Designing a CD Augmentation for Mobile Phones

Niels Henze
University of Oldenburg
Escherweg 2,
26121 Oldenburg, Germany
niels.henze@uni-oldenburg.de

Susanne Boll
University of Oldenburg
Escherweg 2,
26121 Oldenburg, Germany
susanne.boll@uni-oldenburg.de

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Abstract
Interacting with physical CDs can be a very tangible and explorative experience. However, physical objects can't provide access to the digital services we are used to when using with digital music collections. In this paper we develop user interfaces for mobile phones that augment physical CDs to provide access to digital services. The most important functionalities of the music player are derived from a user study. Design sketches for the augmentation shown on the phone's display are collected from 10 participants. Participants' ideas are subsumed by four concepts that are implemented as prototypes for the Android platform.

Keywords
CD, mobile augmented reality, mobile phone, music

ACM Classification Keywords
H.5.m Multimedia Information Systems: Artificial, augmented, and virtual realities.

General Terms
Design, Human Factors, Experimentation

Introduction
Listening to music can be a great experience. We enjoy music in various situations and manifold goals. On-line

music seems to be the future but Compact Discs (CDs) are still very successful compared to on-line music portals. We assume that an important reason is the holistic experience provided by physical audio storage media. CDs in jewel cases, much like records and tapes, are tangible artifacts that communicate their content through expressive album art. Browsing and exploring physical music collections involves the whole human senses and body. We walk from one CD shelf to the other, turn record cardboards in our hands, and sometimes even the smell remind us of the last party.

The way we control music players moved away from this physical experience. MP3, according digital audio player, and portable music player made it very simple to access an unlimited music collection. However, Eggen, for example, reports that selecting from a large collection of music may put cognitive load on the listener that can negatively influences enjoyment of the actual content [1]. Since listening to music is highly emotional we assume that efficiency and effectiveness are not the only important factors for music player. However, digital audio player such as Winamp and Amarok also provide access to additional services as for example lyrics and information about the artists that are usually not available if using CDs.

In this paper we develop user interface designs for a mobile music player that is based on physical CDs. Music playback controls and additional services are presented through an augmentation on a mobile phone's display. The most important functions to control music playback are derived from interviews. To design the visual augmentation we collected design sketches from 10 participants. Based on the results four concepts were derived and prototypically implemented.

Related work

Music has received a great share of attention in the human-computer interaction community and in particular the tangible interaction community. Some attention has been given to the creation of music e.g. the system developed by Paradiso et al. [2] that uses the location of tagged objects to compose music.

A couple approaches have been proposed to control basic music player functions, such as play, stop, and volume) with a tangible object. Kranz et al. [3] developed *GestureCube* a tangible cube to remote control a digital media player. Different poses of the cube and gestures performed with the cube are used to control nine different functions. Alonso and Keyson [4] develop a similar cube that, in addition, provides information via speech and by illuminating the cube's sides. From the user study that compared the system with an Apple iPod they conclude that a balance should be found between ergonomic and hedonic qualities towards improving overall appeal. In addition, they propose to utilize multiple cubes each connected to specific music content.

An approach that connects tangible objects to *very* specific content is utilizing physical CDs. Zhang et al. [5] proposed to control music playback by equipping physical objects with RFID tags. Music can be selected by putting an RFID equipped CD on the RFID reader. Another approach that utilizes physical CDs has been proposed by Masui et al. [6]. Music is selected by putting a CD on the *MouseField* and functions can be invoked by moving the CD to control a graphical user

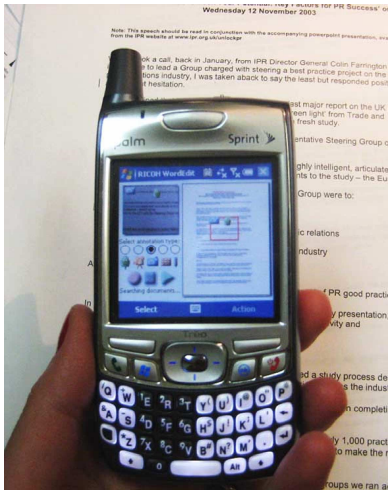


Figure 1. Live text recognition and augmentation on a mobile phone [8].

interface, however, the authors did not conduct a formal evaluation of the system.

