



Article

Understanding and Creating Spatial Interactions with Distant Displays Enabled by Unmodified Off-The-Shelf Smartphones

Teo Babic ^{1,*}, Harald Reiterer ² and Michael Haller ³

¹ BMW Group, 80788 Munich, Germany

² HCI Group, University of Konstanz, 78457 Konstanz, Germany

³ Media Interaction Lab, Free University of Bozen-Bolzano, 39100 Bozen-Bolzano, Italy

* Correspondence: teo.babic@bmw.de

Abstract: Over decades, many researchers developed complex in-lab systems with the overall goal to track multiple body parts of the user for a richer and more powerful 2D/3D interaction with a distant display. In this work, we introduce a novel smartphone-based tracking approach that eliminates the need for complex tracking systems. Relying on simultaneous usage of the front and rear smartphone cameras, our solution enables rich spatial interactions with distant displays by combining touch input with hand-gesture input, body and head motion, as well as eye-gaze input. In this paper, we firstly present a taxonomy for classifying distant display interactions, providing an overview of enabling technologies, input modalities, and interaction techniques, spanning from 2D to 3D interactions. Further, we provide more details about our implementation—using off-the-shelf smartphones. Finally, we validate our system in a user study by a variety of 2D and 3D multimodal interaction techniques, including input refinement.



Citation: Babic, T.; Reiterer, H.; Haller, M. Understanding and Creating Spatial Interactions with Distant Displays Enabled by Unmodified Off-The-Shelf Smartphones. *Multimodal Technol. Interact.* **2022**, *6*, 94. <https://doi.org/10.3390/mti6100094>

Academic Editors: Wei Liu, Jan Auernhammer, Takumi Ohashi, Di Zhu and Kuo-Hsiang Chen

Received: 30 August 2022

Accepted: 2 October 2022

Published: 19 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: smartphone; 3D interaction; mobile devices; distant displays; augmented reality; refinement techniques

1. Introduction

For decades, researchers investigated the way how we are interacting with distant displays by using mobile devices [1–3]. Initially, users controlled distant displays via buttons on a remote controller or later a touchscreen device. Here users mainly relied on uni-modal input made by their fingers by pressing buttons or swiping on a touchscreen. Besides uni-modal input with a single device [4,5], many researchers proposed multi-modal input modalities by using multiple tracking devices, e.g., using smartphone touch combined with glasses for eye-gaze input [6]. This allowed researchers to design even more powerful multi-modal interaction possibilities. The nature of the system as a whole changed as well. From a controller device-oriented system to an environment tracking-oriented system where users were observed in their space through which they can freely move while their movements were recognized and considered in the interaction. Hereby users could start performing spatial interactions with the distant display which was aware of their physical position relative to the distant display and the detailed motions of various body parts. On the negative side of creating such spatial interaction systems was however the fact that the overall systems also became more complex. Enabling inputs that relied on touch and motion of multiple body parts (i.e., as hand [7], body [8], head [6] or eyes [6]) required the augmentation of the user with handheld devices (e.g., smartphones, controller wands), wearables (e.g., glasses, bands, markers), or augmentation of the room with motion-tracking camera systems.

Our main motivation in this work is to find a new user tracking approach that enables a multi-modal interaction combining touch, hand-input, body-input, head-input, and gaze-input while completely removing the need for complex tracking devices. The particular

problem, we aim to solve for the distant display domain is the strong unavoidable bond between the richness of spatial interactions, which results in a hardware complexity. For decades, this dilemma has already limited how we design distant display interaction; it prevents many in-lab systems from being deployed for in-the-wild tests and makes it impossible for end-users to start using and learning novel interactions which have been proven beneficial many times in research setups (see Figure 1).

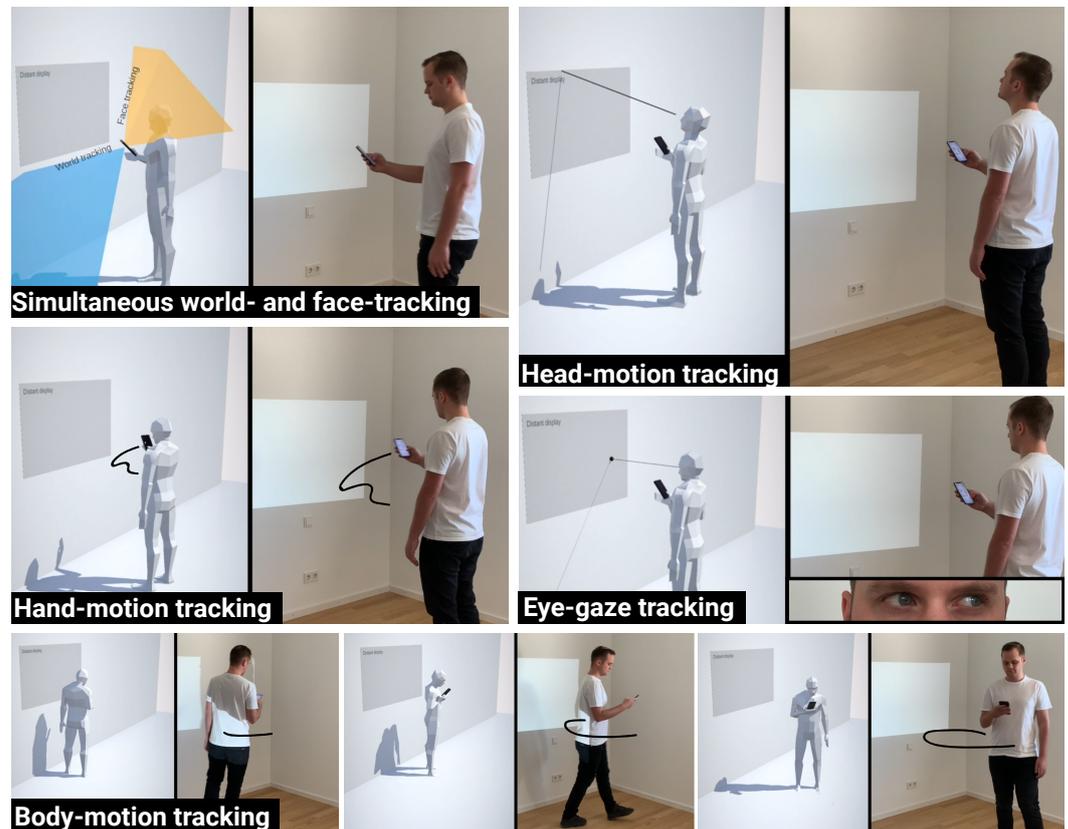


Figure 1. TrackPhone uses simultaneously the front and rear camera of the smartphone for tracking of the absolute world-space pose of the user’s device (hand), body, head, and eyes (gaze). Combined with touch inputs, this enables, after a simple app download, any smartphone user to perform powerful multi-modal spatial interactions for distant displays.

In this paper, we first unify the fragmented research under the umbrella of distant display and smartphone-based interaction. To present a unified overview and inform future research, we analyzed over 400 papers from this domain, synthesizing the state of this research field. We provide an overview of the enabling tracking hardware, input modalities, and interaction techniques of research and commercial systems. We continue by discussing key challenges of the domain and outlining which interactions should be included in an essential “must-have” bundle of interaction techniques for future distant display scenarios. Based on these pre-requisites, we present our approach for user-motion tracking using the SLAM technology already available in billions of off-the-shelf smartphones. By using both the front and rear camera of the smartphone simultaneously, we can enable multi-modal interactions by combining *touch* inputs with real-time world-space tracking of the user’s *hand*, *body*, *head* motion, and *gaze*. Without no additional external tracking devices, users can easily interact with the 3D space and use a rich set of different interaction modalities. In this paper, we further present the results of an initial user study, evaluating our tracking system with a variety of multi-modal 2D/3D interaction techniques (i.e., touch, ray-casting, and virtual-hand inputs), including a dual-precision approach for input refinement. The results reveal that participants achieve very satisfactory performance in 2D as well as

3D interactions, that the refinement techniques can improve the pointing accuracy up to three times, and presents a deeper look into the performance of a wide spectrum of multi-modal interaction techniques. We conclude by presenting demo applications, where we look beyond the input possibilities of our system and showcase how we can easily use it to enhance the 3D user interface output on a 2D display by user-perspective rendering.

In summary, the main contribution of this work can be summarized as follows. We present

- A comprehensive taxonomy and key research objectives addressing tracking devices, input modalities, and interaction techniques used for smartphone-based distant display interactions.
- A novel, solely smartphone-based approach for user-motion tracking, capable of enabling multi-modal interaction by touch and user's hand, body, head as well as gaze tracking.
- Two user studies that initially validate our novel tracking approach and compare the performance of various multi-modal interaction techniques, including a refinement mode, in 2D, and 3D distant display interaction.
- Some demo applications that showcase the versatile input capabilities of our approach as well as enhancing 3D output.
- Future work directions that are specific for smartphone-based tracking, however, still remain unaddressed.

2. Taxonomy

The goal of this taxonomy is to contribute a comprehensive analysis of the design space, where smartphones, with and/or without additional tracking devices, are used to drive the interaction with a distant display. For both new and experienced researchers, we provide a short overview of possible tracking techniques, different input modalities, and interaction techniques, spanning from 2D to 3D distant display interactions. Further, we discuss the trade-offs and impacts of the tracking hardware on the creative process of interaction design and the resulting interaction techniques. Based on our taxonomy, we outline key objectives that we address in this work.

We build upon other related taxonomies, all with their own specialized focus covering: cross-device interaction [9], sensing techniques for mobile devices [10], gaze [11] and augmented reality [12,13] enabled mobile devices, proxemic relationships [14], and spatial input by handhelds [15–17].

2.1. Designing the Taxonomy

In order to create the main corpus of relevant publications, we conducted a systematic search in the field of HCI and analyzed about 405 papers of the last 50 years. Our search included terms: *phone, smartphone, mobile, handheld, controller, spatial, vertical, spatially-aware, cross-device, distant, public, situated, remote, large, pervasive, wall, display, interaction, interfaces*, as well as their acronyms. The papers that got included in our corpus had to be concerned with the interaction tasks or techniques and tracking technology for people and/or devices. By looking at references within our corpus as well as using our own expertise, we identified additional articles (which is a common strategy for survey and taxonomy papers [9]). We tagged all papers from the initial corpus for the usage of a *distant display* or a *smartphone* in their interactive setups. The result was three subsets of papers, including

- Seventy-three papers using a smartphone together with a distant display, optionally with additional tracking devices.
- One-hundred-and-fifty-seven papers using a distant display with other input devices than smartphones (e.g., wands, smartwatches, lighthouse trackers).
- One-hundred-and-seventy-five papers using the smartphone alone (e.g., on-phone user interfaces) or with other devices (e.g., eye-trackers, pens, head-mounted displays, other mobile devices).

Based on our focus, we examined the found 73 papers using a smartphone and a distant display (concurrently) in our taxonomy. Of course, we also considered older mobile devices (without touchscreens) and tablets that were used with a distant display. In the next step, we tagged all these papers for the *tracking hardware*, *input modalities*, and *interaction techniques* (see Tables S1 and S2). We would like to acknowledge that our references are not an exhaustive listing of all the papers in the context of distant display and smartphone research, but instead, they represent a curated subset of the most relevant ones.

2.2. Categorization

Generally, we often divide tracking systems for enabling distant display interactions into two categories: *inside-out* and *outside-in*. In inside-out tracking, the tracking cameras or sensors are located on the device being tracked (e.g., head-mounted display worn by the user, smartphone held by the user) while in outside-in the sensors are placed in a stationary location in the user's environment observing the user or the tracked device from distance (e.g., a Microsoft Kinect on Nintendo Wii sensor placed on the TV looking in the direction of the user). By analyzing all the related papers, we agree with Brudy et al. [9], that outside-in tracking provides high-fidelity information as opposed to a more light-weight low-fidelity inside-out tracking setup. The main reason is that for inside-out systems, the reliable 3D tracking of the smartphone and the inclusion of the user as part of the sensing is still a major challenge [9]. However, none of the inside-out papers addressed this issue so far.

Our analysis also revealed that we can group papers using smartphones to enable distant display interactions, into two categories:

- **Inside-out:** interactions, enabled solely by using the technology, embedded in off-the-shelf smartphones.
- **Hybrid:** interactions, enabled by using inside-out tracking capabilities of smartphones, which get further enhanced by additional tracking devices, i.e., tracking devices attached to the smartphone, worn by the user, or placed in the room.

2.3. Motivation behind Tracking Systems

The compelling idea of “bring-your-own-device” [18] enabling out-of-the-box interactions assures that inside-out smartphone tracking remains a constant focus of many researchers in the domain of distant displays [19]. They challenge the smartphone-embedded technology, to enable innovative 2D [5] and 3D [20,21] interaction techniques, by thoughtfully steering between harsh constraints of the given tracking hardware, only to prevent the need for instrumenting the environment with cameras and buying or wearing dedicated devices.

