Intl. Journal of Human–Computer Interaction, 30: 446–458, 2014

ISSN: 1044-7318 print / 1532-7590 online
DOI: 10.1080/10447318.2014.880142

Snap Target: Investigating an Assistance Technique for Mobile Magic Lens Interaction With Large Displays

Matthias Baldauf and Peter Fröhlich
FTW Telecommunications Research Center Vienna, Vienna, Austria

Modern handheld devices can act as "magic lenses" for public displays and enable camera-based real-time interaction with their contents, thus allowing for manifold interactive applications in public space. To learn more about the characteristics of common techniques and to provide guidelines for new ones, a comparative user study was conducted. With regard to two basic task types (selection and translation) and two typical devices (smartphone and tablet), three interaction techniques in-depth were evaluated: direct touch-based input (Touch), cross-hair targeting (Target), and a new technique based on a dynamic cross-hair on the screen that snaps to visible nearby screen objects (Snap Target). The study results indicate that successful interaction with the Touch technique strongly depends on the usage context: Although Touch enables fast selection, it incurs many errors when small targets have to be selected on a smartphone. Target supports such difficult selections better, but Snap Target proved to be most robust in both investigated task types. Also, users stated that they felt best supported by the latter technique. Cross-hair-based techniques, especially Snap Target, were found to be well suitable for scenarios where several device types need to be supported under similar conditions. Implications and further work are discussed.

1. INTRODUCTION

Over the last few years, we can observe a steadily increasing penetration of public displays in urban environments. According to a recent market analysis by 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. Thus, the deployed digital signage products including traditional flat screens, kiosk terminals, and entire media facades will furthermore shape the appearance of our urban surroundings and will form a new information and advertising medium.

Although interactive applications might enable entirely new use cases and business models, most of these appliances are so far restricted to noninteractive content, either due to missing input hardware (e.g., traditional flat screens and media facades) or due to their physical inaccessibility (e.g., mounted beyond shopping windows or at ceilings). Personal mobile devices have been identified as promising remote controls for enabling the interaction with interactive content for public displays, and several respective mobile interaction techniques have been introduced so far by academia and industry.

One intuitive and appealing recent interaction concept is markerless magic lens interaction through a touch-based smartphone as introduced by Baldauf, Fröhlich, and Reichl (2010) and Boring, Baur, Butz, Gustafson, and Baudisch (2010). In analogy to mobile Augmented Reality (AR) applications, the user may target her smartphone at the distant screen and observe this real-world scene captured by the built-in camera on the mobile display. Using optical recognition algorithms, strokes on the display can then be mapped to the corresponding positions on the remote screen and respective mouse actions can be triggered (Figure 1). To mitigate drawbacks such as inaccuracies due to device shake, temporarily freezing, and zooming, the current content of the mobile display have been proven to be beneficial advancements (Boring et al., 2010). However, the alternatives presented and investigated so far have several shortcomings: Both freezing and zooming the camera view hamper and delay truly spontaneous magic lens interaction with nearby displays, as additional explicit device interactions are required. Further, zooming results in the typical magnified camera view, which is more sensitive to unintended hand movement; freezing is feasible only for static display content as it decouples the capturing of the display from the actual interaction. Hence it is well-suited for grabbing content from the display for later usage but not feasible for an on-site multiplayer game with moving targets to be selected, for example. Both studied techniques thus limit the compelling experience of interacting with the environment through a mobile camera in real-time and constrain more sophisticated interactive applications.

To address these flaws, we revisited pointer techniques with roots in traditional graphical user interfaces. In the remainder of this article, we introduce two adapted alternative techniques for touch-based magic lens interaction with displays and investigate their performance with regard to the original direct *Touch* technique. The first alternative is the *Target* technique, which exploits a cross-hair symbol intended for a more precise aiming

---

Address correspondence to Matthias Baldauf, FTW Telecommunications Research Center, Vienna, Austria, Donau-City-Strasse 1, 1220 Vienna, Austria. E-mail: baldauf@ftw.at

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hihc.

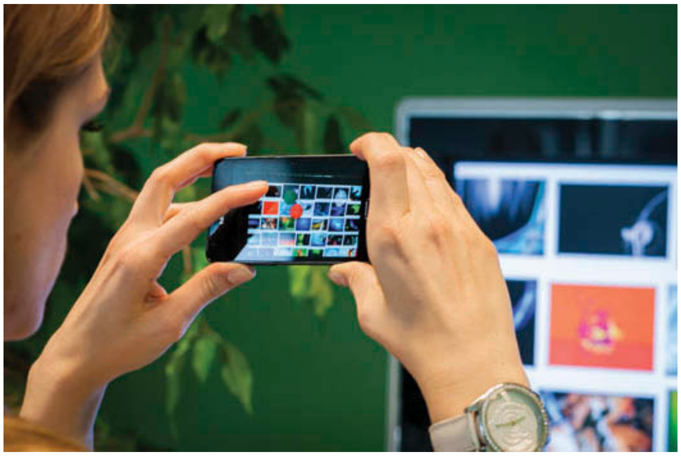

FIG. 1. A user performs a test task with the smartphone utilizing the magic lens interaction approach.

at objects-of-interest shown on the remote display. This cross-hair-based target approach is a common control technique in traditional AR applications with visual overlays but so far has not been considered for interaction with markerless remote displays. Inspired by target acquisition assistance techniques for mouse-controlled desktop applications, our second alternative, the *Snap Target* is a further development of the aforementioned target technique. It features a dynamic cross-hair, which aims at assisting various pointing tasks by snapping to nearby objects targeted on the distant display. Although related techniques have been proposed for graphical user interfaces on desktop systems, this article investigates the adapted mobile version in the field of magic lens interaction for the first time. For both cross-hair-based techniques, the item targeted through the cross-hair can be selected by touching the mobile display at any position.

To investigate the original and the newly proposed techniques, we conducted an in-depth comparative study. Our central research questions were which (dis)advantages the new techniques provide over the pure touch technique in terms of completion times, error rates, and overall user satisfaction with regard to fundamental tasks such as selection and translation as well as what impact a so far neglected factor, namely, the varying (display) size of mobile devices, has on the preferable interaction technique.

2. RELATED WORK

In this section, we give an overview of related work in the research fields that the present work builds on.

2.1. Mobile Augmented Reality

The term *augmented reality* refers to augmenting real-world scenes captured by a camera with superimposed virtual information. Applied on a smartphone, such a device can be used as a “see-through device” and act as a “magic lens” (Bier, Stone, Pier, Buxton, & DeRose, 1993) for targeting and identifying objects of interest. Respective sensor-based approaches utilizing GPS and compass have been available on mobile devices for several years in the form of real-world browsers such as Wikitude (http://www.wikitude.com). However, more accurate vision-based mobile AR solutions based on computer vision algorithms have hit the mass market only recently since AR researchers have made tremendous advances achieving interactive frame rates on off-the-shelf smartphones. Today, sophisticated engines such as Vuforia (http://www.qualcomm.com/solutions/augmented-reality) are publicly available, and human–computer interaction researchers have started to define usability priciples for traditional mobile AR applications (cf. Ko, Chang, & Ji, 2013).

