

From Touchpad to Smart Lens: A Comparative Study on Smartphone Interaction with Public Displays

Matthias Baldauf, FTW Telecommunications Research Center Vienna, Vienna, Austria

Peter Fröhlich, FTW Telecommunications Research Center Vienna, Vienna, Austria

Jasmin Buchta, FTW Telecommunications Research Center Vienna, Vienna, Austria

Theresa Stürmer, FTW Telecommunications Research Center Vienna, Vienna, 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 new approaches for mobile-display interaction, inter-technique comparisons of respective methods are scarce. The authors conducted an experimental user study on four currently relevant mobile-display interaction techniques ('Touchpad', 'Pointer', 'Mini Video', and 'Smart Lens') and learned that their suitability strongly depends on the task and use case at hand. The study 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. Examples for the application of these differentiated results to real-world use cases are provided.

Keywords: Interaction Techniques, Mobile Device, Public Display, Remote Control, Touchscreen

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 triple in the next few years and 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 promising input devices for such remote systems. With their steadily expanding set of features

DOI: 10.4018/jmhci.2013040101

such as built-in sensors, high quality cameras, and increasing processing power, they enable several advanced techniques to interact with large public displays.

Ballagas et al. (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 et al., 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 2), 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, yet direct *Smart Lens* technique (Figure 4) 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 select a specific screen point by touching the mobile device display. Since this direct technique works directly 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) 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 3) 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.

Figure 1. Indirect techniques: Touchpad. Touchpad and Mini Video are not sensitive to device movement.

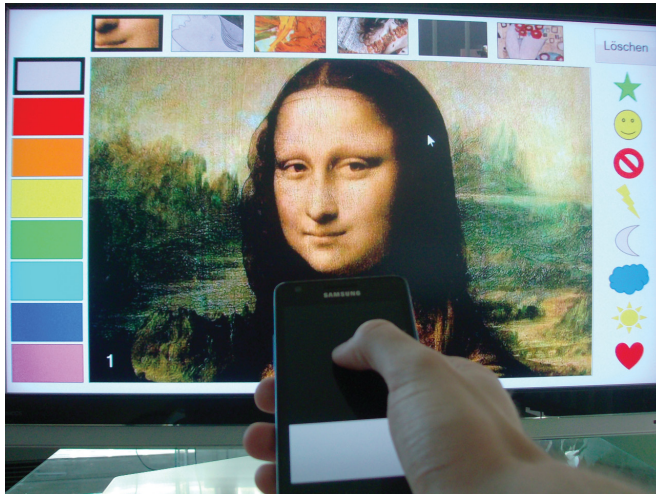


Figure 2. Indirect techniques: Pointer. Pointer and Smart Lens are orientation-aware techniques.

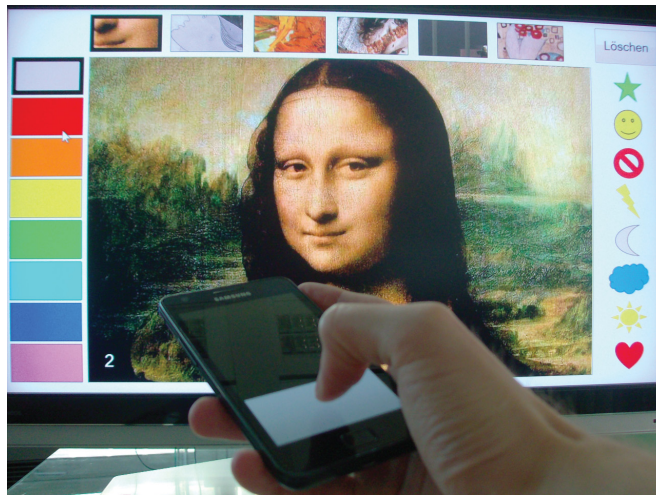


Table 1 shows the four distinct interaction techniques we explore in detail according to the traditional dimension of directness and the novel dimension of orientation-awareness.