Looking closer at hybrid tracking solutions, we noticed that researchers used mostly a closed hardware infrastructure, which is not always easily accessible to the mainstream. However, they contribute very important findings regarding the performance and experience of various uni-modal as well as novel multi-modal interaction techniques for distant displays [6,22,23].

2.4. Tracked Input Modalities

Using only the smartphone to enable the interaction with a distant display allows users to use their fingers for touch inputs and movement of their device-holding hand for hand motion input. These device-tracking inputs are enabled by the device's touchscreen [5] and pose tracking (i.e., device tilting [1,24] as orientation, motion as translation [3,25], or pose as translation + orientation [2]), while touch-input and tilting the device represent the majority of the past as well as recent papers, we can see a trend towards utilizing SLAM [26] that is natively implemented on modern smartphones, e.g., ARCore [27], ARKit [28]. These libraries show how SLAM can be used for precisely tracking the absolute world-space pose of the smartphone (i.e., the motion of the user's device-holding hand) in distant display scenarios [2]. However, approaches that would extend currently known

device-tracking inputs for body-tracking of the user's body, head pose as well as gaze still remain undiscovered (see Supplementary Materials).

Hybrid setups are as well often used for tracking the smartphone pose [22,29] by room-based cameras [30] or smartphone attached lasers [31–33]. Furthermore, they enable many body-tracking input modalities as head pose [6,23], body pose [34–36] and eye-gaze [37,38] tracking, using glasses- [38] or room-based trackers [39]. Although smartphones SLAM can already sufficiently track the world-pose of the smartphone, making external tracking devices obsolete in that regard, outside-in systems still lead in enabling tracking of multiple body parts of the user [9].

Objective 1 of 2: Maximize the number of enabled input modalities: Enable world-scale smartphone-based motion tracking of the user's touch, hand, head, body and eye-gaze.

2.5. Interaction Techniques

Choosing the right interaction techniques for inside-out systems is still a challenge, even for experts in this field. The main body of inside-out approaches investigated and compared different touch [23,40,41], ray-casting [19,24], plane-casting [21,42] and translation-based virtual-hand [2] or peephole [4,43] interactions, ranging from 2D pointing [19], collaboration [20], to 3D object manipulation [41]. Due to the limited input modalities, researchers need to make significant cuts into the effectiveness of their interaction techniques, although they explain that another technique would be more sufficient, however, it would require external trackers. For example, in a recent paper, Siddhpuria et al. found no practical way to detect the absolute orientation of the smartphone relative to the distant display by using IMU [19]. Therefore, they needed to use "fix origin" ray-casting [24], instead of the preferable real-world absolute one. Such compromises create a captivating pull for an increasing hardware complexity and put researchers into a difficult decision between rich interaction techniques versus setup simplicity.

Hybrid setups showed numerous distant display use cases, where knowing the body-position of the user, or the looking-direction can enable highly efficient interaction techniques (e.g., ego-centric [44], head-pointing [23], or gaze+touch [6,38]). Hybrid tracking also allowed researchers to investigate more complex user interfaces, as 3D data visualizations, which require multiple input modalities that simultaneously enable many degrees to control, for 3D object translation, rotation, scale, and also 3D camera viewpoint controls [45,46]. Furthermore, a historic overview of hybrid systems points out the importance of the smartphone's touchscreen. Due to the Midas problem [6] present in input modalities that have only one input state (e.g., hand, head, gaze, or body motions), the touchscreen played an essential role in reliably segmenting these into meaningful interaction techniques. This made smartphones to an indispensable device even in many systems with external tracking. In our taxonomy, we separated the input modalities between the two tracking categories to show the difference in the role of touch in inside-out and hybrid setups. This is mainly that inside-out systems are much more touch heavy compared to hybrid systems which focus more on the motion of the device in 3D space, body interactions and head or gaze inputs, and use touch only for clicking and clutching.

2.6. Hardware Complexity

In inside-out systems, researchers challenged the boundaries of each new smartphone adapted technology, they enabled inputs by using hardware keys [47], joysticks [25], inertial measurement unit (IMU) [1], touchscreens [5], cameras used for optical flow [3] or SLAM algorithms (i.e., simultaneous localization and mapping) [2] for distant display interaction. Researchers using hybrid setups put enormous effort into in-lab hardware setups to investigate spatial interactions with distant displays. For example, in a recent work, Hartmann et al. used six Kinect cameras (each connected to a PC), and a ten-camera Vicon motion tracking system for real-time tracking of the pose of the smartphone and the user's head

(users additionally wore a hat with IR markers) [48]. These are significantly complex hardware setups for tracking the smartphone and head pose, however very common among researchers in the domain.

From the large body of papers using outside-in trackers (e.g., vision-based lighthouse trackers), we can note that these are often in-lab systems; consequently, they are often more complex to set them up and/or replicate the results. They also often require a lot of dedicated space as they are not fully mobile, need to be calibrated carefully, require additional computational units, use a limited tracking range and field-of-view, and need to be used in a controlled environment with a defined number of users, which need to move accordingly (preventing user and camera occlusion). On a technical level, spatial interaction research needs a more practical solution for testing and refinement for situations outside the lab, to support wider-scale use and in-the-wild deployments. Researchers in academia and industry have begun to point out and tackle this infrastructure problem, but no efforts have been strongly focused on minimizing the hardware setup to enable an out-of-the-box experience [9].

Objective 2 of 2: Minimize the need for outside-in tracking for distant display scenarios: Propose an inside-out smartphone-based alternative with sufficient and comparable input capabilities.

2.7. Summary

We can see how the division of tracking technology in inside-out and hybrid setups, influences the smartphone-based distant-display interaction on many levels. None of the existing works was able to bridge the gap between the two streams of research, by providing a sufficient inside-out solution in terms of the number of tracked input modalities as well as precision. Using a smartphone for enabling distant display interactions is therefore still mainly limited to device-tracking inputs and lacks the inclusion of body-tracking inputs. Based on our analysis, we believe that the essential bundle of interaction techniques for future distant displays should consist of: touch inputs and device (hand), head, gaze, and body pose tracking. Being able to track all these input modalities without any external devices by only using an off-the-shelf smartphone would enable many powerful spatial interactions, discovered over decades of in-lab research, for everyone (e.g., head or gaze pointing, virtual-hand or peephole interactions, body-centered inputs). First steps were already made for using smartphone-based world-tracking in the domain of distant displays [2], handheld AR [49–51], head-mounted displays [52] as well as using face-tracking in the domain of cross-device [53] and on-phone interactions [54]. Using simultaneous world- and face-tracking on off-the-shelf smartphones, however, still remains unaddressed, since just recently the first examples of the technology were featured for handheld AR use-cases [55].

3. TrackPhone

In this paper, we also want to present the smartphone application TrackPhone as an iOS ARKit solution, based on our preliminary work [56], where we used two smartphones, combined in a single phone case, to mimic simultaneously world- and face-tracking before it was officially released (see Figure 2). By that, we could start initially implementing our tracking framework that allowed world-space phone (hand), head, and body pose tracking. In that paper, one phone performed world-to-device tracking (i.e., smartphone pose), while another phone was used as the device-to-user tracking (i.e., head and body pose). For TrackPhone, we adapted the two-phone approach to a single-phone implementation using the latest ARKit releases that natively support simultaneous world- and face-tracking. Furthermore, we extended our tracking framework for eye-gaze tracking. Hereby other researchers can access the single or a two-phone implementation, which is still relevant for Android phones since ARCore still does not support simultaneous tracking and therefore requires a workaround.

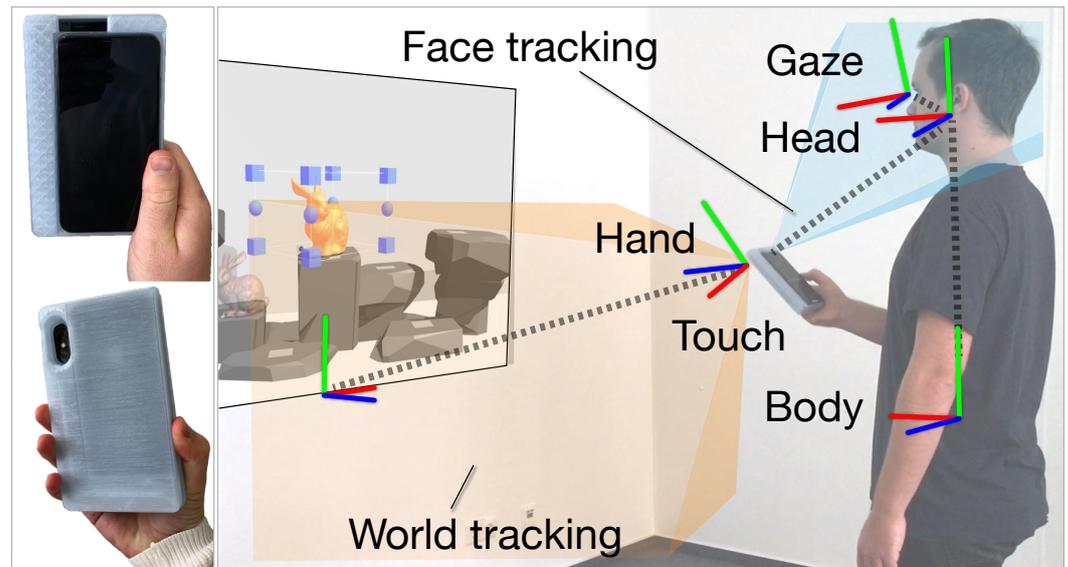


Figure 2. First prototype consisting of two smartphones used for the initial implementation of our tracking framework (left). TrackPhone tracking principle using simultaneously the front and rear camera of the smartphone for tracking touch and the absolute world-space pose of the user's device (hand), body, head, and eye-gaze (right).

TrackPhone provides tracking data that can be described in arbitrary real-world space and scale, but also as motions relative to a real-world object, for example, a distant display. To provide motion relative to a real-world object this object needs to be registered by the image tracking feature of ARKit. This procedure needs to be conducted only once since the tracking information gets saved for further usage on the device's internal world map. Otherwise, no calibration or other preparation steps are required. Users only need to download the app to their own smartphone and they can use it out-of-the-box. In our overall system, the smartphone app, sends all the tracked and processed user's inputs to the distant display computer which renders the user interface.

In total TrackPhone enables simultaneous real-time tracking of the 6DOF world-space position and orientation of the smartphone (i.e., hand) and user's head, body, and eye (i.e., gaze) motions. The touchscreen of the smartphone can be used for clutching, tapping, dragging, or other touch inputs as usual. The body pose is calculated based on the head and smartphone pose and represents a single 3D point in space—the body pose (*attached anonymized preliminary work*). This is a common approach to detect user motion or travel in a distant display scenario. For example, in outside-in systems (e.g., Vicon, Optitrack) users often wore head-mounted IR markers which represented the body's position and orientation [35,57].

When performing simultaneous face- and world-tracking from a single handheld device, we need to be aware of an accumulated tracking noise error, since any tracking noise present in world-to-device tracking will be added additionally to the device-to-user tracking noise. For example, if the user has a jittery hand, this affects the smartphone pose as well as the head, body, or eye pose. Therefore, we use the 1 Euro filter to minimize the tracking noise [58,59].

4. Evaluation

To validate our novel smartphone-based tracking approach we performed a user study to provide quantitative data about the performance of our system. Based on related work on smartphone-enabled distant display interaction, we found studies comparing: touch and head-pointing [19], touch and ray-casting [19,23], touch and a virtual-hand [29,60], and touch and gaze [38]. Therefore, we performed a unified study comparing touch, ray-casting, virtual-hand, head-pointing, against each other to gain further insights into the performance

of multi-modal interaction techniques. Since eye-tracking just got recently implemented for simultaneous world- and face-tracking use, we were not able to include gaze in this study as well. Similarly as in other works we included a dual-precision approach [6,23,61], where one primary modality or input mode is used for coarse (“suggesting”) and a second for fine-grained inputs (“refinement”). This is especially practical for selecting very small targets. Since most of the related studies investigate the previously mentioned techniques with 2D user interfaces, we included also a 3D user interface task for the distant display. We conducted two studies, in the first study, we primarily focused on a 2D pointing and selection task, while in the second study, we focused on 3D selection and manipulation (6DOF docking task). In both studies, we particularly focus on finding the overall best primary and refinement techniques, but also the best combination between them.

4.1. Apparatus

In both studies, we used TrackPhone in a projector-based setup using an Epson EH-TW5650 with a projected image size of 170×95.5 cm (1920×1080 px) as a distant display. As in similar studies were instructed participants to stand on a marked spot, centered and 2 m away from the projected image [19,23,62].

