Pocket Transfers: Interaction Techniques for Transferring Content from Situated Displays to Mobile Devices

Ville Mäkelä[1,2], Mohamed Khamis[2], Lukas Mecke[2,3], Jobin James[1], Markku Turunen[1], Florian Alt[2,3]

[1]University of Tampere, Finland
[2]LMU Munich, Germany
[3]Munich University of Applied Sciences, Germany

{ville.mi.makela, jobin.james, markku.turunen}@sis.uta.fi,
{mohamed.khamis, lukas.mecke, florian.alt}@ifi.lmu.de

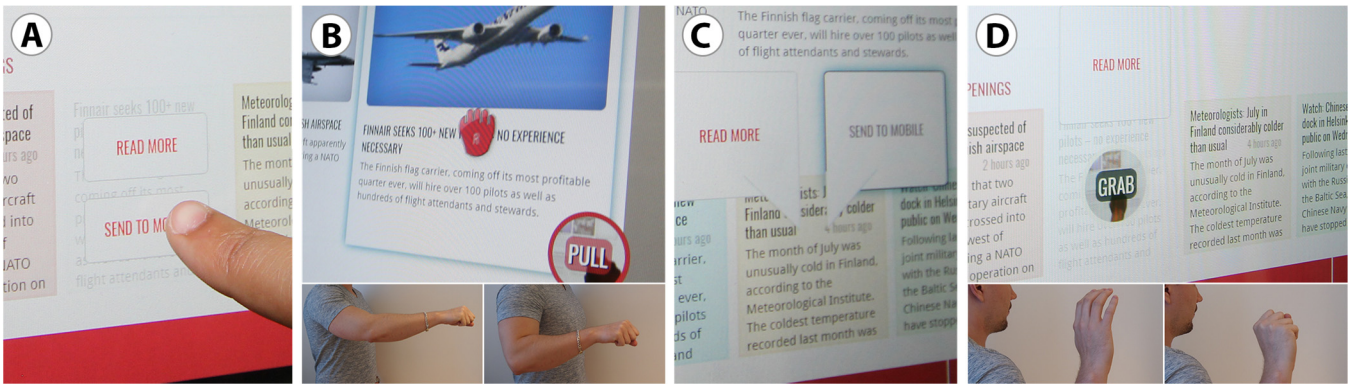

Figure 1. Pocket transfer techniques allow transferring content from a situated display to a personal mobile device that remains in a pocket. A) Touch: tapping an item opens a menu, tapping "Send to mobile" transfers the item. B) Mid-air gestures: pointing at an item, grabbing it, and pulling towards the user transfers the item. C) Gaze: gazing at an item for 1 second opens a menu, and gazing at "Send to mobile" for 1 second transfers the item. D) Multimodal: looking at an item and grabbing in mid-air transfers the item.

ABSTRACT

We present Pocket Transfers: interaction techniques that allow users to transfer content from situated displays to a personal mobile device while keeping the device in a pocket or bag. Existing content transfer solutions require direct manipulation of the mobile device, making interaction slower and less flexible. Our introduced techniques employ touch, mid-air gestures, gaze, and a multimodal combination of gaze and mid-air gestures. We evaluated the techniques in a novel user study (N=20), where we considered dynamic scenarios where the user approaches the display, completes the task, and leaves. We show that all pocket transfer techniques are fast and seen as highly convenient. Mid-air gestures are the most efficient touchless method for transferring a single item, while the multimodal method is the fastest touchless method when multiple items are transferred. We provide guidelines to help researchers and practitioners choose the most suitable content transfer techniques for their systems.

CHI 2018, April 21–26, 2018, Montréal, QC, Canada.

ACM ISBN 978-1-4503-5620-6/18/04...$15.00.
DOI: http://dx.doi.org/10.1145/3173574.3173709

Author Keywords

Public displays; content transfer; cross-device interaction; mid-air gestures; gaze; multimodal; ubiquitous computing.

ACM Classification Keywords

H.5.2 Information interfaces and presentation (e.g., HCI): User Interfaces – *Input devices and strategies*

INTRODUCTION

There are many situations in which users want to transfer content from public displays to personal mobile devices. For example, passersby in a hurry might want to grab a news article from a display to read on their smartphones later; or a user might want to know more about an advertised product but do so in private rather than in front of the display; or a tourist might want to transfer a real-time map of a city's public transportation system at the airport.

We envision that as an increasing number of situated displays appear in urban areas, more opportunities for transferring content to personal mobile devices for later consumption will arise. However, the vast majority of existing methods require users to look at and hold their mobile device in their hands to transfer content to it. This may be undesired in situations where passersby carry other items (coffee, suitcases, etc.), where they first need to take the smartphone out of their pocket, or where a transfer is only a side task (for example, as users skim through multiple news articles, they want to occasionally transfer one to their mobile device) and thus not interrupt browsing.

As a solution, we propose *Pocket Transfers*, an approach where (1) the mobile device can remain in the user's pocket or bag throughout the interaction, and where (2) a set of different interaction techniques distributed across several modalities are supported.

The mobile device can remain in the user's pocket by utilizing a mobile application and a location-tracking solution [17]: people in the space are automatically paired with their mobile device based on location data. Hence, users can interact with a display without touching their mobile device, and the system will know which device to send the content to. Although this feature was proposed in previous work [17], only mid-air gestures for transferring content in this way were evaluated. Therefore, it is unclear how different techniques and modalities using this approach fare in comparison. We believe different modalities are preferred in different settings, as one must cater to the current situation and type of content, the user's privacy needs, and the amount of content the user wants to take away. Furthermore, it is unclear how such techniques fare against a baseline condition, such as QR code scanning, where users need to take the phone out of their pocket to transfer content.

In this work, we introduce several novel content transfer techniques that allow users to keep the mobile device in their pocket during the transfer (Figure 1). Supported interactions include touch, mid-air gestures, gaze, and a multimodal technique combining mid-air gestures and gaze. We also added support for transferring content with QR codes. QR codes require manipulation of the mobile device, and due to their familiarity and ease of use, they work as a suitable baseline.

We conducted a user study in which 20 participants experienced and evaluated all five techniques. We used a novel approach in our study; rather than only considering scenarios where the user is already at the display, participants completed tasks that covered the full interaction process, including walking to and from the display, as well as possible preparations for the interaction. In addition, we included two task types, to accommodate for both short and long interaction sessions. This way, we reached more ecologically valid findings, allowing fair and truthful comparison between the techniques and modalities, as we also factor in the so-called *hidden costs* for interaction, unlike most existing studies.

Our research is driven by the following questions:

- What is the performance and user experience of pocket transfer techniques? What are the positive and negative aspects of each technique?
- How useful is it to keep the mobile device in a pocket with each technique?
- Are different techniques preferred based on the length of the interaction, or the presence of other people?

[1] We use a capital letter to distinguish our techniques from modalities: Touch, Mid-air gestures, Gaze, and Multimodal.

Our primary novel findings are a) all pocket transfer techniques are fast, and b) users highly appreciate being able to keep the recipient device in their pocket regardless of modality. Touch[1] and Mid-air gestures are the fastest techniques for transferring a single content item, and all techniques are seen as suitable for single-item scenarios. Touch and Multimodal are the fastest techniques when transferring multiple items, and are also the most favored. All pocket transfer techniques are acceptable when no other people are around; however, Gaze is the most favored when others are present.