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. Based on the findings in the presented evaluation we conclude with recommendations in the final section.

Figure 3. Direct techniques: Mini Video. Touchpad and Mini Video are not sensitive to device movement.

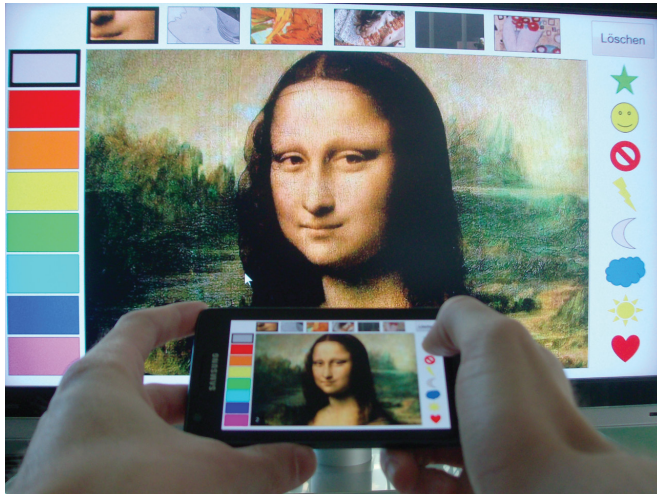


Figure 4. Direct techniques: Smart Lens. Pointer and Smart Lens are orientation-aware techniques.

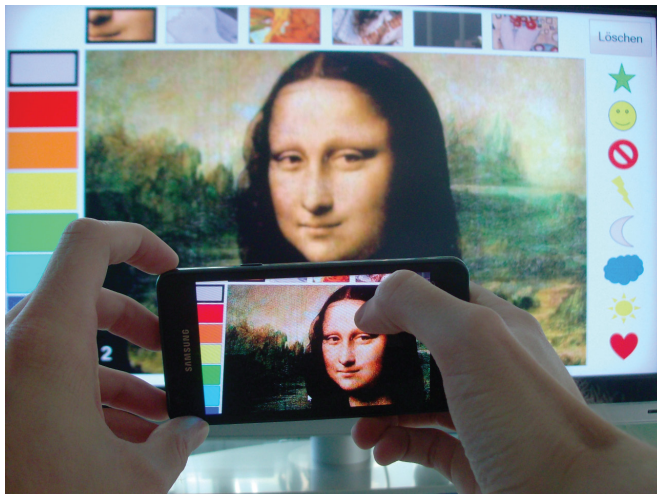


Table 1. Classification of the compared techniques according to the dimensions directness and orientation-awareness

	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 et al. (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 et al., 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 et al. (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 et al. (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, 2010) 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 et al. (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 and 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 et al. (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, 2012) 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 et al., 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 et al., 2005). Pears et al. (2009) introduce the idea of dynamic markers in form of four green boxes. Their preliminary non-comparative user studies with four and ten participants shows 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 et al. (1992) who introduced this concept for remote controlling industrial machines over video. Boring et al. (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 et al. (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 et al. (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 touch-screens 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 et al. (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 5). 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 5 and Figure

Figure 5. A participant using the Smart Lens technique during the free interaction phase of our lab study



6 through 8) 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.

Touchpad

Using this technique the mouse cursor on the remote display could be controlled through respective finger gestures on the mobile touch-screen 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). 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.

Figure 6. For each interaction technique, participants solved three different tasks with increasing complexity: Targeting, selecting the red circle