4.2. Investigated Interaction Techniques

The overall goal was to find interaction techniques that require little physical overhead and which provide high accuracy while interacting with very small target objects. In the following, we discuss four different interaction techniques (see Figure 3):

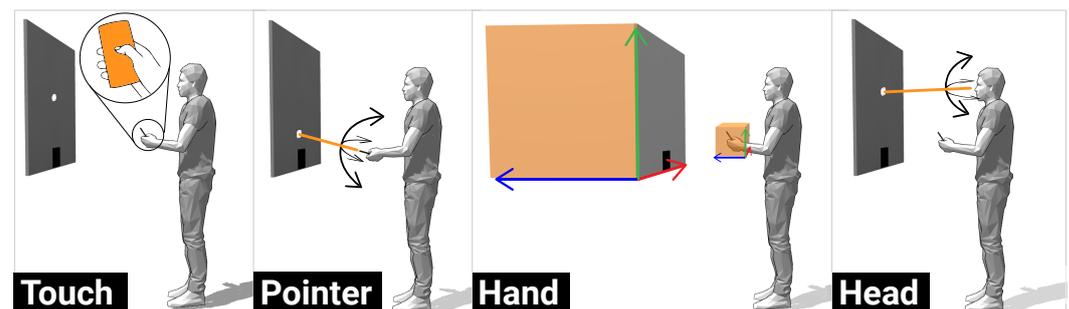


Figure 3. Study interaction techniques: touch, pointer, hand and head.

Touch: Touch maps relative user’s finger movements on the smartphone’s touchscreen to the distant display’s cursor [19,20]. The space of the distant display ($170 \text{ cm} \times 95.5 \text{ cm}$) is mapped directly to the touch area on the phone, which spans over the entire width ($6.9 \text{ cm}/1242 \text{ px}$) and using the according height ($3.9 \text{ cm}/698 \text{ px}$). The primary mode uses a 1:1 mapping, while in the refinement mode, a 2:1 mapping has been used, as proposed by Kytö et al. [61], requiring twice as much finger movement to move the cursor for the same amount as in the primary mode.

Pointer: Pointer is a ray-casting technique [19,63], where the ray originates from the smartphone and points forward. This absolute pointing principle is used in the primary mode, while the refinement mode is based on relative pointing using a ratio of 2:1. This means that starting with the current cursor position when switching to the refinement mode, only a half-degree per physically performed full-degree in smartphone rotation is applied to the ray’s orientation update.

Hand: For mapping hand motions we use a virtual-hand metaphor [2,64,65], where a virtual hand (i.e., cursor) in 2D/3D is controlled directly by the user’s hand motion. The hand’s position is tracked within a virtual control space of a $50 \times 20 \times 20 \text{ cm}$, which, in primary mode, directly corresponds to the display space of $170 \times 95.5 \times 95.5 \text{ cm}$ ($1920 \times 1080 \times 1080 \text{ px}$). Like in the previous techniques we switched to a 2:1 relative mapping for the refinement mode.

Head: The head-pointing technique works similarly to the pointer technique, with the difference that the head position is the origin of the ray and the ray direction is defined by the head's forward direction [6,23]. The refinement mode works similarly to the pointer technique.

4.2.1. 3D Interaction

While all four interaction techniques can be used for 2D user interfaces, only *Hand* directly includes 3D capabilities. The other three techniques can be extended by positioning the cursor on the x- and y-axis as defined by the interaction technique in 2D and controlling the cursor's z-axis following the principle of the Hand technique (i.e., fishing reel metaphor [66]). When using the *Pointer* technique, for example, users are tilting the device to move the cursor on the x/y plane and moving their hand forward and backward to adjust the cursor on the z-axis.

4.2.2. Touchscreen Interaction

To select a 2D target or grab a 3D object, we used the touchscreen of the device. The selection was implemented as a short finger tap (touch down-up < 250 ms, as in [19]) and was set as the fundamental selection method due to its fast, reliable, and deliberate interaction. Additionally to the tap, we also implemented a grab mechanic (touch down > 250 ms) for 3D object manipulation. As we also used a tap to switch between the primary and refinement technique, the touchscreen was therefore horizontally split evenly into two areas, as in [67]. The bottom area was used for object interaction while the upper area was used to activate the refinement mode.

4.3. Participants

In total, 18 paid volunteers (7 female, $M = 28.9$ ($SD = 6.0$) years) participated, selected from different departments of the local organization. 6 participants had intermediate experience with device ray-casting techniques (Nintendo Wii).

4.4. Study 1: 2D Pointing

The first experiment investigated the performance of TrackPhone and all four interaction techniques (*Touch*, *Pointer*, *Head*, *Hand*) while using them in a 2D pointing task in combination with a distant display. We cross-combined all techniques with each other so that each possible combination of primary technique and optional refinement technique was considered. In this study, we used explicit target selection, triggered by a touchscreen tap.

4.4.1. Task

Our study design was based on the experiment of Vogel et al. [62] and participants were required to select a circular target with a diameter of 5 cm using a cursor. The cursor was represented as a circle (diameter of 1 cm), and changed to a 2 cm large crosshair in the refinement technique. Targets were positioned randomly across the whole distant display, while the distance from the previous to the next target was always fixed at 50 cm. Before the study, we asked participants to align the cursor with the target as precisely as possible and to do so as quickly as possible. To discourage excessive effort on accuracy at the expense of time, a limit of 5 s (based on pilot study results) was placed on each trial.

4.4.2. Design

A repeated measures within-subject design was used with two factors: *primary technique* (*Touch* | *Pointer* | *Hand* | *Head*) + *refinement technique* (*None* | *Touch* | *Pointer* | *Hand* | *Head*). Due to the incompatibility of the two techniques, the pair of *Pointer+Head* was excluded. For each technique, participants completed a block of 20 target selections (in addition to 5 blocks for training). Overall, each participant completed 19 techniques × 20 blocks = 380 trials. We randomized the order of the techniques for each participant.

For each trial, we measured the time of the selection error, which is defined by the difference between the cursor and target position upon selection. After each technique, participants were asked to provide their subjective-feedback by commenting on the currently experienced primary and refinement technique regarding ease-of-use, physical, and mental demand. The subjective-feedback phase was also used as a break from the interaction to mitigate physical fatigue.

4.4.3. Results

We performed a repeated-measures ANOVA ($\alpha = 0.05$) for both time and error. When the assumption of sphericity was violated (tested with Mauchly’s test), we used the Greenhouse–Geisser corrected values in the analysis. The post hoc tests were conducted using pairwise t-tests with Bonferroni corrections. Time and error analyses included only successful target selections (6782 of 6840 total trials, or 99.2%). For the analysis, we considered all instances where a particular technique occurred as the primary mode or refinement mode, respectively. Note that all pairwise comparisons presented below are on a significance level of $p < 0.001$, unless noted differently.

Time

We found a significant main effect on time for the primary modes ($F_{3,957} = 362.39, p < 0.001$), refinement modes ($F_{4,1276} = 174.30, p < 0.001$), and the interaction between the primary*refinement—interaction techniques ($F_{12,3828} = 93.19, p < 0.001$), see Figure 4.

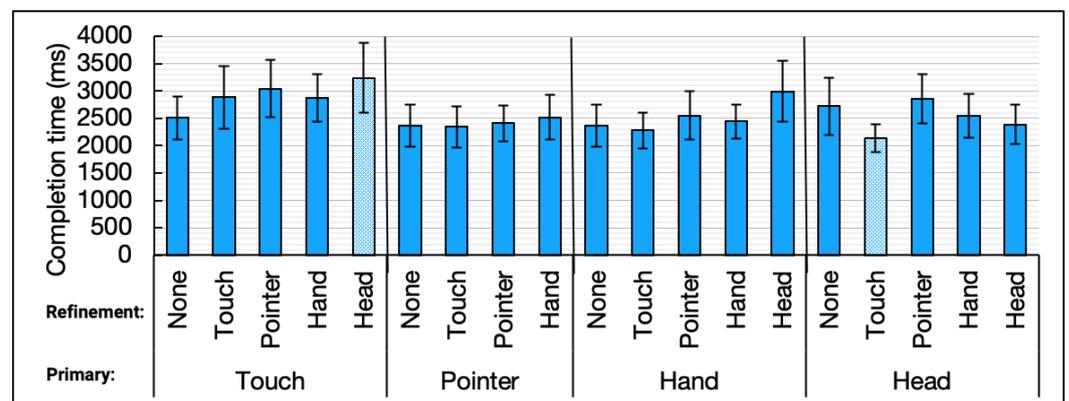


Figure 4. Mean times with different interaction techniques. Error bars represent the standard deviations and the pattern filled bars annotate the fastest and slowest technique.

For the primary mode, which supported participants by moving the cursor quickly over a large distance, we found that the *Pointer* was the fastest mode, followed by *Hand* and *Head* (equally fast), while *Touch* was the slowest, see Figure 5. For the refinement mode, which is used to refine the cursor’s final position *Touch* was the fastest mode, followed by *None*, *Hand*, *Pointer*, and *Head*, which were all significantly different from each other. We can see that some interaction techniques are great for fast and coarse pointing but are too slow for fine adjustments. *Touch* and *Pointer* are clear examples of that, while *Pointer* is the fastest as a primary technique its the second slowest for refining, and while, *Touch* is the slowest for coarse pointing it is the fastest for refining.

Study	Metric	Mode	Technique					Faster / more accurate					Slower / more inaccurate				
			Touch	Pointer	Hand	Head	None	1st	2nd	3rd	4th	5th	1st	2nd	3rd	4th	5th
2D	Time (ms)	Primary	2908 (511)	2408 (373)	2525 (404)	2528 (397)	-	Pointer	Hand/Hand	Touch	-	-	-	-	-	-	-
		Refinement	2410 (382)	2713 (433)	2594 (387)	2873 (517)	2489 (421)	Touch	None	Hand	Pointer	Head	-	-	-	-	-
	Error (cm)	Primary	0.80 (0.40)	0.79 (0.41)	0.89 (0.48)	1.00 (0.55)	-	Touch/Pointer	Hand	Head	-	-	-	-	-	-	-
		Refinement	0.69 (0.38)	0.87 (0.46)	0.75 (0.37)	0.93 (0.49)	1.14 (0.62)	Touch	Hand	Pointer/Head	None	-	-	-	-	-	-
3D	Time (s)	Primary	-	23.3 (8.06)	22.3 (7.64)	28.4 (9.59)	-	Pointer/Hand	Head	-	-	-	-	-	-	-	
		Refinement	-	23.5 (8.07)	26.5 (9.65)	26.9 (8.36)	23.0 (7.73)	Pointer/None	Hand/Head	-	-	-	-	-	-	-	

Figure 5. The study results containing the times and errors for the different interaction techniques in primary as well as refinement mode.

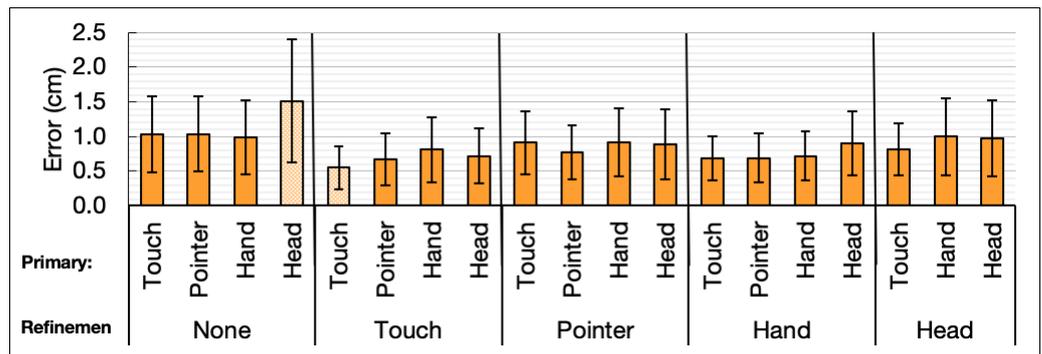


Figure 7. Mean pointing error with different interaction techniques. Error bars represent the standard deviations and the pattern-filled bars annotate the most and the least accurate technique.

In terms of accuracy, we need to point out that we were positively surprised by the overall accuracy of TrackPhone. Independently of which technique was used, participants achieved a precision level of less than 1.14 cm.

The study showed, that both *Touch* and *Pointer* were the most accurate primary modes ($p = 0.57$), followed by *Hand* and *Head*. However, unsurprisingly the refinement mode was more important for accuracy since it is used to correct the pointing error. The results show that *Touch* was the most accurate refinement mode, followed by *Hand*, *Pointer* and *Head* which were not significantly different, and *None* as the most inaccurate. From all 19 interaction techniques, we can see that techniques where *Touch* is used for refinement are highly accurate, with *Touch-Touch* being the most accurate of all.

Overall, our results show that a refinement technique can improve the accuracy for up to three times, if we take the example of *Head-None* and *Head-Touch*, see Figure 8. Our results support findings from Kytö et al. [61], who showed that refinement in head-mounted AR can improve accuracy for nearly five times, as well as the findings from Šparkov et al. [68], Chatterjee et al. [69] and Jalaliniya et al. [70], who found a threefold accuracy improvement for combining eye-gaze with hand-gestures or head-pointing for distant displays.

