain features and compares them with a corresponding database. This requires an authoring process which leads to a potentially huge set of pictures stored in the database which are linked to relevant information. This leads then again to a huge demand of processing power for the comparison of the pictures taken with the ones stored in the database. To reduce this demand some projects store also the location and orientation of the user in order to limit the search space [11, 15].

The advantage of the contextual bookmark system when compared with existing approaches is the possibility to take a bookmark of a diverse variety of objects like a bookmark of a movie trailer shown on a public display, person met at a conference or a poster at a bus stop. Further supporting the proposed system is the management of these contextual bookmarks that is comparable with the bookmark functionalities we know from Internet browsers. So it is possible to store them, group them or to exchange them which is, to our knowledge, not supported by other systems. Also, the Contextual Bookmark system uses bookmarks in combination with other services like buying a licence for a movie advertised by a poster and interaction with a TV in order to play the movie.

3 Architecture

Existing approaches focus on the recognition of specific real world objects. Thus, they are limited to very rare interaction scenarios. The concept of contextual bookmarks, however, covers a much wider spectrum. One key assumption is that users will not accept an interaction that is limited to certain types of real world objects like advertisement pictures in magazines or movie trailers. Our vision is to build a system that addresses a wide spectrum of real world objects as well as not limiting to specific use cases so a flexible concept and architecture is needed. The architecture presented here follows the concept of contextual bookmarks previously described. It supports different recognition techniques and can be extended to collect more information about the current context where the bookmark has been created. As shown on Fig. 1, the basic architecture can be divided into five components:

real world object, personal interaction device, recognition, storage, and playback.

The *personal interaction device* is the main mobile component that centralizes the interaction between the user and the system. It is used to manage bookmarks and is also responsible for collecting context information during the creation of the contextual bookmark, such as for pictures taken by an integrated camera, the user's location, and the time.

In order to recognize real world objects, a digital representation of these objects must be modeled. In addition to these real world objects (advertisement posters, buildings or even persons), immaterial content such as video content shown on public displays or music played on a sound system must also be taken into account.

For the recognition of different kinds of real world objects a *recognition component* that encapsulates multiple recognition techniques suited to the actual object, is needed. For the implementation of this component, two potential solutions are to be discussed: the recognition component could either be located on the mobile device itself (distributed approach) or on a backend recognition server (centralized approach). Since the recognition of real world objects in general is a computing-intensive task, especially regarding the recognition of multimedia content, a backend recognition server is necessary. On the other hand, simple tasks like the recognition of RFID-tagged objects require no special processing power and thus could be performed on the mobile device.

Once the contextual bookmark has been identified using the collected context and the result from the image recognition component, it can be used to access any related digital content. But, in order to provide universal access to the bookmarks from different devices, a centralized persistent storage is necessary. This functionality is

provided by a *storage component* that mediates the sharing of contextual bookmarks between different users.

Besides sharing bookmarks with other users, making use of a Contextual Bookmark also means accessing the digital items connected to it. Different physical objects can be augmented with different types of digital items. As a result, a bookmark created from a scientific poster at a conference could not only provide publications found in the references, but also video sequences related to the topic. Consequently, the associated content displayed on the user's mobile device and/or on any nearby devices such as a public display is handled by the *playback component*.

Before video sequences associated with the bookmark are transferred *on demand* from an online repository to the device, the content must be adapted to the device's resolution, bandwidth, processing and storage capabilities. This means that only the references of contextual bookmarks are exchanged, while the content itself is obtained from a central repository and adapted to the current device's capabilities. This not only reduces network traffic needed for the actual exchange, but it also ensures an optimal viewing experience for the current device.

A concrete implementation of the main components, i.e. the personal interaction device as well as the recognition and storage components, is described in more detail in Sect. 4. This implementation is capable of recognizing posters and snapshots taken from movie trailers being shown on a public display. The resulted contextual bookmark then allows retrieving additional information of digital content connected to the real world object.

4 Components

Based on the architecture defined in Sect. 3 a concrete instantiation of this architecture is implemented. The central point of this implementation is an application for mobile phones that enables users to create, manage, and share contextual bookmarks. The recognition component is implemented on a dedicated server that can recognize videos and posters. For the presentation of videos, a video player was implemented. A web server with a database backend is used for persistent storage and sharing of bookmarks. These components are described in the following.