Whereas directly touching superimposed icons on the smartphone display is the default selection technique in aforementioned real-world browsers, cross-hairs are also commonly applied for targeting purposes in mobile AR applications. Examples include early scientific work with head-mounted displays (Reitmayr & Schmalstieg, 2004), casual games for early mobile phones such as Mozzies (https://www.icg.tugraz.at/daniel/HistoryOfMobileAR), sophisticated shooter games like ARhrrrr! (http://ael.gatech.edu/lab/research/handheld-ar/arhrrrr), and more recent academic work on new mobile AR features (Sukan & Feiner, 2010). In the meanwhile, Rohs, Oulasvirta, and Suomalainen (2011) also validated prediction models for AR pointing based on Fitts’s law, an established model for estimating the time of a cursor pointing task.

Despite the increasing popularity of mobile AR applications, little research was done toward new advanced interaction techniques (cf. Billinghurst, Kato, & Myojin, 2009; Harviainen, Korkalo, & Woodward, 2009). Only a few comparative studies on respective interaction techniques are available. Henrysson, Billinghurst, and Ollila (2005) compared a cross-hair-based targeting technique with traditional button presses for positioning a virtual 3D object in a mobile AR application. The cross-hair technique turned out to be significantly faster. Participants further described it as very enjoyable but less precise than using buttons. More recently, Hürst and van Wezel (2011) explored touch and cross-hair for selecting and translating a virtual object in a hand-held AR application. In this study, touch outperformed the cross-hair with regard to completion time in object selection; however, the cross-hair worked better for translations over wide distance when the object was out of the original viewing window. Drawbacks of this study include the laggy sensor-based prototype implementation relying on accelerometer and compass data from the device (cf. related publicly available implementations such as Wikitude) instead of applying a more accurate state-of-the-art visual object recognition as well as the dwell time of 1.25 s for actually selecting the object behind the cross-hair (indicated by a progress bar) what is not suitable for advanced magic lens interaction with animated display content featuring moving elements. Further, the study considers only smartphones as client devices and provides no suggestions for improving these two traditional techniques.

2.2. Mobile Screen Interaction

Mobile devices that are wirelessly connected to large public displays have been identified as feature-rich enablers for screen interaction (Ballagas, Borchers, Rohs, & Sheridan, 2006). Common indirect approaches utilize various techniques to remotely control the mouse cursor shown on the large screen. RemoteCommander by Myers, Stiel, and Gargiulo (1998) first used a device touchscreen as a touchpad for shifting the remote mouse cursor. More recent research includes work by McCallum and Irani (2009), who proposed the combination of absolute and relative cursor positioning through a mobile phone's touchscreen and proved its effectiveness. Their ARC-Pad supports absolute jumps of the mouse cursor to corresponding screen locations by tapping as well as fine-grained relative movements by sliding the finger. Today, related smartphone applications are publicly available (e.g., Logitech's Touch Mouse; Logitech, 2010). Also the keys, the joystick, or built-in acceleration sensors of a device can be used for moving the remote cursor (cf. Boring, Jurmu, & Butz, 2009). Approaches relying on dedicated additional hardware include the detection of laser pointers, specialized 3D input devices, and the recognition of hand gestures (cf. Vogel & Balakrishnan, 2005) or of a mobilephone (Miyaoku, Higashino, & Tonomura, 2007) through external tracking systems. Mobile camera-based approaches that recognize screen content are of special relevance because they allow for direct absolute interaction with the remote system, which was proven to be superior over relative approaches in terms of task completion times (Baldauf, Fröhlich, Buchta, & Stürmer, 2013). Early work exploited visual markers for facilitating the camera-based detection of selecting screen objects (Ballagas, Rohs, & Sheridan, 2005; Jeon, Hwang, Kim, & Billinghurst, 2010; Pears, Jackson, & Olivier, 2009). Boring et al. (2010) introduced the idea of markerless magic lens interaction and presented a mobile prototype for touch interaction with multidisplay 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. Herbert, Pears, Olivier, and Jackson (2011) presented a related user study conducted with a basic nonmobile 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 frames per second. The presentation of a related mobile prototype by Baldauf et al. (2010), which touch-enables arbitrary display content using natural image features, did not report on a user study.

No further improvements for magic lens interaction with large displays have been investigated. Further, no studies on the respective impact of differently sized mobile devices have been conducted so far, leaving the suitability of alternatives for increasingly used tablet computers unclear.

2.3. Target Acquisition Assistance

Target acquisition is a central task in modern graphical user interfaces. For desktop systems, a number of pointing assistance techniques have been investigated over the last years to cope with an increasing number of control elements in complex user interfaces. Beneficial approaches include *Bubble Targets*, which dynamically extend their effective size as the mouse cursor approaches (Cockburn & Firth, 2003; McGuffin & Balakrishnan, 2002). Successful attempts improving the actual mouse cursor include a larger activation area instead of a single-pixel hotspot (Kabbash & Buxton, 1995) and *Bubble Cursors* (Grossman & Balakrishnan, 2005) dynamically resizing this range.

More suitable for our purposes, we consider snap controls (cf. Bier & Stone, 1986), which convey the experience of "gravity" or "magnetism" by snapping the mouse cursor to gravity-active elements, a common feature in illustrators and desktop publishing software. A related concept from graphical user interfaces are Sticky Icons (Worden, Walker, Bharat, & Hudson, 1997), which reduce the gain ratio of the mouse cursor when moved over them. In their most extreme extent, such techniques lead to the concept of Object Pointing (Guiard, Blanch, & Beaudouin-Lafon, 2004), a mouse cursor directly jumping between the targets. For pointing in 2D environments, object pointing can be beneficial. However, its performance is dependent on the target density.

For mobile finger-operated touchscreens, target acquisition assistance is even more relevant (cf. Lewis, Commarford, Kennedy, & Sadowski, 2008). Touching small targets with the relatively large finger while even occluding the actual target results is difficult (the "fat finger" problem). Early research in this field proved the higher accuracy of the Lift Off strategy (the last contact point on the touchscreen is used as selection point) over the Land On strategy (the first contact determines the selection point) for small targets (Sears, 1991; Sears & Shneiderman, 1989) and identified minimum target sizes (e.g., Douglas, Kirkpatrick, & MacKenzie, 1999). More recent work focused on the requirements of different user groups such as elderly people (cf. Hwangbo, Yoon, Jin, Han, & Ji, 2013) or investigated improvements of traditional selection methods for touchscreens. For example, Cross-Keys (Albinsson & Zhai, 2003) extend a traditional cross-hair with four direction buttons for precise placement through discrete button taps, Precision-Handles (Albinsson & Zhai, 2003) apply an additional on-screen handle for fine-grained continuous selection of small screen areas. Improved visualizations for target selection include the application of fish-eye views (Olwal & Feiner, 2003) and the Shift technique (Vogel & Baudisch, 2007), which shows screen content occluded by a touching finger in a callout and allows the fine-grained selection by a pointer through finger movements. In a similar vein, TapTap (Roudaut, Huot, & Lecolinet, 2008) allows for a more precise touch on a scaled-up callout shown after the first finger press. In Escape (Yatani, Partridge, Bern, & Newman, 2008), targets can be selected by gestures cued by icon positions and appearance. Although these techniques showed advantages over basic touching, they all require additional interactions, thus lengthen the duration of the operation and thus are not feasible for magic lens interaction with dynamic content.

In the context of magic lens interaction with large displays, where we envision (multiuser) applications without traditional mouse cursors, we see the snapping feature as a promising advancement for pointing through a cross-hair.

