Emerging Perspectives on the Design, Use, and Evaluation of Mobile and Handheld Devices

Joanna Lumsden
School of Engineering and Applied Science, Aston University, UK

A volume in the Advances in Wireless Technologies and Telecommunication (AWTT) Book Series

Information Science REFERENCE
An Imprint of IGI Global

Managing Director: Lindsay Johnston
Managing Editor: Austin DeMarco
Director of Intellectual Property & Contracts: Jan Travers
Acquisitions Editor: Kayla Wolfe
Development Editor: Caitlyn Martin
Cover Design: Jason Mull

Published in the United States of America by
Information Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Library of Congress Cataloging-in-Publication Data

CIP Data Pending
ISBN: 978-1-4666-8583-3
eISBN: 978-1-4666-8584-0

This book is published in the IGI Global book series Advances in Wireless Technologies and Telecommunication (AWTT) (ISSN: 2327-3305; eISSN: 2327-3313)

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

Chapter 11
Investigating Serendipitous Smartphone Interaction with Public Displays

Matthias Baldauf
Vienna University of Technology, Austria

Peter Fröhlich
Austrian Institute of Technology (AIT), Austria & FTW Telecommunications Research Center, Austria

ABSTRACT

Today's smartphones provide the technical means to serve as interfaces for public displays in various ways. Even though recent research has identified several approaches for mobile-display interaction, inter-technique comparisons of respective methods are scarce. In this chapter, the authors present an experimental user study on four currently relevant mobile-display interaction techniques ('Touchpad', 'Pointer', 'Mini Video', and 'Smart Lens'). The results indicate that mobile-display interactions based on a traditional touchpad metaphor are time-consuming but highly accurate in standard target acquisition tasks. The direct interaction techniques Mini Video and Smart Lens had comparably good completion times, and especially Mini Video appeared to be best suited for complex visual manipulation tasks like drawing. Smartphone-based pointing turned out to be generally inferior to the other alternatives. Finally, the authors introduce state-of-the-art browser-based remote controls as one promising way towards more serendipitous mobile interactions and outline future research directions.

INTRODUCTION

Digital signage technology such as public displays and projections are starting to become omnipresent in today's urban surroundings. According to ABI Research (2011), the global market for such installations will reach almost $4.5 billion in 2016 indicating their increasing potential. However, typical public displays in the form of LCD flat screens are a passive medium and do not provide any interaction possibilities for an interested passerby. As our steady companions, smartphones have been identified as

DOI: 10.4018/978-1-4666-8583-3.ch011

promising input devices for such remote systems. With their steadily expanding set of features such as built-in sensors, high quality cameras, and increasing processing power, they enable several advanced techniques to interact with large public displays.

Ballagas, Borchers, Rohs, & Sheridan (2006) investigated the available input design space and came up with different dimensions for classifying existing mobile/display interaction techniques. E.g., they suggest distinguishing between relative and absolute input commands as well as between continuous and discrete techniques. A continuous technique may change an object position continually, using a discrete technique the object position changes at the end of the task. Another commonly used dimension is the type of directness of a technique. A direct technique allows for the immediate selection of a favored point on the screen through the mobile device, traditionally using a graphical approach. In contrast, indirect approaches make use of a mediator, typically an on-screen mouse cursor which can be controlled through the mobile device.

Following an early classification of interaction techniques (Foley, Wallace, & Chan, 1984) we extend this smartphone/display interaction design space by the dimension of *orientation-awareness* taking into account the increasing popularity of mobile gesture-based applications. In case of an orientation-aware technique the position and/or orientation of the mobile device affects the interaction with the screen. In contrast, orientation-agnostic approaches are not sensitive to device movement.

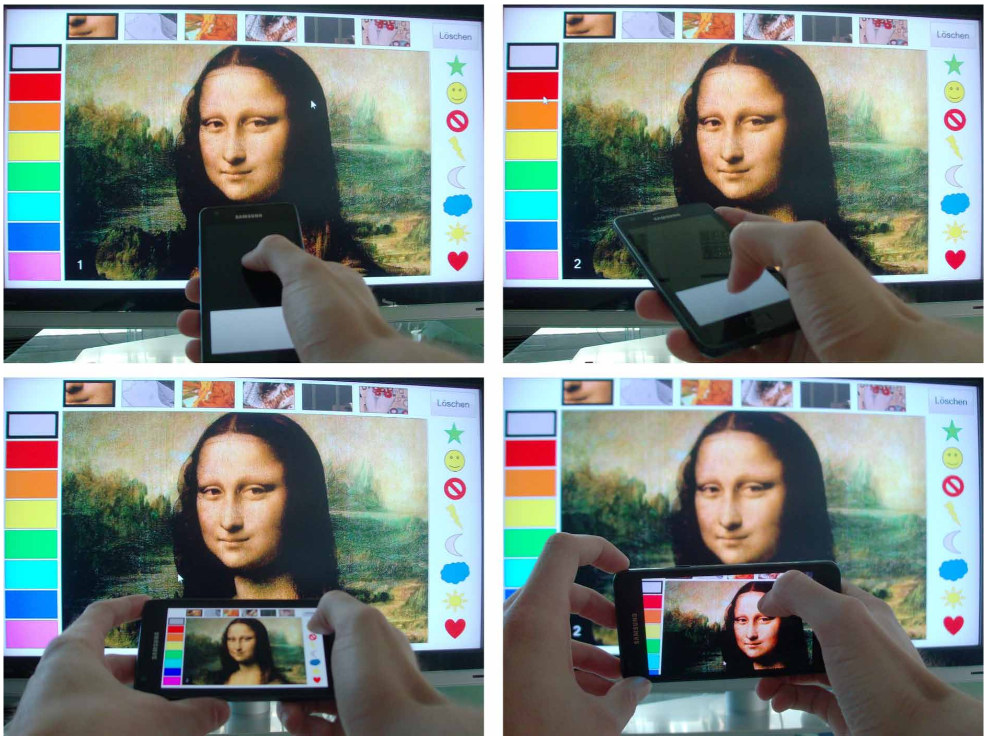
To learn more about upcoming orientation-aware interaction techniques and to evaluate their suitability for spontaneous interaction with public displays in comparison to established techniques, we selected four recent techniques for an in-depth comparative study. We decided to choose two novel orientation-aware interaction techniques which are gaining increasing attention in industry and academia. These techniques became feasible on smartphones only recently due to advances in mobile device technology. Respective implementations have not been scientifically compared with existing more established techniques so far. Thus their actual benefits in terms of performance and user acceptance have not been proven by now.

The first orientation-aware technique, the *Pointer* (Figure 1, top right), is made possible due to gyroscopes integrated into mobile devices of the latest generation. Inspired by a laser pointer, this technique enables the control of the mouse cursor by tilting and thus literally pointing towards the favored display location with the mobile device. The second orientation-aware *Smart Lens* technique (Figure 1, bottom right) enables screen interaction over the live video of the smartphone. By targeting respective areas of the remote screen through the built-in camera users may directly select a specific screen point by touching the mobile device display. Since this technique works on the device's live video, it inherently offers a zoom feature by reaching out and moving the device closer to the display and vice versa.

As more established techniques for our comparison we chose two orientation-agnostic interaction approaches with implementations already publicly available in mobile application stores. These two techniques represent respective counterparts to the abovementioned novel ones according to the dimension of directness. The indirect *Touchpad* technique (Figure 1, top left) makes use of a common interaction style and exploits the touchscreen of the mobile device in analogy to the touchpad of a notebook computer: strokes on the touchscreen are reflected by respective mouse cursor movements on the remote screen. Finally, *Mini Video* (Figure 1, bottom left) represents an orientation-agnostic direct interaction technique showing a cloned miniature view of the large display on the mobile device. Touches on the smartphone display are directly mapped to corresponding large display coordinates.

Table 1 shows the four distinct interaction techniques we explore in detail according to the traditional dimension of directness and the novel dimension of orientation-awareness.

Figure 1. The four compared interaction techniques include two indirect techniques, Touchpad (top left) and Pointer (top right), and two direct techniques, Mini Video (bottom left) and Smart Lens (bottom right). While Pointer and Smart Lens are orientation-aware techniques, Touchpad and Mini Video are not sensitive to device movement.



In the remainder of this paper we compare and discuss these four interaction styles in depth. We present a comprehensive user study designed to explore the advantages and disadvantages of these techniques with regard to different use cases and eventually derive concrete recommendations. Further, we demonstrate how each of these interaction styles may be realized through plain Web technologies in a browser to avoid the installation of a native mobile application and have them ready at hand when needed. We conclude with results from a first study exploring the users' acceptance of such browser-based remote controls and an outlook on promising future research directions.

Table 1. Classification of the compared techniques according to the dimensions directness and orientation-awareness

	Indirect	Direct
Orientation-agnostic	Touchpad	Mini Video
Orientation-aware	Pointer	Smart Lens

RELATED WORK AND RESEARCH HYPOTHESES

In this section we overview related work and identify shortcomings of previous research. Based on this literature review and own experiences with the abovementioned publicly available applications we formulate the research hypotheses for each of the techniques to be evaluated.

Touchpad

One of the first applications utilizing the *Touchpad* technique is *RemoteCommander* by Myers, Stiel, & Gargiulo (1998). The researchers connected several Palmpilot PDAs to a PC in the context of a cooperative work scenario. By stroking on the main display of the Palmpilots the PC's mouse cursor could be controlled. Like on today's notebook touchpads the absolute position on the touch surface was irrelevant but the movement across the device screen was mapped to an incremental movement across the PC screen. Clicking was possible by tapping on the screen while a separate software button toggled dragging mode. While it has been shown that such relative position controls perform better than rate control devices like a joystick (Card, English, & Burr, 1978; Douglas & Mithal, 1994) a crucial issue is 'clutching', i.e., lifting the finger and repositioning it to avoid running out of the input area. The overall completion time increases when clutching becomes more frequent (MacKenzie & Oniszczak, 1998). When the technique is used for distant large screens with high resolutions this drawback may be reinforced since the potential position distances extend while the input area remains constant.

In the meanwhile, the Touchpad technique has been adapted for smartphones, e.g., *Logitech Touch Mouse* for iPhone (Logitech, 2010).

Hypothesis 1a. Due to the mentioned clutching effect occurring for high resolution screens we expect the *Touchpad* to perform worse than the direct techniques in terms of task completion time.