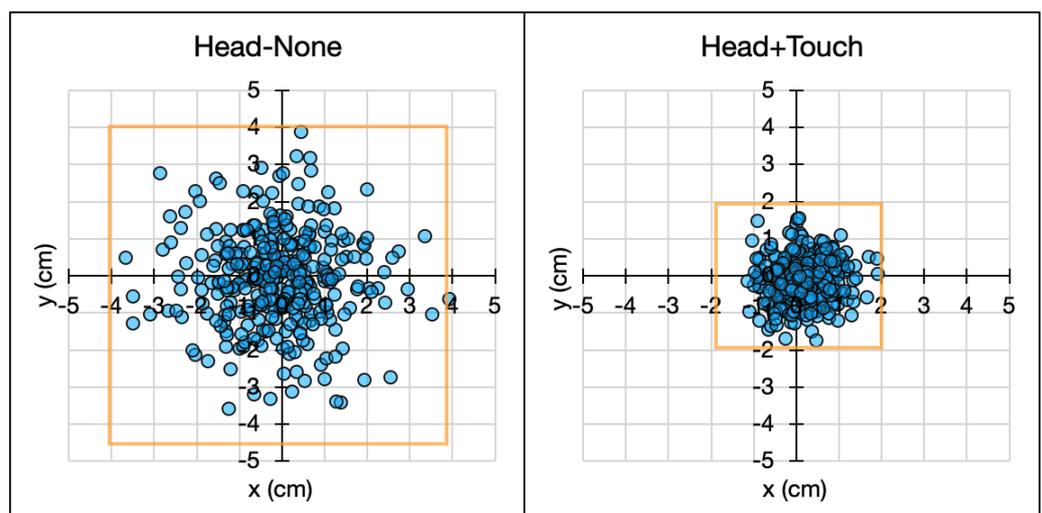


Figure 8. All selection points, showing the selection accuracy of the *Head-None* (no refinement) and *Head-Touch* (with refinement) technique. The coordinate system represents the distance from the exact target centre.

We can see that even in interactive systems where only one input modality is provided offering a refinement mode based on this modality, can significantly improve accuracy, as shown by *Head-Head*, *Touch-Touch*, *Pointer-Pointer*, *Hand-Hand*, and *Head-Head* results,

compared to their *-None counterparts. In a multi-modal system, we can achieve even higher accuracy by mixing modalities (e.g., *Head-Head* vs. *Head-Touch*). Similar results were also found by Kytö et al. [61], who found that head-head was more accurate than any other technique.

Qualitative Results

All 18 participants expressed that the combinations *Head-Touch*, *Pointer-Touch* and *Pointer-Pointer* are the best. This positive feedback was supported by comments such as P9: “*Head-Touch, this is really good, since I only need to look at the target and the cursor is already there, then I just fine-tune. For bigger targets, you would not even need a cursor.*” and P12: “*Pointer-Pointer, I like this one, since you do not need to do any touch actions, besides the tap.*” Participants further provided valuable comments explaining that touch can feel slower compared to non-touch techniques: P16: “*Touch-Touch it’s very precise, however, I need to perform multiple swipes—which makes me slow.*” They also pointed out that switching between modalities, for the refinement, can reduce performance and be mentally demanding: P9: “*Pointer-Hand, the input movement here feels the same as Pointer-Pointer, although it’s harder—I would rather stay in the same modality as Pointer-Pointer.*”, P13: “*Head-Hand, I am slow, since I need to stop rotating my head and start moving my hand. This makes me slow, and I need to actively think about doing it also.*” Finally, many participants were positively surprised about the performance of only the primary techniques, used without refinement: P6: “*Hand-Only, this is surprisingly precise.*”

4.5. Study 2: 3D Selection and Manipulation

In the second experiment, we investigated the performance of TrackPhone and our interaction techniques (*Pointer*, *Head*, *Hand*), while using them in a 3D docking task in combination with a distant display. We excluded the *Touch* techniques from this study, since the touch interaction conflicts with the 3D manipulation, where touch is required for “grabbing” 3D objects. As in the previous study, we cross-combined all primary techniques with all refinement techniques, providing us with an overall of 11 interaction techniques.

4.5.1. Task

Our study is based on the experiments of [29,71,72], where participants were required to perform a 6DOF docking task—selecting a 3D object in 3D space and aligning it with a target object matching position, rotation, and scale.

For rotating and scaling the 3D object, we used 3D widgets [73] since it is an intuitive and conventional method to manipulate 3D objects [64], while participants moved the 3D cursor, it was automatically pre-selecting the closest 3D object, or manipulation widget, as proposed by Baloup et al. [63]. The pre-selected object or widget then had to be selected. Once selected, the 3D object had to be translated, with the translation being directly mapped to the cursor. In addition, users had to select the 3D widgets to rotate/scale the object accordingly. We used axes separation [74,75] on the 3D widgets, by which the cursor’s up/down movements were scaling the 3D object uni-formally. Left/right movements were used to rotate the object around the y-axis. Similar to the first study, the cursor changed from a sphere (1 cm diameter) in the *primary mode*, to a 3D crosshair (3 × 3 × 3 cm) in the *refinement mode*. For each trial, the position of the 3D target was randomly assigned from a pre-defined list of all possible positions. The list was generated before the study and included 3D coordinates which all had a different 3D position on all axes. The rotation and scale of the 3D targets were randomized for each trial, in a way that it was always at least 90–270° differently rotated and at least 10-30 cm differently scaled than the previous target. The docking task was successfully completed once the position on each axis matched with <2 cm, the rotation differed by <4°, and the scale for <2 cm. We defined these parameters based on a previous pilot study and would like to highlight that this requires very high precision—when fulfilling these constraints, the two objects were visually perfectly aligned. Again, we asked participants to align the two 3D objects as precisely as possible and to do

so as quickly as possible. To discourage excessive effort on accuracy at the expense of time, a limit of 40 s was placed on each trial (see Figure 9).

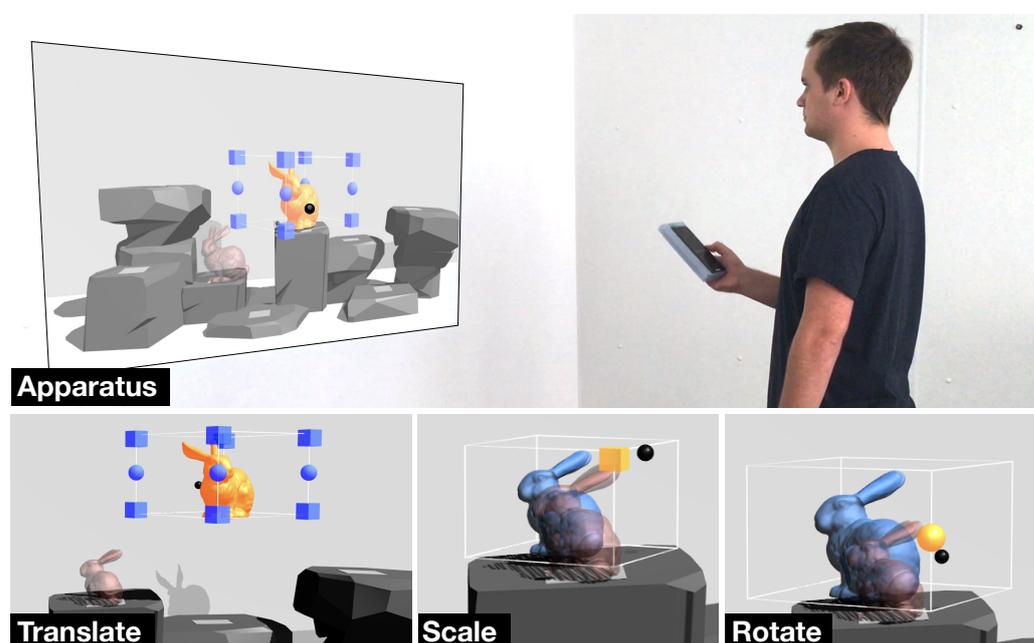


Figure 9. Real-world study apparatus for the 2D/3D study. Required interaction in the 3D study, translate, scale, and rotate.

4.5.2. Design

A repeated measures within-subject design was used with two factors: *primary technique* (*Pointer* | *Hand* | *Head*) + *refinement technique* (*None* | *Pointer* | *Hand* | *Head*). Due to the incompatibility of the two techniques, the combination *Pointer+Head* was excluded. For each technique, participants completed a block of 3 (in addition to 3 blocks for training) docking tasks. Summarizing, each participant completed 11 techniques \times 3 blocks = 33 trials. We randomized the instruction order of the techniques for each participant.

We measured the time of each trial. After each technique, participants were asked to provide their subjective-feedback by commenting on the experienced primary and refinement technique regarding ease-of-use, physical, and mental demand. The subjective feedback phase was also used as a break from the interaction to mitigate physical fatigue.

4.5.3. Results

We conducted a repeated-measures ANOVA ($\alpha = .05$) for the time. As described in the first study also here we tested our data for the assumption of sphericity and performed post hoc tests. The time analyses included only successful target selections (588 of 594 total trials, or 98.9%). Similarly to the first study, we considered all instances, where a particular technique occurred as the primary mode or refinement mode, respectively. Note that all pairwise comparisons presented below were on a significance level of $p < 0.001$ unless noted differently.

Time

We found a significant main effect on time for between the primary techniques ($F_{2,100} = 26.626, p < 0.001$), refinement techniques ($F_{3,150} = 7.535, p < 0.001$), the interaction between the primary*refinement techniques ($F_{6,300} = 2.339, p < 0.032$).

Regarding the primary techniques, we found that the *Pointer* and *Hand* conditions were equally fast and both faster than *Head*, see Figures 5 and 10. For the refinement techniques, we found that *Pointer* and *None*, were the same and faster than *Hand* and *Head*, which were also equal.

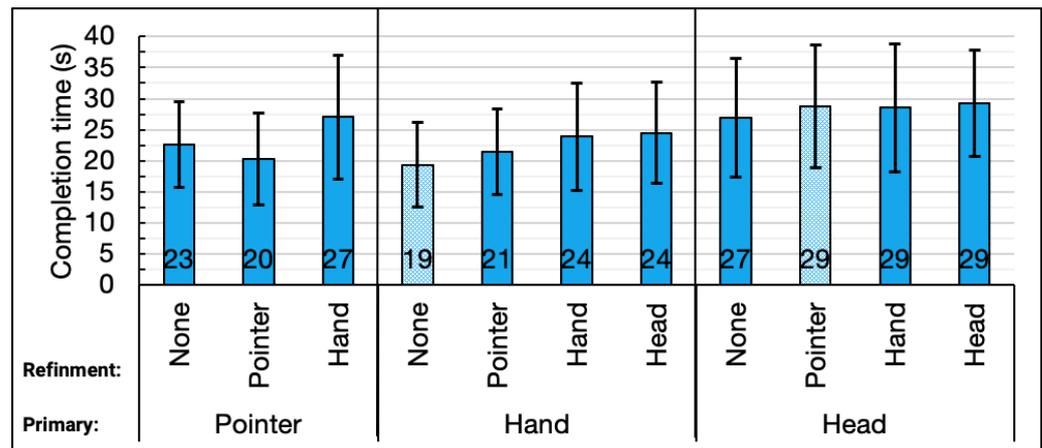


Figure 10. Mean times with different interaction techniques. Error bars represent the standard deviations and the pattern-filled bars annotate the fastest and slowest technique.

Looking again at the interaction between primary*refinement techniques, see Figure 6 (• dots), we see that *Hand* is a particularly fast technique. It does not only provide the fastest primary mode, by which participants used to quickly select and translate the 3D object, but it also provided enough accuracy to be the overall fastest technique even without any refinement (*Hand-None*). In general, we were surprised about the performance of TrackPhone and our techniques even when used without any refinement (**-None*). Although the task required a very precise 3D alignment of an object, participants easily managed it. From these results, we can conclude that the primary techniques alone allowed users enough precision to align the 3D object fast. However, we also need to take into account that the 3D widgets, object pre-selection, and separation of axes for manipulation, helped a lot to achieve such positive results.

Subjective Feedback

Participants expressed positive feedback for *Hand-None* and *Hand-Hand* by comments as P6: “*Hand-Hand, It’s very accurate, even without refinement, I would like that the refinement would be optional—triggered only once needed*”. *Pointer-Hand* was also preferred by a few participants, with arguments such as P10: “*Pointer-Hand, I like this, the other way around is however very bad*”.

5. Discussion

Overall, we can summarize that our smartphone-based tracking approach enabled interaction techniques to achieve very satisfactory performance for 2D as well as 3D interactions, even without refinement and without requiring any complex tracking hardware. This was also frequently pointed out, by many participants, which explained that the basic primary techniques are in many cases already accurate enough.