4.1 Personal interaction device

The developed mobile application to create contextual bookmarks that represents the personal interaction device component is based on Java ME, the most common platform for mobile phones (e.g. the Nokia N95 equipped with an integrated GPS receiver and a high resolution camera shown in the left of Fig. 2). The developed application provides the user with functionalities to create, use, and share Contextual Bookmarks. The application starts with the main menu shown in the middle of Fig. 2. Using the

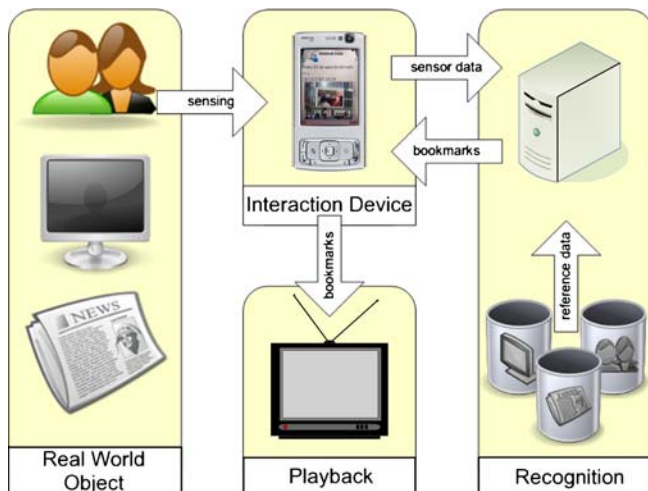


Fig. 1. The central components and their interaction with the contextual bookmark approach

main menu the user can access the different functionalities of the application. When the defined bookmark function is selected the image of the phone's integrated camera is shown on the display. The user can then point at a video or poster he or she wants to bookmark. The application creates a snapshot of the current context containing, in particular, a photo, the time, and the location. The bookmark is sent to a recognition server that recognizes the bookmarked content. According to the recognized content, the phone receives a thumbnail of the content as well as a name and a link to an according trailer and video.

Created bookmarks are stored on the phone and synchronized with a storage component for backup purposes and to exchange bookmarks. Selecting the browse function from the main menu the user can scroll through a list of created bookmarks shown in the right of Fig. 2. Selecting a bookmark the user can view the received thumbnail and description. In addition, according videos and trailers can be watched if available. To share bookmarks they can be sent via Short Messaging Service. Bookmarks can also be sent to playback components implementing the necessary protocol using an IP connection. The output devices can then be remotely controlled.

4.2 Real world objects

The implementation supports the recognition of analogue posters as well as videos as a representative for digital objects. The core of the recognition process is the comparison of a photo taken by the mobile application with the video and or poster. Since photos of real world objects are taken by the user it can be assumed that the characteristics of these photos vary. We expect that typical examples for photographed videos and posters are shown in Fig. 3. Even if the quality of these examples is high, the poster is distorted and a shadow darkens the center of the poster. In addition, the poster is curled and surrounded by the street

setting. Furthermore, it can not be ensured that the photo is in landscape or portrait format. The poster might be rotated, partially occluded, and varied by lighting conditions. The resolution of mobile phone cameras is low and so too is the resolution of the photos. Characteristics of videos shown on public displays vary even more because displays differ in terms contrast, resolution, and color fidelity.

To show content on public displays an application has been developed that shows videos based on a simple playlist. To enable the recognition component to use the temporal and location context prior to the recognition of videos, the public display application provides an interface that offers the display's position and identifiers of the played videos. Each public display registers at the recognition server. Thus, the recognition server is informed what content was displayed on which public screen.