3. RESEARCH HYPOTHESES

Based on the literature review and own experiences with related AR applications and tested devices, we formulate the following research hypotheses. Hypotheses 1 address the differences between Touch and Target:

H1a: As fully direct interaction technique we expect Touch to outperform Target in terms of completion times.
H1b: We assume that Target has a lower error rate on the smartphone for small targets than Touch suffering from the fat finger problem.
H1c: We expect user satisfaction for Target for small targets through the smartphone target significantly higher.

Hypotheses 2 refer to the advantages of Snap Target over Target:

H2a: Due to the reduced distance, we expect Snap Target to outperform Target in terms of completion times.
H2b: We assume that Snap Target leads to fewer errors than Target due to experiences reported in related work on target acquisition.
H2c: We expect users to feel better supported by Snap Target than with Target.

Finally, Hypotheses 3 make assumptions on the impact of different mobile devices:

H3a: Due to the larger display, we expect the tablet to benefit more from the Touch technique.
H3b: Because smartphones are smaller and lighter than tablets, we assume that they are better suited for (Snap) Target requiring moving and aiming with the device.

4. METHOD

We designed a comparative user study to validate the aforementioned research questions. The study was conducted at FTW's user experience lab featuring recent observation equipment. We invited 30 users (17 male, 13 female) between 21 and 63 years of age ($M = 36.7$, $Mdn = 33$) from our test person database. As we knew from previous user tests, 20 of the participants owned a smartphone and/or a tablet. Concerning their experience with touchscreens, the participants stated a mean of 1.8 on a scale from 1 (*very skillful*) to 5 (*very awkward*). 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 received a voucher for a consumer electronics store as remuneration.

Each participant was asked to use the three different interaction techniques to solve two different tasks, both on a smartphone and a tablet (Table 1), while standing upright at a

TABLE 1
During Each Test, Each Interaction Technique Was Applied on Each Device to Solve Tasks on the Screen

Device	Technique	Task	Trials
Smart-phone	*	Game	*
	Touch	Selection	2 dist. × 2 sizes × 4 orient
		Translation	2 dist. × 2 sizes × 4 orient
	Target	Selection	2 dist. × 2 sizes × 4 orient
		Translation	2 dist. × 2 sizes × 4 orient
	Snap target	Selection	2 dist. × 2 sizes × 4 orient
		Translation	2 dist. × 2 sizes × 4 orient
Tablet	*	Game	*
	Touch	Selection	2 dist. × 2 sizes × 4 orient
		Translation	2 dist. × 2 sizes × 4 orient
	Target	Selection	2 dist. × 2 sizes × 4 orient
		Translation	2 dist. × 2 sizes × 4 orient
	Snap target	Selection	2 dist. × 2 sizes × 4 orient

Note. To avoid learning and preference effects, the orders of the techniques and the devices were systematically varied between the test users.

distance of 1 m in front of the large display, a typical scenario in a shopping street. Simulating this real-life situation, we did not specify the detailed device usage and let the subjects decide, for example, whether they prefer holding the smartphone with one or two hands. After having completed all trials of one task, we had participants rate the support of the technique on the current device for the solved task on a 5-point Likert scale. After having solved both tasks with a specific technique, participants answered a questionnaire assessing the overall characteristics of the technique and its suitability for usage on the respective device. In contrast to generic usability surveys, we designed the questionnaire according to the one proposed by Douglas et al. (1999), which explicitly aims at assessing devices and interaction techniques for remote pointing tasks. It includes questions concerning the mental and physical effort, the perceived accuracy and speed, the experienced fatigue of fingers and arms, and the difficulty and joy of usage, each on a 5-point Likert scale (the used questionnaire is available as the appendix). Finally, we asked the participants to rank the techniques according to their preference when they had used all three techniques with a specific device. A questionnaire regarding demographic data and mobile device usage was answered before the device change in the middle of the test to allow for a longer break for relaxing arms and fingers. This overall test procedure took about 75 min.

4.1. Setup

As public display we used a Philips Cineos flat-screen TV with a screen diagonal of 47 in. (119 cm). A desktop computer hosted both the large screen via HDMI as well a normal TFT monitor via VGA for the test manager. There, our custom study application showed a window for configuring the tests while it presented the chosen task in full-screen on the large display to the participant. This application acted as a server receiving simple input commands from the connected mobile devices via TCP Sockets to trigger mouse actions.

Our tested mobile devices included a Samsung Galaxy S2 smartphone and an Acer Iconia A501 tablet. The smartphone features a 4.3-in. capacitive touchscreen with a resolution of 800 × 480 pixels and weighs approximately 116 g. The touch display of the tablet is 10.1 in. with a resolution of 1280 × 800 pixels, and its weight is about 700 g. Both devices are powered by Android, allowing us to use the same mobile application for both the smartphone and the tablet. For realizing the visual detection and tracking of the screen content we make use of the aforementioned Vuforia toolkit, an advanced AR library optimized for mobile devices. When the mobile is targeted at the screen, the task background images shown on the large display are visually recognized and continuously tracked. Based on the calculated transformation matrix, the mobile display coordinates can be mapped to corresponding screen coordinates. These are then sent in form of a simple remote control protocol over TCP sockets via WiFi to the remote computer where our study application triggers the respective mouse actions. The application could be easily modified to initiate the corresponding application functions directly (without the simulation of mouse clicks) when receiving a respective remote control command from a mobile device. Thus, all techniques are also feasible for multiuser scenarios.

4.2. Techniques

The following magic lens interaction techniques were compared during the study.

Touch. As first interaction technique we chose the purely touch-based form of magic lens interaction. A user points the device toward the screen and uses her finger to touch controls on the remote screen viewed through the plain camera viewfinder (Figure 2a). Due to its higher accuracy proven in previous studies (Sears, 1991; Sears & Shneiderman, 1989), we implemented the *Lift off* strategy, that is, the last contact point with the touchscreen determines the selection position and triggers the action. For dragging operations, the user needs to touch the item of interest and either shift it on the display or move the entire device while keeping the finger pushed down.

Target. The Target technique tries to avoid the inherent occlusion of the actual target by the user's finger of the Touch technique. This technique applies a centered cross-hair superimposed over the plain camera viewfinder for targeting controls on the remote screen (Figure 2b). When the cross-hair center is moved over a control, it is highlighted to indicate a potential action. In analogy to the Touch technique, a selection is triggered by pressing and releasing the touch display, a translation action is performed by pressing the display, moving the device to position the cross-hair over the destination and releasing. However, the location of the touch on the mobile display is irrelevant because the trigger point is given through the cross-hair, allowing the user to use any comfortable display area for touching.

Snap target. The Snap technique is based on the Target technique featuring the equal highlighting. However, instead of its static cross-hair, it applies a "smart," dynamic cross-hair (Figure 2c). Following related support mechanisms for mouse usage on desktop systems, the snap technique exploits knowledge about potential targets: When the screen center (i.e., the initial position of the cross-hair) is brought closer than 50 pixels (measured in pixels of the large display to compensate for varying distances from the display and different pixel densities of mobile devices) to a target, this dynamic cross-hair moves over the center of this target (i.e., snaps to the target). In case of several targets within this range, the cross-hair snaps to the closest one. The range was defined during a series of pretests. Moving farther away from this target (i.e., out of its tolerance area), the cross-hair returns to its original centered position. Again, clicking and dragging is achieved by touching the mobile display at any position because the target is determined by the (this time perhaps corrected) position of the cross-hair.