Our studies showed that by using multi-modal refinement techniques, we can improve the pointing accuracy up to three times, beyond the accuracy of the primary techniques, without compromising time. We found that some techniques are better suited for primary (e.g., *Head, Pointer*) and others for refinement (e.g., *Touch*). An in-depth comparison showed that by carefully combining different multi-modal techniques, we can create high-performing techniques, which are faster and at the same time also more accurate, than all other combinations out of the same modalities, e.g., *Head-Touch* for 2D interaction.

We can report that *Head-Touch* was overall the best 2D interaction technique as it was both the fastest and the most accurate, *Head* and *Pointer* was the best primary techniques and *Touch* was the best refinement technique. In 3D interaction, the *Hand-None* technique was the fastest, *Hand* and *Pointer* were the fastest primary techniques, and *None* and *Pointer* the fastest refinement techniques.

We also learned that using the same modality for primary interaction and refinement is a good option (e.g., *Pointer-Pointer* or *Hand-Hand*). First, this works also for uni-modal systems and second, as pointed out by our participants, switching between different kinds of 3D motions done by different parts of the body (e.g., *Pointer-Hand*, *Head-Pointer* or *Hand-Pointer*), can cause mental efforts and slow them down. Therefore, in some situations, a uni-modal combination is more appropriate.

6. Applications

In this section, we present application scenarios that show the benefits of TrackPhone in for interaction with 3D content on a 2D display. Besides the object manipulation options discussed and evaluated in the study, we also explore other means of interactions that become possible through the available tracking and interaction possibilities.

The 3D Studio is an application designed to showcase the capabilities of the interaction techniques for pointing, selection, and manipulation of 3D content as seen in the second study in a more crowded setting. As shown in Figure 11a,b, users can freely rearrange and alter the interior of a room using it as a playground to test the different interaction techniques we presented in this work.

Besides this, we were also interested in how our system could be used to actively enrich the user-perception of a displayed 3D scene [76]. Due to the lack of depth-cues 3D objects/scenes can just poorly be displayed on flat 2D screens. User-perspective rendering addresses this issue but usually requires complex hardware setups making such 3D experiences inaccessible for the main-stream [77]. Due to its head-tracking capabilities TrackPhone enables user-perspective rendering without any additional hardware and thus can significantly improve the 3D interaction and experience, see Figure 11c,d.

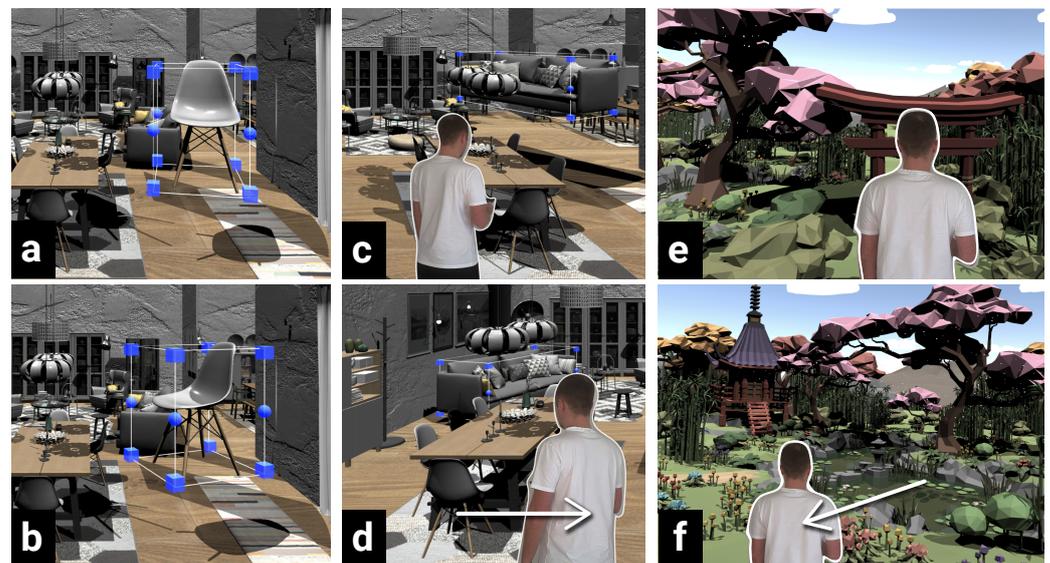


Figure 11. Applications to showcase the various utilities of TrackPhone. 3D Studio for testing interaction techniques by rearranging and manipulating furniture (a,b). Head-tracking used to create a parallax effect (c,d). Open game scene to use body and head motion for travel (e,f).

We were further interested if the body and head motion, we capture, can be used for navigation in a virtual world. In an outdoor game scene, we map the user's movement in physical space to the virtual avatar, see Figure 11e,f. This enables users to travel through the scene [78] and adapt their view, which in combination with the user-perspective rendering from before provides an even greater sense of immersion, even if the movement is restricted by the constraints of the physical room in the current implementation.

7. Conclusions and Future Work

In this paper, we first presented a taxonomy of works where smartphones were used to enable distant display interaction and exposed the open challenges of the domain. Based on the key challenges as the hardware complexity and the need for rich spatial interactions, we presented a new smartphone-based tracking approach, which in contrast to previous systems, enables the transfer of beneficial interactions only known from complex smartphone + X setups to a smartphone-only system. By using the front and back cameras of the smartphone simultaneously, we enabled world-scale motion tracking of the user's hand, head, body, as well as eye-gaze. Next, we presented a user study, to validate our tracking approach and beyond that give further insights into a variety of multi-modal 2D and 3D interaction techniques. In summary, we have found out that

- The interaction techniques enabled by our smartphone-based tracking approach achieving very satisfactory performance for 2D as well as 3D interactions, even without refinement and without requiring any complex tracking hardware.
- Using multi-modal refinement techniques, we can improve the pointing accuracy up to three times, beyond the accuracy of the primary techniques, without compromising time.
- Some concrete techniques are better suited for primary (e.g., *Head*, *Pointer*) and other for refinement (e.g., *Touch*).
- By carefully combining different multi-modal techniques, we can create concrete high-performing techniques, which are faster and at the same time also more accurate than uni-modal techniques (e.g., *Head-Touch* in 2D interaction).
- Certain combinations are less intuitive to use and can cause mental efforts and slow the interaction down (e.g., *Pointer-Hand* or *Head-Pointer*).
- Using the same modality for primary interaction and refinement, in case that only a uni-modal system can be provided, is as well a good option for improving the interaction performance (e.g., *Pointer-Pointer* or *Hand-Hand*).

Finally, we demonstrated several demo applications that show our approach in action. Overall, we presented a powerful new tool that can change the way how researchers develop, distribute (e.g., app store), test, and finally how we will interact in the future using our own smartphones with distant displays. Furthermore, we showed how state-of-the-art smartphone technology can be used to implement interactions far beyond touch or device-tilting interactions, still mainly considered today.

For future work, we would like to explore how our approach can be improved by supporting additional sensory modalities such as haptic [79] and auditory feedback, or input modalities as speech-input [55,80]. Furthermore, we would like to study the gaze ability of our system and body interactions, such as walk-by [37] and ego-centric inputs [81]. Finally, we plan to investigate other smartphone AR features as novel ways to generate and share the world-map (tracking features) among users by multi-user or online stored world-maps (e.g., cloud anchors), as well as other user tracking features like front and rear-camera skeleton tracking, 3D object scanning.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/mti6100094/s1>, Table S1. Taxonomy showing which input modalities and interaction techniques can be enabled by the inside-out (smartphone only) tracking approach for distant display. Indicated with • are tracking abilities enabled by TrackPhone and which currently lack from known related approaches. Indicated with •• are tracking abilities enabled by TrackPhone, which currently lack from known related approaches and which we investigated in our user study. Table S2. Taxonomy showing which input modalities and interaction techniques can be enabled by the hybrid (smartphone + additional hardware) tracking approach for distant display. On the bottom, we show related scenarios that could also benefit from our taxonomy and findings. References [80–255] are cited in the supplementary materials.

Author Contributions: Supervision, H.R. and M.H.; Writing original draft, T.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Dachsel, R.; Buchholz, R. Natural throw and tilt interaction between mobile phones and distant displays. In Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems—CHI EA'09, Boston, MA, USA, 4–9 April 2009; ACM Press: New York, NY, USA, 2009; p. 3253. [\[CrossRef\]](#)
2. Babic, T.; Reiterer, H.; Haller, M. Pocket6: A 6DoF Controller Based On A Simple Smartphone Application. In *Proceedings of the Symposium on Spatial User Interaction*; SUI'18; ACM Press: New York, NY, USA, 2018; pp. 2–10. [\[CrossRef\]](#)
3. Ballagas, R.; Rohs, M.; Sheridan, J.G. Sweep and point and shoot: Phocam-based interactions for large public displays. In *Proceedings of the CHI'05 Extended Abstracts on Human Factors in Computing Systems*; CHI EA'05; ACM Press: New York, NY, USA, 2005; pp. 1200–1203. [\[CrossRef\]](#)
4. Boring, S.; Baur, D.; Butz, A.; Gustafson, S.; Baudisch, P. Touch projector: Mobile interaction through video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; CHI'10; ACM Press: New York, NY, USA, 2010; pp. 2287–2296. [\[CrossRef\]](#)
5. McCallum, D.C.; Irani, P. ARC-Pad: Absolute+relative cursor positioning for large displays with a mobile touchscreen. In Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology—UIST'09, Victoria, BC, Canada, 4–7 October 2009; ACM Press: New York, NY, USA, 2009; p. 153. [\[CrossRef\]](#)
6. Stellmach, S.; Dachsel, R. 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, Paris, France, 27 April–2 May 2013; CHI'13; ACM Press: New York, NY, USA, 2013; pp. 285–294. [\[CrossRef\]](#)
7. Mäkelä, V.; Khamis, M.; Mecke, L.; James, J.; Turunen, M.; Alt, F. Pocket Transfers: Interaction Techniques for Transferring Content from Situated Displays to Mobile Devices. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; CHI'18; ACM Press: New York, NY, USA, 2009; pp. 1–13. [\[CrossRef\]](#)
8. Kister, U.; Reipschläger, P.; Matulic, F.; Dachsel, R. BodyLenses: Embodied Magic Lenses and Personal Territories for Wall Displays. In Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces—ITS'15, Madeira, Portugal, 15–18 November 2015; ACM Press: New York, NY, USA, 2015; pp. 117–126. [\[CrossRef\]](#)
9. Brudy, F.; Holz, C.; Rädle, R.; Wu, C.J.; Houben, S.; Klokmose, C.N.; Marquardt, N. Cross-Device Taxonomy: Survey, Opportunities and Challenges of Interactions Spanning Across Multiple Devices. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019; CHI'19; ACM Press: New York, NY, USA, 2019; pp. 1–28. [\[CrossRef\]](#)
10. Hinckley, K.; Pierce, J.; Sinclair, M.; Horvitz, E. Sensing techniques for mobile interaction. In Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology—UIST'00, San Diego, CA, USA, 6–8 November 2000; ACM Press: New York, NY, USA, 2000; pp. 91–100. [\[CrossRef\]](#)
11. Khamis, M.; Alt, F.; Bulling, A. The past, present, and future of gaze-enabled handheld mobile devices: Survey and lessons learned. In Proceedings of the 20th International Conference on Human–Computer Interaction with Mobile Devices and Services—MobileHCI'18, Barcelona, Spain, 3–6 September 2018; ACM Press: New York, NY, USA, 2018; pp. 1–17. [\[CrossRef\]](#)
12. Schmalstieg, D.; Höllerer, T. *Augmented Reality: Principles and Practice*; Addison-Wesley Usability and HCI Series; Addison-Wesley: Boston, MA, USA, 2016.
13. Wagner, D.; Schmalstieg, D. History and Future of Tracking for Mobile Phone Augmented Reality. In Proceedings of the 2009 International Symposium on Ubiquitous Virtual Reality, Guangju, Korea, 8–11 July 2009; pp. 7–10. [\[CrossRef\]](#)
14. Marquardt, N.; Diaz-Marino, R.; Boring, S.; Greenberg, S. The proximity toolkit: Prototyping proxemic interactions in ubiquitous computing ecologies. In Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology—UIST'11, Barbara, CA, USA, 16–19 October 2011; ACM Press: New York, NY, USA, 2011; p. 315. [\[CrossRef\]](#)
15. Hinckley, K.; Pausch, R.; Goble, J.C.; Kassell, N.F. A survey of design issues in spatial input. In *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology, Association for Computing Machinery*; UIST'94; ACM Press: New York, NY, USA, 1994; pp. 213–222. [\[CrossRef\]](#)
16. Foley, J.D.; Wallace, V.L.; Chan, P. The human factors of computer graphics interaction techniques. *IEEE Comput. Graph. Appl.* **1984**, *4*, 13–48. [\[CrossRef\]](#)
17. Ullmer, B.; Ishii, H. Emerging frameworks for tangible user interfaces. *IBM Syst. J.* **2000**, *39*, 915–931. [\[CrossRef\]](#)
18. Ballagas, R.; Rohs, M.; Sheridan, J.G.; Borchers, J. BYOD: Bring Your Own Device. In Proceedings of the Workshop on Ubiquitous Display Environments, Ubicomp, Nottingham, UK, 7–10 September 2004.
19. Siddhpuria, S.; Malacria, S.; Nancel, M.; Lank, E. Pointing at a Distance with Everyday Smart Devices. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; CHI'18; ACM Press: New York, NY, USA, 2018; pp. 1–11. [\[CrossRef\]](#)