4.3 Recognition

Retrieving a video or poster based on an image taken by a mobile phone's camera is the core of the matching process. For large collections of physical objects containing thousands of posters and videos, recognition based on the content alone is hardly possible. Thus, the number of potential media items is reduced using the user's context, in particular, his or her position and the time when the bookmark is created. After reducing the number of potential objects the photo of the object is compared with descriptors of the objects. For this comparison, the scale-invariant feature transform algorithm developed by Lowe [10] is used. SIFT is invariant or robust against most of the variations described in Sect. 4.2. Based on SIFT, algorithms can be developed that are robust against partial occlusion. A Best-Bin-First [4] algorithm based on KD-trees can be used to match an image with a small number of objects

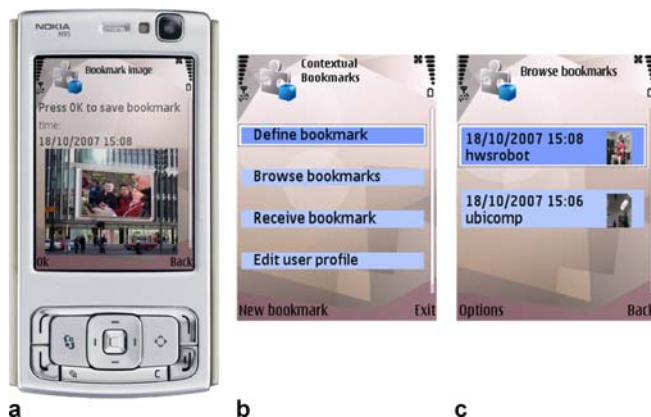


Fig. 2. **a** Nokia N95 running the contextual bookmark prototype, **b** Main menu of the mobile application, **c** Bookmark management view containing two bookmarks



Fig. 3. **a** Photo of a poster used to create a bookmark of the Simpsons movie, **b** Photo of a public display used to create a bookmark of the movie

by comparing the SIFT keypoints of the image with SIFT keypoints of the objects. For the recognition of posters we implemented standard techniques similar to the approach from Liu et al. [9]. However, the complexity of this approach increases linearly with the number of objects and videos consist of about 1500 frames per minute. With a naive approach each frame is treated as an individual object. Thus, we enhance the algorithms as described in the following.

Videos are analyzed offline taking the temporal characteristics of videos into account. First, the SIFT keypoints are extracted from each frame. Most frames have similar proceeding and successive frames and thus the according SIFT keypoints are similar as well. Similar keypoints are merged to a single keypoint reducing the number of keypoints drastically. Merged keypoints allow more robust matching, in particular for photos with long exposure time. After processing all frames a KD-tree for storing the merged keypoints is created that is later used to recognize photographed videos. If the recognition component receives a photo, SIFT keypoints are extracted and the process iterates through the list of videos. For each of the photo's keypoints the keypoint's two nearest neighbor keypoints are computed. According to the distances between the photo's keypoint and the two neighbors, it is decided if the keypoint is a match. Further heuristics, such as using all of the video's keypoints only once, increase the processes performance and robustness. The process determines the object with the highest number of matches and transmits according links back to the mobile phone.

4.4 Storage

As soon as the requested object is determined by the recognition component, the according contextual bookmark is transmitted to the storage component. This component is implemented using a MySQL database with an Apache Web-Server frontend. Since the interface is based on standard web technologies, accessing the storage component from different devices and applications can be implemented easily. Each bookmark is stored with a unique identifier inside the database. This identifier serves as a handle that is used by the mobile application to exchange bookmarks over low bandwidth connections such as the Short Messaging Service. Applications receiving the handle can request the database to retrieve the according contextual bookmark.

4.5 Playback

To use different content types linked by a contextual bookmark, playback of these links is encapsulated in the playback component. This component is realized as a software player derived from the Java Media Framework, which ensures platform independence. The same component is

used on the mobile phone as well as on the home theater PC. The playback component uses a toolkit for open human-computer interaction with ambient services that abstracts from concrete input devices. The network transparent protocol based on XML-RPC enables to use the same user interface to control playback on the phone as well as playback on the TV at home.

5 Communication

This section describes the communication between the components described in Sect. 4. With regard to the initially presented scenario, three fundamental interactions with contextual bookmarks are presented: creation of a contextual bookmark, exchange of Contextual Bookmarks, and accessing content associated with a Contextual Bookmark. The following subsections illustrate the communication between the different components that is established for each of the three cases focussing on the creation and exchange of bookmarks.

5.1 Creation of a contextual bookmark

Figure 4 shows the messages exchanged between the mobile phone that demands recognition the user's context and the recognition component. The recognition component only addresses videos and posters presented near the user's location.