Our contribution in this paper is twofold. First, we present the design and evaluation of four pocket transfer techniques. We show that all of them are fast and convenient, and present strengths and weaknesses for each technique as well as guidelines to help researchers and practitioners decide which modalities to use in their content transfer systems. Second, we contribute a novel user study design, wherein we factor in the preparation for, and halting of, the interaction. We argue this approach results in higher ecological validity, and we encourage researchers to utilize a similar approach in future studies.

RELATED WORK

Ng et al. [22] present a survey on screen-smart device interaction (SSI), which also covers methods for content transfer. Two general method types are recognized: vision-based and radio-based. Of vision-based methods, QR codes have been used actively [1,10,22], which allow smartphone users to scan a code using the device-integrated camera, to receive content such as a link to a website. Of radio-based methods, near-field communication (NFC) technology has been utilized for content transfer [2,7,22,26]. For instance, Hardy and Rukzio [7] attached individual NFC tags behind each content item, thereby allowing items to be transferred by touching the corresponding item with a mobile device. Broll et al. [2] used a somewhat similar method for more advanced interactions, such as dragging-and-dropping, by allowing users to select actions on the mobile.

Langner et al. [13] presented techniques to share content between a display and a mobile device, using a combination of spatial interaction and mobile touch screen interaction. Each technique was designed to cater to a different situation, based on, e.g., distance to the display and the number of items being transferred.

Turner et al. presented numerous interaction techniques for content transfer combining gaze and touch [29,30,31]. For instance, using their Eye Pull, Eye Push concept [29], users can select an item on a display by looking at it, and transfer it by swiping down on their mobile device.

The solutions above require manipulation of the recipient device. In particular, for frequent users of content transfer fea-

tures, it is worth investigating techniques that allow the device to remain wherever it is being kept. For instance, many people carry their mobile device in a handbag, and taking it out may take time and feel cumbersome.

Mäkelä et al. closed this gap with their SimSense smart space system, with which users could keep their mobile device in their pocket and use mid-air gestures to transfer content from a distance. Using gestures for this purpose was found to provide a good user experience [17]. In particular, Mäkelä et al. compared two mid-air gestures for the same purpose, focusing on single content item transfers [17].

Building on the concept of enabling content transfer without taking a mobile device out of the pocket, we investigate how this approach can be extended to multiple modalities, so as to cater to the diverse situations in which users encounter public displays and want to take away information. In particular, we introduce three techniques in addition to mid-air gestures, and evaluate as well as compare them in a user study. Additionally, and unlike previous work, we compare our techniques to QR code scanning as a baseline. Each Pocket Transfer technique has its own strengths, and therefore our set of techniques cover a wide range of settings and use cases. In addition, we compare the techniques in two different content transfer scenarios: single-item and multi-item transfers.

IMPLEMENTATION

We extend the SimSense system, which allows seamless transferring of content from an information display to mobile devices. Users are automatically paired with their mobile devices when entering the space. Consequently, content transfer can begin right away without a separate setup, and the mobile device does not need to be interacted with at all. In particular, in this work we extend SimSense, which originally enabled interaction via mid-air gestures [17], to support touch, mid-air gestures, gaze, a multimodal combination of mid-air gestures and gaze, and QR code scanning.

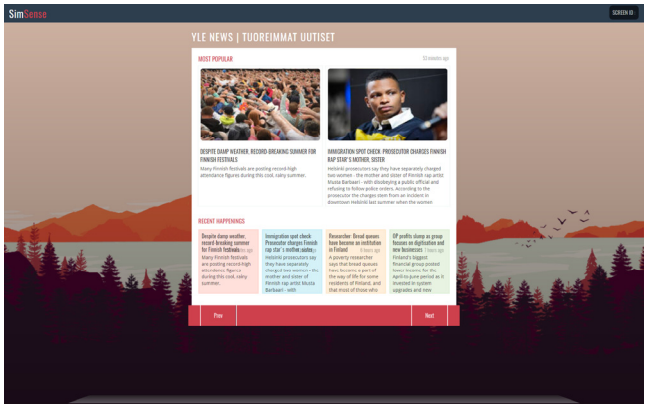

Figure 2. The main screen, displaying two popular articles on top, and four recent articles below. Navigation buttons for changing news feeds are located at the bottom.

The system displays content from external sources. Although a variety of different content, even applications, could be transferred, in this version we included content from popular news portals. Users can switch between news feeds and explore the content in more detail (Figure 2).

User-Mobile Pairing

To enable pocket transfers, users are automatically paired with their mobile device, provided they have the related mobile application installed. The location of mobile devices in the space is determined via Bluetooth beacons, and the user's location is determined via a Kinect sensor. Users and mobiles with matching locations are paired. Consequently, users can transfer content using the proposed interaction techniques, without ever touching the recipient device.

The method works with multiple simultaneous users. However, a practical limitation of our current implementation is the maximum number of people the Kinect sensor can track. Also in very crowded spaces people may be so close to each other that reliable pairing may not be possible. This component of our system is independent and can be independently upgraded or replaced without affecting the rest of the system. For example, another approach for pairing users with mobiles is by comparing the accelerometer readings from the mobile device to movements of the user [35].

Mobile Application

The application is implemented in Android. It utilizes Bluetooth to communicate its location to the system. Receiving content results in a notification along with vibration and sound effects. Opening the app or tapping on the notification shows a scrollable list of all the transferred content.

In our study, we focused on different techniques for transferring content to the mobile device. Participants did not need to interact with the transferred content. The mobile device stayed in their pockets and provided tactile and auditory feedback whenever content was successfully transferred.

Content Transfer Techniques

The extended SimSense system supports five different interaction schemes for content transfer. When content transfer is triggered, the transferred item on the screen is enlarged as if coming out of the screen, accompanied with sound effects. This applies to all pocket transfer techniques.

Next, we describe all five techniques used in this study to transfer content from a display to a mobile device. Our multimodal technique is novel; to the best of our knowledge, gaze and mid-air gestures were never utilized together for content transfer. While mid-air gestures were used for content transfer before [17], this is the first comparison between mid-air gestures, gaze, and touch for content transfer purposes. It is unclear how this novel context affects the performance and user experience of the techniques, especially in relation to each other. We argue this is also valuable outside the context of seamless content transfer, as we are one of the few who extensively evaluate different modalities for the same purpose. Therefore, this study serves as an overview on the individual strengths of said modalities. We discuss technique-specific implications in the following subsections.

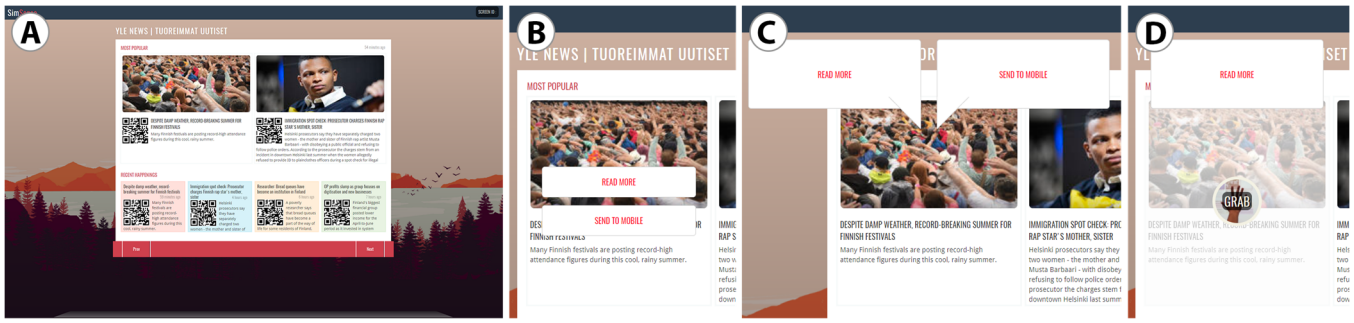

Figure 3. Interface differences. A) QR: codes are displayed on the main page. B) Touch: the menu is overlaid on the item. C) Gaze: the menu is positioned above the element. D) Multimodal: instructions for grabbing are displayed when the item is gazed at.

