

Display Pointing – A Qualitative Study on a Recent Screen Pairing Technique for Smartphones

Matthias Baldauf
Telecommunications Research Center FTW
Vienna, Austria
baldauf@ftw.at

Markus Salo
Department of Computer Science and Information Systems
University of Jyväskylä, Finland
markus.t.salo@jyu.fi

Stefan Suetter
Telecommunications Research Center FTW
Vienna, Austria
suetter@ftw.at

Peter Fröhlich
Telecommunications Research Center FTW
Vienna, Austria
froehlich@ftw.at

In this paper, we investigate display pointing as a recent mobile pairing technique for public screens enabling the simple targeting of the remote display with the mobile device. We evaluate a functional prototype in a user study emphasizing the crucial first connection phase independent from any consequent interaction style. In the presented initial study, we deliberately focus on the qualitative issues of this technique and identified a list of practical and fun aspects. According to the results, the display pointing approach is perceived as a (partly too) fast, robust and fascinating pairing technique which has the potential to gamify the connection process.

Mobile interaction; public displays; visual recognition; magic lens.

1. INTRODUCTION

With the on-going penetration of digital signage technology such as large public displays in urban surroundings, researchers and practitioners increasingly explore ways to exploit these installations and enable more compelling content with interactive features. Over the last few years, researchers have come up with lots of innovative interaction techniques how mobile devices can be used as remote controls for interactive applications (Ballagas et al. 2006).

Surprisingly, the majority of this work assumes that the address or identifier of the distant screen is known to the mobile device and thus the required wireless connection can be established transparently to the user. However, just this first step before the actual remote control of the distant display is a crucial factor in today's urban environments instrumented with lots of public displays and other networked objects. The ease and speed of setting up a data connection between the mobile device and the environment is considered a "key design consideration to create a low threshold and allow for highly serendipitous interactions" (Ballagas et al. 2006).

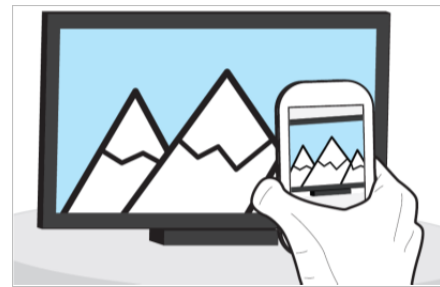


Figure 1: Selecting a display for further mobile interaction by pointing the smartphone at a screen.

This paper compares a mobile pointing gesture towards a display (Figure 1) to alternative mobile physical interaction techniques such as touching an NFC tag or capturing a QR code. In contrast to typical performance measures, we investigate how users perceive qualitative issues such as practical and fun aspects in this first study.

2. RELATED WORK

This section gives a brief overview on relevant previous work on qualitative evaluations in the field

of mobile interaction, camera-based interaction with displays as well as approaches for connecting mobile devices and displays.

2.1 Qualitative evaluation of mobile interaction techniques

Generally, numerous researchers have identified two fundamental aspects of user perceptions as antecedents of quality, value and satisfaction: utilitarian and hedonic aspects (cf. Hassenzahl and Tractinsky 2006; O'Brien 2010). In short, utilitarian aspects refer to traditional usability, practicality and instrumental quality, while hedonic aspects refer to enjoyment, fun, and aesthetics (Hassenzahl and Tractinsky 2006; O'Brien 2010). In respect of these, different pointing sub-techniques have been evaluated differently: laser pointers have strengths in utilitarian aspects such as simplicity and reliability, contradictory to capturing visual markers (Broll et al. 2009). However, hedonic aspects such as fun and novelty seem to be the strength of both laser pointers and visual markers (Rukzio et al. 2007). Overall, pointing techniques seem to have advantages over Bluetooth-based scanning techniques and manually entering pairing information, but users have not appreciated pointing – at least the older sub-techniques – as much as the NFC-based touching technique (cf. Broll et al. 2009; Rukzio et al. 2007; Mäkelä et al. 2007).