The result of the analysis is sent back to the mobile device that stores the result in the local memory as a Contextual Bookmark. Concurrently, the contextual bookmark is also stored on the central storage component for later access as demanded by the user, for example if she needs to transfer the bookmarks to another mobile device or exchange between different mobile devices.

5.2 Exchange of bookmarks

As illustrated by Fig. 5, a contextual bookmark is exchanged between two mobile devices by sending the contextual bookmark's ID via the GSM network. The receiving device queries the central storage component for the transmitted ID and downloads the contextual bookmark via the web server.

5.3 Showing digital content associated with a contextual bookmark

When accessing the digital content behind a contextual bookmark, the application checks whether the content is protected by DRM and if a licence has to be acquired first. If this is the case, an additional step for payment is needed. The next choice consists in selecting the output device. The content can either be shown on the mobile device itself or be sent to a larger display. The mobile

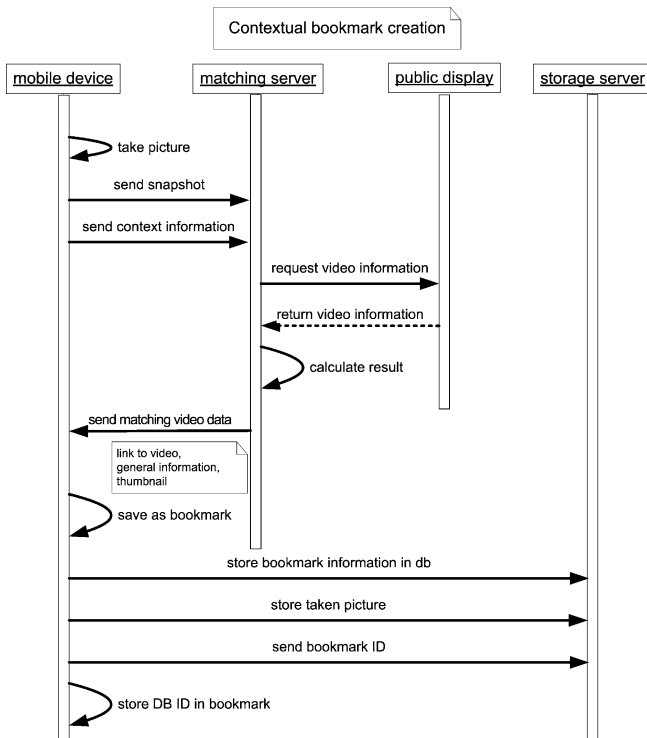


Fig. 4. Sequence diagram for the creation of contextual bookmarks

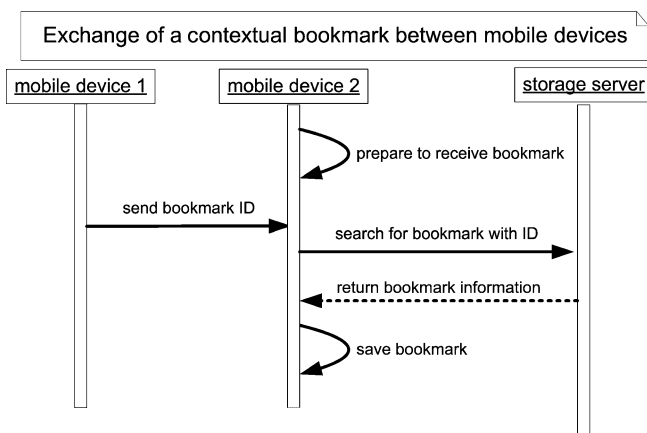


Fig. 5. Sequence diagram for the exchange of contextual bookmarks

device then acts as a remote control for this display. As described in Sect. 3, the content itself is not transferred, but a reference to it. Since the large screen provides different capabilities with respect to display quality, the content sent to this device can be adapted to its capabilities. For the mobile device, the same counts; the device receives content fitting to its capabilities concerning display quality and resolution as well as memory space.

6 Lessons learned

During the development, demonstration and evaluation of the first version of the contextual bookmark prototype, we gathered new insights into the related research questions and were able to analyze implementation questions in a more detailed way. The most important ones will be discussed in the following.