QR Codes

We utilize QR codes as a baseline for the study. QR code scanning represents a more traditional way of transferring content, as it requires users to hold and interact with, the mobile device, and offers a simple and familiar content transfer method. At the same time, QR codes are streamlined in that they also do not require explicit connection to the target display – content can be transferred directly. Moreover, QR codes have been utilized as a baseline in previous work [1].

QR codes for each item are readily displayed in the default view of the application (Figure 3A), therefore users do not need to interact with the screen at all. In this condition, the UI is widened to accommodate for the additional space that the QR codes require.

In the study, QR codes were scanned with a third-party Android application. This results in a link to the original content item, clicking on which directly opens the article on a web browser. Although not needed for our study, users could cycle through the feeds and inspect content via touch interactions similar to Touch, which is explained next.

Touch

With Touch, users can tap on items to bring up a menu with two actions (Figure 3B). Tapping on "Send to mobile" will transfer the item to the mobile device. The other action opens the item in a detailed view. Users can also access the menu and transfer the item from the detailed view.

Mid-Air Gestures

We utilize the same approach as Mäkelä et al. [17] for mid-air gestures. Users transfer content with the grab-and-pull gesture, wherein users point to a content item on the screen, grab it, and pull it towards themselves to transfer the item to their mobile device (Figure 1B). Pointing and grabbing is visualized via an on-screen cursor. Contextual feedback is provided on the screen when a transferable item is hovered over, and when an item is grabbed.

Users can navigate between feeds using point-and-dwell on the navigational buttons at the bottom. Content items can be opened in a detailed view via point-and-dwell, in which content transfer is also allowed using the grab-and-pull gesture.

Gaze

With Gaze, we utilize dwell-time to trigger selections. Gazing on a content item will bring up a menu similar to that of the touch condition. However, as eye tracking is occasionally inaccurate [28] and suffers from the Midas Touch issue [9,28], in this condition the action buttons are larger and appear on top of the item (Figure 3C). This was done to a) avoid the menu blocking the content of the item being gazed at, and b) to ease gaze selection by avoiding intersecting elements.

Navigational buttons are also triggered with gaze-dwell, during which time the button fills up with a different color to visualize dwell time. The dwell time for all triggers is 1 second, decided by a pilot test and related work [15,20].

Multimodal

Multimodal combines gaze and mid-air gestures. Users transfer content by looking at an item, and doing a grab gesture (forming a fist) in mid-air (Figure 1D). Due to the grab gesture working as a confirmation for a transfer, no dwell time is needed - instructions for grabbing appear in the middle of an item immediately when the user looks at it (Figure 3D). The grab can be done with either hand and in any position, although for stable recognition we recommended that study participants raise their hand at shoulder height for grabbing. Other interactions in this condition, however, work with gaze-dwell, similar to the gaze condition.

Current research presents very few multimodal systems that combine gaze and mid-air gestures. Some solutions exist for desktop type tasks [3] and multi-screen interactions [6]. However, to our knowledge, this is the first time such a multimodal technique has been used for content transfer, and evaluated for use in public and semi-public spaces.

STUDY

To evaluate the five content transfer techniques presented above, we recruited 20 participants to carry out content transfer tasks with each of them. We addressed two use cases of different length:

- **Single-item transfer.** The user passes by the display, sees an interesting item (e.g., a news article), quickly transfers it to their mobile device, and leaves the scene.
- **Multi-item transfer.** The user transfers several items in a row to their mobile device while passing by.

We hypothesized that preferences towards the techniques might differ based on whether the user intends to transfer one or several content items. This hypothesis is supported by Mackay [14]: they compared techniques in a desktop environment, and found that the efficiency of, and preferences towards, their tested techniques were dependent on the exact task at hand. Prior work has developed different techniques for single and multi-item transfers before [13]; however, to our knowledge, we are the first to evaluate a set of techniques equally with both use cases.

In both use cases, participants walked to the display from a marked area to interact, and finished the task by walking to another marked area on the other side of the display (Figure 4). We did this for two reasons. First, walking to and from the display resembles real-life situations. Users are rarely readily *at* the display – instead, they are walking past it and must deviate from their course to reach the display [34]. Doing this in the tasks makes participants better equipped to evaluate the techniques in a real context. Second, the distance to the display varies between techniques, which contributes to the overall performance and experience. For instance, we assume Touch would be faster than Gestures in terms of interaction time; however, it is unclear whether Touch would actually be faster when accounting for the time it takes to walk up to the display as opposed to mid-air gestures with which one can interact from a distance. Therefore, it makes sense to measure the full duration of the use cases when comparing the interaction techniques.

Due to the experimental setup of the system and the study with a multitude of sensors and cameras, we carried out the study in an office-like environment wherein we had full control of the setup. For instance, for the multimodal condition, users needed to stand relatively close to the display to be recognized by the eye tracker. Due to this, the Kinect sensor needed to be positioned further back (behind and above the display) for it to reliably see the user and recognize the grab gestures. We did not want to use head-mounted eye trackers as external equipment might hinder the user experience.

Participants

We recruited 20 participants (7 females) between 19 and 29 years of age (M = 24.7, SD = 2.7). All participants had normal or corrected-to-normal vision. Sixteen participants were bachelor or master level students, three were PhD students, and one was an IT consultant.

Participants answered statements about being familiar with QR codes as well as the remaining modalities on a 7-point Likert scale (1 = "strongly disagree"; 7 = "strongly agree"). Participants stated being very familiar with QR codes and touch (md = 7), somewhat familiar with gaze (md = 5), neutral with mid-air gestures (md = 4), and unfamiliar with combinations of gaze and gestures (md = 2.5).

Apparatus

We set up the system in an office-like space. The full setup is described in Figure 4. The display, a 24" full HD touch screen, was positioned on top a shelf roughly at eye level. The eye tracker (Tobii REX) was taped on the display right below the viewport. The Microsoft Kinect One sensor was attached to a tripod and positioned above and behind the display roughly at the height of 2 meters. This was mandatory for the multimodal condition, as both the eye tracker as well as the Kinect needed to see the user simultaneously.

We taped two cross-shaped markers on the floor to indicate the start and end position for the tasks. The markers were positioned 4 meters from each other, so that the line between the markers was 2 meters from the display. Additionally, we recorded the study session with a video camera.

