

Shoot & Copy: Using Mobile Phones for Accessing Information on Large Displays

Sebastian Boring¹, Manuela Altendorfer², Gregor Broll¹,
Otmar Hilliges¹, Andreas Butz¹

University of Munich, Media Informatics,
Amalienstr. 17, 80333 Munich, Germany

¹{sebastian.boring, gregor.broll, otmar.hilliges, andreas.butz}@ifi.lmu.de,
²altendor@cip.ifi.lmu.de

Abstract. We present *Shoot & Copy*, an information transfer technique that does not require visual codes that interfere with shown content on large displays. By using *Shoot & Copy* users can capture an arbitrary sub-region of a large screen that contains pieces of data. The captured image is then analyzed and a reference to the corresponding data or the data itself is sent back to the mobile phone. When the user wants to view the captured information in more detail, the *Shoot & Copy* prototype allows retrieving the data on a standard personal computer or viewing data directly on the mobile phone. With this technique, displayed content in our everyday life is easily memorable by users. In addition, supplementary information can be transferred to the mobile phone.

1 Introduction

Electronic displays become increasingly available at low cost and can thus be found in our everyday life. Besides monitors attached to personal desktop computers, people experience a variety of different display technologies that range from their personal television screens up to large public displays mounted in airports or train stations. These devices usually do not have any direct input capabilities for several reasons. First, the need to interact directly with the surface might be superfluous due to other indirect input opportunities such as mice and keyboards. Second, large public displays are usually mounted in a way that they are unreachable (e.g. to protect them against damage) and thus only allow viewing information. Third, as displays are getting bigger, one cannot interact with the entire screen surface. Hence, users cannot interact with displayed information on those surfaces.

As camera-equipped phones become more and more available and accompany their users most of the time, people use them as their external visual memory by taking pictures of interesting things. Furthermore, today's built-in cameras are capable of taking high-resolution images resulting in highly detailed information. While taking a picture of the content of interest on a large screen is suitable for observing the information, interaction is still impossible. The *Shoot & Copy* prototype evaluates the captured image in order to transfer the content directly onto the phone for detailed viewing or further manipulation.

2 Related Work

Recognizing visual markers and their encrypted codes using image processing is an established technology for the identification of various objects in the real world [7][8]. As mobile technology advances, the step from desktop systems such as the AR Toolkit [1] to mobile devices has been made. Similar to visual code systems, other visual marker recognition systems have emerged in the last years. Among them are Semacode [9], QR Code [6], ShotCode [11] or BeeTagg [3]. Most of these standards use encrypted code values as input for mobile applications such as Semapedia [10].

In addition, (semi) markerless systems have been developed by other researches. *Point & Shoot* [2] uses the built-in camera to move small objects on a large screen. Once the user presses the phone's joystick, a grid of visual code markers is temporally superimposed over the content on the display in order to determine the phone's position and orientation. *Sweep* is a marker-free technique that uses optical flow detection by capturing successive images in order to move a cursor on a large screen.

While markers are easy to handle on a technical level, they only encrypt a small amount of information (e.g. a numerical identifier) that can nevertheless be used as input for interaction with different services. As they are not human readable, the technology for communicating information via visual patterns needs to blend in with the media that display them – the entire screen becomes the marker.

3 The Concept of *Shoot & Copy*

Our concept is designed to make technology invisible. Thus, it does not use standardized visual markers or barcodes. Instead, the entire screen acts as a single marker without being recognized as such by the user. Hence users are only exposed to familiar technology, such as a desktop screen and a camera-equipped mobile phone.

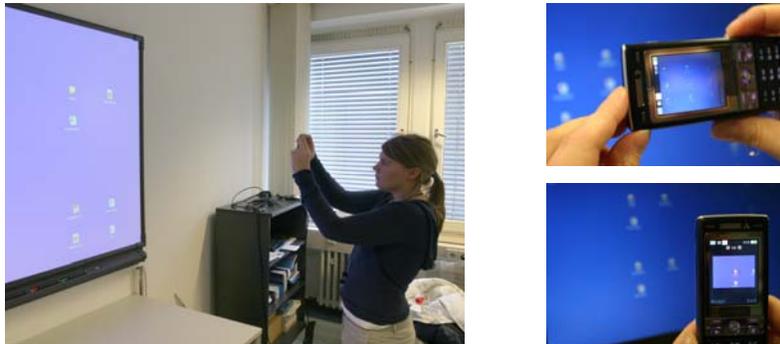


Figure 1. Left shows the interacting with a large display using *Shoot & Copy*. Right shows the interaction with a laptop (top) and the displayed results (bottom).