Hypothesis 1b. Mouse-like pointing techniques have been shown to be very accurate (cf. Card et al., 1978). Thus, we expect that the *Touchpad* outperforms all other techniques in terms of accuracy.

Pointer

Pointing gestures in various forms are a heavily investigated technique for interaction with large screens and projections. For most studies researchers have been using custom hardware such as laser pointers extended with hardware buttons while the position of the laser point has been detected by cameras and means of computer vision. For example, Myers, Bhatnagar, Nichols, Peck, Kong, Miller, & Long (2002) compared different ways to hold laser pointer devices. The handheld device with a built-in laser turned out to be the fastest and most stable since due to its size it could be held with both hands. In a second study they found out that a traditional mouse suffers from fewer errors than the laser pointer approach.

An early study evaluating sensor-based pointing with mobile devices was conducted by MacKenzie and Jusoh (2001). In their comparison study, the two early off-the-shelf remote pointing devices demonstrated 32% and 65% worse performance than the standard desktop mouse used as a base-line condition. In a more recent study, Boring, Jurmu, & Butz (2009) compared a related, yet relative accelerometer-based technique called 'tilt' with a traditional joystick approach for controlling a mouse cursor on a remote screen. In contrast to a natural absolute pointing gesture, the device needed to be tilted to the left or right for moving the cursor horizontally. The results show that tilt technique performs better in terms of selection time but suffers from a higher error rate than the orientation-agnostic joystick technique. A comparable interaction technique is now publicly available in the app *Mobile Mouse Pro* (RPATech, 2014) which supports the touchpad technique as well.

Hypothesis 2a. In analogy to the aforementioned related tilt technique we assume that the even more natural *Pointer* approach has a lower task completion time than the *Touchpad*.

Hypothesis 2b. Based on the reported high error rates for pointing, we expect the *Pointer* to be less accurate than the alternative techniques.

Mini Video

The idea of *Mini Video* refers back to the *Worlds in Miniature* metaphor introduced by Stoakley, Conway, & Pausch (1995). In their virtual reality system, users are not only able to manipulate the virtual life-sized objects but also work with them using a handheld miniature model superimposed over the viewport. Related handheld concepts for large screen interaction have been presented e.g. by Kruppa & Krüger (2003) who suggest to display an abstract representation of the image shown on the large display on the mobile device for simple touchscreen interaction. Myers et al. (2002) introduce *Semantic Snarfing*, a combination of pointing and visual feedback where the targeted area of interest from the big screen is copied to the handheld device for more precise interaction. Their study shows that the direct interaction with a 'smartboard' (i.e., touching it with the hand) outperforms the remote pointing techniques in terms of both completion time and error rate.

The miniature technique in the current context - when directly interacting with copied content on a mobile device using its touchscreen - is obviously related to basic research on touchscreen interaction. Early touchscreen research (Greenstein, 1997) recommends a minimum button width of 22 mm. Relevant recent research investigating proper sizes and locations of so-called soft buttons on smartphones includes work by Lee and Zhai (2009) who showed that the performance of finger-operated touchscreen soft buttons deteriorated when the size of the button falls below a certain fraction of the finger width. Significantly poor performance of small touch keys in terms of success rate and the number of errors has also reported e.g. by Park, Han, Park, & Cho (2008) who compared soft buttons with 4 mm and 10 mm width.

Today, several commercial mobile applications for remotely controlling desktop systems such as *TeamViewer* (TeamViewer, 2014) adopted the miniature video technique.

Hypothesis 3a. Due to its direct touch-approach we assume *Mini Video* to outperform all indirect techniques in terms of completion times.

Hypothesis 3b. According to previous touchscreen research, we expect *Mini Video* to suffer from high error rates for small targets.

Smart Lens

Exploiting the built-in camera, a smartphone can be used as a 'see-through device' (Bier, Stone, Pier, Buxton, & DeRose, 1993) for targeting and identifying objects of interest, e.g., to infer related information about them or interact with them. Early work investigating such 'smart lenses' for interacting with screens exploited visual markers shown on the display (e.g., Ballagas, Rohs, & Sheridan, 2005). Pears, Jackson, & Olivier (2009) introduce the idea of dynamic markers in form of four green boxes. Their preliminary non-comparative user studies with four and ten participants show that the system is easy to use, however, does not give any detailed performance insights.

The idea of fully markerless live video interaction through a mobile device is inspired by early work by Tani, Yamaashi, Tanikoshi, Futakawa, & Tanifuji (1992) who introduced this concept for remote controlling industrial machines over video. Boring, Baur, Butz, Gustafson, & Baudisch (2010) presented a respective mobile prototype for touch interaction with multi-display environments. They evaluated four design alternatives and could show that an automatic zooming feature and temporary freezing the live video enhances the overall performance of the technique. In general, the technique suffered from a higher completion times and failures at decreasing target sizes. However, they did not compare this novel technique with established remote interaction approaches such as Mini Video. Baldauf, Fröhlich, & Reichl (2010) introduced a related fully functional prototype which touch-enables arbitrary display content using natural image features but did not report on a user study. Herbert, Pears, Olivier, & Jackson (2011) presented a related user study conducted with a very basic prototype involving a webcam instead of a touch-sensitive smartphone. The authors compared four different technical settings and found that high scores for responsiveness, accuracy, and ease of use were given for the alternative providing the highest frame rate of three fps.

Despite the actuality of this novel interactive smart lens approach, respective comparisons with alternative screen interaction techniques are missing so far.

Hypothesis 4a. As a direct interaction technique we expect the *Smart Lens* to perform similar than the *Mini Video* technique in terms of completion times.

Hypothesis 4b. Due to its orientation-aware nature, we assume the pure *Smart Lens* to be less accurate than the *Mini Video* technique.

METHOD

To address these hypotheses we designed an experimental laboratory study. The 24 participants (12 female, 12 male) were aged between 23 and 65 (mean = 34.5 and median = 31.5). As remuneration, each participant received a voucher for a consumer electronics store. 19 users regularly used a smartphone. On average, participants rated their experience with touchscreens with 4 (‘good’) on a five-point scale. Five participants stated they have used a mobile remote control application for presentation software before, two for remotely maintaining a computer. We deliberately aimed at arranging a well-balanced user group in terms of sex, age, and technology affinity and experience to gain generalizable results.

Each participant used each technique to perform three different types of tasks. After each task type, participants stated to which extent they felt supported by the technique in the respective task. The order of techniques was systemically varied to avoid learning and preference effects. Having completed all three types of tasks for a technique, participants were asked to respond to a questionnaire proposed by Douglas, Kirkpatrick, & MacKenzie (1999) to rate their experience with the technique. In contrast to general usability surveys, this questionnaire was designed explicitly for assessing devices and interaction techniques for remote pointing tasks and thus includes relevant questions concerning the mental and physical effort, the subjective perception of accuracy and operation speed as well as the experienced fatigue of fingers, the wrist, the arm and the shoulder. In the last study phase, the *Free Interaction* phase, participants were allowed for free experimentation with the techniques in the context of a painting application. The test, which took about two hours, closed with a final interview.

Experiment Setup

The hardware setup for our user study consisted of a *Philips Cineos* flat screen TV with a screen diagonal of 47 inch (119 cm) and a screen resolution of 1600x900 pixels acting as the public display and a *Samsung Galaxy S2* smartphone (see Figure 2). This device is equipped with a 4.3 inch touch display with a resolution of 480x800, an 8 megapixel camera at the back and several built-in sensors such as accelerometers and a gyroscope. Via HDMI the flat screen TV was connected to a notebook running an application custom-designed for our experiment. It consists of two windows: the actual task window displayed on the flat screen TV in full screen mode (cf. Figures 2 and 3) and a simple console for the test manager shown on the notebook screen. Here, the test manager could enter the user identifier, select the mode (training vs. test), specify the technique to be used and select, start, and stop the tasks.

When receiving the smartphone from the test manager, participants were asked to stand upright in front of the screen or use a barstool at a distance of 1.5 meters from the large screen. They were free how to hold the device and whether to extend or bend the arm. The mobile study application installed on the smartphone was connected to the notebook via WiFi for exchanging remote control commands using a simple custom protocol via TCP. The graphical interface of the mobile application showed a main menu with four buttons labeled with the four techniques for selecting the respective technique. For most of the test time this menu was disabled since the current technique was remotely configured by the test manager, i.e., before a new task type was started the mobile application switched to the respective interaction technique as remotely specified by the test operator. Only during the final *Free Interaction* phase, this menu was enabled and thus users were allowed to freely switch between the available interaction techniques. We used this *Free Interaction* phase to observe spontaneous interactions without performance pressure, and to gain qualitative feedback on the four techniques.

Figure 2. A participant using the Smart Lens technique during the Free Interaction phase of our lab study.