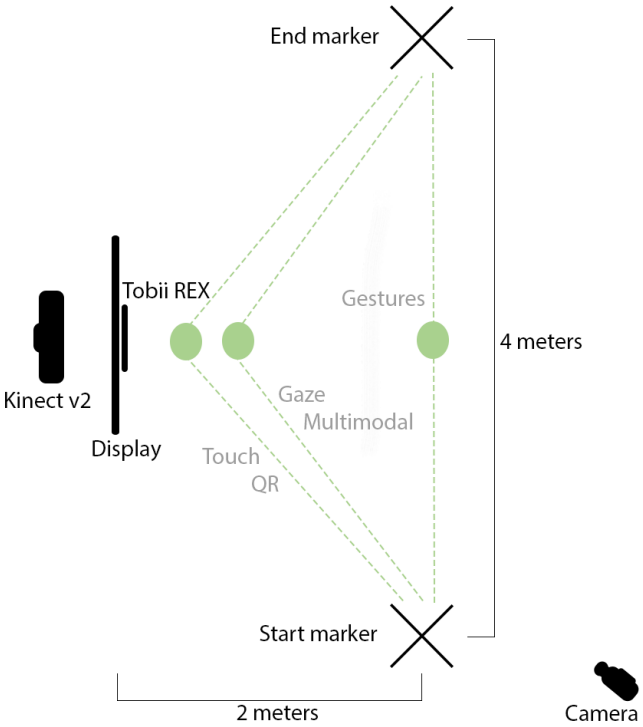

Figure 4. Study setup. Green, dotted lines represent the walking paths with each condition.

Procedure

All 20 participants went through the following procedure. Study sessions lasted between 50 and 75 minutes. First, the participant filled a consent form and a background questionnaire. The participant was then explained the study.

A mobile device, Nexus 5 with the Android application installed, was handed to the participant, and they were instructed to put the device in their trousers' pocket. The mobile device was not directly needed during the tasks for interaction (except during the QR condition), but rather for receiving tactile and auditory feedback when content is received. This was done to indicate to the participant that the transfer was successful.

We explained that content transfer will be approached through two use cases, both of which revolve around a realistic scenario, wherein they are walking past an interactive display and decide to transfer content for later consumption. For this, the participant was requested to maintain a quick, natural walking pace, and to keep it consistent across the techniques. The order of conditions was balanced with using a Latin Square. Participants went through the following process five times, once for each technique:

1. *Practice phase*. The participant was positioned on the interaction area (green ellipses in Figure 4), and any calibrations needed were conducted (e.g., for eye tracking). For Mid-air gestures, participants were positioned between the markers, 2 meters from the screen. For Gaze and Multimodal, the interaction area slightly varied between participants and was defined during calibration. On average, distance to the screen was around 80 cm as recommended by the manufacturer[2]. For QR and Touch, users were free to interact from whichever distance was comfortable. In the practice phase, the participant was asked to transfer a randomly highlighted item (visualized with thick, red borders) to the mobile device without prior instructions. The researcher gave instructions during the practice phase when necessary.
2. *Single-item use case*. The participant was positioned on the start marker. The task was to start walking when a randomly highlighted item appears, walk to the specified interaction area, transfer the highlighted item using the active technique, and continue to the end marker. This task was repeated five times.
3. *Multi-item use case*. The participant was asked to repeat the task, but this time, transfer five highlighted items in a sequence instead of just one before continuing to the end marker. The next highlight on the screen would appear after the previous one was transferred. The task similarly began from the start marker, and finished at the end marker. This task was repeated twice.

This procedure resulted in 15 content transfers (excluding practice) for each technique, totaling up to 75 transfers per participant. We concluded with a questionnaire and a semi-structured interview. Due to some of the study sessions taking a long time, 15 out of 20 participants were interviewed.

Due to its different nature, some special features applied in the QR condition. A successful transfer task included scanning the correct code and opening the contained link in a browser, which could be done with a button press in the QR app. For a realistic scenario, participants were asked to either put the phone in their pocket or hold it with their hand lowered prior to each task. Participants could lift the phone and open the QR app as soon as they started walking. In practice, participants were ready to scan the code by the time they reached the display. Also, participants could continue from the display to the end marker right after scanning the code, i.e., they could open the link in a browser while walking.

Limitations

This study was conducted in a controlled environment instead of a public setting. Hence, it could be argued whether participants were equipped to evaluate their usage of the proposed techniques in public and semi-public settings. However, all participants had experience with various interactive public displays, especially with those employing touch. We believe this prior experience makes the participants well equipped to evaluate their use of the proposed techniques in such situations. Furthermore, we alleviate this problem by conducting the study in an office-like environment, and by introducing realistic scenarios wherein users walked past the display and stopped to interact before continuing forward.

RESULTS

We first present results on the performance of the techniques, including task completion times as well as error rates. Then, we present user feedback and technique preferences in different situations based on the questionnaire and interview.

Performance

We measured full task completion times, including walking to and from the display. Duration was measured manually from the video recordings: the task began when participants started moving from the start area (lifted their foot), and ended when their foot touched the end area. Given that the videos were recorded at 25 FPS, the margin for error with manual measuring was roughly one frame (40 milliseconds).

In addition, we used interaction logs to measure individual selection times from when a highlighted item appeared on the screen, until the user had sent the corresponding item to the mobile device. For this measurement, we only used the last four selections from the multi-item tasks. This was done because the first highlight in each task appeared when the user was standing on the start marker. To exclude the walking time, we did not account for single-item tasks nor for the first highlight of the multi-item tasks.

We removed instances from the analysis wherein noticeable technical issues were encountered. For instance, the Kinect sensor was not always stable and occasionally performed poorly in recognizing the grab gesture (this was almost entirely specific to few select participants with e.g. very reflective clothing). Similarly, for both gaze and multimodal conditions, the eye tracker sometimes did not start tracking even when participants were standing at a correct spot. Hence, we excluded roughly 7% of the data from the analysis.

Completion times for single-item and multi-item tasks as well as individual selection times are presented in Figure 5.

[2] https://www.tobiipro.com/learn-and-support/learn/steps-in-an-eye-tracking-study/run/how-to-position-the-participant-and-the-eye-tracker/

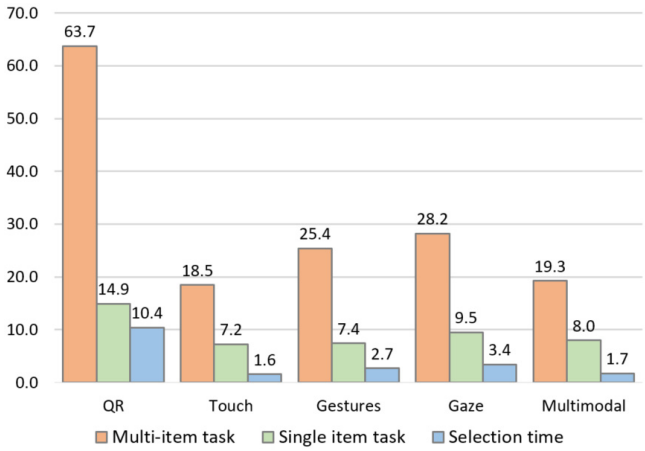

Figure 5. Completion times for single item and multi-item tasks with each technique.

For *single-item tasks*, Touch and Mid-air gestures were the fastest, followed by Multimodal, Gaze, and finally, QR codes. A repeated measures ANOVA with a Greenhouse-Geisser correction revealed a significant main effect of the used technique on completion time ($F(1.281, 108.896) = 114.448$, $p < 0.0005$). Post-hoc analysis with Bonferroni correction showed significant differences in completion time between all pairs ($p < 0.0005$) except between Touch and Mid-air gestures ($p = 1.000$), and between Mid-air gestures and Multimodal ($p = 0.14$).