6.1 Extensibility

One way of extending the functionality of the contextual bookmark system is through the exploration of recognition algorithms that are also applicable to content different from image data. Context information also includes the location at which a contextual bookmark is created as well as user preferences, which could help to optimize the search strategy and the resulting presentation which could be oriented at a user model. The fundamentals for these extensions exist in the current implementation.

An alternative to a centralized service that contains all the multimedia content is the distribution of the content among different devices. This solves the problem of possible bottlenecks in situations where a huge amount of content is requested from and processed by the repository server. This approach assumes that the mobile device is equipped with sufficient storage and adequate data communication mechanisms. On the other hand, the exchange of contextual bookmarks would become much more time-consuming since the users have to wait for the whole data transfer.

6.2 Network integration and scalability

The current prototype works on the basis of a wireless network in which the snapshots and information are exchanged and requests as well as responses are sent. The only access to the GSM/UMTS network is established when a certain bookmark ID is sent to another mobile phone. The current system is not yet scalable to a very large system of public displays and screens due to the limitation of the context information to the time frame in which the snapshot is taken by the mobile phone's camera. With that timeframe information, all available video contents on all public displays must be analyzed. Future developments will enrich the gathered context information by localization data such that the public display from which the snapshot was taken can precisely be determined. Future versions will also be capable of connecting to different networks such as GSM network or UMTS by making use of ad-hoc networking mechanisms.

6.3 Usability

The contextual bookmarks showcase provides an excellent hands-on experience because people can feel a real inter-

action taking place which is tangible as well as feasible in real-time. The interaction still has room for improvement regarding the payment system and specific menu entries. The interaction could be enriched by more functionality such as a browser-like interface to access the information associated with a bookmark in an easier way. At the current stage of development, exactly one query result is returned by the recognition component, the others are discarded. The presentation of several results that exceed a certain threshold in the form of a ranking is a concept to be explored in the future. It has to be clarified if users really want to choose among different search results or whether they preferred one fixed outcome that is directly

illustrating examples of capturing an interesting poster or a movie trailer from a public display showed how one can easily bookmark real world entities and receive a mapping to digital content. The poster can be mapped to an on-line ticket store or the movie trailer can be watched on the home TV.

The implementation of contextual bookmarks will be extended by further recognition techniques to validate our architecture and address more types of real world objects. In particular, location-based searches [1] could be one pillar that provides context-sensitive digital services. We also work on the integration of a mobile guidance system and wearable interfaces in the contextual bookmark system. Furthermore we will evaluate our system with users in order to compare it with existing approaches, to find issues in our current implementation and to get further insight of the users' needs.

7 Conclusion

In this paper, we presented our approach to contextual bookmarks that allows the capture of objects in the real world using a camera phone and mapping the photo to a digital concept that is related to the digital object. The

Acknowledgement This paper is supported by the European Community in the project InterMedia. We thank Rob Hess for his implementation of the SIFT algorithm.