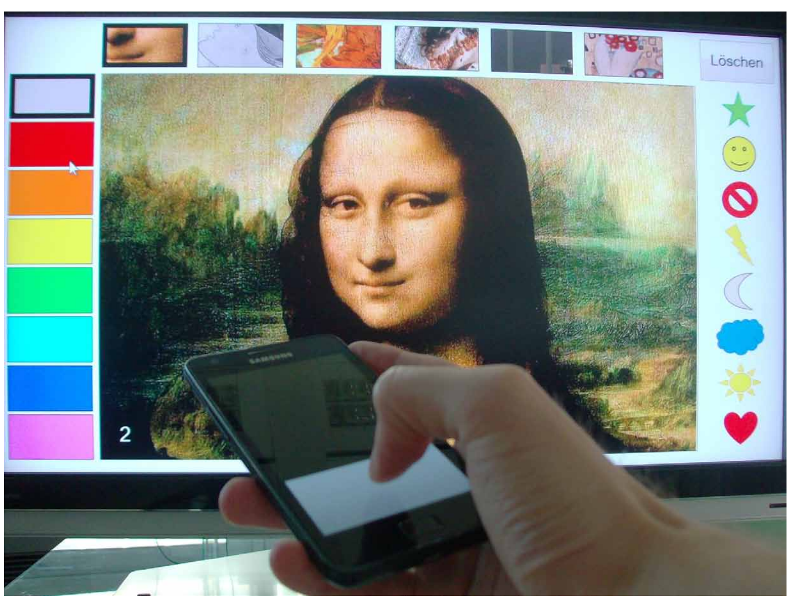
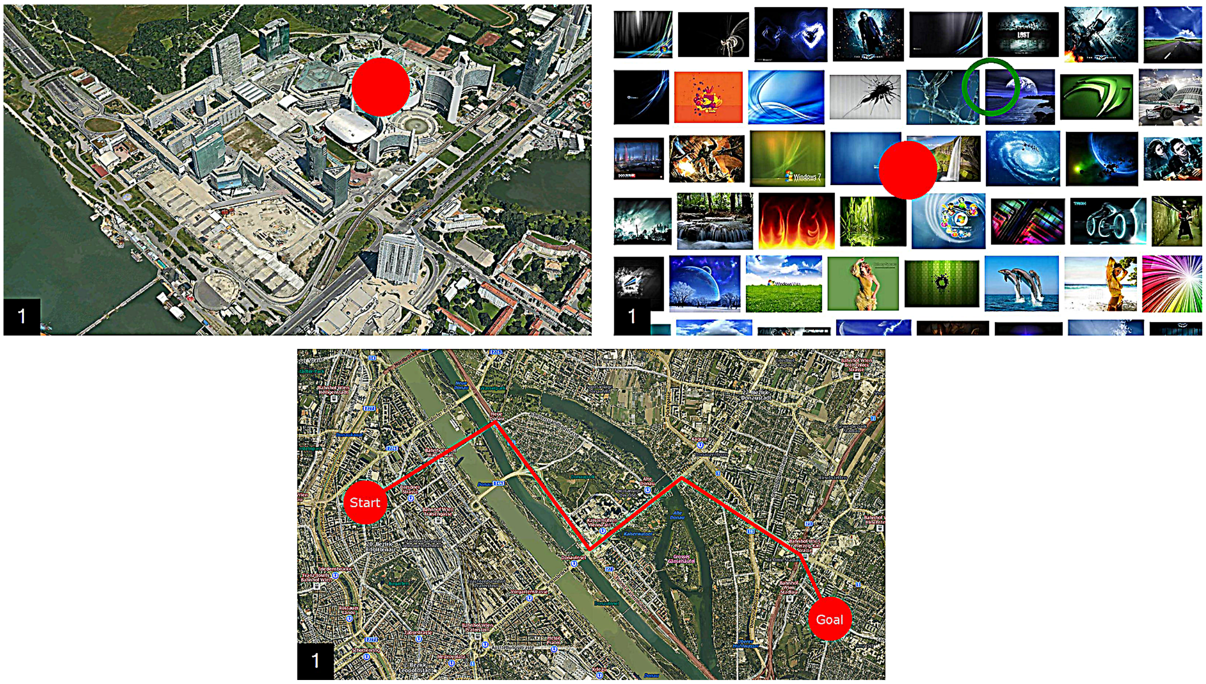


Figure 3. For each interaction technique, participants solved three different tasks with increasing complexity: Targeting (top left), i.e., selecting the red circle, Drag'n'Dropping (top right), i.e., moving the red circle over the green destination circle, and Drawing (bottom), i.e., tracing the red path from start to end point.



Touchpad

Using this technique the mouse cursor on the remote display could be controlled through respective finger gestures on the mobile touchscreen device. The largest portion of the mobile screen served as a touchpad while at the bottom a soft button allowed for triggering the action of the left mouse button (Figure 1, top left). Following the configuration of other researchers (e.g., Boring et al., 2009), we used the typical CD (control/display) ratio of 1, i.e., a panning gesture on the smartphone over a distance of 10 pixels moves the mouse cursor accordingly 10 pixels on the large screen. The multi-touch capability of the used smartphone enables both panning on the touchpad area and pushing the soft button at the same time.

Pointer

We utilized the device orientation for positioning the remote mouse cursor in analogy to a laser pointer. We exploited the built-in gyroscope and accelerometer for determining changes of the device orientation and applied a complementary filter combining a low- and high-pass filter for reducing noise in the raw sensor data. Based on knowledge of the user's distance from the screen we could calculate absolute cursor positions from the orientation changes. Before using the *Pointer* a short 'calibration' was necessary, i.e., participants had to point towards the display center. The graphical interface resembled the touchpad technique with its button for triggering a mouse button action (Figure 1, top right).

Mini Video

Usage of the *Mini Video* technique was enabled by streaming the content of the large display to the mobile device. The video stream was then scaled down to fit the display size of the smartphone, i.e., the mobile device showed a cloned view of the large screen (Figure 1, bottom left). Taps onto the smartphone display could be directly mapped to mouse actions at the corresponding position of the large screen.

Smart Lens

For enabling the *Smart Lens* interaction technique (Figure 1, bottom right), we chose a lightweight implementation: when the user touched the smartphone screen, the current camera frame was scaled, compressed, and transmitted to the notebook application where the frame was mapped to the actual screen content using natural image features. The derived transformation matrix was then used to convert the position of consequent touch actions to actual display coordinates in order to trigger the corresponding mouse action.

Task Types

Each participant was asked to perform three different task types per technique as shown in Table 2. We chose the types *Target* and *Drag'n'Drop* as traditional pointing tasks (MacKenzie, Sellen, & Buxton, 1991; Kabbash, MacKenzie, & Buxton, 1993) and extended them with the more recent *Draw* (cf. Pears et al., 2009; Herbert et al., 2011) resulting in three task types with increasing complexity. The order of the techniques was systemically varied to avoid learning and preference effects. Before testing a new task type with a new technique the users went through a training phase to get used to the new task type and technique until they felt comfortable for the test. For the test situation, users were asked to complete the trials as fast and accurately as possible. (Un)successful actions were indicated by audio signals. For each task trial we logged all input actions to calculate the completion time, the accuracy as well as the error rate. For increased ecological validity we chose a suitable background image for each of the first three task types. The final task type *Free Interaction* was a more informal one where users could freely experiment with the techniques and report on their experiences.

Targeting

In this task type, participants selected a set of targets in form of red circles (Figure 3, top left). Before the next target was displayed, a 'Start' button in the screen center needed to be clicked (cf. Douglas et al., 1999). Overall, 32 distinct targets were shown in randomized order: two different target sizes (radius of 40 and 80 pixels on the display, translated to a diameter of 5mm and 10mm on the *Mini Video* view) at two different distances from the screen center (150 and 320 pixels) in eight different orientations (0° to 315° in steps of 45°). As background image we chose a 3D city environment to mimic the selection of building parts. Completion time was measured between push of the 'Start' button and the moment of target selection. We captured the selection accuracy by measuring the distances of screen selections to the correct target in pixels, and by counting the number of errors (i.e., missing the shown target).

Table 2. Each participant used each technique (in varied orders) to solve three types of tasks: target, drag'n'drop, and draw. In the final Free Interaction phase participants could freely switch between techniques

Technique	Task Type	Trials
Touchpad	1.Target	2 dist. x 2 sizes x 8 orient.
	2.Drag'n'Drop	2 dist. x 2 sizes x 4 orient.
	3.Draw	4 paths
Pointer	1.Target	2 dist. x 2 sizes x 8 orient.
	2.Drag'n'Drop	2 dist. x 2 sizes x 4 orient.
	3.Draw	4 paths
Mini Video	1.Target	2 dist. x 2 sizes x 8 orient.
	2.Drag'n'Drop	2 dist. x 2 sizes x 4 orient.
	3.Draw	4 paths
Smart Lens	1.Target	2 dist. x 2 sizes x 8 orient.
	2.Drag'n'Drop	2 dist. x 2 sizes x 4 orient.
	3.Draw	4 paths

*Free Interaction.

Drag'n'Dropping

In this task type, participants were asked to drag a red circle from the screen center and drop them onto the green target destinations (Figure 3, top right) simulating a photo gallery. Also this task consisted of 16 trials while the target and destination was varied by two target sizes, two different distances and four different orientations. Data logging was started when the red circle was dragged for the first time. A trial was completed when the target's center was placed inside the destination circle. Completion time and selection accuracy were derived as in the targeting task, and errors were counted for not hitting the target (unsuccessful dragging) or not dropping it within the destination. The dropping accuracy referred to the distance between target and destination center.

Drawing

In the Drawing task participants had to trace four given paths from the start to the end circle (Figure 3, bottom) on a 2D map. The complexity of the paths steadily increased from two up to five straight path segments. The user's actions were logged beginning with the task start (i.e., when they started to draw within the start circle) until task completion (i.e., when having arrived within the end circle). The average drawing accuracy was calculated a posteriori by determining the shortest distance to the path for each drawn point.

Free Interaction

In this final task type users were allowed to freely switch between interaction techniques. They were asked to create their own art collage using a simple painting application shown in Figure 2. This application was designed to combine the formerly performed tasks: users were able to push buttons to select a drawing color and chose a famous painting to be used as a collage background (*targeting*) as well as to drag cliparts (*drag'n'dropping*) and paint (*drawing*) onto the collage.

RESULTS

The analyzed interaction logfile included more than 170,000 lines. For the below reported statistical analysis, the dataset was consolidated by deriving meaningful values (e.g., accuracy parameters, mean duration per trial, and number of errors per trial), and it was aggregated by averaging per test person. For the analysis of main and interaction effects, ANOVAs for repeated measures with the factors technique (4), target size (2) and distance target (2) were calculated (normal distribution was evaluated by means of Kolmogorov Smirnov tests). In case of a rejected sphericity assumption, the degrees of freedom were corrected by means of a Greenhouse & Geisser estimate. Pair-wise comparisons used Bonferroni corrected confidence intervals to maintain comparisons against $\alpha=0.05$. Errorbars in graphs represent a 95% confidence interval.

Targeting

- **Completion Time:** The results including all target sizes and distances (see 'overall' bars in Figure 4, left), indicate that using the Pointer technique took most time (M=3750 ms, SD=1081ms), followed by Touchpad (M=2973, SD=830 ms). Selection time was lowest with Miniature Video (2123 ms, SD=1109 ms) and Smart Lens (M=2075 ms, SD=833ms). Post-hoc pairwise tests reveal that these two techniques are the only ones not significantly differing from each other (all others $p<0.008$). An interaction effect was identified for target size and technique, $F_{2.1,48.7}=6.319$, $p<0.01$. When comparing selection time results for large and small target sizes (Figure 4, left), the relative profile did not strongly differ: Mini Video and Smart Lens were fastest, followed by Touchpad, and then Pointer. The direct techniques (Mini Video and Smart Lens) gained relatively more from larger target sizes than the indirect techniques (Touchpad and Pointer). We as well found an interaction effect for distance and technique, $F_{1.9,44.3}=18.972$, $p<0.01$. Comparing short and long distances did as well not materialize in a strong change of the overall relative profile.