Shoot & Copy can be divided into two major phases, the capture phase and the access phase (a more detailed description can be found in [4]). In the capture phase, the user takes a picture of the desired region on a display (see Figure 1). The image is

then sent to the display's host computer, which compares the image to its own screen content and identifies the originally captured region. Once the region has been identified, the application takes the icon closest to the region's center. The information represented by this icon (e.g., a text document, image, music or video file) is then placed on a web server while the corresponding URL is sent back to the user's mobile phone. Alternatively, the original file could be sent directly to the mobile phone where it can then be viewed or played back by the user at any time.

In the access phase, the user places the mobile phone next to a Bluetooth enabled computer. The list of references stored on the mobile phone is then sent to an application running on the computer, where it is shown as an interactive list of icons and corresponding URLs. Images, text documents, music files and videos are shown (or played back) within the browser, while complex files, such as ZIP archives or other binary files, need to be downloaded by the user. If the file has been sent directly to the mobile phone, the user can access it using the phone's built in media browser.

4 Implementation of the *Shoot & Copy* Prototype

In order to test the *Shoot & Copy* technique, we built a fully functional prototype using a state of the art mobile phone (Sony Ericsson K800i). Its camera provides high-quality pictures (3.2 mega pixels). The shown content is a standard desktop as known from today's operating systems. The wireless communication between the phone and a server for image processing can be realized using Bluetooth or GPRS. For our demo, we use a direct Bluetooth connection to a laptop computer nearby.



Figure 2. Left shows the prototype with one possible match. Right shows the interaction with two identical constellations. The magnifications denote the capturing process (top) and the results shown to the user (bottom).

The entire interaction process can be divided into four phases. First the user needs to capture an image that will then be sent to display's host computer. In the second phase, the display server begins with the identification of the captured icons using image processing algorithms. During this phase, the system identifies the constellation of icons in the captured image. It then searches for similar configurations on the original screen. If this step results in more than one match, the system will use further techniques, such as color differencing, in order to identify the correct constellation. Subsequently, the results will be sent back to the user. If there is only one match, it

will be displayed directly. If the system was not able to distinguish between several hits, the user is able to select between different matches using the media browser. The displayed images are portions of the original screen content in order to allow the direct association with the desired content (see Figure 2).

5 Summary and Future Work

The *Shoot & Copy* demo consists of both the capture and the access interface. It incorporates a large front-projected screen, a laptop representing the user's home computer and a mobile phone with a built-in camera for high quality images. The total time between taking an image and receiving the reference is still close to ten seconds but will decrease due to increasing computational power and higher transmission bandwidths. In addition, some of the processing code can be implemented on the mobile phone once their processing power increases.

6 Acknowledgments

This work was funded by "Deutsche Forschungsgemeinschaft" (DFG) under a young investigator award to Andreas Butz, the SMS project (funded by the European Union) and the Bavarian Government. We also thank Sony Ericsson for providing a K800i mobile phone. Finally we thank all participants for their time and patience.

References

1. AR Toolkit, <http://www.hitl.washington.edu/artoolkit/>
2. Ballagas, R., et al., *Sweep and Point & Shoot: Phonecam- Based Interactions for Large Public Displays*, In CHI '05 extended abstracts, New York, NY, USA, ACM, 2005
3. BeeTag, <http://www.beetag.com/>
4. Boring, S., et. al., *Shoot & Copy: Phonecam-Based Information Transfer from Public Displays onto Mobile Phones*, In Proceedings of Mobility 2007, Singapore, ACM, 2007
5. Broll, G., et al., *Supporting Mobile Service Usage through Physical Mobile Interaction*, In Proceedings of PerCom 2007, White Plains, NY, USA, IEEE, 2007
6. International Organization for Standardization: Information Technology – Automatic Identification and Data Capture Techniques – Bar Code Symbology – QR Code. ISO/IEC 18004, 2000
7. Kato, H., Billingham, M., *Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System*, In Proceedings of IWAR 99, San Francisco, CA, USA, IEEE, 1999, 85-94
8. Rohs, M., Gfeller, B., *Using Camera-Equipped Mobile Phones for Interacting with Real-World Objects*, In Proceedings of UCS 2004, Tokyo, Japan, Revised Selected Papers, LNCS 3598, 2005, 74-89
9. Semacode, <http://semacode.org/>
10. Semapedia, <http://www.semapedia.org>
11. ShotCode, <http://shotcode.org>