References

- Ahlers, D., Boll, S.: Location-based web search. In: Scharl, A., Tochtermann, K. (eds.) *The Geospatial Web. How Geo-Browsers, Social Software and the Web 2.0 are Shaping the Network Society*. Springer, London (2007)
- Ailisto, H., Korhonen, I., Plomp, J., Pohjanheimo, L., Strömmer, E.: Realising physical selection for mobile devices. In: *Proceedings of Physical Interaction – Workshop on Real World User Interfaces in Conjunction with Mobile HCI*, pp. 38–41. University of Munich, Munich (2003)
- Ballagas, R., Borchers, J., Rohs, M., Sheridan, J.: The smart phone: a ubiquitous input device. *IEEE Pervasive Comput. J.* **5**, 70–77 (2006)
- Beis, J., Lowe, D.: Shape indexing using approximate nearest-neighbour search in high-dimensional spaces. In: *Conference on Computer Vision and Pattern Recognition (Puerto Rico)*, pp. 1000–1006. IEEE Computer Society (1997)
- Fitzmaurice, G.: Situated information spaces and spatially aware palmtop computers. *Commun. ACM* **36**(7), 39–49 (1993)
- Föckler, P., Zeidler, T., Brombach, B., Bruns, E., Bimber, O.: PhoneGuide: museum guidance supported by on-device object recognition on mobile phones. In: *Proceedings of the International Conference on Mobile and Ubiquitous Computing*, pp. 3–10. ACM, New York, NY (2005)
- Fowler, G.: QR codes: In Japan, Billboards Take Code-Crazy Ads to New Heights. *Wall Street Journal* (10.10.2005). http://online.wsj.com/article/SB112888593628963880.html?mod=telecommunications_primary_hs
- Fritz, G., Seifert, C., Luley, P., Paletta, L., Almer, A.: Mobile vision for ambient learning in urban environments. In: Attewell, J., Savill-Smith, C. (eds.) *Mobile Learning Anytime Everywhere*, pp. 77–80. Learning and Skills Development Agency, London (2004)
- Liu, Q., McEvoy, P., Lai, C.: Mobile camera supported document redirection. In: *Proceedings of the ACM International Conference on Multimedia*, pp. 791–792. ACM, New York, NY (2006)
- Lowe, D.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **60**(2), 91–110 (2004)
- Nokia Corporation: When “now” just isn’t soon enough. Using camera phones to put mobile Net users in touch with the information they want, easily and instantly. Nokia (10/2007). http://www.nokia.com/NOKIA_COM_1/Press/twWln/press_kit/Point_&Find_Press_Backgrounder_October_2007.pdf. Accessed 15 April 2008
- Rekimoto, J., Nagao, K.: The world through the computer: computer augmented interaction with real world environments. In: *Proceedings of the 8th ACM Symposium on User Interface Software and Technology*, pp. 29–36. ACM, New York, NY (1995)
- Rukzio, R., Broll, G., Leichtenstern, K., Schmidt, A.: Mobile interaction with the real world: an evaluation and comparison of physical mobile interaction techniques. In: *Proceedings of the European Conference on Ambient Intelligence*, pp. 1–18. Springer, Berlin Heidelberg (2007)
- Semapedia.org: Semapedia – Hyperlink your World. www.semapedia.org. Accessed 15 April 2008
- Simon, R., Fröhlich, P., Obernberger, G., Wittowetz, E.: The point to discover GeoWand. In: *Adjunct Proceedings of the 9th International Conference on Ubiquitous Computing*, pp. 3–10. Springer, Berlin Heidelberg (2007)
- Välkkynen, P., Korhonen, I., Plomp, J., Tuomisto, T., Cluitmans, L., Ailisto, H., Seppä, H.: A user interaction paradigm for physical browsing and near-object control based on tags. In: *Proceedings of Physical Interaction – Workshop on Real World User Interfaces in conjunction with MobileHCI*, pp. 31–34. University of Munich, Munich (2003)



NIELS HENZE is working as a research associate at the Institute OFFIS in Oldenburg, Germany. He is involved in European research projects such as ENABLED and InterMedia. His research interests are multimodal interaction and interaction with real world objects.



RENÉ REINERS is working as a research assistant at the Fraunhofer Institute for Applied Information Technology (FhG FIT) in Sankt Augustin, Germany. He is involved in the EU research projects InterMedia and HYDRA. His interests are the exploration of supportive tools for rapid prototyping and the design of user interfaces in the field of human-computer interaction.



XAVIER RIGHETTI is currently working as a research assistant and pursues his Ph.D. at VRLab, the Virtual Reality Laboratory of the Ecole Polytechnique Fédérale de Lausanne (EPFL). His main research topic in human-computer interaction includes the development of tangible user interfaces in the context of wearable computing.



ENRICO RUKZIO is working as an academic fellow and lecturer in the Computing Department at Lancaster University. Enrico's research interests are physical mobile interactions and applications as well as context-aware mobile services. He founded the PERMID (Pervasive Mobile Interaction Devices) and MIRW (Mobile



Interaction with the Real World) workshop series and is co-author of more than 40 international publications.

SUSANNE BOLL is professor for media informatics and multimedia systems in the department of computing science at the University of Oldenburg, Germany. She is also a member of the scientific board of the OFFIS Institute for Information Technology, where she heads the multimedia and Internet information services division. Focuses of her research are semantic information retrieval, mobile multimedia systems, and intelligent multimodal user interfaces.