- **Errors:** Overall results including all distances and target sizes (Figure 4, middle) show that with Touchpad almost no errors were made (0.06 errors per trial). For the other three techniques, we observed significantly more errors than Touchpad ($p<0.0017$), with 0.3 to 0.4 errors per trial (no significant difference among these three techniques). A significant interaction effect between technique and target size has been found, $F_{1.6,36.7}=10.9$, $p<0.01$. For small targets Touchpad had significantly fewer errors than all other techniques ($p<0.0017$), but for large targets also Smart Lens and Mini Video achieved very few errors (only Pointer was significantly higher, $p<0.008$). There was no interaction effect of technique and distance.
- **Accuracy:** Overall, selection accuracy was pretty similar among the four techniques (Figure 4, right); no significant differences were identified, except that selections with the Pointer were significantly less accurate (32 pixels distance from the target center) than the others (~ 28 pixels). We found an interaction effect between technique and target size, $F_{2.2,51.6}=11,9$, $p<0.01$. While large targets were less accurately selected with the Pointer compared to Miniature Video and Smart Lens ($p<0.0017$, no other significant pairwise differences), small targets were most accurately selected with Touchpad ($p<0.0017$, no other significant differences). We did not find an interaction effect of distance and technique.

Figure 4. Results of targeting task: completion times (left), errors (middle), and accuracies (right) per trial of the four techniques, separately presented for small, overall (all target sizes), and large target sizes

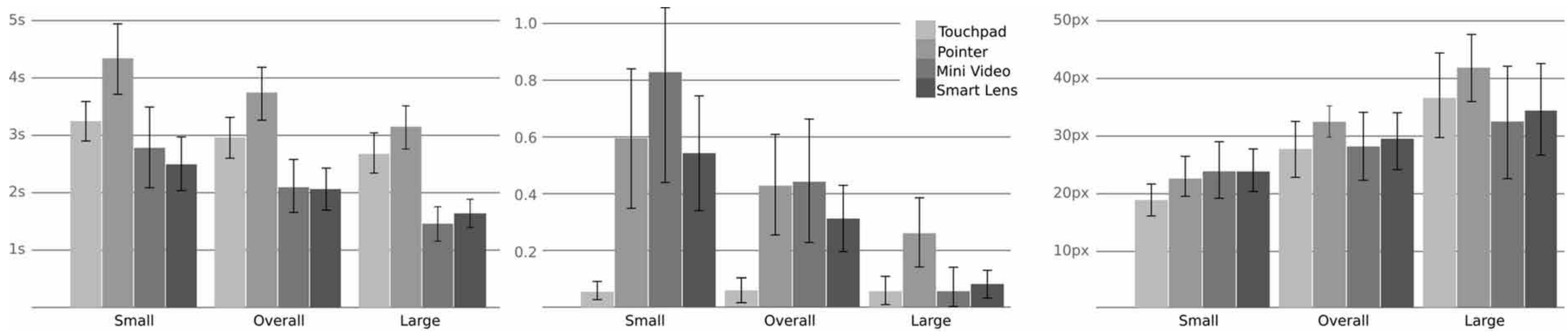
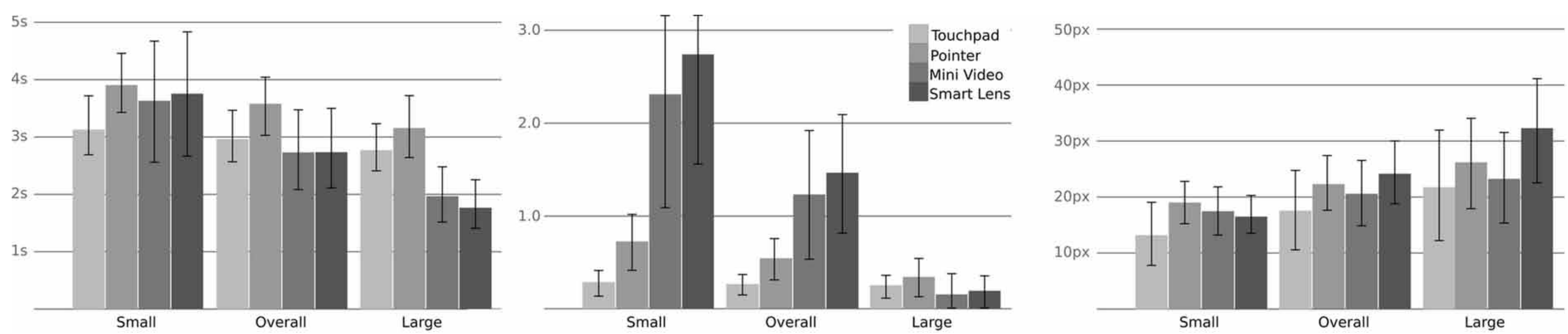


Figure 5. Results of drag'n'drop task: completion times (left), errors (middle), and accuracies (right) per trial of the techniques, separated by small, overall (all target sizes), and large target sizes



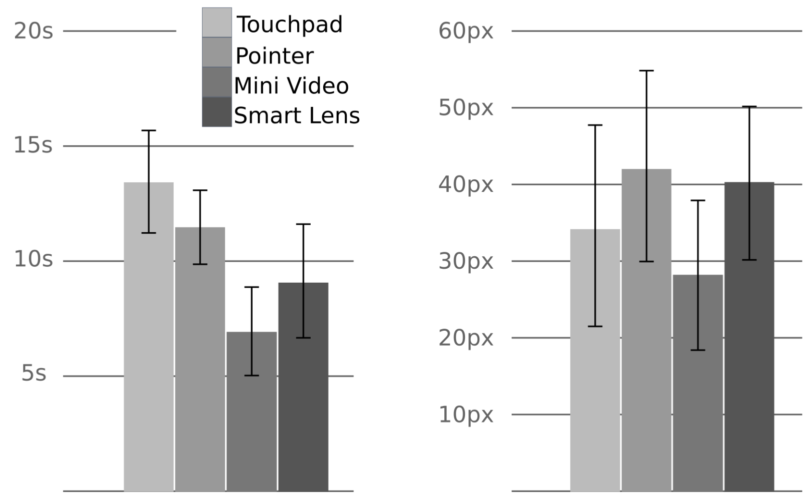
Drag'n'Dropping

- **Completion Time:** Overall results including all distances and target sizes (Figure 5, left) indicate that mean durations were quite similar among the techniques, only Pointer had relatively longer selection times (significantly longer than Touchpad and Miniature Video, $p<0.008$). We identified a significant interaction effect of target size and technique, $F_{2.2,51.5}= 6.5$, $p<0.01$. Larger targets enabled quicker selection by both direct techniques (Miniature Video and Smart Lens) than both indirect techniques (Touchpad and Pointer, $p<0.0017$), but with small targets this effect was not observed (no significant differences between none of the direct and indirect techniques). We did not find a significant interaction effect of target distance and technique.
- **Errors:** Overall results including all distances and target sizes (Figure 5, middle) indicate that Touchpad achieved the smallest mean number of errors per trial (M=0.27). These were significantly fewer errors than for Smart Lens, $p<0.0017$), but also Pointer had few errors (significantly lower than Smart Lens, $p<0.008$). An interaction effect of technique and target size was found, $F_{1.5,34.2}=15.7$, $p<0.01$, but not for technique and target distance. Drag'n'drop with small targets resulted in a steady increase of errors per trial from Touchpad (M=0.29, SD=0.32), Pointer (M=0.72, SD=0.71) and Mini Video (M=2.30, SD=2.89) to Smart Lens (M=2.73, SD=2.79), all differences were significant ($p<0.008$), except between Mini Video and Smart Lens. By contrast, with large targets, all techniques had similarly few errors (mean errors per trials were 0.15 to 0.33, no significant pairwise difference was obtained).
- **Accuracy:** Overall drag'n'drop accuracy (Figure 5, right) was highest with Touchpad (significantly smaller distance to the target center than Pointer and Smart Lens, $p<0.0017$) and lowest with Smart Lens (significantly higher distance than Mini Video and Touchpad, $p<0.008$). We found an interaction effect of target size and technique on accuracy, $F_{3,69}=9.2$, $p<0.01$. With small targets, Touchpad was most accurate, as it yielded a significantly smaller mean distance to the target center than Pointer and Mini Video, $p<0.008$ (no other pairwise differences significant). The most outstanding result here was that Smart Lens was significantly less accurate than all other techniques ($p<0.008$, no other pairwise differences significant). We did not find an interaction effect for target distance.

Drawing

- **Completion Time:** There was a main effect of technique on completion time, $F_{3,69}=26.089$, $p<0.01$. Figure 6 (left) indicates that drawing with Mini Video was significantly faster than with any other technique (M=6.9 sec, SD=4.6 sec), followed by Smart Lens (M=9.1 sec, SD=5.9 sec), Pointer (M=11.4 sec, SD=3.8 sec) and Touchpad (M=13.5 sec, SD=5.4). All pairwise differences were significant, except between Smart Lens and Pointer. We could not identify an interaction between path complexity and technique.
- **Accuracy:** We did neither identify a significant main effect of the techniques on drawing accuracy, nor an interaction effect with path complexity. Due to a generally high variance, only one pairwise difference was significant: drawing with Miniature Video was more accurate than with Smart Lens, $t_{23}=-2.717$, $p<0.017$ (see also Figure 6, right). Analysis of errors is not reported, as it is not applicable due to the nature of the task.