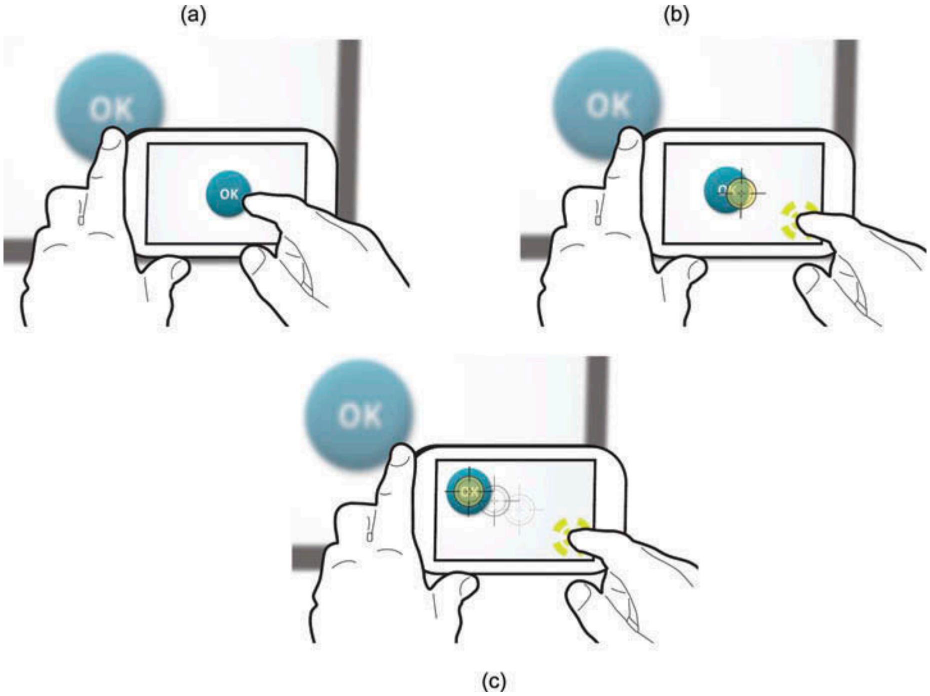

FIG. 2. We compared three interaction techniques: touching a visible object-of-interest (a), touching anywhere on the mobile display while targeting the object through a centered cross-hair (b), and a dynamic cross-hair which snaps to nearby objects (c).

4.3. Tasks

We asked the participants to use these three interaction techniques on a smartphone and a tablet to perform two traditional pointing tasks on the remote screen: selection and translation (cf. MacKenzie, Sellen, & Buxton, 1991). To get used to the different techniques and the current device, each participant started with a simple game. Each following task was preceded by a short training phase limited to 3 min where users could exercise until either they felt well-prepared for the actual test or the training period has elapsed. For the test situation, users were asked to complete the trials as fast and accurately as possible. Having completed the tasks with the first device, the test was repeated with the second device (Table 1). Overall, each subject completed 96 task trials. To avoid learning and preference effects, the orders of the techniques and the devices was systematically varied: We started with the conditions in Table 1 and created respective permutations of the techniques for the next five users, the following six users, then repeated these procedures but started with the tablet, and so forth. (Un)successful actions were indicated by audio signals. For each task trial we logged all input actions to calculate the completion time (including the time to correct errors) and the error rate. For increased ecological validity we chose a suitable background image for each task with strong features to ensure a robust visual recognition. We deliberately used tasks without moving elements, as these dynamics would represent another study variable and would increase the experiments complexity.

Game. At the beginning and when users switched from the smartphone to the tablet (or vice versa), participants could experiment with the new device using all three techniques for a simple shooter game depicted in Figure 3a. We deliberately integrated this informal “game task” to let the subjects experience the involved devices and interaction techniques in an enjoyable playful manner with no relevance for the evaluated study tasks. For example, participants used this phase to find convenient ways to hold and control the device depending on the technique. The game showed a space scenario with a flying saucer appearing at different positions and sizes with the goal to “shoot” (i.e., select) it. After each attempt the alien vanished and changed its location and size. When the alien was successfully hit, a game counter in the upper right corner was increased. The success of an action was indicated by a suitable audio signal. The test manager remotely switched to the next interaction technique when the participants agreed to have understood the current technique and felt comfortable with the device, and finally stopped the game and continued with the following tasks.

Selection task. The goal of the selection task was to hit a set of red circles as depicted in Figure 3b. Each trial consisted of 16 distinct targets consecutively displayed in randomized order. These were varied between small and large target sizes (radius of 40 and 80 pixels), small and large distances from the screen center (210 and 450 pixels), and four orientations with regard to the screen center (45°, 135°, 225°, 315°). To address the impact of target density, three gray fake targets were displayed

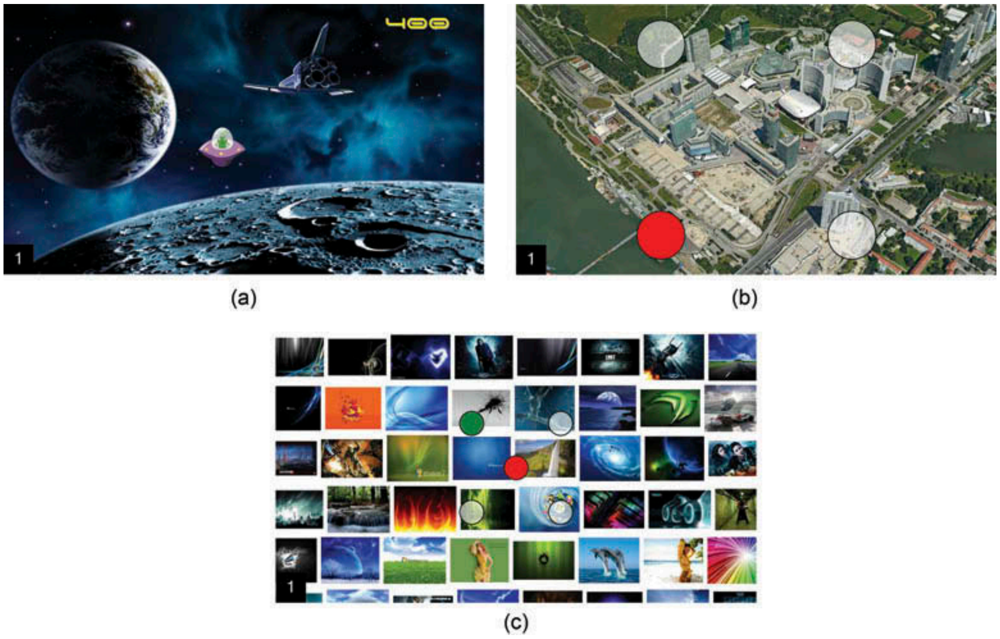

FIG. 3. Our study consisted of three task types: a simple game for getting used to the techniques and devices (a), a selection task where users had to pick the red circle (b), and a translation task where the red circle needed to be moved over the green circle (c).

in addition to the actual red target. These had the same size and distance from the screen center as the correct target and were arranged according to the remaining three orientations. For example, Figure 3b shows a trial configuration with large item sizes at large distance with the actual red target at the lower left (orientation of 135°). As background image we used a snapshot of a 3D environmental model to imitate the selection of a map part. Another typical selection scenario for a public display would be an interactive product presentation controlled through navigation buttons.

To avoid stressfully rushing through the targets, each trial needed to be triggered by the user by pushing a "start" button in the center of the screen. Time was measured from hitting this button until the successful selection of the appearing red circle. Missing the red circle was counted as an error.