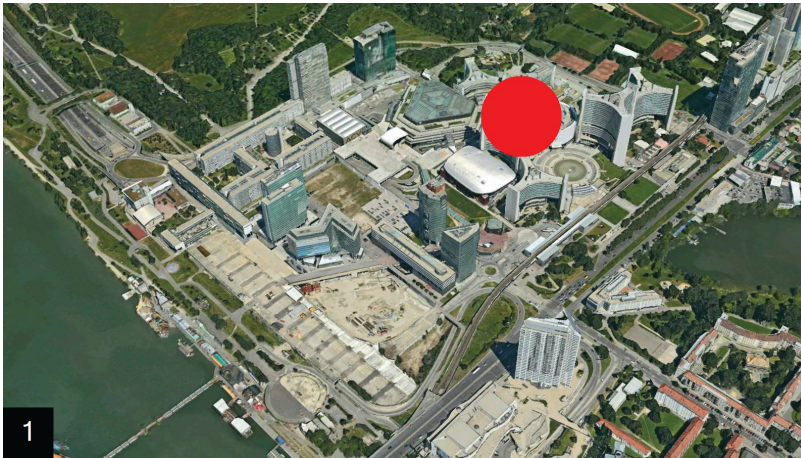
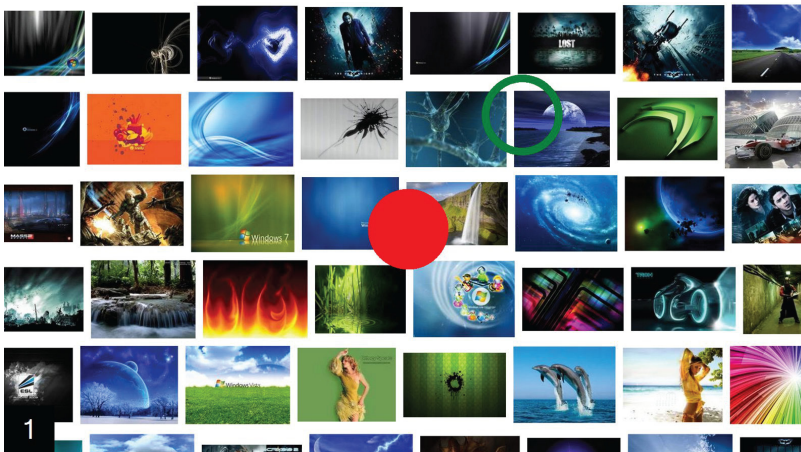


Figure 7. For each interaction technique, participants solved three different tasks with increasing complexity: Drag 'n' Dropping, moving the red circle over the green destination circle

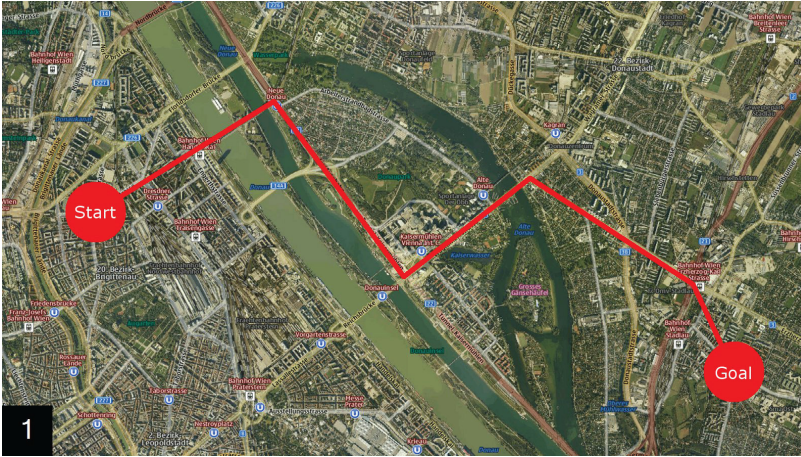


Pointer

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

Figure 8. For each interaction technique, participants solved three different tasks with increasing complexity: Drawing, tracing the red path from start to end point



center. The graphical interface resembled the touchpad technique with its button for triggering a mouse button action (Figure 2).

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 3). 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 4), 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 (cf. MacKenzie et al., 1991; Kabbash et al., 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

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	

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 6). 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).

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 7) 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 8) 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 5. 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. Pairwise comparisons used Bonferroni corrected

confidence intervals to maintain comparisons against $\alpha=0.05$. Errorbars in graphs represent a 95% confidence interval.