Much like recent visual search engines for mobile phones, such as SnapTell[1] or Google's Goggles[2], Tsai et al. [7] developed a system that enables users to take a photo of a CD and receive a thumbnail and the option the play an according music sample. Partly because these systems rely on a server to recognize the CD they cannot provide instant feedback. Browsing a physical CD collection is thus hindered. A system developed for printed text instead of CDs that provides direct feedback has been demonstrated by Erol et al. [8] (see Figure 1). The system, running on a mobile phone, recognizes preprocessed text using the phone's camera. The phone's display presents a dynamic overlay on top of the camera image that highlights annotated regions (e.g. audio comments or videos). Revising recent work on object recognition for mobile phones (e.g. [9,10]) we assume that recognizing any 2D object (such as physical CDs) and estimate its pose to present an aligned augmentation will be feasible in the very near future.

User studies

We intend to develop a system that enables fast and simple browsing of physical music collections. Unlike stationary systems that force the user to bring the physical CD to a specific device, mobile phone based applications enables users to freely explore CD collections anywhere. Due to the inherent size restrictions of mobile phones' displays the amount of interaction controls and information that can be visualized is more restricted than on desktop devices. We conducted two user studies to derive the most important playback controls and the visual design of an augmentation for CDs presented on the phone's display that will be described in the following.

Determine a music-playback function-set

Even though mobile music player are almost pervasive we found little work that shows which functions are how important. Thus the aim of the first study was to determine the most important functions to control music playback. 11 functions were derived from the user interface of common digital music player and mobile music player. We asked 10 participants (5 female, Ø age=30) to rate the importance of each of these 11 functions for the intended use case on a 5 point Likert scale. The results are outlined in the table below.

Function	Ø importance	SD
Next track	4,7	0,2
Start	4,5	0,5
Increase volume	4,2	0,6
Decrease volume	4,0	0,9
Previous track	4,0	0,7
Pause	3,6	2,3
Stop	3,3	1,3
Repeat	3,2	2,2
Wind back	2,4	2,3
Fast forward	2,1	1,7
Shuffle	2,1	2,5

Most results, e.g. that start is rated more important than stop, are not surprising. The outcome is consistent with a function-set that Kranz et al. [3] derived from a user study (even though, it is not completely clear how they gained their results).

[1] SnapTell – Visual product search: http://www.snaptell.com

[2] Google goggles: http://www.google.com/mobile/goggles

Figure 2. Picture of a mobile phone provided to the participants to sketch a UI design.

Sketching UI designs

Controlling music playback is surely an important aspect of exploring music collections, however, our aim is to also provide the user with additional services such as lyrics and background information. Thus, we conducted a second study to explore design alternatives for the visual information presentation on a mobile phone's display.

We asked 10 participants (2 female, Ø age=24) to sketch a visual interface for the system. The task was to draw on a sheet of paper containing a mobile phone that shows an image of one or two CDs on its screen (see Figure 2). The picture of the phone had the same size as the physical device. Participants were asked to consider that the device has a touchscreen. They were free to use multiple sheets of paper to sketch different ideas or discard drafts. 16 sketches have been collected from the participants in total.

We classify the sketches by the way playback controls, services, and information are presented, how the recognized CDs are highlighted, and what user input is needed to access services.

Playback controls, services, and information can be presented aligned with the phone (on top of the camera image) or aligned with the CD (inside the camera image). Participants designed solutions using both approaches and, in addition, hybrid solutions that present some information aligned with the phone and others aligned with the CD. All sketches contain icons that represent the availability of some sort of service (e.g. 'W' for Wikipedia). Touching one of these icons invokes the respective functionality. If sketches contained text with CD's metadata (e.g. title of an album and the year of its release) the text was always aligned with the phone.

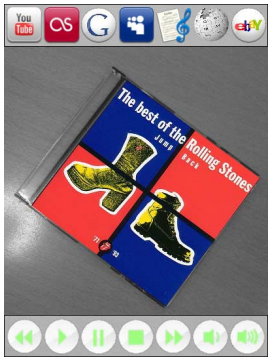
Participants designed two approaches to highlight the recognized CDs. Highlighting recognized CDs was proposed for sketches that presented information aligned with the phone, in particular. Three participants suggested that the camera image should be presented in grayscale and only the recognized CDs should be colored. Three different participants proposed to draw a colored rectangle around the respective CDs.

For most sketches the functionalities are accessible with a single touch. This was achieved by directly connecting a function with a touchable icon. E.g. a typical play button directly attached to the top of a CD together with other playback controls. A two-stage process was proposed by two participants for selecting a CD if multiple CDs are recognized at a time. This was considered necessary if the icons are not aligned with the CD but with the phone. Two participants intend to make some functionality (e.g. a link to Youtube) accessible by a form of sub-menu that pops up if the user touches a menu icon.

Derived designs

From the user studies we derived four screen designs that are described in the following. Based on the first user study and the related work we reduced the playback controls to the 7 highest rated functions. For all designs the CD could either be highlighted by a rectangle around the phone or by discoloring the background and drawing an image of the CD on top of the physical CD's position. We assume that the second approach makes it more obvious if and which CD is recognized and choose this approach for most designs.