Translation task. In the translation task, participants had to drag the red circle from the screen center and drop it over the green destination (Figure 3c). In analogy to the selection task we had a total of 16 trials (with the same variations of sizes, distances, and orientations). Also for the translation task, three fake targets were inserted to examine the impact of item density (Figure 3c). Given the distance of 1 m to the screen, both the red circle and the green destination were visible through the viewfinder of the smartphone and the tablet, that is, when placed appropriately, moving the device was not necessary during the Touch technique. The background mimicked a photo gallery with typical drag-and-drop functionality for resorting images.

We started data logging when the red item was successfully selected for the first time. Time was measured until the red item was moved over the green circle, that is, the center of the red item was inside the destination. In this task, we counted both missing the red circle and mistakenly dropping as an error.

5. RESULTS

For each task type and measure we ran a generalized linear model repeated measures analysis with SPSS to analyze main and interaction effects, as well as to derive pairwise differences (based on Bonferroni-adjusted confidence intervals). Error bars in the figures indicate 95% confidence intervals.

5.1. Task Completion Times

Selection. Figure 4 (top, left) summarizes overall completion times of the three investigated techniques for the selection task. On the top right, the results for the two devices, differentiated by target size ("Size") are shown. The figure is not additionally differentiated by target distance ("Distance"), as this factor did not have a notable impact on performance of the techniques (see the report of insignificant interaction effects of Distance × Technique next). Overall task completion time of the three techniques ranged between 2.02 s and 2.30 s, but no significant difference was obtained, $F(2, 56) = 0.17$, $p = .85$.

We also did not observe a significant main effect for device (tab vs. smartphone), $F(1, 28) = 0.56$, $p = .47$. However, large targets were selected significantly faster than small targets ($M = 2.56$ vs. 1.94), $F(1, 28) = 0.157$, $p < .001$. Also, with regard to distance, the selection of close targets was significantly faster than selection of distant targets ($M = 2.00$ vs. $M = 2.50$), $F(1, 28) = 16.927$, $p < .001$. We found a significant interaction effect

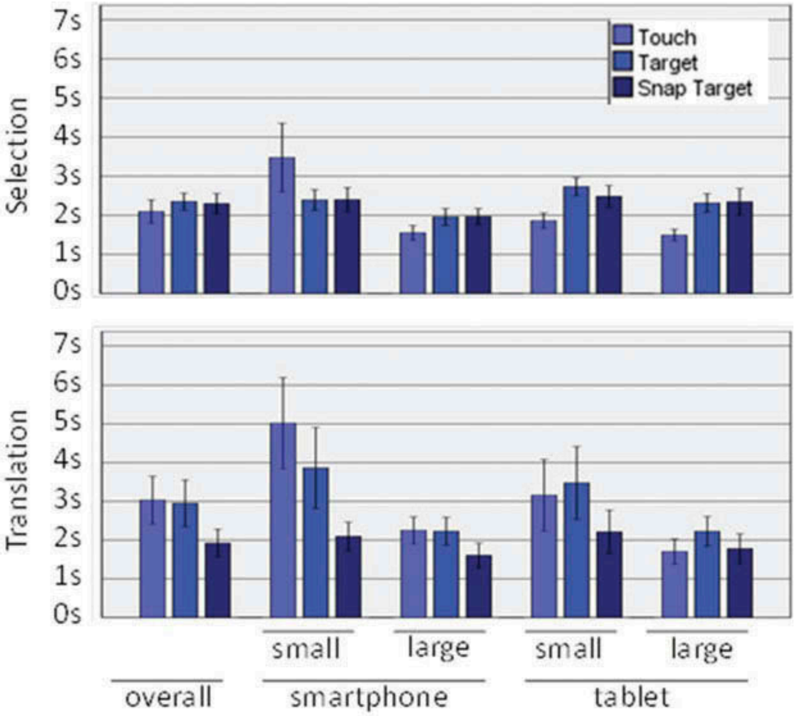

FIG. 4. Task completion time of the three techniques in seconds (top: selection task, bottom: translation task).

of size with technique, but we did not observe a significant interaction effect of distance with technique, $F(2, 56) = 0.44$, $p = .002$; $F(2, 56) = 0.44$, $p = .64$. Most notably, we obtained a significant interaction effect of technique, device, and size, $F(2, 56) = 4.72$, $p = .013$. Indeed, Figure 4 clearly indicates that the touch technique was always fastest (pairwise differences, $p < .05$), except when used on a smartphone with small targets.

Translation. Figure 4 (bottom) shows that due to its higher complexity, completing the translation task naturally took more than one second longer than the selection task. Also, more differentiated results for the translation task have been obtained, which are supported by significant main effects for all our investigated experimental factors technique, device, size, and distance, $F(2, 58) = 16.73$, $p < .001$; $F(1, 29) = 9.25$, $p = .005$; $F(1, 29) = 35.01$; $F(1, 29) = 49.96$, $p < .001$.

In contrast to the selection task, Touch as well as Target were significantly slower than Snap ($M = 2.9$s vs. $M = 1.9$s, $p < .001$). Furthermore, dragging took longer on the smartphone than on the tablet ($M = 2.82$s vs. 2.41s, $p < .01$). Smaller target sizes were significantly slower than larger target sizes ($M = 3.30$s vs. 1.94s, $p < .001$). Also translation toward distant targets was significantly slower than to close targets ($M = 3.16$ vs. $M = 2.10$, $p < .001$).

Again, we obtained a significant interaction of size with technique and no significant interaction of distance with technique, $F(2, 58) = 12.01$, $p < .001$; $F(2, 58) = 1.23$, $p < .30$. As with selections, we found a significant interaction effect of device, technique, and target size, $F(2, 58) = 3.51$, $p = .50$. We can clearly see in the figure that durations differed most strongly between techniques in case of small target sizes on the smartphone.

5.2. Errors

Selection. We found a significant main effect for our investigated experimental factors Technique, $F(2, 58) = 31.02$, $p < .001$; Device, $F(1, 29) = 22.73$, $p < .001$; Target Size, $F(1, 29) = 65.77$, $p < .001$; and Target Distance, $F(1, 29) = 14.19$, $p < .028$. Touch attained more errors than Target and Snap Target, respective pairwise differences were significant ($M = 3.591$ vs. $M = 0.84$ and $M = 0.90$, $p < .001$). Selections on the smartphone had on average many more errors than selections on the tablet ($M = 2.28$ vs. $M = 0.98$). However, as is indicated by Figure 5 (top), the dominating effect is the exceptionally high number of 10.3 errors when Touch is used with smartphones and small targets; each other condition has fewer than three errors. Indeed, a significant interaction effect of technique, target size, and device was found, $F(2, 58) = 137.5$, $p < .001$. We again did not find a significant interaction of distance and technique, $F(2, 58) = 3.16$, $p = .42$.