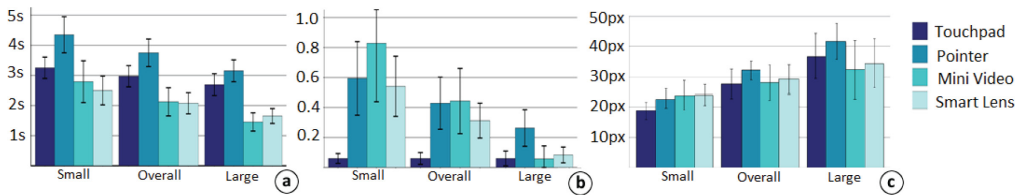
Targeting

- **Completion time:** The results including all target sizes and distances (see 'overall' bars in Figure 9a), indicate that using the Pointer technique took most time ($M=3750$ ms, $SD=1081$ ms), 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, $SD=833$ ms). 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 9a), 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 9b) 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 fewest errors than all other techniques ($p<0.0017$), but

Figure 9. Results of targeting task: (a) completion times, (b) errors, and (c) accuracies per trial of the four techniques, separately presented for small, overall (all target sizes), and large target sizes



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 9c); 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.

Drag'n'Dropping

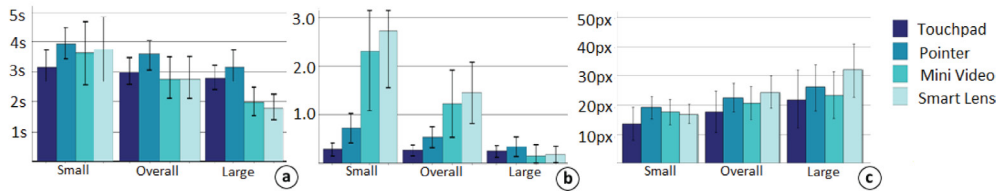
- Completion time:** Overall results including all distances and target sizes (Figure 10a, center) indicate that mean durations were quite similar among the techniques, only Pointer had relatively longer selection times (significantly longer than Touchpad and Miniature Video, $p < 0.008$).

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 10b) 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 10c) 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

Figure 10. Results of drag'n'drop task: (a) completion times, (b) errors, and (c) accuracies per trial of the techniques, separated by small, overall (all target sizes), and large target sizes



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 11a 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 11b). Analysis of errors is not reported, as it is not applicable due to the nature of the task.

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 12). 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).

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 ($p<0.0017$)).

Figure 11. Results of drawing task: (a) completion times and (b) accuracy in mean distance from line of the four drawing techniques

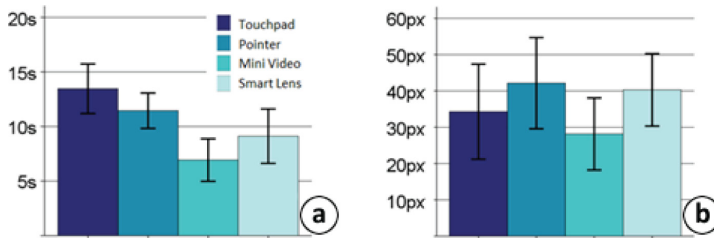
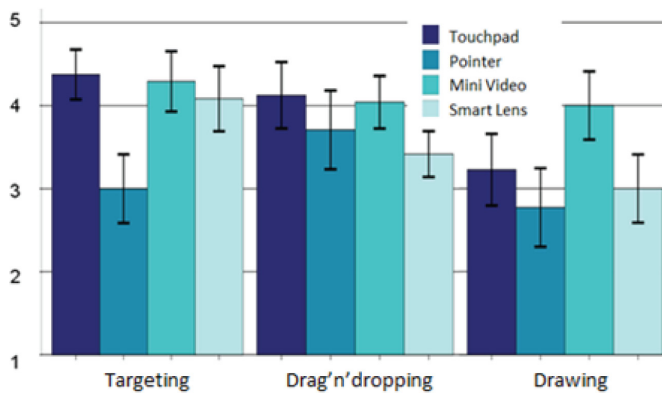


Figure 12. Perceived task support by technique per task (five-point rating scale, 1 = no task support, 5 = very high)



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 AND CONCLUSION

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

Hypothesis 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.