For *multi-item tasks*, Touch and Multimodal were the fastest, followed by Mid-air gestures, Gaze, and QR codes. A repeated measures ANOVA with a Greenhouse-Geisser correction similarly revealed a significant main effect of the used technique on completion time ($F(1.417, 33.997) = 158.312$, $p < 0.0005$). Post-hoc analysis with Bonferroni correction showed significant differences in completion time between all pairs ($p < 0.0005$), except between Touch and Multimodal ($p = 1.000$) and between Mid-air gestures and Gaze ($p = 1.000$).

When only accounting for selection time, Touch and Multimodal were the fastest, followed by Mid-air gestures, then Gaze, and finally, QR codes. A repeated measures ANOVA with a Greenhouse-Geisser correction similarly revealed a significant main effect of the used technique on completion time ($F(2.391, 270.131) = 717.243$, $p < 0.0005$). Post-hoc analysis with Bonferroni correction showed significant differences in completion time between all pairs ($p < 0.0005$) except between Touch and Multimodal ($p = 1.000$).

Error rates were low across all conditions. As an error, we considered transferring the wrong item, i.e., not the one that was highlighted. Error rates were as follows: QR codes (5.8%), Multimodal (1.5%), Mid-air gestures (0.9%), Gaze (0.0%), and Touch (0.0%). The higher error rate of QR codes is explained by the QR app automatically scanning codes that came to its view, sometimes resulting in an incorrect code being scanned as the user was moving the phone to the target. It is likely that QR codes in general have a lower error rate.

Usefulness and Preferences of the Techniques

Preferences and evaluations of the techniques are presented in Figure 6. Across all pocket transfer techniques, the ability to keep the device in a pocket was evaluated highly useful (md = 7). Similarly, all pocket transfer techniques were rated suitable for transferring content between situated displays and mobile devices (md = 6-7). Although QR codes were also rated suitable for content transfer (md = 6), a Mann-Whitney U test revealed it was rated significantly lower than the pocket transfer techniques ($p < 0.05$).

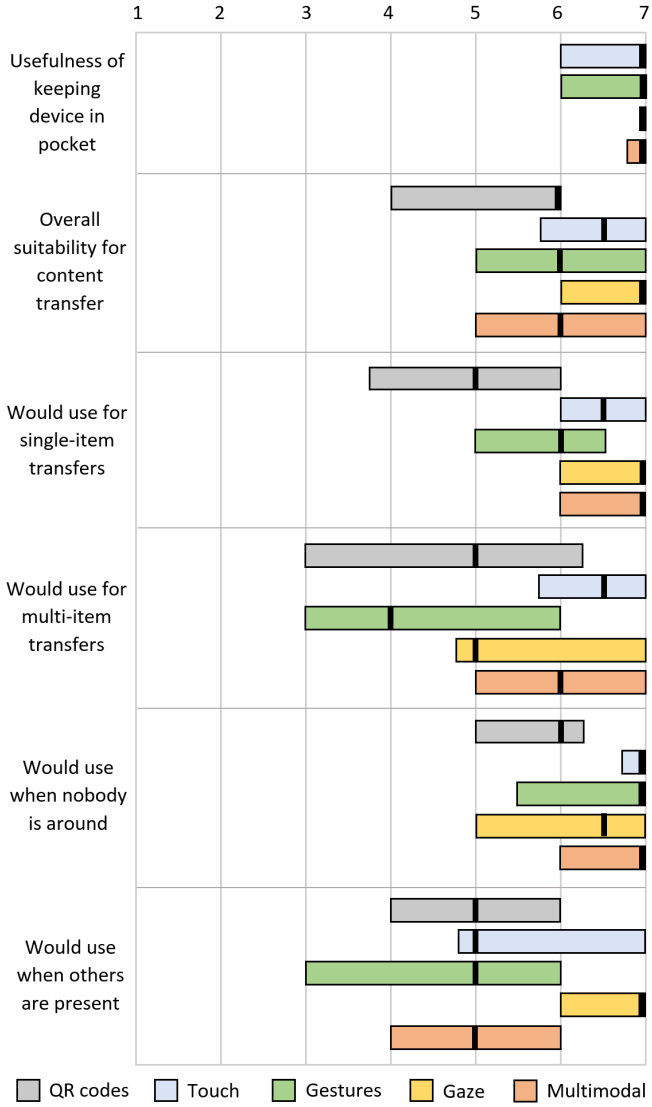

Figure 6. Boxplots for statements regarding all content transfer techniques. Boxes represent inner quartiles, and the middle lines represent medians.

For *single-item transfers*, Gaze and Multimodal were rated the most desirable techniques, followed by Touch and Gestures, and lastly, QR codes. A Mann-Whitney U test revealed a significant difference between QR codes and all other techniques, as well as between Gestures and Gaze, and Gestures and Multimodal ($p < 0.05$). For *multi-item transfers*, the most desired technique was Touch, followed by Multimodal,

Gaze, QR, and Gestures. A Mann-Whitney U test revealed a significant difference between QR codes and Touch, Touch and Gestures, and Multimodal and Gestures ($p < 0.05$).

For situations where *no other people are present*, all techniques were rated suitable. The most preferred techniques were Touch, Multimodal, and Mid-air gestures. A Mann-Whitney U test revealed a significant difference between QR and Touch, and QR and Multimodal ($p < 0.05$). For situations where *other people are present*, Gaze was clearly preferred (md = 7), and a Mann-Whitney U test revealed a significant difference between Gaze and all other techniques ($p < 0.05$).

Interview Results

We interviewed 15 participants to further assess their opinions on and experiences with the proposed techniques.

Participants were seemingly positive about all pocket transfer techniques. 14 out of 15 interviewed participants explicitly described keeping the phone in a pocket or bag as useful and convenient. The remaining participant mentioned that he holds his phone all the time anyways and hence failed to see the benefit for himself. However, it is notable that the benefit is not only about where the device is being held, as P20 elaborated: *"It's not only about not having to pull it out of your pocket. It's also about not having to do anything with it, like start an application. So, it doesn't really matter if I have the phone in my pocket or in my hand, it still makes the interaction straightforward"*. Moreover, two female participants noted that they occasionally carry their mobile device in a large handbag, and that they must specifically look for the device, which is time-consuming and tedious.

We asked participants to describe each technique in their own words:

QR codes received more negative feedback than the pocket transfer techniques. 10 participants explicitly mentioned that having to pull out the phone to interact is a negative trait. QR codes were further described as tedious (5/15), error-prone (5/15), and tiring (3/15). Among the positive aspects were that it is familiar (6/15) and easy to use (4/15).

Touch was described as easy to use (4/15), fast (3/15), and natural (3/15). Among the negative traits, the most notable were that it is "boring" or "nothing new" (8/15), easy to observe by others (6/15), unhygienic (6/15), and that one needs to get close to the display to interact (6/15). However, Touch was favored due to its familiarity and its prevalence in public displays. 11 out of 15 participants reported that they would expect a display to work with touch, and that they would expect to know how to use it right away.