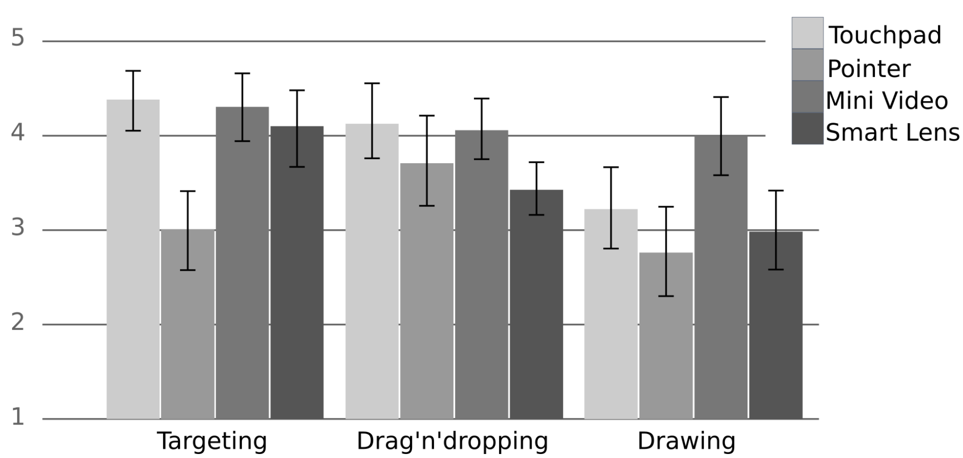
Figure 6. Results of drawing task: completion times (left) and accuracy in mean distance from line (right) of the four techniques



Subjective Ratings

When participants were asked after each completed experimental block (e.g., Targeting with Smart Lens) how much they felt supported by the respective interaction technique, different preference profiles were observed (see Figure 7). For targeting, Pointer was rated significantly lower than all three other techniques: the mean rating score for Pointer was 3, whereas it was approximately 4.2 for the other techniques ($p<0.008$, no other pairwise differences were significant). For drag'n'drop, pointer did not any more significantly differ from the other techniques; here Smart Lens was rated lowest (significantly lower than Mini Video, $p<0.008$, no other pairwise differences detected). For drawing, Mini Video achieved highest mean scores; these were significantly higher than Smart Lens and Pointer, $p<0.008$ (no other pairwise differences significant).

Figure 7. Perceived task support by technique per task (five-point rating scale, 1 = no task support, 5 = very high)



Techniques also had an effect on perceived mental effort, $F_{3,69}$=7.7, p<0.01 and on physical effort, $F_{3,69}$=7.7, p<0.01. While Pointer was experienced as most mentally demanding (significantly more than Touchpad and Mini Video, p<0.0017), Smart Lens was rated as most physically demanding (significantly more than all other techniques (p<0.0017).

BEHAVIORAL OBSERVATIONS AND USER COMMENTS

We analyzed qualitative behavioral observation notes and thinking aloud protocols from the free interaction phase, to gain a more detailed understanding of user performance and experience.

- **Touchpad:** Overall, the touchpad technique received favorable comments, as it relies on a well-known interaction metaphor known from many everyday tasks. A frequently observed problem was the unwilling activation of the software buttons next to the touchpad. Many users would have preferred a hardware button which could then be identified by touch and which would then relieve users from frequent switches of visual attention between smartphone and screen.
- **Pointer:** The general concept of pointing was positively acknowledged by many users, as it was considered 'intuitive' and enabled to keep visual attention focused on the screen. However, two 'practical' problems related to contemporary accelerometer- and gyroscope-based mobile orientation sensing severely hindered user performance and satisfaction. First, the short calibration necessary for absolutely aligning the mobile device to the screen was often not considered acceptable, especially to users with low smartphone experience. Second, for many users, exact control of the movement sensor was difficult. To 'stabilize' the mobile against unintended movement deviations, they sometimes applied creative strategies, such as firming up the elbow on the body or the second arm.
- **Mini Video:** Was easily learnable by participants, as it corresponded with everyday smartphone touch interactions. In turn, the usual drawbacks with touchscreen smartphones were experienced as well: small targets could not be easily selected and fingers were hindering visibility of display contents (especially for users with little smartphone experience). Further advantages were that the device could be held in the hand in different ways. We often observed the user's strategy to first identify a starting point at the mobile touchscreen and then to continue interacting while watching the large screen.
- **Smart Lens:** Was generally seen as an interesting novel interaction technique by many users and was often described as fascinating and technically intriguing. However, one strong drawback for many users was to hold the hand stable and precise enough to select targets. Another problem within the given long testing situation was arm and shoulder fatigue.

DISCUSSION

In this section, we refer back to our research hypotheses and discuss the described study results with regard to our expectations.

Touchpad

Both hypotheses 1a and 1b concerning the *Touchpad* are confirmed. As hypothesized, the technique performs well in terms of accuracy, however, it is slower than the direct techniques.

While in targeting tasks the direct visual techniques perform similarly in general, *Touchpad* revealed its strength for precisely selecting small targets. Also for drag'n'drop tasks, the touchpad works well for small targets. The technique further suffers from fewer errors for these two task types outperforming the direct techniques (80% less errors than second technique in targeting, 50% less errors in drag'n'dropping). Unexpectedly, the touchpad technique shows no significant advantage over the alternatives for drawing tasks. We explain this performance with the observed overshooting effect of the technique due to two reasons. The CD ratio of 1 enables fine-grained movement but also leads to inaccuracies by less experienced users. Another group consisting of younger subjects (partly with video gaming experience) expected faster reaction times and did not compensate for the slight delays naturally occurring from network transmission. While this effect does not show noticeable impact on target and drag'n'drop tasks, it grows apparent for sensitive drawing tasks.

Our results show that the *Touchpad* technique is significantly slower than the direct techniques for targeting and drawing. While it is about 50% slower than the second slowest technique in targeting, it performs almost 100% worse in drawing tasks. We thus conclude that the Touchpad technique is not convenient for drawing tasks. While the technique justifies its long completion times with high accuracy for targeting and drag'n'dropping, it does not show a significant improvement in accuracy for drawing. Thus, *Touchpad* is well-suited for remotely controlling traditional graphical user interfaces with small control elements on public displays avoiding the creation of an adapted interface. Another example are applications demanding for high dropping accuracy such as a puzzle game for precisely placing tiles or a collaborative art application involving the movement of images.

Pointer

Our hypothesis 2a is rejected: The completion time of the Pointer technique turns out to be the worst of the four techniques for targeting and drag'n'dropping (being 30% and 26% slower than the third-ranked technique). Only in drawing tasks it outperforms the touchpad technique. In contrast, hypothesis 2b concerning the general accuracy is confirmed: The pointer technique is dominated by the alternatives in all evaluated tasks but drag'n'drop where we found a slight advantage to the *Smart Lens* technique.

Even though previous studies (e.g., MacKenzie & Jusoh, 2001) already detected drawbacks of sensor-based pointing we did not expect this bad performance for the smartphone pointing approach. As reasons we identified both implementation details and technical limitations. Since the pointer technique was prototyped for high sensitivity to enable fine-grained operations, it was less tolerant for unintended movement and hand jitter. Second, despite applying an angle complementary filter, the pointer was affected by gyroscope drift over time hardening a precise intuitive control. Finding a compromise between error tolerance and sensitivity under these circumstances is challenging and might prevent successful smartphone-based pointing solutions for accuracy-demanding tasks in the near future.

Mini Video

Hypthesis 3a had expected the direct *Mini Video* approach to outperform the indirect techniques in terms of completion times what was not fully verified by our results. Both direct techniques generally outperform the remaining techniques in targeting tasks, while this is only true for large targets in drag'n'drop tasks. For the drag'n'dropping of smaller targets, the non-visual indirect *Touchpad* technique shows slight (non-significant) advantages in completion time. For drawing, the *Mini Video* technique has the significantly shortest completion time being 24% faster than the second-ranked *Smart Lens* technique.

The expected high error rate for small targets (hypothesis 3b) is confirmed by the results for both targeting and drag'n'dropping. In terms of accuracy, *Mini Video* significantly outperforms its direct competitor, the *Smart Lens*, for drag'n'dropping and drawing. For such precise complex interaction, the orientation-agnostic approach of the *Mini Video* technique ignoring device movement is a benefit. Thus, *Mini Video* seems to be perfectly suited for quick selecting tasks on public displays such as choosing a product to gain further information about it or targeting games expecting a fast reaction. However, *Mini Video* requires adapted user interfaces with large controls to reduce error rates and avoid user frustration. Further, promising *Mini Video* use cases are urban art applications allowing for collaborative drawing and applications including free-hand selection of areas by tracing.

Smart Lens

The study results confirm our hypotheses 4a: The *Smart Lens* performs similarly well as the *Mini Video* approach in terms of good completion time for all three tasks. Concerning the overall accuracy the results support our hypothesis 4b: while the technique's accuracy is comparable to the accuracy of the *Mini Video* in targeting tasks, its accuracy for drag'n'dropping and drawing is significantly lower than the accuracy of the *Mini Video*. Further, we could observe the (not significant) tendency of a lower error rate for small targets in the targeting task what we ascribe to the inherent zooming opportunity of the technique.

From these study results, we conclude that this pure form of the *Smart Lens* technique is well-suited for spontaneous targeting tasks with short interaction periods. A special advantage can be supposed for smaller-sized targets. As we expected, this implementation offers no benefit for more complex tasks involving several working steps. The advantage of the zoom does not become manifest in the more complex tasks due to the technique's sensitivity to hand jitter and unintended device movement during longer dragging and drawing tasks. Features such as temporarily freezing the live video (such as suggested by Boring et al., 2010) could be applied to stabilize control over mobile live video. However, in the context of public displays, they would hamper truly spontaneous interaction and limit the experience of live interaction. An alternative is the integration of an intelligent cross-hair which is aware of control elements targeted through the camera viewfinder and facilitates selecting them (Baldauf & Fröhlich, 2014).

TOWARDS SERENDIPITOUS INTERACTION

The presented study was deliberately conducted in a controlled lab environment to investigate the strengths and weaknesses of the chosen mobile-display interaction techniques in detail. However, making these input techniques available in the field and providing successful and user-accepted interactive services, poses several new research challenges. Besides the investigation of effective and entertaining remote controls and appealing and useful applications, a key design challenge is providing a low threshold of use in order to attract passers-by and make them actual users (Ballagas et al., 2006).