LIMITATIONS AND FURTHER WORK

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 vision-based pointing technique 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.

In the future, we plan to investigate promising combinations of these basic techniques taking advantage of the single techniques' identified strengths. One concept which was imagined by several study participants includes a hybrid video approach where the *Smart Lens* is used as initial technique – while the view switches to the *Mini Video* as soon as the display has been recognized. One promising combination of an indirect and a direct technique is using *Mini Video* or *Smart Lens* for selecting a screen object which can then be translated by the *Touchpad* technique for more precise control.

REFERENCES

- Baldauf, M., Fröhlich, P., & Reichl, P. (2010). Touching the untouchables: Vision-based real-time interaction with public displays through mobile touchscreen devices. In *Proceedings of the 8th International Conference on Pervasive Computing*, Helsinki, Finland.
- Ballagas, R., Borchers, J., Rohs, M., & Sheridan, J. (2006). The smart phone: a ubiquitous input device. *Pervasive Computing, IEEE*, 5(1), 70–77. doi:10.1109/MPRV.2006.18.
- 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 SIGGRAPH* (pp. 73–80). ACM.
- Boring, S., Baur, D., Butz, A., Gustafson, S., & Baudisch, P. (2010). Touch projector: Mobile interaction through video. In *Proceedings of the CHI* (pp. 2287–2296). 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 OZCHI* (pp. 161–168). ACM.
- Broll, G., Reithmeier, W., Holleis, P., & Wagner, M. (2011). Design and evaluation of techniques for mobile interaction with dynamic nfc-displays. In *Proc. TEI* (pp. 205–212). ACM.
- 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, 601–613. doi:10.1080/00140137808931762.