Translation. Also here, all investigated factors Technique, Device, Target Size, and Target Distance had significant main effects, $F(2, 58) = 19.81$, $p < .001$; $F(2, 58) = 11.263$, $p = .002$; $F(1, 29) = 57.64$, $p < .001$; $F(1, 29) = 10.67$, $p = .003$. Figure 5 (bottom) indicates that the pattern of the errors occurring during the translation task is similar to the selection results

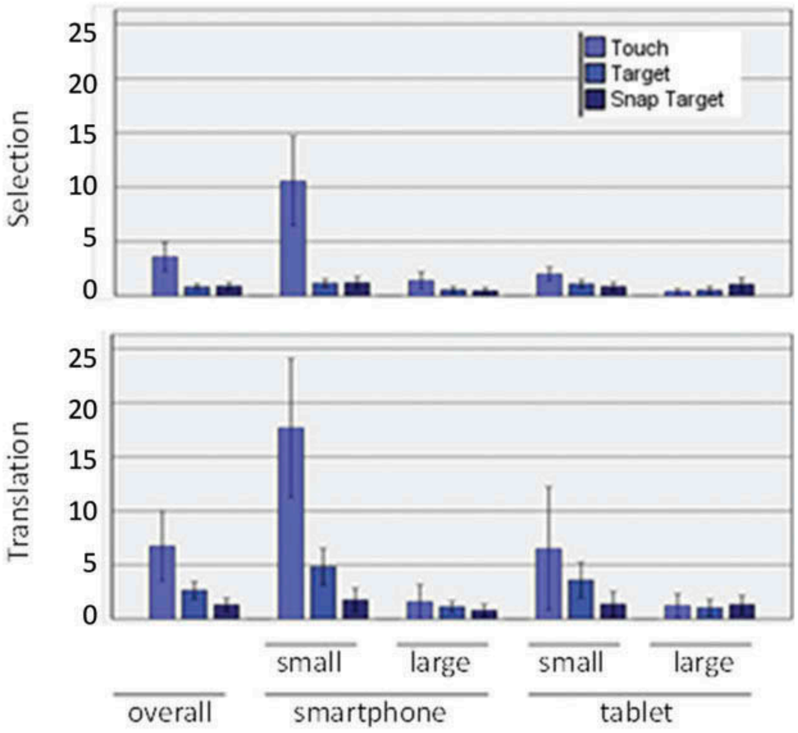

FIG. 5. Mean number of errors per trial (top: selection task, bottom: translation task).

just presented, but due to the higher complexity and duration of this task, the absolute number of errors was higher. Pairwise comparisons show that Touch also here had significantly more errors than the other two techniques ($M = 6.72$) but that in this task additionally Target had significantly more errors than Snap Target ($M = 2.62$ vs. $M = 1.41$, $p < .05$). Again, the most prominent effect is the interaction of Technique × Device × Target size, supporting the impression that by far the most errors occurred when dragging small targets with the touch technique on the smartphones.

5.3. Inquiry

Preference ranking. The participants' final mean ranking scores presented in Figure 6 (with values ranging from 1 to 3) show a similar preference profile for the three techniques throughout devices, target sizes, and tasks. Correspondingly, both in the selection as well as in the translation task a significant main effect for Technique was found, $F(2, 58) = 18.85$, $p < .001$; $F(2, 58) = 28.53$, $p < .001$, but not for Device and Size. Pairwise comparisons for both tasks show that Touch and Target received similar mean ranking scores ($M \sim 1.8$, no significant difference), and that Snap was preferred to the other two techniques ($M = 2.50$, $p < .05$).

Again, we see a significant interaction effect of Device, Technique, and Size, $F(2, 58) = 7.12$, $p = .002$. When looking at Figure 6, we can see that Touch was ranked similarly high as Target but that it was ranked significantly lower than Target when selecting small target sizes with the smartphone (pairwise difference, $p < .05$).

Perceived support ratings. The ratings about perceived support that had been given after having used the techniques for each task support the preference results: For both task types, Snap Target was rated highest, followed by Target and subsequently by Touch, $F(2, 58) = 9.06$, $p < .001$; $F(2, 58) = 33.48$, $p < .001$. All pairwise differences were significant, except that Target and Touch did not differ in selection tasks. However, in contrast to the very homogeneous results in the questionnaire and preference rankings, we found in both task types an interaction effect of Technique and Device, $F(2, 58) = 6.69$, $p = .002$; $F(2, 58) = 6.50$, $p = .003$. Similarly to the performance data just presented, Touch received relatively better scores with the tablet as with the smartphone.

Questionnaire. We found significant main effects of Technique for every scale of our questionnaire ($p < .01$). The results were consistent to the ranking results: Pairwise comparisons reveal that Snap Target was qualified as by far the most positive with regard to all scales. Touch and Target were close to each other, whereas Touch often received slightly worse ratings ($p < .05$; the difference was not significant for mental demand, physical demand, arm fatigue, difficulty, and enjoyment). With regard to the comparison of smartphone against tablet, the only significantly differing effects were observed on the physical demand and arm fatigue scales. Tablet usage was experienced as more physically demanding than the smartphone ($M = 2.41$ vs. 3.00), $F(1, 29) = 13.62$, $p < .001$, and as inducing more arm fatigue ($M = 3.17$ vs. $M = 2.60$), $F(1, 29) = 20.31$, $p < .001$.

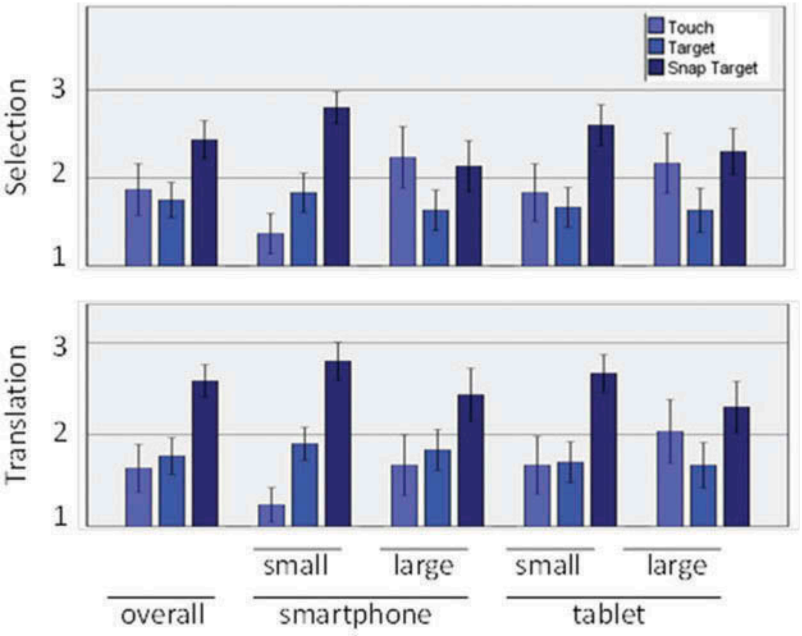

FIG. 6. Mean ranking scores for the three techniques (top: selection task, bottom: translation task).

Observations and participant comments. During or after usage, many participants commented that Touch was intuitive for them, and they felt that the opportunity to immediately press onto the target was comfortable. The major problem that was reported referred to situations in which smartphones had to be used for selecting small targets. Some mentioned that potentially touch registration is too accurate, giving too little tolerance to slightly misplaced touches. Two participants who had problems with the small targets were observed to use the mouse cursor on the large screen (our study prototype triggered mouse actions at the remote screen) as an aid to see the offset of their touch attempt. The tablet was found to be well suitable for selecting, because the hand does not fully cover the viewable screen.

Target was especially intriguing for younger participants, as it was seen to be innovative. The color feedback for selection was regarded useful; only in rare cases was it also qualified as distracting from the actual selection task. Participants with little touchscreen experience often pressed the screen very keenly and thereby inadvertently displaced the cross-hair. For the smartphone, the cross-hair was relatively large in comparison to the screen and thus subjects could partly have difficulties in seeing through the mobile phone and maintaining orientation and reference to the large screen. Tab appeared to be more difficult for the older participants, as compared to the smartphone technique. We made the interesting observation that a smaller portion of users were intuitively tempted to move the device not over the target element for acquisition but to the opposite direction.