Connecting to the Public Screen

One central cornerstone for such serendipitous mobile interactions with a public display is the ease and speed of setting up a data connection between the smartphone and the environment (Ballagas et al., 2006). Today, several methods for transmitting the required pairing information are realizable. Users may simply enter a unique identifier displayed by the screen, may use Bluetooth for scanning their surroundings for screens and selecting desired ones from a result list, may photograph a QR (Quick Response) code or touch an NFC (Near-Field Communication) tag. Instead of a separate pairing (e.g., by touching an NFC tag) and controlling (e.g., by using the Mini video technique) phase, interaction styles which feature an inherent pairing procedure seem favorable. For example, using advanced computer vision and an intelligent back-end handling several screens, the Smart Lens technique can exploit the actual screen content as a visual marker for detecting and identifying the screen (Baldauf, Fröhlich, & Lasinger, 2012; Boring et al., 2010) and thus enable a pairing process transparent to the user.

In a recent user study (Baldauf, Salo, Suette, & Fröhlich, 2013), these five pairing approaches (the four established ones plus the novel camera-based detection named 'display pointing') were compared. In general, the subjects stated to prefer long-distance techniques (such as scanning and display pointing) over techniques requiring immediate vicinity to the screen. Relevant reasons of the participants included not to be forced to walk to the display and being able to literally operate from a position in the background without being immediately recognizable as the one responsible for the actions on the public display. Explicitly scanning for screens and manually entering identifiers was perceived by the participants as easy to learn and to use, yet as less innovative and fun. Surprisingly, capturing QR codes turned out to be difficult for a considerable number of users and lowered the learnability. NFC, by contrast, received good ratings in general, however, suffered from the general preference for distant techniques in the context of public displays. Display pointing was proved to be a good combination of a directional technique supporting distant pairing while being efficient and fun.

Providing the Remote Control Software

Another important, yet often neglected cornerstone for lowering the threshold of use is the efficient provisioning of the mobile application. As long as there are no standardized protocols for remotely controlling a public screen through a smartphone or even a market-leading app, passers-by will probably not have a respective application installed on their devices. However, finding, downloading and installing a suitable application is cumbersome and time-consuming and represents another major obstacle for spontaneously using a mobile device as a remote control on-the-go.

One promising approach to avoid the installation of a native mobile application for the user, is the realization of respective mobile interaction techniques through modern Web technologies, thus only requiring a state-of-the-art mobile Web browser. With the arrival of the latest version of the Web markup language, HTML5, developers can use of large set of advanced multimedia features and access mobile hardware: respective Web applications may read the acceleration sensors of a device, exploit its camera(s), use advanced drawing operations and establish additional network connections for streaming data via so-called WebSockets.

For research purposes, we re-implemented the four evaluated input techniques Touchpad, Pointer, Mini Video and Smart Lens using pure Web technologies and conducted a first informal study to learn about the user acceptance of such web-based remote controls. In the following, we describe our technical implementations for each technique, highlight in-sights from the development and practical tests and report on the user feedback.

Touchpad

Since basic touch events such as the start, movement and end of a touch point are supported by all modern mobile browsers, a web-based version of the touchpad can be realized easily. Our current implementation is based on a simple ‘div’ element which registers occurring strokes and forwards the relative mouse movements through a WebSocket connection to the server hosting the public screen with the actual application. Additional buttons allow for sending an action command triggering mouse clicks at the server-side.

Pointer

Similar to the Touchpad, the Pointer can be implemented with Web technologies without major problems. Access to raw sensor data of built-in accelerometers and gyroscopes is standardized and supported by state-of-the-art mobile browsers. The algorithms for calculating mouse cursor coordinates from changes in the device orientation can be easily ported to JavaScript. For these calculations, the mobile device needs to be aware of its resolution. We solved this issue by sending a short welcome message to the mobile client as soon as the connection is established. Among other parameters, this welcome message contains the height and width of the remote display. Communicating with the server is realized again through WebSockets.

Mini Video

This technique requires the detection of touch gestures (which were briefly described above in the Touchpad section) and the display of images dynamically received from the server. The server-side part of our web-based Mini Video is slightly modified in contrast to the native app. Again, screenshots of the current display are continuously taken and compressed as JPEG images. However, we finally convert these binary images to base64-encoded text strings to send them over a WebSocket connection and to later exploit the so-called inline presentation of images in a Web browser: instead of specifying a path to a locally stored picture for an image element, we can directly pass the base64-encoded screenshot as source value as soon as we have received it from the server. Further, we can exploit available image scaling features to eventually enlarge the screenshot to fill the entire browser viewport (Figure 8, left).

Smart Lens

Also the technically most sophisticated remote control, the Smart Lens, can be replicated using state-of-the-art Web technologies. For our prototype (Figure 8, right) we made use of the so-called WebRTC standard (Web Real-time Communication) for accessing the built-in camera(s) from within the browser. While showing a basic camera preview is well-supported, accessing the current video frame in a stan-

Figure 8. Also the two more complex remote controls can be realized with pure state-of-the-art Web technologies: the Mini Video (left) displaying screenshots of the remote screen and the Smart Lens (right) accessing the mobile camera.

dard file format upon touch is currently feasible through a work-around: The frame can be drawn on an invisible canvas element whose content can be fetched as a compressed JPEG image. Finally, this JPEG frame is sent together with the touch coordinates over a binary WebSocket for further processing at the platform side, similar to our original native prototype.

User Feedback

To gain first insights into the general reception and acceptance of such web-based remote controls, we invited 30 test persons to our user experience lab. The group comprised 15 male and 15 female participants who were aged between 19 and 50 years (mean = 31.4, median = 30.0).

In a setup similar to the one from the previously described formal user study, the participants experimented with several web-based remote controls in a pointing test and a video game and answered a short questionnaire afterwards. First, we asked whether and why they noticed that they were using a Web application. Second, we wanted to know whether they would prefer a native app or a Web application for the use case of controlling an application on a remote screen or whether the technical realization does not make any difference to them. Again, we also asked for their reasons.

8 participants stated that they recognized the remote controls as Web applications because they had noticed the browser window with its address bar and tabs. Surprisingly, 22 test persons did not actively notice that they were using a Web application in a browser. These users were mainly less technology-affine and probably less interested in the underlying technology and thus did not pay attention to such details. A further explanation could be the visual appearance of our remote controls which does not resemble typical layouts of Web pages with lots of text, images, or hyperlinks. Not a single participant mentioned performance issues related to the Web-based realization, one even appreciated “the quick response times”.

Concerning the second question, 14 participants stated that they would prefer a native app and 3 participants would favor a mobile Web application. For the remaining 14 participants the technical realization was not relevant.

8 participants who preferred a native solution mentioned data connectivity as a reason. They assumed that once the application is downloaded and installed it can be used offline. This is true for the majority of applications but is no valid argument for the remote control use case. Here, a wireless connection is required anyway to exchange commands between the mobile device and the remote system.

The three participants who were in favor of a Web application explained that no download and installation through an application store is required, what would also save storage space on their mobile devices. One tech-savvy person said that she would prefer the Web application also due to security considerations. She considered this approach more trustworthy since she might have a look at the source code and see, for example, what information is transmitted – a true argument which obviously is only of relevance for users with programming expertise.

For the undecided group, the main argument was the functionality: as long as the application works properly, they do not care about the technical realization.

While only three participants of our study explicitly preferred a Web application, overall 17 of our 30 participants would accept purely Web-based solutions, a promising result given the current popularity of native apps.

Restrictions

Throughout the preparations of the prototypes and the conduct of the study, we noticed a few peculiarities and limitations of our Web-based remote controls in direct comparison to their native counterparts. First, providing a fully app-like experience is difficult since Web applications cannot change to full-screen mode (i.e., without the typical browser controls) by themselves. Modern mobile browsers support full-screen applications, however, switching to full-screen mode is only possible through an explicit user interaction (such as hitting a button) as well as the confirmation of the consequent information message.

Such an explicit permission is also required when accessing the camera for realizing the Smart Lens, a clear drawback with regard to the envisioned serendipitous interaction. Further, at the time of writing, the web-based Smart Lens suffers from several mobile browser limitations and incompatibilities. For example, current mobile browsers differ in ways of selecting the preferred camera (in contrast to typical WebRTC applications like video conferencing, we need to make sure that the rear camera is used for the Smart Lens) and partly restrict specifying the video frame size (what can lead to bad results of the image matching algorithms).

Further, Web applications are justifiably not allowed to change settings of the operating system or the Web browser. For example, a Web application has no control over the display rotation. While a developer of a native application may determine that his application uses either portrait or landscape mode, a Web application is dependent on the corresponding settings for the browser. In case of the pointer technique, this is disadvantageous: expressive gestures with strong tilt movements may lead to a re-orientation of the display (e.g., switching from landscape to portrait mode). As a consequence, the displayed button appears at another display location and typically blindfold touches by the user fail.

In a similar vein, a Web application is not able to hide or disable any system buttons or keys. Thus, users can hit such buttons by mistake, especially in fast-paced games, and unintendedly close the browser. In case of a game this is particularly frustrating since it may lead to the death of the player character, or similar.

FUTURE RESEARCH DIRECTIONS

In this section, we summarize some emerging trends in the context of mobile interaction with public displays and outline promising future research directions.

Multi-User/Purpose Interaction Models

Public screens are typically not dedicated to one application but show different kinds of applications based on a timer. Consequently, the requirements of an interactive public display concerning the suitable type of remote control, the number of concurrent users, etc. would constantly change according to the currently active application. From a technical point of view, we need generic frameworks to handle such varying requirements: to enable different smartphone remote controls and dynamically provide the correct one, to extend and limit the number of users as well as to intelligently manage available applications based on enhanced criteria such as active users instead of simple schedules. Concerning user experience, open research questions include how to efficiently communicate the current criteria for interaction or how to deal with varying social aspects (e.g., when switching from a competitive to a collaborative game, cf. Luojus et al., 2013).

Simplified Pairing

As mentioned above, enabling serendipitous interaction is a key success factor for developing user-accepted mobile-enhanced applications for public displays. Selecting the screen to interact with on the mobile device and pairing the mobile device with this screen must be as simple and easy as possible. Aforementioned research investigated and compared several established ways of connecting a mobile device with nearby appliances, yet technological progress continuously enables new respective approaches. For example, recent low-powered wireless transmitters such as Apple's *iBeacons* could be used by a smartphone application to spatially sense an environment for screens, make recommendations to passers-by and even establish a connection to a screen based on proximity. This approach would reduce explicit user interactions during the pairing process, however has not yet been studied.