2.2 Camera-based display interaction

Early research work investigating the camera-based interplay of mobile devices and large displays exploited visual markers for facilitating the optical recognition of targeted screen areas (cf. Ballagas et al. 2005; Madhavapeddy et al. 2004). Boring et al. (2010) introduced the idea of markerless smart lens interaction and presented a 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. The presentation of a related fully functional prototype which touch-enables arbitrary display content using natural image features did not report on a user study (Baldauf et al. 2010).

While these examples utilize camera-based approaches for special interaction styles and for dedicated tasks on a remote display, none of the previous work focused on the investigation of a generic technique for setting up the required data connection.



Figure 2: A test person trying to establish a connection to one of the study screens by display pointing.

2.3 Connecting to public displays

As already mentioned before, techniques for establishing a data connection to the screen (i.e. the hosting computer) are a rather neglected aspect in previous HCI research. While most of the introduced camera-based interaction approaches include implicit connection mechanisms, respective research proposing and investigating dedicated novel connection techniques is scarce.

In early work, visual markers were not only used to identify the screen area a mobile device was targeted at but also to encode the computer's address (Ballagas et al. 2005, Madhavapeddy et al., 2004). Other researchers propose location-based connection techniques, e.g. the mobile device connects to a screen computer when entering the room (Dachsel and Buchholz, 2009). However, such approaches are not feasible for densely instrumented multi-display environments.

Scan-based approaches e.g. based on Bluetooth and descriptive device names (Baldauf et al. 2010) could solve this issue but are time-consuming and thus might be cumbersome for the user. Techniques for visual markerless screen recognition are considered in the aforementioned work of Boring et al. (2010) for realizing the smart lens-based interaction approach. However, none of these researchers studied the user-related characteristics of the actual connection technique



a) Display pointing



b) QR code



c) NFC tag

Figure 3: In our user study we explored the characteristics of display pointing (a) by relating it to two traditional connection techniques, capturing a QR code (b) and touching an NFC tag (c).

independent from any task-related interaction with the remote display.

3. DISPLAY POINTING

In the following, we describe the recent display pointing technique and give insights about our prototypical implementation.

3.1 Interaction technique

Typically, a user interested in one of the interactive features presented by a nearby public display would open the respective application on her smartphone and then point her device towards the respective screen. As soon as the display (or its content, respectively) is recognized, the mobile application switches to the actual control mode. A wide range of such remote control techniques for smartphones have been investigated (cf. Ballagas et al. 2006). Typical examples with publicly available implementations include indirect interaction techniques such as controlling the remote mouse cursor by tilting gestures or display strokes and direct interaction techniques with graphical representations of the large display's content on the mobile which forwards touches on interactive elements to the public display.

Dependent on the type of interactive application, the user can decide to manually return to connection mode by pushing a back button and start connecting to another display. Alternatively, the mobile application can switch back to connection mode automatically after the interactive task (such as participating in a contest) has been completed.

3.2 Prototypical implementation

We implemented a functional mobile prototype to study the concept of display pointing. Typically, the detection of pointing gestures and the identification of their actual target are implemented following either a sensor-based approach using a built-in

GPS receiver, a compass, accelerometers and/or a gyroscope or a camera-based approach utilizing optical recognition methods. While sensor-based approaches are highly sensitive to the precision of the localization system and thus work well for identifying large outdoor objects such as buildings through a pointing gesture, their accuracy is not sufficient for interacting with smaller or even mobile objects. Thus, we decided on a camera-based implementation for our prototype.

To reduce prototyping efforts, we made use of *Qualcomm's Vuforia* toolkit, an advanced mobile augmented reality library offering the abovementioned operations optimized for mobile devices. For realizing our display pointing prototype, we only made use of the recognition feature. As reference targets we selected a couple of photos with strong features that can be used as background images for respective applications run on public displays. Each of the chosen background images is associated with a unique identifier for a specific screen. The resulting mobile application features the immediate recognition of these images in continuous camera mode without the need for the user to explicitly take a photo or press a button in any other way.

4. EVALUATION

To gain first user feedback on the display pointing technique, we designed and conducted a user study. We invited 32 participants (13 females) aged between 23 and 66 (mean=36.8 and median=31.5). All but two elderly test persons had previous experience with smartphones. For them, we gave short introductions to get familiar with basic touchscreen interaction.