20. Grandi, J.G.; Debarba, H.G.; Nedel, L.; Maciel, A. Design and Evaluation of a Handheld-based 3D User Interface for Collaborative Object Manipulation. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems—CHI'17, Denver, CO, USA, 6–11 May 2017; ACM Press: New York, NY, USA, 2017; pp. 5881–5891. [[CrossRef](#)]
21. Pietroszek, K.; Wallace, J.R.; Lank, E. Tiltcasting: 3D Interaction on Large Displays using a Mobile Device. In Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology—UIST'15, Charlotte, NC, USA, 11–15 November 2015; ACM Press: New York, NY, USA, pp. 57–62. [[CrossRef](#)]
22. Langner, R.; Kister, U.; Dachsel, R. Multiple Coordinated Views at Large Displays for Multiple Users: Empirical Findings on User Behavior, Movements, and Distances. *IEEE Trans. Vis. Comput. Graph.* **2019**, *25*, 608–618. [[CrossRef](#)] [[PubMed](#)]
23. Nancel, M.; Chapuis, O.; Pietriga, E.; Yang, X.D.; Irani, P.P.; Beaudouin-Lafon, M. High-precision pointing on large wall displays using small handheld devices. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI'13; ACM Press: New York, NY, USA, 2013; p. 831. [[CrossRef](#)]
24. Pietroszek, K.; Kuzminykh, A.; Wallace, J.R.; Lank, E. Smartcasting: A discount 3D interaction technique for public displays. In Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures the Future of Design—OzCHI'14; ACM Press: New York, NY, USA, 2014; pp. 119–128. [[CrossRef](#)]
25. Boring, S.; Jurmu, M.; Butz, A. Scroll, tilt or move it: Using mobile phones to continuously control pointers on large public displays. In Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7; OZCHI'09; ACM Press: New York, NY, USA, 2009; pp. 161–168. [[CrossRef](#)]
26. Durrant-Whyte, H.; Bailey, T. Simultaneous Localisation and Mapping (SLAM): Part I The Essential Algorithms. *IEEE Robot. Autom. Mag.* **2006**, *13*, 99–110. [[CrossRef](#)]
27. Google. ARCore Overview | Google Developers. Available online: <https://developers.google.com/ar/discover> (accessed on 4 May 2022).
28. Apple. ARKit—Augmented Reality. Available online: <https://developer.apple.com/augmented-reality/arkit/> (accessed on 4 May 2022).
29. Bergé, L.P.; Dubois, E.; Raynal, M. Design and Evaluation of an “Around the SmartPhone” Technique for 3D Manipulations on Distant Display. In Proceedings of the 3rd ACM Symposium on Spatial User Interaction—SUI'15; ACM Press: New York, NY, USA, 2015; pp. 69–78. [[CrossRef](#)]
30. von Zadow, U. Using Personal Devices to Facilitate Multi-user Interaction with Large Display Walls. In Proceedings of the Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology; UIST'15 Adjunct; ACM Press: New York, NY, USA, 2015; pp. 25–28. [[CrossRef](#)]
31. Myers, B.A.; Bhatnagar, R.; Nichols, J.; Peck, C.H.; Kong, D.; Miller, R.; Long, A.C. Interacting at a distance: Measuring the performance of laser pointers and other devices. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; CHI'02; ACM Press: New York, NY, USA, 2002; pp. 33–40. [[CrossRef](#)]
32. Seifert, J.; Bayer, A.; Rukzio, E. PointerPhone: Using Mobile Phones for Direct Pointing Interactions with Remote Displays. In Proceedings of the Human-Computer Interaction—INTERACT 2013; Lecture Notes in Computer Science; Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 18–35. [[CrossRef](#)]
33. Zizka, J.; Olwal, A.; Raskar, R. SpeckleSense: Fast, precise, low-cost and compact motion sensing using laser speckle. In Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology—UIST'11; ACM Press: New York, NY, USA, 2011; p. 489. [[CrossRef](#)]
34. Jansen, Y.; Schjerlund, J.; Hornbæk, K. Effects of Locomotion and Visual Overview on Spatial Memory when Interacting with Wall Displays. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems—CHI'19; ACM Press: New York, NY, USA, 2019; pp. 1–12. [[CrossRef](#)]
35. Mayer, S.; Lischke, L.; Grønbaek, J.E.; Sarsenbayeva, Z.; Vogelsang, J.; Woźniak, P.W.; Henze, N.; Jacucci, G. Pac-Many: Movement Behavior when Playing Collaborative and Competitive Games on Large Displays. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems—CHI'18; ACM Press: New York, NY, USA, 2018; pp. 1–10. [[CrossRef](#)]
36. Mayer, S.; Schwind, V.; Schweigert, R.; Henze, N. The Effect of Offset Correction and Cursor on Mid-Air Pointing in Real and Virtual Environments. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems—CHI'18; ACM Press: New York, NY, USA, 2018; pp. 1–13. [[CrossRef](#)]
37. Khamis, M.; Hoesl, A.; Klimczak, A.; Reiss, M.; Alt, F.; Bulling, A. 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; Association for Computing Machinery: New York, NY, USA, 2017; pp. 155–166. [[CrossRef](#)]
38. Turner, J.; Alexander, J.; Bulling, A.; Gellersen, H. Gaze+RST: Integrating Gaze and Multitouch for Remote Rotate-Scale-Translate Tasks. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems—CHI'15; ACM Press: New York, NY, USA, 2015; pp. 4179–4188. [[CrossRef](#)]
39. Liu, C.; Chapuis, O.; Beaudouin-Lafon, M.; Lecolinet, E.; Mackay, W.E. Effects of display size and navigation type on a classification task. In Proceedings of the 32nd annual ACM conference on Human factors in computing systems—CHI'14; ACM Press: New York, NY, USA, 2014; pp. 4147–4156. [[CrossRef](#)]
40. Debarba, H.; Nedel, L.; Maciel, A. LOP-cursor: Fast and precise interaction with tiled displays using one hand and levels of precision. In Proceedings of the 2012 IEEE Symposium on 3D User Interfaces (3DUI), Costa Mesa, CA, USA, 4–5 March 2012; pp. 125–132. [[CrossRef](#)]

41. Vinayak.; Ramanujan, D.; Piya, C.; Ramani, K. MobiSweep: Exploring Spatial Design Ideation Using a Smartphone as a Hand-held Reference Plane. In *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction—TEI'16*; ACM Press: New York, NY, USA, 2016; pp. 12–20. [[CrossRef](#)]
42. Katzakis, N.; Teather, R.J.; Kiyokawa, K.; Takemura, H. INSPECT: extending plane-casting for 6-DOF control. *Hum.-Centric Comput. Inf. Sci.* **2015**, *5*, 22. [[CrossRef](#)]
43. Ball, R.; North, C.; Bowman, D.A. Move to improve: Promoting physical navigation to increase user performance with large displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI'07*; ACM Press: New York, NY, USA, 2007; pp. 191–200. [[CrossRef](#)]
44. Rädle, R.; Jetter, H.C.; Schreiner, M.; Lu, Z.; Reiterer, H.; Rogers, Y. Spatially-aware or Spatially-agnostic? In *Elicitation and Evaluation of User-Defined Cross-Device Interactions. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems; CHI'15*; ACM Press: New York, NY, USA, 2015; pp. 3913–3922. [[CrossRef](#)]
45. Daiber, F.; Speicher, M.; Gehring, S.; Löchtfeld, M.; Krüger, A. Interacting with 3D Content on Stereoscopic Displays. In *Proceedings of the Proceedings of The International Symposium on Pervasive Displays; PerDis'14*; ACM Press: New York, NY, USA, 2014; pp. 32–37. [[CrossRef](#)]
46. Lopez, D.; Oehlberg, L.; Doger, C.; Isenberg, T. Towards An Understanding of Mobile Touch Navigation in a Stereoscopic Viewing Environment for 3D Data Exploration. *IEEE Trans. Vis. Comput. Graph.* **2016**, *22*, 1616–1629. [[CrossRef](#)] [[PubMed](#)]
47. Hardy, R.; Rukzio, E. 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 Press: New York, NY, USA, 2008; p. 245. [[CrossRef](#)]
48. Hartmann, J.; Vogel, D. An Evaluation of Mobile Phone Pointing in Spatial Augmented Reality. In *Proceedings of the Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems; CHI EA'18*; ACM Press: New York, NY, USA, 2018; pp. 1–6. [[CrossRef](#)]
49. Anderegg, R.; Ciccone, L.; Sumner, R.W. PuppetPhone: Puppeteering virtual characters using a smartphone. In *Proceedings of the 11th Annual International Conference on Motion, Interaction, and Games—MIG'18*; ACM Press: New York, NY, USA, 2018; pp. 1–6. [[CrossRef](#)]
50. Wacker, P.; Nowak, O.; Voelker, S.; Borchers, J. ARPen: Mid-Air Object Manipulation Techniques for a Bimanual AR System with Pen & Smartphone. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems—CHI'19*; ACM Press: New York, NY, USA, 2019; pp. 1–12. [[CrossRef](#)]
51. Wacker, P.; Wagner, A.; Voelker, S.; Borchers, J. Heatmaps, Shadows, Bubbles, Rays: Comparing Mid-Air Pen Position Visualizations in Handheld AR. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems—CHI'20*; ACM Press: New York, NY, USA, 2020; p. 11.
52. Mohr, P.; Tatzgern, M.; Langlotz, T.; Lang, A.; Schmalstieg, D.; Kalkofen, D. TrackCap: Enabling Smartphones for 3D Interaction on Mobile Head-Mounted Displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems—CHI'19*; ACM Press: New York, NY, USA, 2019; pp. 1–11. [[CrossRef](#)]
53. Voelker, S.; Hueber, S.; Holz, C.; Remy, C.; Marquardt, N. GazeConduits: Calibration-Free Cross-Device Collaboration through Gaze and Touch. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; CHI'20*; ACM Press: New York, NY, USA, 2020; pp. 1–10. [[CrossRef](#)]
54. Voelker, S.; Hueber, S.; Corsten, C.; Remy, C. HeadReach: Using Head Tracking to Increase Reachability on Mobile Touch Devices. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; CHI'20*; ACM Press: New York, NY, USA, 2020; pp. 1–12. [[CrossRef](#)]
55. Mayer, S.; Laput, G.; Harrison, C. Enhancing Mobile Voice Assistants with WorldGaze. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; CHI'20*; ACM Press: New York, NY, USA, 2020; pp. 1–10. [[CrossRef](#)]
56. Babic, T.; Perteneder, F.; Reiterer, H.; Haller, M. Simo: Interactions with Distant Displays by Smartphones with Simultaneous Face and World Tracking. In *Proceedings of the Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems; CHI EA '20*; ACM Press: New York, NY, USA, 2020; p. 1–12. [[CrossRef](#)]
57. Jakobsen, M.R.; Hornbæk, K. Is Moving Improving?: Some Effects of Locomotion in Wall-Display Interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems—CHI'15*; ACM Press: New York, NY, USA, 2015; pp. 4169–4178. [[CrossRef](#)]
58. Casiez, G.; Roussel, N.; Vogel, D. 1 € filter: A simple speed-based low-pass filter for noisy input in interactive systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; CHI'12*; ACM Press: New York, NY, USA, 2012; pp. 2527–2530. [[CrossRef](#)]
59. Casiez, G.; Roussel, N.; Vogel, D. 1€ Filter. Available online: <https://crystal.univ-lille.fr/~casiez/1euro/> (accessed on 28 April 2020).
60. Jain, M.; Cockburn, A.; Madhvanath, S. Comparison of Phone-Based Distal Pointing Techniques for Point-Select Tasks. In *Proceedings of the Human-Computer Interaction—INTERACT 2013; Lecture Notes in Computer Science; Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 714–721. [[CrossRef](#)]*
61. Kytö, M.; Ens, B.; Piumsomboon, T.; Lee, G.A.; Billinghurst, M. Pinpointing: Precise Head- and Eye-Based Target Selection for Augmented Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems—CHI'18*; ACM Press: New York, NY, USA, 2018; pp. 1–14. [[CrossRef](#)]