Advanced Web-Based Interaction

Our prototypical web-based implementations of the remote control techniques prove the general feasibility of modern Web technologies for ubiquitous computing applications. However, we consider these basic demonstrators as just the beginning. For example, the trivial Web version of the Mini video technique continuously transmits full screenshots of the remote display, a more sophisticated version could utilize advanced video codecs to enable smooth video streaming within the Web browser. Further, the Smart Lens implementation sends the current camera frame to the server executed on a powerful notebook computer for processing. However, first performant computer vision libraries (cf. Tracking.js, 2014) for Web-based applications emerged recently and enable remarkable multimedia applications within modern (mobile) Web browsers. Applying this technology for mobile Web applications such as our interactive Smart Lens will enhance response times, improve the overall user experience and enable a couple of novel use cases.

Integration of Wearables

Wearable gadgets such as data glasses projecting digital information directly in the user's field of vision or smart watches enabling access to digital services at the user's wrist have been recently becoming available on the mass market. While passers-by have to take their smartphones out of the pockets to start interacting with remote screens, wearable technology is literally ready at hand, yet its feasibility and potential for such interactive networked services are largely unexplored. Data glasses such as Google Glass seem especially promising for visual interaction approaches such as the presented Mini video and Smart Lens techniques. A suitable related application is the 'Augmented Video Wall', which visually overlays a public screen with personalized content (Baldauf & Fröhlich, 2013). Due to their small touch displays, gestural interactions with a remote application based on accelerometer data could be explored for smart watches, e.g., pointing and throwing gestures or wiping commands.

Multi-Modal Screen Interaction

Today, users are often equipped with several mobile devices such as a smartphone, a tablet computer, and even wearable gadgets. At the same time, public screens increasingly feature advanced input hardware such as multi-touch displays and (depth) cameras. Future research should explore novel beneficial combinations of the available interaction styles and respective applications. Examples include the seamless transfer of media files from a mobile device to a public screen where they can be collaboratively edited on the large touch display and sent back to the mobile. A more advanced transfer example is the selection of a file on a mobile device and its precise positioning on the remote screen through analyzing the user's gaze through a screen-mounted camera (Turner, Bulling, Alexander, & Gellersen, 2013). A public poll application may prompt passers-by to quickly cast their opinion through a touch display, while in-depth follow-up questions appear on their smart watches as soon as they left the screen.

Privacy-Awareness

Mobile interaction with wearables and public screens will always have to be highly aware of individual and societal privacy demands. One of the promises of smartphone-screen interactions could be that more person-specific aspects (such as opinion polls, individual preferences or expressions) could be handled on the personal device, and those targeting larger groups of spectators could be provided on the public screen. The requirements for such concepts have started to be explored only recently (e.g., Baldauf, Suette, Fröhlich, & Lehner, 2014). An important challenge for research is to enable people to spontaneously engage with their personal interests, motivations and opinions, while still preserving seamless means to protect them from threats.

CONCLUSION

We presented an extensive comparison of recent, not yet compared smartphone techniques for interaction with public displays, with regard to three generic task types.

Regarding the interpretation and practical application, we would like to note that results were not necessarily only a function of the technique's overall concept, but also of certain technical limitations inherent in today's available sensor and network technology. This applies especially for the *Pointer* technique, which was often hard to use due to the mentioned sensor inaccuracies. Irreducible transmission delays are inherent to all wireless network-based remote interaction techniques and occurred for all evaluated techniques to the same extent. However, in our study they grew most apparent for the *Touchpad* technique where users expected a completely simultaneous motion of the remote mouse cursor even for very quick operations. The development of customized laboratory implementations such as a pointing technique based on external vision-based tracking may have alleviated some of these restrictions, but in turn the generalizability for today's widespread smartphone usage would then have been lowered.

Summarizing our study results, none of the orientation-aware techniques could generally outperform its orientation-agnostic counterpart with regard to mere performance measures. However, the *Pointer* and the *Smart Lens* were often described as intuitive and fascinating and the *Smart Lens* showed beneficial peculiarities for special task instances. Thus, we encourage the deeper investigation of the impact of orientation-awareness by exploring stabilizing techniques to cope with hand jitter while still preserving real-time interaction. Further, we suggest replication studies to re-evaluate the recommendations of the investigated techniques (especially the *Pointer* technique) following performance improvements on the consumer market, which may continuously dissolve today's technological limitations.

Moreover, we introduced browser-based remote controls as a strategy towards more serendipitous mobile interactions. Our Web versions of the four evaluated interaction techniques reveal the current pitfalls of respective realizations during development and usage, yet demonstrate the overall applicability of state-of-the-art Web technologies for such ubiquitous computing applications. Together with the encouraging first feedback of test persons, these results are promising and can help to create user-accepted and easy-to-use smartphone remote controls and corresponding interactive services for public screens.

REFERENCES

Baldauf, M., & Fröhlich, P. (2013). The augmented video wall: Multi-user ar interaction with public displays. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13), 3015-3018*. New York: ACM. doi:10.1145/2468356.2479599

Baldauf, M., & Fröhlich, P. (2014). Snap target: Investigating an assistance technique for mobile magic lens interaction with large displays. *International Journal on Human Computer Interaction /30*(6), 446-458.

Baldauf, M., Fröhlich, P., & Lasinger, K. (2012). A scalable framework for markerless camera-based smartphone interaction with large public displays. In *Proceedings of the first International Symposium on Pervasive Displays (PerDis'12)*. New York: ACM. doi:10.1145/2307798.2307802

Baldauf, M., Fröhlich, P., & Reichl, P. (2010). Touching the untouchables: Vision-based real-time interaction with public displays through mobile touchscreen devices. In *Adj. Proceedings of the Eighth International Conference on Pervasive Computing* (Pervasive'10) (LNCS), (vol. 6030). Berlin: Springer.

Baldauf, M., Salo, M., Suette, S., & Fröhlich, P. (2013). The screen is yours - comparing handheld pairing techniques for public displays. In *Proceedings of the Fourth International Joint Conference on Ambient Intelligence* (AmI'13). Berlin: Springer.

Baldauf, M., Suette, S., Fröhlich, P., & Lehner, U. (2014). Interactive opinion polls on public displays - Studying privacy requirements in the wild. In *Proceedings of 16th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI'14)*. New York: ACM.

Ballagas, R., Borchers, J., Rohs, M., & Sheridan, J. (2006). The smart phone: a ubiquitous input device. IEEE *Pervasive Computing*, 5(1), 70-77.

Ballagas, R., Rohs, M., & Sheridan, J. (2005). Sweep and point and shoot: phonecam-based interactions for large public displays. In CHI '05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05), 1200-1203. New York: ACM. doi:10.1145/1056808.1056876

Bier, E. A., Stone, M. C., Pier, K., Buxton, W., & DeRose, T. D. (1993). *Toolglass and magic lenses: the see-through interface*. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques (SIGGRAPH'93)*, 73-80. New York: ACM. doi:10.1145/166117.166126

Boring, S., Baur, D., Butz, A., Gustafson, S., & Baudisch, P. (2010). *Touch projector: mobile interaction through video*. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'10)*, 2287-2296. New York: ACM.

Boring, S., Jurmu, M., & Butz, A. (2009). Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays. In *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction (OZCHI'09)*, 161-168. New York: ACM. doi:10.1145/1738826.1738853

Broll, G., Reithmeier, W., Holleis, P., & Wagner, M. (2011). *Design and evaluation of techniques for mobile interaction with dynamic NFC-displays*. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction (TEI'11)*, 205-212. New York: ACM. doi:10.1145/1935701.1935743

Card, C., English, W., & Burr, B. (1978). Evaluation of mouse, rate controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics*, 21(8), 601–613.

Douglas, A., & Mithal, A. (1994). The effect of reducing homing time on the speed of a finger-controlled isometric pointing device. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'94)*, 411-416. New York: ACM.

Douglas, S. A., Kirkpatrick, A. E., & MacKenzie, I. S. (1999). *Testing pointing device performance and user assessment with the ISO 9241, part 9 standard*. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI'99)*, 215-222. New York: ACM. doi:10.1145/302979.303042

Foley, J. D., Wallace, V. L., & Chan, P. (1984). The human factors of computer graphics interaction techniques. *IEEE Computer Graphics and Applications*, 4(11), 13-48.

Greenstein, J. S. (1997). Pointing Devices. Handbook of Human-Computer Interaction, 1317-1348. New York: Elsevier Science.

Herbert, L., Pears, N. E., Olivier, P., & Jackson, D. (2011). *Mobile device and intelligent display interaction via scale-invariant image feature matching*. In *Proceedings of the first International Conference on Pervasive and Embedded Computing and Communication System (PECCS'11). SciTePress 2011.*

Jota, R., Nacenta, M. A., Jorge, J. A., Carpendale, S., & Greenberg, S. (2010). A comparison of ray pointing techniques for very large displays. In Proceedings of Graphics Interface (GI'10), 269–276. Toronto: Canadian Information Processing Society.

Kabbash, P., MacKenzie, I. S., & Buxton, W. (1993). *Human performance using computer input devices in the preferred and non-preferred hands*. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems (CHI'93)*, 474-481. New York: ACM. doi:10.1145/169059.169414

Kruppa, M., & Krüger, A. (2003). Concepts for a combined use of personal digital assistants and large remote displays. In Proceedings of Simulation und Visualisierung (SIMVIS'03), 349-361. Magdeburg: SCS Publishing House e.V.

Lee, S., & Zhai, S. (2009). The performance of touch screen soft buttons. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'09)*, 309-318. New York: ACM. doi:10.1145/1518701.1518750

Logitech Blog. (2010). *New: Logitech touch mouse turns your iphone or ipod touch into a wireless trackpad and keyboard*. Retrieved October 10, 2014, from http://blog.logitech.com/2010/01/29/

Luojus, P., Koskela, J., Ollila, K., Mäki, S.-M., Kulpa-Bogossia, R., Heikkinen, T., & Ojala, T. (2013). Wordster: *Collaborative versus competitive gaming using interactive public displays and mobile phones*. In *Proceedings of the 2nd ACM International Symposium on Pervasive Displays (PerDis'13)*, 109-114. New York: ACM. doi:10.1145/2491568.2491592

MacKenzie, I. S., & Jusoh, S. (2001). *An evaluation of two input devices for remote pointing*. In Proceedings of the 8th IFIP International Conference on Engineering for Human-Computer Interaction (EHCI'01), Springer Lecture Notes in Computer Science, Vol 2254, 235-250. Heidelberg, Berlin, New York: Springer. doi:10.1007/3-540-45348-2_21

MacKenzie, I. S., & Oniszczak, A. (1998). A comparison of three selection techniques for touch pads. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'98)*, 336-343. New York: ACM.