Mid-air gestures were described as useful since a display can be accessed from a distance (7/15), "cool" (5/15), fast (3/15), and fun (3/15). However, participants were worried about using mid-air gestures in public (7/15). Nonetheless, some participants thoroughly enjoyed using gestures. Although gestures have been previously found "fun" in a variety of contexts, such as co-operative tasks [10] and gaming [4], in our study participants made more explicit remarks, like those reported by Mäkelä et al. [17], as P19 demonstrated: *"Gestures were cool, I felt like in Minority Report. It feels a little bit like magic. I really liked the fun factor and the novelty."*

Gaze was described as fast (8/15), private (6/15), "cool" (4/15), and natural (4/15). 6 participants explicitly mentioned that they "liked Gaze a lot". 3 participants mentioned that gaze interaction gets tiring after some time. Three participants mentioned that it is practical that Gaze is completely hands-free. P15 noted that the hands-free characteristic goes particularly well with pocket transfers: *"I liked Gaze the most since it's hands-free. You don't need to use any part of your body at all. It was a really great experience."*

Multimodal was described as fast (6/15), fun (5/15), and useful (3/15). 7 participants explicitly mentioned that they "liked Multimodal a lot". No commonly shared negative traits were identified. P17 summarized the technique: *"Multimodal, I like it the best. It was fast, accurate, and it was also fun to use it. No downsides."*

Finally, we asked if participants had any worries related to the technology and the techniques that allow transferring content to a personal device that remains in a pocket. Three participants were generally worried about shoulder-surfing, i.e., others seeing what content they are interested in. However, all three mentioned that they would not be worried if they were using gaze. Another three were worried about data security in some form. Two participants wondered if the system could be exploited to share malicious content.

DISCUSSION

All pocket transfer techniques reached fast completion times in both single-item and multi-item scenarios, and achieved a high user experience. As study participants pointed out, keeping the mobile device in a pocket is very useful with all techniques (md = 7), and 14 out of 15 interviewees explicitly remarked that this feature is useful and convenient. In addition, 10 out of 15 interviewees described QR codes as cumbersome due to requiring manipulation of the mobile device.

Some existing content transfer studies report selection times that are comparable to those of the pocket transfer techniques [5,21]. However, the strength of pocket transfer techniques is in that the *preparation* for the interaction is greatly mitigated, and therefore we argue that pocket transfers would outperform these techniques in a real situation.

It is worth noting that techniques that require holding the mobile device allow for other interactions that pocket transfers could not achieve; however, we argue that for one-way transfers, especially if such transfers are done frequently, our proposed techniques outperform other current solutions. We also note that despite somewhat negative feedback, the benefit of QR codes is that scanning a code with a mobile device is not tied to any system or infrastructure. Therefore, QR codes may be useful in one-time use scenarios, wherein users might not bother installing a mobile application to enable pocket transfer interactions.

Based on the study results and the discussion above, we formulate our first design implication:

Design Implication 1: Pocket transfer techniques are fast and convenient regardless of modality, and should be considered especially for frequent users when designing content transfer systems.

We also want to make a larger point regarding evaluation of interaction techniques. In this study, we utilized study tasks which included participants approaching and leaving the display, in addition to the actual content transfer, and included both single and multi-item transfer tasks. In other words, we accounted for the *preparation* for the interaction as well as the immediate steps *after* the interaction. Most existing studies leave these phases out of their tasks and therefore the evaluation of their techniques. For instance, content transfer studies use tasks that only begin when the user is already in position, holding the mobile device, and ready to interact, therefore not accounting for the time and effort it takes to reach this state in the first place [e.g., 2,7,16,17,29].

We argue that our dynamic study tasks have two significant implications. First, this approach is a viable way to fairly compare techniques that span across different modalities, and techniques that might require different preparational actions. For instance, using this approach, we discovered that while the selection times with mid-air gestures were expectedly slower than with touch, mid-air gestures reached comparable speed for single-item transfers because the distance to the display was different between the techniques. Second, including the full process results in a more realistic user experience and therefore more ecologically valid feedback.

The importance of such approaches is further highlighted when we move towards more seamless interactions with technology. The advantages of future interaction techniques do not necessarily lie in the so-called direct interaction phase [18,33], but rather, in alleviating the steps to *prepare* for the interaction, or even skipping them completely.

Based on the discussion above, we formulate a recommendation for future user studies:

Study Design Recommendation: Interaction techniques should be evaluated with various realistic tasks that include preparation for, and halting of, the interaction, especially when different modalities are compared.

Next, we summarize and discuss the results for each pocket transfer technique.

Touch

Touch was the fastest technique in both single-item and multi-item scenarios as well as individual selection times. Despite it being the most traditional way of interaction, many users felt most comfortable using Touch, primarily attributing it to stability and familiarity. Participants also stated that they would simply assume that an interactive display would work by touching it.

Many users felt Touch makes it easy to observe what content is being transferred. Many also made remarks about not wanting to touch a potentially dirty display, which has been reported by previous work as well [25].

Touch was evaluated very suitable for both single-item and multi-item transfers. However, due to the threat of shoulder-surfing, participants were somewhat worried about using Touch in public when other people are present.

Design Implication 2: Touch should be used when the display is reachable and when familiarity and efficiency is important, or when it is unclear how the display will be primarily used.

Mid-Air Gestures

Mid-air gestures were, together with Touch, the fastest technique for single-item transfers. While being slower in selection time, Gestures greatly benefit from not having to walk up to the display. With only 2 meters from the display, Gestures already reached comparable efficiency with Touch. To our knowledge, we are the first to reach such an estimate on a distance threshold, after which Gestures would become the most efficient interaction technique for quick sessions. This benefit was also noted by participants in the interview.

However, mid-air gestures were not seen as suitable for long interactions as the other techniques. Acquiring the target with mid-air gestures is slower than with the other techniques, and users may also suffer from fatigue in prolonged interactions [8]. Consequently, the benefit of not having to walk to the display diminishes in longer interactions. However, with a few participants, the performance of the Kinect sensor was unstable, resulting in jittery interaction, which was reflected in their feedback.

Gestures were also not seen as suitable for very crowded spaces. Contrary to Touch, wherein users were worried about others seeing what content they interact with, with Gestures, they worried more about drawing attention to themselves, as already pointed out by earlier work [23].

Design Implication 3: Mid-air gestures should be used in calm spaces where people are not always around, where people are expected to transfer single items, or where the display is not along the primary walking paths.

Gaze

In contrast to prior work that found gaze faster than many other modalities [24,27], Gaze was the slowest pocket transfer technique. This is likely due to the uniqueness of the content transfer context, in which users need to position themselves within the tracking area, and signal the transfer command in two steps. Changing the dwell time would present a trade-off between accuracy and transfer time. Nonetheless, Gaze was most commonly perceived as being fast. We attribute this to the nature of gaze dwell – users do not necessarily perceive “looking” as interaction [4]. Participants evaluated

Gaze as more suitable for short than long interactions, as using Gaze for an extended period can be tiring [11].

Gaze was perceived very suitable for public spaces (md = 7), performing significantly better than the other techniques in this regard (md = 5). As participants pointed out, interacting with gaze does not look any different from simply observing the display, creating a stronger sense of privacy. A related benefit of gaze is that it is completely hands-free, even more so when the recipient mobile device can remain in a pocket.

Design Implication 4: Gaze should be used in crowded spaces where sensitive content might be available (e.g., selections might imply political interests [32], or contain personal information), or where users are expected to carry items (e.g., a drink or a bag).

Multimodal

Multimodal was the second slowest pocket transfer technique for single item tasks; however, for multi-item tasks it was the fastest technique together with Touch. Multimodal was evaluated to be a suitable technique for both single-item and multi-item transfers.

Selecting the target with gaze and confirming the transfer with a grab gesture gained positive feedback and performed efficiently. That said, there is much room for improvement. Gesture recognition was not always stable, and participants often had to repeat the grab gesture before it was recognized. Similar to Gaze, Multimodal suffered from the small interaction area, as users had to position themselves carefully. As sensing technologies continue to advance [12], Multimodal has a high potential to be a very fast technique, as even in its current form its performance was comparable to Touch.