62. Vogel, D.; Balakrishnan, R. 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—UIST'05*; ACM Press: New York, NY, USA, 2005; p. 33. [[CrossRef](#)]
63. Baloup, M.; Pietrzak, T.; Casiez, G. RayCursor: A 3D Pointing Facilitation Technique based on Raycasting. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems—CHI'19*; ACM Press: New York, NY, USA, 2019; pp. 1–12. [[CrossRef](#)]
64. LaViola, J.J.; Kruijff, E.; McMahan, R.P.; Bowman, D.A.; Poupyrev, I. *3D User Interfaces: Theory and Practice*; Addison-Wesley: Boston, MA, USA, 2017.
65. Poupyrev, I.; Billingham, M.; Weghorst, S.; Ichikawa, T. The Go-Go Interaction Technique: Non-linear Mapping for Direct Manipulation in VR. In *Proceedings of the The 31st Annual ACM Symposium on User Interface Software and Technology—UIST'96*; ACM Press: New York, NY, USA, 1996; pp. 79–80.
66. Bowman, D.A.; Hodges, L.F. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics; I3D '97*; ACM Press: New York, NY, USA, 1997; p. 35–ff. [[CrossRef](#)]
67. Nancel, M.; Wagner, J.; Pietriga, E.; Chapuis, O.; Mackay, W. Mid-air pan-and-zoom on wall-sized displays. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems—CHI'11*; ACM Press: New York, NY, USA, 2011; p. 177. [[CrossRef](#)]
68. Špakov, O.; Isokoski, P.; Majaranta, P. Look and lean: Accurate head-assisted eye pointing. In *Proceedings of the Symposium on Eye Tracking Research and Applications—ETRA'14*; ACM Press: New York, NY, USA, 2014; pp. 35–42. [[CrossRef](#)]
69. Chatterjee, I.; Xiao, R.; Harrison, C. Gaze+Gesture: Expressive, Precise and Targeted Free-Space Interactions. In *Proceedings of the 2015 ACM International Conference on Multimodal Interaction—ICMI'15*; ACM Press: New York, NY, USA, 2015; pp. 131–138. [[CrossRef](#)]
70. Jalaliniya, S.; Mardanbegi, D.; Pederson, T. MAGIC pointing for eyewear computers. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers; ISWC '15*; ACM Press: New York, NY, USA, 2015; pp. 155–158. [[CrossRef](#)]
71. Besançon, L.; Issartel, P.; Ammi, M.; Isenberg, T. Mouse, Tactile, and Tangible Input for 3D Manipulation. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems—CHI'17*; ACM Press: New York, NY, USA, 2017; pp. 4727–4740. [[CrossRef](#)]
72. Zhai, S.; Milgram, P. Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI'98*; ACM Press: New York, NY, USA, 1998; pp. 320–327. [[CrossRef](#)]
73. Houde, S. Iterative design of an interface for easy 3-D direct manipulation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI'92*; ACM Press: New York, NY, USA, 1992; pp. 135–142. [[CrossRef](#)]
74. Cohé, A.; Dècle, F.; Hachet, M. tBox: A 3D transformation widget designed for touch-screens. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems—CHI'11*; ACM Press: New York, NY, USA, 2011; p. 3005. [[CrossRef](#)]
75. Martinet, A.; Casiez, G.; Grisoni, L. The effect of DOF separation in 3D manipulation tasks with multi-touch displays. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology—VRST'10*; ACM Press: New York, NY, USA, 2010; p. 111. [[CrossRef](#)]
76. Ware, C.; Arthur, K.; Booth, K.S. Fish tank virtual reality. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI'93*; ACM Press: New York, NY, USA, 1993; pp. 37–42. [[CrossRef](#)]
77. Lee, J.C. Hacking the Nintendo Wii Remote. *IEEE Pervasive Comput.* **2008**, *7*, 39–45. [[CrossRef](#)]
78. Ware, C.; Fleet, D. Integrating flying and fish tank metaphors with cyclopean scale. In *Proceedings of the Proceedings Computer Graphics International, Hasselt and Diepenbeek, Belgium, 23–27 June 1997*; IEEE: Piscataway, NJ, USA, 1997; pp. 39–46. [[CrossRef](#)]
79. Prouzeau, A.; Cordeil, M.; Robin, C.; Ens, B.; Thomas, B.H.; Dwyer, T. Scaptics and Highlight-Planes: Immersive Interaction Techniques for Finding Occluded Features in 3D Scatterplots. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems—CHI'19*; ACM Press: New York, NY, USA, 2019; pp. 1–12. [[CrossRef](#)]
80. Ding, Y.; Zhang, Y.; Xiao, M.; Deng, Z. A Multifaceted Study on Eye Contact based Speaker Identification in Three-party Conversations. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*; ACM: New York, NY, USA, 2017; pp. 3011–3021. [[CrossRef](#)]
81. Rädle, R.; Jetter, H.C.; Butscher, S.; Reiterer, H. The effect of egocentric body movements on users' navigation performance and spatial memory in zoomable user interfaces. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces—ITS'13*; ACM Press: New York, NY, USA, 2013; pp. 23–32. [[CrossRef](#)]
82. Hartmann, J.; Holz, C.; Ofek, E.; Wilson, A.D. RealityCheck: Blending Virtual Environments with Situated Physical Reality. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; CHI'19*; ACM Press: New York, NY, USA, 2019; pp. 1–12.
83. Casas, L.; Ciccone, L.; Çimen, G.; Wiedemann, P.; Fauconneau, M.; Sumner, R.W.; Mitchell, K. Multi-reality games: An experience across the entire reality-virtuality continuum. In *Proceedings of the 16th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry—VRCAI'18*; ACM Press: New York, NY, USA, 2018; pp. 1–4. [[CrossRef](#)]
84. Ng, P.C.; She, J.; Jeon, K.E.; Baldauf, M. When Smart Devices Interact with Pervasive Screens: A Survey. *ACM Trans. Multimed. Comput. Commun. Appl. (TOMM)* **2017**, *13*, 1–23. [[CrossRef](#)]

85. Khamis, M.; Alt, F.; Bulling, A. Challenges and design space of gaze-enabled public displays. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct—UbiComp '16*; ACM Press: New York, NY, USA, 2016; pp. 1736–1745. [[CrossRef](#)]
86. Arth, C.; Grasset, R.; Gruber, L.; Langlotz, T.; Mulloni, A.; Schmalstieg, D.; Wagner, D. The History of Mobile Augmented Reality. *arXiv* **2015**, arXiv:1505.01319.
87. Scoditti, A.; Blanch, R.; Coutaz, J. A novel taxonomy for gestural interaction techniques based on accelerometers. In *Proceedings of the 15th International Conference on Intelligent User Interfaces—IUI'11*; ACM Press: New York, NY, USA, 2011; p. 63. [[CrossRef](#)]
88. Kruijff, E.; Swan, J.E.; Feiner, S. Perceptual issues in augmented reality revisited. In *Proceedings of the 2010 IEEE International Symposium on Mixed and Augmented Reality*, Seoul, Korea, 13–16 October 2010; pp. 3–12. [[CrossRef](#)]
89. Rohs, M. *Linking Physical and Virtual Worlds with Visual Markers and Handheld Devices*; ACM Press: New York, NY, USA, 2005; p. 236.
90. Ostkamp, M.; Heitmann, S.; Kray, C. Short-range optical interaction between smartphones and public displays. In *Proceedings of the 4th International Symposium on Pervasive Displays—PerDis'15*; ACM Press: New York, NY, USA, 2015; pp. 39–46. [[CrossRef](#)]
91. Baldauf, M.; Salo, M.; Suetterle, S.; Fröhlich, P. Display pointing: A qualitative study on a recent screen pairing technique for smartphones. In *Proceedings of the 27th International BCS Human Computer Interaction Conference*; BCS-HCI'13; BCS Learning & Development Ltd.: London, UK, 2013; pp. 1–6.
92. Alt, F.; Shirazi, A.S.; Kubitzka, T.; Schmidt, A. 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 Press: New York, NY, USA, 2013; pp. 1709–1718. [[CrossRef](#)]
93. Kray, C.; Nesbitt, D.; Dawson, J.; Rohs, M. User-defined gestures for connecting mobile phones, public displays, and tabletops. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*; MobileHCI'10; ACM Press: New York, NY, USA, 2010; pp. 239–248. [[CrossRef](#)]
94. Myers, B.A.; Stiel, H.; Gargiulo, R. Collaboration using multiple PDAs connected to a PC. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work*; CSCW'98; ACM Press: New York, NY, USA, 1998; pp. 285–294. [[CrossRef](#)]
95. Jann, F.; Reinhold, S.; Teistler, M. Nutzung eines Smartphones als virtuelle Sonde im medizinischen Ultraschalltraining: Six-Degrees-of-Freedom-Tracking mittels ARCore. In *Proceedings of the Proceedings of Mensch und Computer 2019*; MuC'19; ACM Press: New York, NY, USA, 2019; pp. 759–763. [[CrossRef](#)]
96. Raynal, M.; Gauffre, G. Tactile camera vs. tangible camera: Taking advantage of small physical artefacts to navigate into large data collection 2010. In *NordiCHI '10: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*; ACM Press: New York, NY, USA, 2010; p. 10.
97. Fuentes-Pacheco, J.; Ruiz-Ascencio, J.; Rendón-Mancha, J.M. Visual simultaneous localization and mapping: A survey. *Artif. Intell. Rev.* **2015**, *43*, 55–81. [[CrossRef](#)]
98. Pattanakimhun, P.; Chinthammit, W.; Chotikakamthorn, N. Evaluation of mobile phone interaction with large public displays. In *Proceedings of the 2017 14th International Joint Conference on Computer Science and Software Engineering (IJCSE)*, Nakhon Si Thammarat, Thailand, 12–14 July 2017; pp. 1–6. [[CrossRef](#)]
99. Pattanakimhun, P.; Chinthammit, W.; Chotikakamthorn, N. Enhanced engagement with public displays through mobile phone interaction. In *Proceedings of the SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications*; SA'17; ACM Press: New York, NY, USA, 2017; pp. 1–5. [[CrossRef](#)]
100. Liu, C.; Chapuis, O.; Beaudouin-Lafon, M.; Lecolinet, E. CoReach: Cooperative Gestures for Data Manipulation on Wall-sized Displays. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*; CHI'17; ACM Press: New York, NY, USA, 2017; pp. 6730–6741. [[CrossRef](#)]
101. Besancon, L.; Issartel, P.; Ammi, M.; Isenberg, T. Hybrid Tactile/Tangible Interaction for 3D Data Exploration. *IEEE Trans. Vis. Comput. Graph.* **2017**, *23*, 881–890. [[CrossRef](#)] [[PubMed](#)]
102. von Zadow, U.; Siegel, A.; Dachselt, R. Multi-user Multi-device Interaction with Large Displays at the Point of Sale: An Application Case. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces—ITS'15*; ACM Press: New York, NY, USA, 2015; pp. 343–348. [[CrossRef](#)]
103. Song, P.; Goh, W.B.; Fu, C.W.; Meng, Q.; Heng, P.A. WYSIWYF: Exploring and annotating volume data with a tangible handheld device. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems—CHI'11*; ACM Press: New York, NY, USA, 2011; p. 1333. [[CrossRef](#)]
104. Du, Y.; Ren, H.; Pan, G.; Li, S. Tilt & touch: Mobile phone for 3D interaction. In *UbiComp '11: Proceedings of the 13th International Conference on Ubiquitous Computing*; ACM Press: New York, NY, USA, 2013; p. 2.
105. Bragdon, A.; DeLine, R.; Hinckley, K.; Morris, M.R. Code space: Touch + air gesture hybrid interactions for supporting developer meetings. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*; ITS'11; ACM Press: New York, NY, USA, 2011; pp. 212–221. [[CrossRef](#)]
106. Bauer, J.; Thelen, S.; Ebert, A. Using smart phones for large-display interaction. In *Proceedings of the 2011 International Conference on User Science and Engineering (i-USER)*, Selangor, Malaysia, 29 November–1 December 2011; pp. 42–47. [[CrossRef](#)]
107. Paay, J.; Raptis, D.; Kjeldskov, J.; Skov, M.B.; Ruder, E.V.; Lauridsen, B.M. Investigating Cross-Device Interaction between a Handheld Device and a Large Display. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems—CHI'17*; ACM Press: New York, NY, USA, 2017; pp. 6608–6619. [[CrossRef](#)]