4.1 Setup

We prepared two flat television sets and a *Nexus S* smartphone for our study. We aimed at imitating a typical scenario with shopping windows, thus the two displays were facing the same direction, while

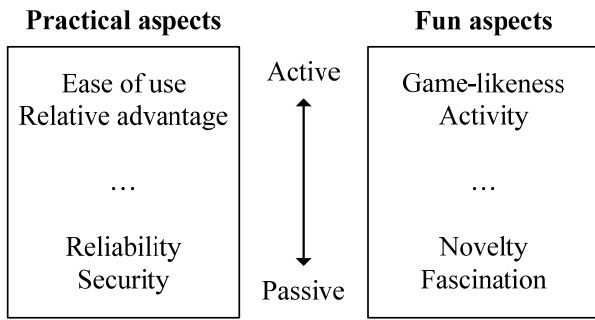


Figure 4: Subcategories of practical and fun aspects on an active-passive continuum.

their distance was about three meters (one screen of the setup is depicted Figure 2). Both were connected to one desktop computer, which executed the backend part of our study application. This program showed exemplary public display applications with the aforementioned background images on each of the two screens: one was a contest of a travel agency, the other one a participation tool for city residents (see Figure 3a).

To facilitate the comparison with established approaches and thus to emphasize the characteristics of the display pointing approach, we extended the described mobile application and integrated support for capturing a QR code and scanning an NFC tag (Figures 3a and 3b). For any of the three mobile interaction techniques, the screen identifier was sent to the backend application via Wifi as soon as a correct screen identifier was detected. Such a “successful” connection was communicated by a visual signal on the respective large screen (the participation screen showed a poll, the travel agency screen a quiz question) and a short vibration of the smartphone.

4.2 Method

The participants used each of the three techniques to establish a connection to the screens. We asked them to switch to the other screen when a connection attempt was successful, i.e. to alternately use the two screens. The order of the techniques was systematically varied to avoid learning or preference effects. We did not restrict the test time but let the users freely experiment with the techniques. Having tested all techniques, we asked each participant for their opinion concerning display pointing.

To explore practical and fun aspects in a relatively new context, we chose to conduct semi-structured qualitative interviews. Since the terminology of utilitarian and hedonic aspects and many of their proposed subcategories such as efficiency, excellence, play and aesthetics (Holbrook 1996),

are not always obvious and similarly understood, we wanted to use ordinary wording in the interview questions. Thus, we asked the participants to tell about practical and fun aspects of display recognition. Additionally, the respondents rated the importance of such aspects, as suggested by Hassenzahl and Roto (2007). We then conducted a content analysis on the transcripts and followed one of the suggested coding (Hassenzahl and Roto 2007) including initial categories from prior studies and added new categories on the grounds of the content when needed. Our aim was to form a categorization that would include all aspects represented in the data and position the subcategories on an active-passive continuum after their characteristics (Holbrook 1996).

5. RESULTS

Each of the 32 participants was able to identify both practical and fun aspects of display pointing with several issues. We organized these statements at first in the two main categories (practical and fun), and then formed four subcategories for both of them. The respondents ranked the importance of the mentioned practical aspects with an average of 4.16 (1=not important at all, 5=extremely important), and similarly fun aspects with an average of 3.13. Only three respondents ranked fun aspects for higher importance than practical aspects, while all the others ranked practical aspects equally or more important than fun aspects. Figure 4 depicts the placement of the subcategories on an active-passive continuum after their characteristics. Each subcategory is described with illustrative quotations as follows.

5.1 Practical aspects

Ease of use

Most of the respondents mentioned that display pointing is easy to use, because of being “simple, straightforward”, “really fast and effortless” and there is “no input required”. The respondents highlight especially the speed, which derives from simplicity. Also, the technique is seen as an intuitive way to connect with the displays. Only one respondent commented negatively on ease of use by concluding that display recognition “seems to make things more complicated”.

Relative advantage

Display recognition has a “significant advantage over the other techniques”, and the respondents see it as an efficient way to connect with displays. Respondents clarified that with display recognition, “it is possible to work from distance” and “each angle is possible in comparison to [other techniques]”. Some respondents stated that the practicality results also from low requirements of

attention and concentration: “less focused on smartphone display, less concentration necessary”.