Similar to Mid-air gestures, participants were somewhat worried about the expressiveness of Multimodal. When using Multimodal, we asked users to raise their hand up to make sure the sensor recognized the grab gesture reliably. With more advanced technology, the multimodal approach could be used in a subtle, unnoticeable manner. For instance, users could make the grab gesture against their upper body to hide it from others.

Design Implication 5: Multimodal should be used when users are expected to transfer multiple items, and when the display is unreachable. In crowded spaces, the design should allow subtle gestures (e.g., against the body) when confirming content transfers to avoid drawing attention.

FUTURE WORK

An interesting proposal for future work is how the described techniques could work in parallel. As we found in this study, numerous factors (personal and external) affect the users' preferences, and therefore multiple techniques should be available. Prior work has already investigated transitioning between mid-air gestures and touch [21]. However, how gaze, and above all, multimodal techniques, could be incorporated without interfering with other techniques would be worthwhile to investigate in the future.

In addition, especially considering automatic user-mobile pairing as well as transferring content from public to personal devices, a multitude of concerns related to privacy, data security, and interaction in public are likely present. We asked participants about their potential worries, and while said topics were raised by a few participants, no shared, major concerns were identified. Nonetheless, we primarily focused on interaction and performance, and therefore any related concerns should receive more attention in the future.

CONCLUSION

We presented Pocket Transfers: interaction techniques that allow content being transferred from a situated display to a personal mobile device, while keeping the mobile device in a pocket or bag throughout the interaction process. In a 20-participant user study, we evaluated four techniques employing touch, mid-air gestures, gaze, and a multimodal technique combining mid-air gestures and gaze, and compared them to QR codes, which served as a baseline condition.

We found that pocket transfers are fast and convenient across different modalities and designs. Users highly appreciate not having the manipulate the mobile device, independent of the technique used. Touch and Mid-air gestures were the fastest techniques for quick interactions wherein only a single content item is transferred. Touch and Multimodal were the fastest techniques for interactions wherein multiple items are transferred. For situations where other people are present, Gaze was the most preferred technique due to its subtlety.

Our work is useful to researchers and practitioners in a multitude of ways. First, we showed that content transfer methods where the recipient device remains in a pocket are generally fast and useful, and are therefore a solid consideration for a variety of content transfer systems. Second, we presented four designs for state-of-the-art pocket transfer techniques employing three different modalities as well as a combination of two modalities. Third, we recognized strengths and weaknesses for each technique, and presented guidelines to help researchers and practitioners choose the most suitable modalities and techniques for their content transfer systems.

Finally, we presented a novel user study design, wherein participants completed tasks that included the full interaction process, including preparation for, and halting of, the interaction. This way, we argue we reached more ecologically valid results. We encourage researchers to utilize such approaches in future studies.

ACKNOWLEDGEMENTS

Work on this project was partially funded by the Bavarian State Ministry of Education, Science and the Arts in the framework of the Centre Digitisation.Bavaria (ZD.B). This research was supported by the Deutsche Forschungsgemeinschaft (DFG), Grant No. AL 1899/2-1.

REFERENCES

1. Florian Alt, Alireza Sahami Shirazi, Thomas Kubitza, and Albrecht Schmidt. 2013. Interaction techniques for creating and exchanging content with public displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13). ACM, New York, NY, USA, 1709-1718. DOI: https://doi.org/10.1145/2470654.2466226
2. Gregor Broll, Wolfgang Reithmeier, Paul Holleis, and Matthias Wagner. 2010. Design and evaluation of techniques for mobile interaction with dynamic NFC-displays. In *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction* (TEI '11). ACM, New York, NY, USA, 205-212. DOI: http://dx.doi.org/10.1145/1935701.1935743
3. Ishan Chatterjee, Robert Xiao, and Chris Harrison. 2015. Gaze+Gesture: Expressive, Precise and Targeted Free-Space Interactions. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (ICMI '15). ACM, New York, NY, USA, 131-138. DOI: http://dx.doi.org/10.1145/2818346.2820752
4. Heiko Drewes. 2010. *Eye Gaze Tracking for Human Computer Interaction*. Ph.D Dissertation. LMU München: Faculty of Mathematics, Computer Science and Statistics, München, Germany.
5. Kathrin Gerling, Ian Livingston, Lennart Nacke, and Regan Mandryk. 2012. Full-body motion-based game interaction for older adults. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '12). ACM, New York, NY, USA, 1873-1882. DOI: http://dx.doi.org/10.1145/2207676.2208324
6. Jeremy Hales, David Rozado, and Diako Mardanbegi. 2013. Interacting with objects in the environment by gaze and hand gestures. In *Proceedings of the 3rd International Workshop on Pervasive Eye Tracking and Mobile Eye-Based Interaction* (PETMEI '13), 1-9
7. Robert Hardy and Enrico Rukzio. 2008. Touch & interact: touch-based interaction of mobile phones with displays. In *Proceedings of the 10th international conference on Human computer interaction with mobile devices and services* (MobileHCI '08). ACM, New York, NY, USA, 245-254. DOI: http://dx.doi.org/10.1145/1409240.1409267
8. Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (CHI '14). ACM, New York, NY, USA, 1063-1072. DOI: http://doi.acm.org/10.1145/2556288.2557130
9. Robert J. K. Jacob. 1990. What you look at is what you get: eye movement-based interaction techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '90), Jane Carrasco Chew and John Whiteside (Eds.). ACM, New York, NY, USA, 11-18. DOI: http://dx.doi.org/10.1145/97243.97246
10. Svebor Karaman, Andrew D. Bagdanov, Lea Landucci, Gianpaolo D'Amico, Andrea Ferracani, Daniele Pezzatini, and Alberto Del Bimbo. 2016. Personalized multimedia content delivery on an interactive table by passive observation of museum visitors. *Multimedia Tools and Applications*, 75(7), 3787–3811. DOI: https://doi.org/10.1007/s11042-014-2192-y
11. Mohamed Khamis, Ozan Saltuk, Alina Hang, Katharina Stolz, Andreas Bulling, and Florian Alt. 2016. TextPursuits: using text for pursuits-based interaction and calibration on public displays. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (UbiComp '16). ACM, New York, NY, USA, 274-285. DOI: https://doi.org/10.1145/2971648.2971679
12. Mohamed Khamis, Axel Hoesl, Alexander Klimczak, Martin Reiss, Florian Alt, and Andreas Bulling. 2017. EyeScout: Active Eye Tracking for Position and Movement Independent Gaze Interaction with Large Public Displays. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (UIST '17). ACM, New York, NY, USA, 155-166. DOI: https://doi.org/10.1145/3126594.3126630
13. Ricardo Langner, Ulrich von Zadow, Tom Horak, Annett Mitschick, and Raimund Dachselt. 2016. Content Sharing Between Spatially-Aware Mobile Phones and Large Vertical Displays Supporting Collaborative Work.In Collaboration Meets Interactive Spaces, Craig Anslow, Pedro Campos, and Joaquim Jorge (eds). Springer, Cham, Switzerland, 75-96. DOI: https://doi.org/10.1007/978-3-319-45853-3_5
14. Wendy E. Mackay. 2002. Which interaction technique works when?: floating palettes, marking menus and toolglasses support different task strategies. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (AVI '02), Maria De Marsico, Stefano Levialdi, and Emanuele Panizzi (Eds.). ACM, New York, NY, USA, 203-208. DOI: http://dx.doi.org/10.1145/1556262.1556294
15. Päivi Majaranta, Anne Aula, and Kari-Jouko Räihä. 2004. Effects of feedback on eye typing with a short dwell time. In *Proceedings of the 2004 symposium on Eye tracking research & applications* (ETRA '04). ACM, New York, NY, USA, 139-146. DOI: http://dx.doi.org/10.1145/968363.968390
16. Ville Mäkelä, Hannu Korhonen, Jarno Ojala, Antti Järvi, Kaisa Väänänen, Roope Raisamo, and Markku Turunen. 2016. Investigating mid-air gestures and

handhelds in motion tracked environments. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays* (PerDis '16). ACM, New York, NY, USA, 45-51. DOI: http://dx.doi.org/10.1145/2914920.2915015