Phone aligned controls



This design aligns the controls to the phone. The playback controls are at the bottom of the screen while the services are at the top. The icons are always at the same position and fade to gray when no CD is visible. If multiple CDs are found in the camera's video only the most prominent CD (the one that takes most screen space) is highlighted.

This approach has the advantage that the icons are relatively big. Since the icons do not change position movements of the phone do not affect the interaction. A drawback is that only one CD is accessible at a time.

CD aligned controls

For this design the controls are attached to the CD. The playback controls are aligned with the bottom of the object while the services are aligned to the top of the object. The controls follow the movement of the CD inside the camera's video. If multiple recognized CDs are visible all CDs are highlighted and each has its own control elements.

An advantage is that multiple CDs are accessible at the same time and screen space is used efficiently. Disadvantages are the icons' small size and that touching them is affected by accidental movement of the phone.

Mixed design

This design has elements aligned to the CD and others aligned to the phone. If a CD is found in the camera's video the CDs title, name of the band and the year of release are presented at the top of the screen. The playback controls are attached to the screen's bottom and fade to gray if no CD is found. The services are attached to the CD's top.

Presumably more frequently used functions are at a fixed position. While additional metadata can be presented more screen space is consumed. In addition, only one CD is accessible at a time.

Two stage approach

This design does not gray out the background but a thin rectangle is drawn around recognized CDs. In addition a star highlights recognized CDs. If multiple CDs are visible at the same time all are highlighted in the same way. A CD can be selected by touching it. All functions are accessible on a view that is shown if a CD is selected.

This design consumes the least screen space and does not alter the camera's video. The clear drawback is that a CD must be selected to before accessing any functionality, however, while interacting with a CD it does not necessarily be in the focus of the camera.

Implementation

Each of the described designs has some advantages but also some obvious disadvantages. In order to conduct formative and subsequent summative evaluations we implemented prototypes of the approaches. We implemented the prototypes for an Android HTC G1 Smartphone[3]. To recognize CDs and estimate their pose the video from the phone is constantly transmitted to a server via WiFi. The server analyses the video and estimates the CDs' pose which is transmitted back to the phone. Image processing is performed at a rate of around 8Hz. The visual overlay is drawn using OpenGL ES. Access to services is implemented for selected CDs only.

Summary and future work

In this paper we described the development of designs for an augmentation of physical CDs with additional services. In two user studies sketches for screen designs were collected and necessary playback controls for music identified. Based on this input four designs have been developed and according prototypes implemented. Main differences between the designs are if the controls are aligned to the physical object or the phone and if functionalities for multiple CD are accessible simultaneously.

The next step is to conduct formative evaluations of the identified designs to identify usability enhancements. Afterwards the approaches have to be compared in a formal user study. In order to develop a complete system, we currently implement an algorithm for mobile phones that recognizes CDs.

[3] Minor changes were applied to the phone's firmware to speed up image handling.

Acknowledgements

This paper is supported by the European Community within the InterMedia project (project No. 038419).

References

[1] J. Eggen: Turn on the base. Technical Note No. 3309, Philips Research Laboratories, Redhill, UK., 1995.

[2] J. A. Paradiso, K. Hsiao, and A. Benbasat: Tangible music interfaces using passive magnetic tags. In: NIME, 2001.

[3] M. Kranz, S. Freund, P. Holleis, A. Schmidt, and H. Arndt: Developing Gestural Input. In: IWSAWC, 2006.

[4] M. B. Alonso and D. V. Keyson: MusicCube: making digital music tangible. In: CHI (extended abstracts), 2005.

[5] N. Zhang, S. Jang and W. Woo: Nomadic Tangible Music Player with RF-enabled Sticker. In: ICAT, 2002.

[6] T. Masui, K. Tsukada, and I. Siio: MouseField: A Simple and Versatile Input Device for Ubiquitous Computing. In: UBICOMP, 2004.

[7] S. S. Tsai, D. Chen, J. P. Singh, and B. Girod: Rate-efficient, real-time cd cover recognition on a camera-phone. In: ACM Multimedia, 2008.

[8] B. Erol, E. Antúnez, and J. J. Hull: HOTPAPER: multimedia interaction with paper using mobile phones. In: ACM Multimedia, 2008.

[9] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg: Pose tracking from natural features on mobile phones. In: ISMAR, 2008.

[10] N. Henze, T. Schinke, and S. Boll: What is That? Object Recognition from Natural Features on a Mobile Phone. In: MIRW, 2009.