Snap target was perceived as very useful, and automatic skipping toward the target was perceived more supportive than the color highlighting. Participants often used this technique very efficiently and did not appear to experience usability problems. However, after a few successful actions several participants committed more overshooting errors. With rising confidence and interaction speed, they tended to move the device too quick and so imprecisely that not even the tolerant Snap Target could compensate for the inaccuracy. Some test participants just used one hand, and they qualified the technique as more comfortable as it did not require major correction movements.

6. DISCUSSION

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

6.1. Touch versus Target

H1a. Our hypothesis concerning the shorter completion times of Touch is only partly fulfilled. For the selection, we found an interaction effect with the target sizes. The direct Touch technique significantly outperforms Target for large targets both on devices and for large sizes on the smartphone. Touch is only slower than Target for small target sizes on a smartphone. Whereas these appear larger on the tablet and can easily be touched, more time is required for precisely hitting them on a small display. This represents a major advantage for Touch on tablet-sized displays and is in line with basic touchscreen research suggesting minimum button sizes.

For the translation task we did not find significant differences between the completion times of Touch and Target, and thus we cannot confirm H1a for translation.

H1b. The hypothesis expecting a lower error rate for Target for small targets using a smartphone is confirmed. The found low error rate renders Target a superior alternative for precise pointing, both for selection and translation.

H1c. Finally, the study results confirm this hypothesis: User satisfaction for interacting with small targets through the smartphone with Target was significantly higher than with Touch. This can be explained by the aforementioned frustration when trying to touch smaller targets.

6.2. Target versus Snap Target

H2a. The assumption that Snap Target outperforms Target in terms of completion times is not fully supported by the study results. The durations are very similar; only for the translation task we could observe significant advantages for Snap Target on the smartphone. Although for simple selection tasks within the small scope of a remote display the "shortcut" taken by the Snap Target is not so gainful, this technique facilitates and thus speeds up the final placement during the translation task. This trend can be observed for both devices; however, it becomes significant for the smartphone. We consider Snap Target to be particularly useful for such small screen usage situations.

H2b. We expected Snap Target to lead to fewer errors than Target due to its "smart" aiming approach. This hypothesis can be confirmed only for the translation task, not for the simpler selection task. However, also here the gain by Snap Target compared to Target is much smaller than expected. Although the Snap Target had been designed as an assistance feature, it may obviously easily lead to wrong selections. As noted during our observations, this is especially true for applications with very dense content (in our study represented by the combination of large targets with short distances) as well as for elderly users being less dexterous in using a mobile AR application. Further, we observed technology-affine users relying on the Snap Target and speeding up after first successful operations, which in some cases led to errors due to a lack of attention.

H2c. This hypothesis is confirmed: Users felt significantly better supported by Snap Target than with Target during both tasks. In comparison, Touch was rated similarly to Snap Target only for large targets (however it was rated inferior for translating those using a smartphone).

6.3. Smartphone versus Tablet

H3a. The hypothesis that the tablet benefits more from the Touch technique is confirmed by our study. Especially for small target sizes, the completion times are significantly lower for both tasks. We attribute this result to the prototype implementation: For ensuring equal testing conditions for both devices we developed one application executable on both smartphones and tablets. For reasons of compatibility such mobile applications are scaled up for running full screen also on tablets. In the context of magic lens interaction, the resulting effect can be compared to a zoom feature (Boring et al., 2010), however, without the more sensitive viewfinder known from a traditional camera zoom function. For related touch-based interaction applications, we suggest to follow this approach, which has proven to be very beneficial. We consider the reduced video quality in the context of magic lens interaction with large displays acceptable, as this use case does not focus on high-quality visual overlays or similar.

H3b. We assumed that handling the smaller smartphone with the (Snap) Target technique would result in a better performance than using the heavier tablet—what is rejected by the results. The results for applying (Snap) Target on the smartphone are very similar to the ones utilizing the tablet. We ascribe this outcome to the specific attributes of magic lens interaction with displays through: The actual targets are located within a very small scope (i.e., the distant display), and, as observed during our lab study, only little device movement is required when using the cross-hair techniques. One could expect that such interdevice differences become relevant when targets are more widely distributed and thus extensive device movement is necessary (e.g., within real-world AR browsers).

7. CONCLUSIONS AND OUTLOOK

In this article, we presented a comparative user study assessing advanced techniques for magic lens interaction with public displays through mobile devices. In contrast to previous work, the three techniques—Touch, Target, and the newly investigated Snap Target—support remotely controlling dynamic applications. We evaluated each technique on a smartphone and a tablet with regard to the basic task types selection and translation.

In our comparison, Touch showed its strength for simple selection tasks with large targets. We therefore conclude that it is well suited for quick, spontaneous interaction with public displays, given that the content is especially designed to be shown in shopping windows at a reasonable size. As a typical use case example, we consider interactive advertisements with large presentations of products to be selected and simple casual games with large elements for multiuser scenarios.

The Target approach proved to perform robust also for smaller screen elements with low error rates. This attribute makes it a highly suitable candidate for interacting with content that has not been adapted for being shown on public displays or contains small elements in general. Examples include schedules and information tables at railway stations and airports as well as advanced games with lots of tiny moving objects. If besides selection tasks repositioning operations are required, Snap Target is beneficial and should be applied. For either cross-hair-based interaction technique, we recommend the device-based highlighting feature, which is appreciated by users. Further, we suggest simple visual feedback on the remote screen during the user's action on the mobile touchscreen: Using the Lift Off strategy on the mobile device, the current position determined on the distant screen can be indicated, for example, by a mouse cursor, the moment the user puts his finger on the mobile touchscreen.

Although performance of Touch seems to be more dependent on the mobile display size, we found out that targeting techniques are less influenced by this parameter. Based on the similar results for the devices, we conclude that (Snap) Target is a suitable technique for cases where smartphones and tablets should be supported and the same conditions need to be provided for both device classes.

In future work, limitations of this study, representing an initial effort to investigate assistance techniques for magic lens interaction, need to be addressed. Obviously, like its mouse cursor counterparts the evaluated Snap Target is dependent on the target density and the activation area (the snapping range). We therefore argue for exploring this relation and suitable values in more depth based on available previous research. As a complementary aspect we suggest to also investigate and assess concepts such as area cursors and self-adapting activation areas for novel mobile use cases such as magic lens interaction, but also traditional smartphone-based AR applications. Further, related concepts could be investigated for the Touch interaction with selected positions corrected transparently to the user according to the closest target. In the present study, we deliberately omitted such implicit assistance approaches without any visual feedback.

More general, additional research is needed to understand the fundamental applicability of Fitts's law and respective adapted models (cf. Rohs et al., 2011) for touch interactions on a remote display through a mobile magic lens considering new influencing variables such as the size of the mobile device and the distance to the remote screen.

We consider the proposed *Snap Target* technique just the tip of the iceberg. With the present article we would like to encourage more intensive research on novel advanced interaction techniques and potential assistance features for enabling the creation of more sophisticated but easy-to-use mobile AR applications in the future.

FUNDING

This work has been carried out within the project PRIAMUS financed by FFG and A1. The Competence Center FTW Forschungszentrum Telekommunikation Wien GmbH is funded within the program COMET Competence Centers for Excellent Technologies by BMVIT, BMWA, and the City of Vienna. The COMET program is managed by the FFG.