Reliability

Some respondents appreciate the reliability and lack of errors in connecting with the displays from various angles and distances. Many respondents did not refer to reliability at all, and actually, there were few respondents who perceived “[other techniques] more reliable”. For example, “if two displays are located close to each other, one could select the wrong display by mistake”. Additionally, there are contextual factors that affect reliability. For example, “there could be problems with the sun or [using it] in darkness”.

Security

A share of the respondents commented on security-related issues. For a positive note, a respondent described that there are “no risks for the smartphone”. However, respondents felt the display recognition was “[even] too fast and too simple” to enable a conscious and confirmed connection. On another similar note, one respondent felt insecure if “seeing something and logging in happens immediately – it’s awful”. One respondent took such thoughts a bit further: “Scary, maybe also Facebook profiles could be searched in a similar way”. However in general, perceived benefits of display recognition seemed to outweigh such perceived risks.

5.2 Fun aspects

Game-likeness

A part of the respondents perceived the technique as a game-like, highly active, and playful way to connect with displays. For example, one respondent told “it reminds me of playing a computer game”. Display recognition has the possibility to gamify the activity of connecting with the displays, since it can “make [the activity of connecting] playful”. It may also enable other game-like characteristics, since it may e.g. lead to a “sense of achievement” or rouse “hunting instinct in finding out where information is hidden”.

Activity

Respondents also considered the technique fun because of its active nature without any references to gaming or playfulness. This category is perhaps best described by the following quotation: “the fun factor [is there]: I actively do something and then something happens”. A few respondents referred to enjoyment of taking photos: “it’s fun to take photos, establishing a connection itself rather isn’t fun” and “it is fun, since there are always new motives, like taking a photo”.

Novelty

Some of the respondents passively appreciated the novelty and innovativeness of the technique. It was

considered fun “maybe because it’s the most advanced technique for me” and “because it’s something new, gives pleasure – don’t know whether this is still the case in several years”. As the latter part of the previous comment reveals, indeed, the novelty value might not last too long, since some respondents felt the technique “will become a routine quite fast”.

Fascination

The display recognition technique “fascinated” and “amazed” a few respondents. Such fascination may be at its strongest level during the first impressions of the technique. One of the respondents described the technique even magical: “I don’t know how it works, it’s somehow magical, wow. [It has the] lowest technological impression, and it arouses curiosity”. Similar to the novelty value, some respondents expect fascination to decrease over time.

6. DISCUSSION

For the practical aspects, the ease of use was the primary category for our participants’ comments emphasizing the simplicity and speed of display pointing. The statements highlight the importance of these factors for setting up connections with public displays and make display pointing in its present form a suitable candidate for enabling the targeted “*serendipitous interactions*”. As a potential limitation, it can be stated that recognition times could increase with a growing number of displays. However, we assume that smart location-based prefetching mechanisms and improved image recognition algorithms enable similarly quick response times also for multiple displays.

Another crucial element turned out to be the perceived robustness of a connection technique: With the evaluated version of display pointing, participants appreciated the technique’s robustness with regard to different distances and viewing angles. While it is no problem to use the technique in darkness, the comment concerning reflections hindering the correct recognition of the display is correct: this is an obvious limitation of the camera-based pointing implementation.

As a side effect, the very fast and robust detection resulted in a lack of control for some participants, since it may easily lead to wrong selections, especially in case of environments densely equipped with displays. As potential improvements we propose either the manual confirmation of a detected display, e.g. by pushing a newly appeared “connect” button, or a short dwell time when pointing towards the screen, e.g. indicated by a progress bar once the screen is detected. However, to not reduce the speed advantage for

experienced users who do not have any problems with the quick response times, such assistant confirmation features should be optional and deactivatable.

While we had expected that the importance of the practical aspects would be rated higher than the one of the fun aspects, the still rather high rating for the fun aspects was quite surprising. One could assume that fun aspects would play a negligible role in this context of screen connection, what was not the case. In concrete terms, display pointing seems to have the possibility to gamify the rather normal activity of connecting to a display without any explicitly created game elements. Thus, application developers could take advantage of such implicit game-likeness instead of promoting pure novelty and fascination of the technique, which may decrease rather quickly over time, as described by several people.