MacKenzie, I. S., Sellen, A., & Buxton, W. (1991). A comparison of input devices in elemental pointing and dragging tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'91)*, 161-166. New York: ACM Press.

Myers, B. A., Bhatnagar, R., Nichols, J., Peck, C. H., Kong, D., Miller, R., & Long, A. C. (2002). Interacting at a distance: measuring the performance of laser pointers and other devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'02)*, 33–40. New York: ACM. doi:10.1145/503376.503383

Myers, B. A., Stiel, H., & Gargiulo, R. (1998). *Collaboration using multiple pdas connected to a pc*. In *Proceedings of the 1998 ACM conference on Computer supported cooperative work (CSCW'98)*, 285-294. New York: ACM. doi:10.1145/289444.289503

Park, Y. S., Han, S. H., Park, J., & Cho, Y. (2008). Touch key design for target selection on a mobile phone. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services (MobileHCI'08)*, 423-426. New York: ACM. doi:10.1145/1409240.1409304

Pears, N., Jackson, D., & Olivier, P. (2009). Smart phone interaction with registered displays. *IEEE Pervasive Computing*/8/(2), 14-21.

ABI Research. (2011). *Digital signage revenue to approach $4.5 billion in 2016*. Retrieved October 10, 2014, from https://www.abiresearch.com/press/digital-signage-revenue-to-approach-45-billion-in-

Roudaut, A., Huot, S., & Lecolinet, E. (2008). *Taptap and magstick: improving one-handed target acquisition on small touch-screens*. In *Proceedings of the working conference on Advanced visual interfaces (AVI'08)*, 146-153. New York: ACM. doi:10.1145/1385569.1385594

Stoakley, R., Conway, M. J., & Pausch, R. (1995). *Virtual reality on a wim: interactive worlds in miniature*. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'95)*, 265-272. New York: ACM Press/Addison-Wesley Publishing Co. doi:10.1145/223904.223938

Tani, M., Yamaashi, K., Tanikoshi, K., Futakawa, M., & Tanifuji, S. (1992). *Object-oriented video: interaction with real-world objects through live video*. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'92)*, 593-598. New York: ACM. doi:10.1145/142750.143011

TeamViewer. (2014). *Mobile TeamViewer app*. Retrieved October 10, 2014, from http://www.teamviewer.com/en/download/mobile.aspx Tracking.js. (2014). *Official web page*. Retrieved October 10, 2014, from http://trackingjs.com

Tech, R. P. A. (2014). *Mobile Mouse Pro*. Retrieved October 10, 2014, from http://www.mobilemouse.com

Turner, J., Bulling, A., Alexander, J., & Gellersen, H. (2013). Eye drop: an interaction concept for gaze-supported point-to-point content transfer. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia (MUM'13)*. New York: ACM. doi:10.1145/2541831.2541868

ADDITIONAL READING

Alt, F., Shirazi, A. S., Kubitza, T., & Schmidt, A. (2013). Interaction techniques for creating and exchanging content with public displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'13)*, 1709-1718. New York: ACM. doi:10.1145/2470654.2466226

Ananny, M., & Strohecker, C. (2009). Textales: Creating interactive forums with urban publics. In *Handbook of Research on Urban Informatics*. Hershey, PA: IGI Global. doi:10.4018/978-1-60566-152-0.ch005

Baldauf, M., Lasinger, K., & Fröhlich, P. (2012). Private public screens – detached multi-user interaction with large displays through mobile augmented reality. In *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia (MUM'12)*. New York: ACM. doi:10.1145/2406367.2406401

Bødker, S., Klokmose, C. N., Korn, M., & Polli, A. M. (2014). Participatory IT in semi-public spaces. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction (NordiCHI'14)*, 765-774. New York: ACM.

Brignall, H., & Rogers, Y. (2003). *Enticing people to interact with large public displays in public spaces*. In Proceedings of the ninth IFIP TC13 International Conference on Human-Computer Interaction (INTERACT'03), 17-24. IOS Press.

Cardoso, J., & José, R. (2014). Interaction tasks and controls for public display applications. *Advances in Human-Computer Interaction*, *2014*, 371867. doi:10.1155/2014/371867

Chen, J., Dourish, P., Hayes, G. R., & Mazmanian, M. (2014). From interaction to performance with public displays. *Personal and Ubiquitous Computing*/18/(7), 1617-1629. Springer-Verlag London, UK.

Clinch, S. (2013). Smartphones and pervasive public displays. Pervasive Computing/12/(1), 92-95. IEEE. Fröhlich, P., Oulasvirta, A., Baldauf, M., & Nurminen, A. On the move, wirelessly connected to the world. Communications of the ACM/54/(1), 132-138. New York: ACM.

Gianluca, S., Milano, M., Saldivar, J., Nasir, T., Zancanaro, M., & Convertino, G. (2013). *Agora2.0: enhancing civic participation through a public display.* In Proceedings of the 6th International Conference on Communities and Technologies (C&T'13), 46-54. New York: ACM.

Hardy, R., Rukzio, E., Holleis, P., & Wagner, M. (2010). *Mobile interaction with static and dynamic NFC-based displays*. In Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (MobileHCI'10), 123-132. New York: ACM.

Jeon, S., Hwang, J., Kim, G. J., & Billinghurst, M. (2010). Interaction with large ubiquitous displays using camera-equipped mobile phones. *Personal Ubiquitous Computing/*14/(2), 83–94.

Krüger, A., & Kuflik, T. (2012). *Ubiquitous Display Environments*. Heidelberg, Berlin, New York: Springer. doi:10.1007/978-3-642-27663-7

Kukka, H., Oja, H., Kostakos, V., Gonçalves, J., & Ojala, T. (2013). What makes you click: exploring visual signals to entice interaction on public displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'13)*, 1699-1708. New York: ACM. doi:10.1145/2470654.2466225

Lorenz, A., & Jentsch, M. (2010). The ambient media player: A media application remotely operated by the use of mobile devices and gestures. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia (MUM'10)*, 15:1-15:10. New York: ACM. doi:10.1145/1899475.1899490

Müller, J., Alt, F., Michelis, D., & Schmidt, A. (2010). Requirements and design space for interactive public displays. In *Proceedings of the international conference on Multimedia (MM'10)*, 1285-1294. New York: ACM. doi:10.1145/1873951.1874203

Müller, J., Walter, R., Bailly, G., Nischt, M., & Alt, F. (2012). Looking glass: a field study on noticing interactivity of a shop window. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'12)*, 297-306). New York: ACM.

Natapov, D., Castellucci, S. J., & MacKenzie, I. S. (2009). ISO 9241-9 evaluation of video game controllers. In Proceedings of Graphics Interface 2009 (GI'09), 223-230. Toronto: Canadian Information Processing Society.

Ostkamp, M., Hülsermann, J., Kray, C., & Bauer, G. (2013). Using mobile devices to enable visual multiplexing on public displays: Three approaches compared. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia (MUM'13)*. New York: ACM. doi:10.1145/2541831.2541835

Peltonen, P., Kurvinen, E., Salovaara, A., Jacucci, G., Ilmonen, T., Evans, J., & Saarikko, P. et al. (2008). *It's mine, don't touch!: interactions at a large multi-touch display in a city centre*. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'08)*, 1285-1294. New York: ACM. doi:10.1145/1357054.1357255

Salo, M., Baldauf, M., Fröhlich, P., & Suette, S. (2013). *Peak moments of physical mobile interaction techniques*. In Proceedings of the 19th Americas Conference on Information Systems (AMCIS'13). Chicago: Association for Information Systems 2013

She, J., Crowcroft, J., Fu, H., & Li, F. (2014). *Convergence of interactive displays with smart mobile devices for effective advertising: A survey. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)/10/(2)*. New York: ACM.

Terrenghi, L., Quigley, A., & Dix, A. (2009). A taxonomy for and analysis of multi-person-display ecosystems. *Personal and Ubiquitous Computing*, *13*(8), 583–598. doi:10.1007/s00779-009-0244-5

Vajk, T., Coulton, P., Bamford, W., & Edwards, R. (2008). *Using a mobile phone as a "wii-like" controller for playing games on a large public display* (Vol. 2008, p. 539078). Article, ID: Computer Games Technology.

Walter, R., Bailly, G., & Müller, J. (2013). *StrikeAPose: revealing mid-air gestures on public displays*. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'13)*, 841-850. New York: ACM. doi:10.1145/2470654.2470774

Zhang, Y., Bulling, A., & Gellersen, H. (2013). SideWays: A Gaze Interface for Spontaneous Interaction with Situated Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'13)*, 851-860. New York: ACM. doi:10.1145/2470654.2470775

Zhang, Y., Müller, J., Chong, M. K., Bulling, A., & Gellersen, H. (2014). GazeHorizon: Enabling Passers-by to Interact with Public Displays by Gaze. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'14)*, 559-563. New York: ACM. doi:10.1145/2632048.2636071

KEY TERMS AND DEFINITIONS

Direct Interaction: A mobile interaction style, which enables the one-step specification of a position on the remote screen.

Indirect Interaction: A mobile interaction style, which utilizes a mediator, mostly the remote mouse cursor, to select a screen position.

Mini Video: An interaction technique applying a touch-enabled cloned miniature version of the content of the remote screen.

Orientation-Agnostic Interaction: An interaction style, which does not take the current spatial orientation of the device into account.

Orientation-Aware Interaction: An interaction style, which exploits the current spatial orientation of the device for input purposes.

Pointer: An interaction technique, which mimics a traditional laser pointer and maps the orientation of the device to absolute cursor positions on the remote screen.

Smart Lens: An interaction technique utilizing the device camera for augmenting the live video or allowing for context-sensitive operations.

Touchpad: An interaction technique, which utilizes the touch display of a mobile device as replacement for a traditional trackpad of a laptop computer.