REFERENCES

ABI Research. (2011, May 31). Digital signage review to approach $4.5 billion in 2016. Retrieved from http://www.abiresearch.com/press/digital-signage-revenue-to-approach-45-billion-in-

Albinsson, P.-A., & Zhai, S. (2003). High precision touch screen interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 105–112.

Baldauf, M., Fröhlich, P., Buchta, J., & Stürmer, T. (2013). From touchpad to smart lens—A comparative study on smartphone interaction with public displays. *International Journal on Mobile Human–Computer Interaction*, *5*, 1–20.

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 8th International Conference on Pervasive Computing*.

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

Ballagas, R., Rohs, M., & Sheridan, J. G. (2005). Sweep and point and shoot: Phonecam-based interactions for large public displays. *Extended Abstracts of CHI*, 1200–1203.

Bier, E. A., & Stone, M. C. (1986). Snap-dragging. *Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques*, 233–240.

Bier, E. A., Stone, M. C., Pier, K., Buxton, W., & DeRose, T. D. (1993). Toolglass and magic lenses: The see-through interface. *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, 73–80.

Billinghurst, M., Kato, H., & Myojin, S. (2009). Advanced interaction techniques for augmented reality applications. *Proceedings of the International Conference on Virtual and Mixed Reality*, 13–22.

Boring, S., Baur, D., Butz, A., Gustafson, S., & Baudisch, P. (2010). Touch projector: Mobile interaction through video. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2287–2296.

Boring, S., Jurmu, M., & Butz, A. (2009). Scroll, tilt or move it: Using mobile phones to continuously control pointers on large public displays. *Proceedings of OZCHI*, 161–168.

Cockburn, A., & Firth, A. (2003). Improving the acquisition of small targets. In *Proceedings of the British HCI Conference*, 181–196.

Douglas, S. A., Kirkpatrick, A. E., & MacKenzie, I. S. (1999). Testing pointing device performance and user assessment with the ISO 9241, part 9 standard. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 215–222.

Grossman, T., & Balakrishnan, R. (2005). The bubble cursor: Enhancing target acquisition by dynamic resizing of the cursor's activation area. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 281–290.

Guiard, Y., Blanch, R., & Beaudouin-Lafon, M. (2004). Object pointing: A complement to bitmap pointing in GUIs. *Proceedings of the 2004 Graphics Interface Conference*, 9–16.

Harviainen, T., Korkalo, O., & Woodward, C. (2009). Camera-based interactions for augmented reality. In *Proceedings of the ACE*, 307–310.

Henrysson, A., Billinghurst, M., & Ollila, M. (2005). Virtual object manipulation using a mobile phone. *Proceedings of the 2005 International Conference on Augmented Tele-Existence*, 164–171.

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 1st International Conference on Pervasive and Embedded Computing and Communication Systems, 207–214.*

Hürst, W., & van Wezel, C. (2011). Multimodal interaction concepts for mobile augmented reality applications. In *Advances in multimedia modeling* (Lecture Notes in Computer Science Vol. 6524, pp. 157–167). Berlin, Germany: Springer.

Hwangbo, H., Yoon, S. H., Jin, B. S., Han, Y. S., & Ji, Y. G. (2013). A study of pointing performance of elderly users on smartphones. *International Journal of Human–Computer Interaction*, *29*, 604–618.

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

Kabbash, P., & Buxton, W. (1995). The "prince" technique: Fitts' law and selection using area cursors. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 273–279.

Ko, S. M., Chang, W. S., & Ji, Y. G. (2013). Usability principles for augmented reality applications in a smartphone environment. *International Journal of Human–Computer Interaction*, *29*, 501–515.

Lewis, J. R., Commarford, P. M., Kennedy, P. J., & Sadowski, W. J. (2008). Handheld electronic devices. In C. M. Carswell (Ed.), *Reviews of human factors and ergonomics* (pp. 105–148). Santa Monica, CA: Human Factors and Ergonomics Society.

Logitech. (2010, January 29). New: Logitech touch mouse turns your iPhone or iPod Touch into a wireless trackpad and keyboard [Web log]. Retrieved from http://blog.logitech.com/2010/01/29/

MacKenzie, I. S., Sellen, A., & Buxton, W. (1991). A comparison of input devices in elemental pointing and dragging tasks. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 161–166.

McCallum, D. C., & Irani, P. (2009). ARC-Pad: Absolute+relative cursor positioning for large displays with a mobile touchscreen. *Proceedings of the 22nd annual ACM Symposium on User Interface Software and Technology*, 153–156.

McGuffin, M., & Balakrishnan, R. (2002). Acquisition of expanding targets. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 57–64.

Miyaoku, K., Higashino, S., & Tonomura, Y. (2007). C-Blink: A hue-difference-based light signal marker via a cell phone color display. *International Journal of Human–Computer Interaction*, *23*, 143–161.

Myers, B. A., Stiel, H., & Gargiulo, R. (1998). Collaboration using multiple PDAs connected to a PC. *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work*, 285–294.

Olwal, A., & Feiner, S. (2003). Rubbing the fisheye: Precise touch-screen interaction with gestures and fisheye views. *Proceedings of the 16th annual ACM Symposium on User Interface Software and Technology*, 83–84.

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

Reitmayr, G., & Schmalstieg, D. (2004). Collaborative augmented reality for outdoor navigation and information browsing. *Proceedings of the Symposium Location Based Services and TeleCartography*, 31–41.

Rohs, M., Oulasvirta, A., & Suomalainen, T. (2011). Interaction with magic lenses: Real-world validation of a Fitts' law model. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2725–2728.

Roudaut, A., Huot, S., & Lecolinet, E. (2008). TapTap and MagStick: Improving one-handed target acquisition on small touch-screens. *Proceedings of the Working Conference on Advanced Visual Interfaces*, 146–153.

Sears, A. (1991). Improving touchscreen keyboards: Design issues and a comparison with other devices. *Interacting with Computers*, *3*, 253–269.

Sears, A., & Shneiderman, B. (1989). *High precision touchscreens: Design strategies and comparisons with a mouse* (Tech. rep. CS-TR-2268). College Park: University of Maryland.

Sukan, M., & Feiner, S. (2010). Snapar: Storing snapshots for quick viewpoint switching in hand-held augmented reality. *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*, 273–274.

Vogel, D., & Balakrishnan, R. (2005). Distant freehand pointing and clicking on very large, high resolution displays. In. *Proceedings of the 18th annual ACM Symposium on User Interface Software and Technology*, 33–42.

Vogel, D., & Baudisch, P. (2007). Shift: A technique for operating pen-based interfaces using touch. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 657–666.

Worden, A., Walker, N., Bharat, K., & Hudson, S. (1997). Making computers easier for older adults to use: area cursors and sticky icons. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 266–271.

Yatani, K., Partridge, K., Bern, M. W., & Newman, M. W. (2008). Escape: A target selection technique using visually-cued gestures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 285–294.

ABOUT THE AUTHOR

Matthias Baldauf is a senior researcher and project manager at FTW. Matthias received master's degrees and a Ph.D. in Computer Science from Vienna Technical University and University of Vienna. 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 research area "User-centered Interaction." Peter holds a master's degree in Psychology from University of Salzburg and a Ph.D. in Applied Psychology from University of Vienna. His research interests include UX evaluations, mobile spatial interaction, and automotive user interfaces.