- 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 CHI* (pp. 411-416). 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 CHI* (pp. 215-222). ACM.
- Foley, J. D., Wallace, V. L., & Chan, P. (1984). The human factors of computer graphics interaction techniques. *IEEE Computer Graphics and Applications*, 11, 13-48.
- Greenstein, J. S. (1997). Pointing devices. In *Handbook of human-computer interaction* (pp. 1317-1348). 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 PECCS*.
- Jota, R., Nacenta, M. A., Jorge, J. A., Carpendale, S., & Greenberg, S. (2010). A comparison of ray pointing techniques for very large displays. [Canadian Information Processing Society.]. *Proceedings of the Graphics Interface, 2010*, 269-276.
- 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 INTERCHI '93 Conference on Human Factors in Computing Systems: Bridges Between Worlds* (pp. 474-481). New York, NY: ACM Press.
- Kruppa, M., & Krüger, A. (2003). Concepts for a combined use of personal digital assistants and large remote displays. In T. Schulze, S. Schlechtweg, & V. Hinz (Eds.), *Proceedings of the SIMVIS* (p. 349-361). SCS Publishing House e.V.
- Lee, S., & Zhai, S. (2009). The performance of touch screen soft buttons. In *Proceedings of the CHI* (pp. 309-318). ACM.
- Logitech Blog. (2010). *New: Logitech touch mouse turns your iphone or ipod touch into a wireless trackpad and keyboard*. Retrieved December 02, 2011, from <http://blog.logitech.com/2010/01/29/>
- MacKenzie, I. S., & Jusoh, S. (2001). An evaluation of two input devices for remote pointing. In *Proceedings of the EHCI* (pp. 235-250). Springer-Verlag.
- MacKenzie, I. S., & Oniszczak, A. (1998). A comparison of three selection techniques for touch pads. In *Proceedings of the CHI* (pp. 336-343). ACM.
- MacKenzie, I. S., Sellen, A., & Buxton, W. (1991). A comparison of input devices in elemental pointing and dragging tasks. In *Proceedings of the CHI '91 Conference on Human Factors in Computing Systems: Reaching Through Technology* (pp. 161-166). New York, NY: ACM Press.
- Myers, B. A., Bhatnagar, R., Nichols, J., Peck, C. H., Kong, D., Miller, R., et al. (2002). Interacting at a distance: Measuring the performance of laser pointers and other devices. In *Proceedings of the CHI* (pp. 33-40). ACM.
- Myers, B. A., Stiel, H., & Gargiulo, R. (1998). Collaboration using multiple pdas connected to a pc. In *Proceedings of the CSCW* (pp. 285-294). ACM.
- 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 MobileHCI* (pp. 423-426). ACM.
- Pears, N., Jackson, D., & Olivier, P. (2009). Smart phone interaction with registered displays. *IEEE Pervasive Computing / IEEE Computer Society [and] IEEE Communications Society*, 8(2), 14-21. doi:10.1109/MPRV.2009.35.
- Research, A. B. I. (2011). *Digital signage revenue to approach \$4.5 billion in 2016*. Retrieved December 2, 2011, from <http://www.abiresearch.com/press/3687>
- Roudaut, A., Huot, S., & Lecolinet, E. (2008). Taptap and magstick: improving one-handed target acquisition on small touch-screens. In *Proceedings of the AVI* (pp. 146-153). ACM.
- Stoakley, R., Conway, M. J., & Pausch, R. (1995). Virtual reality on a wim: Interactive worlds in miniature. In *Proceedings of the CHI* (pp. 265-272). ACM Press/Addison-Wesley Publishing Co.
- 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 CHI* (pp. 593-598). ACM.
- TeamViewer. (2012). *Mobile TeamViewer app*. Retrieved July 4, 2012, from <http://www.teamviewer.com/en/download/mobile.aspx>
- Tech, R. P. A. (2010). *Mobile mouse pro*. Retrieved December 2, 2011, from <http://www.mobilemouse.com>

Matthias Baldauf is a senior researcher and project manager at FTW Telecommunications Research Center Vienna. He currently leads the project PRIAMUS which investigates smart lens-based interaction for mobile devices and distant screens. Matthias received master's degrees in Information Systems from University of Vienna in 2004 and in Software Engineering from Vienna University of Technology in 2005. In 2011, he received his PhD in Computer Science from Vienna Technical University with a thesis on advanced user interfaces for mobile urban exploration. Matthias has authored over 30 peer-reviewed scientific publications and has about 10 years of experience in the domain of mobile computing research and development. His research interests include pervasive interaction techniques, context-aware applications and mobile computer vision.

Peter Fröhlich is a senior researcher at FTW and leads the scientific competence area "User-centered Interaction and Communication Economics". His research interests include user experience evaluation, mobile spatial interaction, and automotive user interfaces. Peter holds a master's degree in Psychology from the University of Salzburg (2001) and a PhD in Applied Psychology from the University of Vienna (2007). He has authored more than 50 peer-reviewed scientific papers, and he is a regular organizer, editor and reviewer for renowned conferences and journals, such as the Journal of Personal and Ubiquitous Computing, Mobile HCI, Automotive UI, and CHI.

Jasmin Buchta is a research assistant at FTW where she is responsible for conducting user tests and evaluating and visualizing results. Jasmin pursues her master's degree in psychology at Karl Franzens University Graz. Her research interests include advanced user interfaces and the utilization and analysis of eye movement. She currently writes her master thesis with the title "Visual Perception of Fine Arts: Eye Movements".

Theresa Stürmer is a research assistant at FTW. She supports the design, carrying-out, and analysis of user tests. Theresa studies biology with a focus on anthropology at the University of Vienna. Her research interests include pervasive user interfaces and quality of experience for mobile services. Currently, Theresa writes her bachelor thesis in the field of human ethology.