7. CONCLUSIONS AND OUTLOOK

In this paper, we addressed the challenge of enabling serendipitous interactions with public displays through mobile devices and thus lowering the threshold for using offered interactive applications.

This initial user study on display pointing as a pairing technique focused on qualitative aspects instead of pure performance measures and helped to identify several beneficial aspects of the technique. Hence, the study also contributes to the evaluation of physical mobile interactions, since we managed to reach decent qualitative explanations on why display pointing is (or is not) considered practical or fun. The users considered display pointing remarkably practical due to various reasons. Concerning the fun aspects, which were ranked considerably important among the users, we did explore some previously uncovered issues such as the game-like aspect.

While the presented initial lab study yielded highly promising results for display pointing, future studies are required to explore further characteristics under real-world conditions. Interesting research aspects include whether display pointing leads to more connections by passers-by than alternative connection techniques due to its game-like nature and its applicability and performance at crowded public places.

8. ACKNOWLEDGMENTS

This work has been carried out within the projects *b-Part* and *PRIAMUS* financed by FFG and A1. Further funding was received from Nokia Foundation, the Research and Training Foundation

of TeliaSonera Finland Oyj, and HPY Research Foundation. The Competence Center FTW Forschungszentrum Telekommunikation Wien GmbH is funded within the program COMET by BMVIT, BMWA, and the City of Vienna. The COMET program is managed by the FFG.

9. REFERENCES

- Baldauf, M., Fröhlich, P., and Reichl, P. (2010) Touching the untouchables – vision-based real-time interaction with public displays through mobile touchscreen devices. *Pervasive'10*. ACM.
- Ballagas, R., Rohs, M., and Sheridan, J.G. (2005) Sweep and point and shoot: phonecam-based interactions for large public displays. *CHI '05*, 1200-1203. ACM.
- Ballagas, R., Borchers, J., Rohs, M., and Sheridan, J.G. (2006) The smart phone: a ubiquitous input device. *Pervasive Computing*, vol.5, no.1, 70-77. IEEE.
- Boring, S., Baur, D., Butz, A., Gustafson, S., and Baudisch, P. (2010). Touch projector: mobile interaction through video. *CHI '10*, 2287-2296. ACM.
- Broll, G., Rukzio, E., Paolucci, M., Wagner, M., Schmidt, A., and Hussmann, H. (2009) Perci: pervasive service interaction with the internet of things. *Internet Computing*, 13, 6, 74-81. IEEE.
- Dachselt R. and Buchholz, R. (2009) Natural throw and tilt interaction between mobile phones and distant displays. *CHI '09*, 3253-3258. ACM.
- Hassenzahl, M., and Roto, V. (2007) Being and doing: a perspective on user experience and its measurement. *Interfaces*, 72, 10-12.
- Hassenzahl, M. AND Tractinsky, N. (2006) User experience – a research agenda. *Behaviour & Information Technology*, 25, 2, 91-97.
- Holbrook, M.B. (1996) Customer value: a framework for analysis and research. *Advances in Consumer Research*, 23, 138-142.
- Madhavapeddy, A., Scott, D., Sharp, R., and Upton, E. Using camera-phones to enhance human-computer interaction. *UbiComp'04*.
- O'Brien, H.L. (2010) The influence of hedonic and utilitarian motivations on user engagement: The case of online shopping experiences. *Interacting with Computers*, 22, 5, 344-352.
- Mäkelä, K., Belt, S., Greenblatt, D., and Häkkinen, J. (2007) Mobile interaction with visual and RFID tags: a field study on user perceptions. *CHI '07*, 991-994, ACM.
- Reischach von, F., Michahelles, F., Guinard, D., Adelman, R., Fleisch, E., and Schmidt, A. (2009) An evaluation of product identification techniques for mobile phones. *Interact'09*, 804-816.
- Rukzio, E., Broll, G., Leichtenstern, K. and Schmidt, A. (2007) Mobile interaction with the real world: an evaluation and comparison of physical mobile interaction techniques. *AmI'07*, 1-18.