17. Ville Mäkelä, Jobin James, Tuuli Keskinen, Jaakko Hakulinen, and Markku Turunen. 2017. "It's Natural to Grab and Pull": Retrieving Content from Large Displays Using Mid-Air Gestures. *IEEE Pervasive Computing 16*(3), 70-77. DOI: https://doi.org/10.1109/MPRV.2017.2940966
18. Daniel Michelis and Jörg Müller. 2011. The Audience Funnel: Observations of Gesture Based Interaction with Multiple Large Displays in a City Center. *International Journal of Human–Computer Interaction 27*(6), 562-579. DOI: http://dx.doi.org/10.1080/10447318.2011.555299
19. Meredith Ringel Morris, Anqi Huang, Andreas Paepcke, and Terry Winograd. 2006. Cooperative gestures: multi-user gestural interactions for co-located groupware. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '06), Rebecca Grinter, Thomas Rodden, Paul Aoki, Ed Cutrell, Robin Jeffries, and Gary Olson (Eds.). ACM, New York, NY, USA, 1201-1210. DOI: http://dx.doi.org/10.1145/1124772.1124952
20. Martez E. Mott, Shane Williams, Jacob O. Wobbrock, and Meredith Ringel Morris. 2017. Improving Dwell-Based Gaze Typing with Dynamic, Cascading Dwell Times. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17). ACM, New York, NY, USA, 2558-2570. DOI: https://doi.org/10.1145/3025453.3025517
21. Jörg Müller, Gilles Bailly, Thor Bossuyt, and Niklas Hillgren. 2014. MirrorTouch: combining touch and mid-air gestures for public displays. In *Proceedings of the 16th international conference on Human-computer interaction with mobile devices & services* (MobileHCI '14). ACM, New York, NY, USA, 319-328. DOI: http://dx.doi.org/10.1145/2628363.2628379
22. Pai Chet Ng, James She, Kang Eun Jeon, and Matthias Baldauf. 2017. When Smart Devices Interact With Pervasive Screens: A Survey. *ACM Trans. Multimedia Comput. Commun. Appl.* 13, 4, Article 55 (August 2017), 23 pages. DOI: https://doi.org/10.1145/3115933
23. Mark Perry, Steve Beckett, Kenton O'Hara, and Sriram Subramanian. 2010. WaveWindow: public, performative gestural interaction. In *ACM International Conference on Interactive Tabletops and Surfaces* (ITS '10). ACM, New York, NY, USA, 109-112. DOI: https://doi.org/10.1145/1936652.1936672
24. Ken Pfeuffer, Jason Alexander, Ming Ki Chong, and Hans Gellersen. 2014. Gaze-touch: combining gaze with multi-touch for interaction on the same surface. In *Proceedings of the 27th annual ACM symposium on User interface software and technology* (UIST '14). ACM, New York, NY, USA, 509-518. DOI: https://doi.org/10.1145/2642918.2647397
25. Kathy Ryall, Clifton Forlines, Chia Shen, Meredith Ringel Morris, and Katherine Everitt. 2006. Experiences with and Observations of Direct-Touch Tabletops. In *Proceedings of the First IEEE International Workshop on Horizontal Interactive Human-Computer Systems* (TABLETOP '06). IEEE Computer Society, Washington, DC, USA, 89-96. DOI: https://doi.org/10.1109/TABLETOP.2006.12
26. Khoovirajsingh Seewoonauth, Enrico Rukzio, Robert Hardy, and Paul Holleis. 2009. Touch & connect and touch & select: interacting with a computer by touching it with a mobile phone. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI '09). ACM, New York, NY, USA, , Article 36 , 9 pages. DOI: http://dx.doi.org/10.1145/1613858.1613905
27. Linda E. Sibert and Robert J. K. Jacob. 2000. Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (CHI '00). ACM, New York, NY, USA, 281-288. DOI: http://dx.doi.org/10.1145/332040.332445
28. Sophie Stellmach and Raimund Dachselt. 2013. Still looking: investigating seamless gaze-supported selection, positioning, and manipulation of distant targets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '13). ACM, New York, NY, USA, 285-294. DOI: https://doi.org/10.1145/2470654.2470695
29. Jayson Turner, Jason Alexander, Andreas Bulling, Dominik Schmidt, Hand Gellersen. 2013. Eye Pull, Eye Push: Moving Objects between Large Screens and Personal Devices with Gaze and Touch. In: Kotzé P., Marsden G., Lindgaard G., Wesson J., Winckler M. (eds) Human-Computer Interaction – INTERACT 2013. INTERACT 2013. Lecture Notes in Computer Science, vol 8118. Springer, Berlin, Heidelberg. DOI: https://doi.org/10.1007/978-3-642-40480-1_11
30. Jayson Turner, Andreas Bulling, Jason Alexander, and Hans Gellersen. 2013. Eye drop: an interaction concept for gaze-supported point-to-point content transfer. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia (MUM '13)*. ACM, New York, NY, USA, Article 37, 4 pages. DOI: https://doi.org/10.1145/2541831.2541868
31. Jayson Turner, Andreas Bulling, Jason Alexander, and Hans Gellersen. 2014. Cross-device gaze-supported point-to-point content transfer. In *Proceedings of the*

Symposium on Eye Tracking Research and Applications (ETRA '14). ACM, New York, NY, USA, 19-26. DOI: https://doi.org/10.1145/2578153.2578155

32. Nina Valkanova, Robert Walter, Andrew Vande Moere, and Jörg Müller. 2014. MyPosition: sparking civic discourse by a public interactive poll visualization. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing* (CSCW '14). ACM, New York, NY, USA, 1323-1332. DOI: http://dx.doi.org/10.1145/2531602.2531639
33. Daniel Vogel and Ravin Balakrishnan. 2004. Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users. In *Proceedings of the 17th annual ACM symposium on User interface software and technology* (UIST '04). ACM, New York, NY, USA, 137-146. DOI: http://dx.doi.org/10.1145/1029632.1029656
34. Julie R. Williamson and John Williamson. 2014. Analysing Pedestrian Traffic Around Public Displays. In *Proceedings of The International Symposium on Pervasive Displays* (PerDis '14), Sven Gehring (Ed.). ACM, New York, NY, USA, , Pages 13 , 6 pages. DOI: http://dx.doi.org/10.1145/2611009.2611022
35. Andrew D. Wilson and Hrvoje Benko. 2014. Cross-Motion: Fusing Device and Image Motion for User Identification, Tracking and Device Association. In *Proceedings of the 16th International Conference on Multimodal Interaction* (ICMI '14). ACM, New York, NY, USA, 216-223. DOI: http://dx.doi.org/10.1145/2663204.2663270