108. Boring, S.; Gehring, S.; Wiethoff, A.; Blöckner, A.M.; Schöning, J.; Butz, A. Multi-user interaction on media facades through live video on mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; CHI'11*; ACM Press: New York, NY, USA, 2011; pp. 2721–2724. [[CrossRef](#)]
109. Fan, M.; Patterson, D.; Shi, Y. When camera meets accelerometer: A novel way for 3d interaction of mobile phone. In *MobileHCI '12: Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services Companion*; ACM Press: New York, NY, USA, 2012; p. 6.
110. Henderson, J.; Mizobuchi, S.; Li, W.; Lank, E. Exploring Cross-Modal Training via Touch to Learn a Mid-Air Marking Menu Gesture Set. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services; MobileHCI'19*; ACM Press: New York, NY, USA, 2019; pp. 1–9. [[CrossRef](#)]
111. Speicher, M.; Daiber, F.; Gehring, S.; Krüger, A. Exploring 3D manipulation on large stereoscopic displays. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays—PerDis'16*; ACM Press: New York, NY, USA, 2016; pp. 59–66. [[CrossRef](#)]
112. Baldauf, M.; Adegeye, F.; Alt, F.; Harms, J. Your browser is the controller: Advanced web-based smartphone remote controls for public screens. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays—PerDis'16*; ACM Press: New York, NY, USA, 2016; pp. 175–181. [[CrossRef](#)]
113. Katzakis, N.; Kiyokawa, K.; Takemura, H. Plane-casting: 3D cursor control with a SmartPhone. In *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction—APCHI'13*; ACM Press: New York, NY, USA, 2013; pp. 199–200. [[CrossRef](#)]
114. Graf, H.; Jung, D.K. The Smartphone as a 3D Input Device. In *Proceedings of the IEEE Second International Conference on Consumer Electronics—Berlin (ICCE-Berlin)*, Berlin, Germany, 3–5 September 2012; p. 4.
115. Jeon, S.; Hwang, J.; Kim, G.J.; Billingham, M. Interaction with large ubiquitous displays using camera-equipped mobile phones. *Pers. Ubiquitous Comput.* **2010**, *14*, 83–94. [[CrossRef](#)]
116. Katzakis, N.; Hori, M. Mobile Phones as 3-DOF Controllers: A Comparative Study. In *Proceedings of the 2009 Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing*, Chengdu, China, 12–14 December 2009; pp. 345–349. [[CrossRef](#)]
117. Scheible, J.; Ojala, T.; Coulton, P. MobiToss: A novel gesture based interface for creating and sharing mobile multimedia art on large public displays. In *Proceedings of the 16th ACM International Conference on Multimedia; MM'08*; ACM Press: New York, NY, USA, 2008; pp. 957–960. [[CrossRef](#)]
118. Vincent, T.; Nigay, L.; Kurata, T. Handheld Augmented Reality: Effect of registration jitter on cursor-based pointing techniques. In *Proceedings of the 25th ICME Conference Francophone on l'Interaction Homme-Machine—IHM'13*; ACM Press: New York, NY, USA, 2013; pp. 1–6. [[CrossRef](#)]
119. Zhong, Y.; Li, X.; Fan, M.; Shi, Y. Doodle space: Painting on a public display by cam-phone. In *Proceedings of the 2009 workshop on Ambient media computing—AMC'09*; ACM Press: New York, NY, USA, 2009; p. 13. [[CrossRef](#)]
120. Bellino, A.; Cabitza, F.; De Michelis, G.; De Paoli, F. Touch&Screen: Widget collection for large screens controlled through smartphones. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia—MUM'16*; ACM Press: New York, NY, USA, 2016; pp. 25–35. [[CrossRef](#)]
121. Vepsäläinen, J.; Di Rienzo, A.; Nelimarkka, M.; Ojala, J.A.; Savolainen, P.; Kuikkaniemi, K.; Tarkoma, S.; Jacucci, G. Personal Device as a Controller for Interactive Surfaces: Usability and Utility of Different Connection Methods. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces; ITS'15*; ACM Press: New York, NY, USA, 2015; pp. 201–204. [[CrossRef](#)]
122. Bergé, L.P.; Serrano, M.; Perelman, G.; Dubois, E. Exploring smartphone-based interaction with overview+detail interfaces on 3D public displays. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services—MobileHCI'14*; ACM Press: New York, NY, USA, 2014; pp. 125–134. [[CrossRef](#)]
123. Turner, J.; Bulling, A.; Alexander, J.; Gellersen, H. 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 Press: New York, NY, USA, 2013; pp. 1–4. [[CrossRef](#)]
124. Häkkinen, J.R.; Posti, M.; Schneegass, S.; Alt, F.; Gultekin, K.; Schmidt, A. Let me catch this! experiencing interactive 3D cinema through collecting content with a mobile phone. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; CHI'14*; ACM Press: New York, NY, USA, 2014; pp. 1011–1020. [[CrossRef](#)]
125. Kaviani, N.; Finke, M.; Fels, S.; Lea, R.; Wang, H. What goes where?: Designing interactive large public display applications for mobile device interaction. In *Proceedings of the First International Conference on Internet Multimedia Computing and Service—ICIMCS'09*; ACM Press: New York, NY, USA, 2009; p. 129. [[CrossRef](#)]
126. Langner, R.; Kister, U.; Satkowski, M.; Dachsel, R. Combining Interactive Large Displays and Smartphones to Enable Data Analysis from Varying Distances. In *MultimodalVis '18 Workshop at AVI 2018*; ACM Press: New York, NY, USA, 2018; p. 4.
127. Stellmach, S.; Dachsel, R. Look & touch: Gaze-supported target acquisition. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems—CHI'12*; ACM Press: New York, NY, USA, 2012; p. 2981. [[CrossRef](#)]
128. Sollich, H.; von Zadow, U.; Pietzsch, T.; Tomancak, P.; Dachsel, R. Exploring Time-dependent Scientific Data Using Spatially Aware Mobiles and Large Displays. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces; ISS'16*; ACM Press: New York, NY, USA, 2016; pp. 349–354. [[CrossRef](#)]
129. Kister, U.; Klamka, K.; Tominski, C.; Dachsel, R. GraSp: Combining Spatially-aware Mobile Devices and a Display Wall for Graph Visualization and Interaction. *Comput. Graph. Forum* **2017**, *36*, 503–514. [[CrossRef](#)]

130. Peck, C.H. Useful parameters for the design of laser pointer interaction techniques. In *Proceedings of the CHI'01 Extended Abstracts on Human Factors in Computing Systems*; CHI EA'01; ACM Press: New York, NY, USA, 2001; pp. 461–462. [[CrossRef](#)]
131. Jiang, Z.; Han, J.; Qian, C.; Xi, W.; Zhao, K.; Ding, H.; Tang, S.; Zhao, J.; Yang, P. VADS: Visual attention detection with a smartphone. In *Proceedings of the IEEE INFOCOM 2016—The 35th Annual IEEE International Conference on Computer Communications*, San Francisco, CA, USA, 10–14 April 2016; pp. 1–9. [[CrossRef](#)]
132. Langner, R.; Dachselt, R. Towards Visual Data Exploration at Wall-Sized Displays by Combining Physical Navigation with Spatially-Aware Devices. In *Proceedings of the IEEE VIS 2018 Poster Program*, Berlin, Germany, 21–26 October 2018; p. 2.
133. Le, H.V.; Mayer, S.; Henze, N. InfiniTouch: Finger-Aware Interaction on Fully Touch Sensitive Smartphones. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology—UIST'18*; ACM Press: New York, NY, USA, 2018; pp. 779–792. [[CrossRef](#)]
134. Alvina, J.; Griggio, C.F.; Bi, X.; Mackay, W.E. CommandBoard: Creating a General-Purpose Command Gesture Input Space for Soft Keyboard. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology—UIST'17*; ACM Press: New York, NY, USA; pp. 17–28. [[CrossRef](#)]
135. Gupta, A.; Anwar, M.; Balakrishnan, R. Porous Interfaces for Small Screen Multitasking using Finger Identification. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology—UIST'16*; ACM Press: New York, NY, USA, 2016; pp. 145–156. [[CrossRef](#)]
136. Wagner, D.; Schmalstieg, D.; Bischof, H. Multiple target detection and tracking with guaranteed framerates on mobile phones. In *Proceedings of the 2009 8th IEEE International Symposium on Mixed and Augmented Reality*, Orlando, FL, USA, 19–22 October 2009; pp. 57–64. [[CrossRef](#)]
137. Surale, H.B.; Gupta, A.; Hancock, M.; Vogel, D. TabletInVR: Exploring the Design Space for Using a Multi-Touch Tablet in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems—CHI'19*; ACM Press: New York, NY, USA, 2019; pp. 1–13. [[CrossRef](#)]
138. Pfeuffer, K.; Gellersen, H. Gaze and Touch Interaction on Tablets. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology—UIST'16*; ACM Press: New York, NY, USA, 2016; pp. 301–311. [[CrossRef](#)]
139. Khamis, M.; Baier, A.; Henze, N.; Alt, F.; Bulling, A. Understanding Face and Eye Visibility in Front-Facing Cameras of Smartphones used in the Wild. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems—CHI'18*; ACM Press: New York, NY, USA, 2018; pp. 1–12. [[CrossRef](#)]
140. Zhang, X.; Kulkarni, H.; Morris, M.R. Smartphone-Based Gaze Gesture Communication for People with Motor Disabilities. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*; ACM: New York, NY, USA, 2017; pp. 2878–2889. [[CrossRef](#)]
141. Krafka, K.; Khosla, A.; Kellnhofer, P.; Kannan, H.; Bhandarkar, S.; Matusik, W.; Torralba, A. Eye Tracking for Everyone. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 27–30 June 2016; pp. 2176–2184. [[CrossRef](#)]
142. Mariakakis, A.; Goel, M.; Aumi, M.T.I.; Patel, S.N.; Wobbrock, J.O. SwitchBack: Using Focus and Saccade Tracking to Guide Users' Attention for Mobile Task Resumption. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems—CHI'15*; ACM Press: New York, NY, USA, 2015; pp. 2953–2962. [[CrossRef](#)]
143. Rivu, R.; Hassib, M.; Abdrabou, Y.; Alt, F.; Pfeuffer, K. Gaze'N'Touch: Enhancing Text Selection on Mobile Devices Using Gaze. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–8. [[CrossRef](#)]
144. Lee, J.L.; Kim, S.; Fukumoto, M.; Lee, B. Reflector: Distance-Independent, Private Pointing on a Reflective Screen. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology—UIST'17*; ACM Press: New York, NY, USA, 2017; pp. 351–364. [[CrossRef](#)]
145. Cicek, M.; Xie, J.; Wang, Q.; Piramuthu, R. Mobile Head Tracking for eCommerce and Beyond. *Electron. Imaging* **2020**, *2020*, 303. [[CrossRef](#)]
146. Seyed, T.; Yang, X.D.; Vogel, D. A Modular Smartphone for Lending. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology—UIST'17*; ACM Press: New York, NY, USA, 2017; pp. 205–215. [[CrossRef](#)]
147. Serrano, M.; Lecolinet, E.; Guiard, Y. Bezel-Tap gestures: Quick activation of commands from sleep mode on tablets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI'13*; ACM Press: New York, NY, USA, 2013; p. 3027. [[CrossRef](#)]
148. Gotsch, D.; Zhang, X.; Burstyn, J.; Vertegaal, R. HoloFlex: A Flexible Holographic Smartphone with Bend Input. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems—CHI EA'16*; ACM Press: New York, NY, USA, 2016; pp. 3675–3678. [[CrossRef](#)]
149. Lv, Z.; Feng, S.; Khan, M.S.L.; Ur Réhman, S.; Li, H. Foot motion sensing: Augmented game interface based on foot interaction for smartphone. In *Proceedings of the Extended Abstracts of the 32nd Annual ACM Conference on Human Factors in Computing Systems—CHI EA'14*; ACM Press: New York, NY, USA, 2014; pp. 293–296. [[CrossRef](#)]
150. Yamada, W.; Manabe, H.; Ikeda, D. CamTrackPoint: Camera-Based Pointing Stick Using Transmitted Light through Finger. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology—UIST'18*; ACM Press: New York, NY, USA, 2018; pp. 313–320. [[CrossRef](#)]

151. Bai, H.; Gao, L.; El-Sana, J.; Billinghurst, M. Free-hand interaction for handheld augmented reality using an RGB-depth camera. In *Proceedings of the SIGGRAPH Asia 2013 Symposium on Mobile Graphics and Interactive Applications on—SA'13*; ACM Press: New York, NY, USA, 2013; pp. 1–4. [[CrossRef](#)]
152. Arslan, C.; Rekik, Y.; Grisoni, L. E-Pad: Large Display Pointing in a Continuous Interaction Space around a Mobile Device. In *Proceedings of the 2019 on Designing Interactive Systems Conference; DIS'19*; ACM Press: New York, NY, USA, 2019; pp. 1101–1108. [[CrossRef](#)]
153. Hasan, K.; Ahlström, D.; Kim, J.; Irani, P. AirPanes: Two-Handed Around-Device Interaction for Pane Switching on Smartphones. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*; ACM: New York, NY, USA, 2017; pp. 679–691. [[CrossRef](#)]
154. Lakatos, D.; Blackshaw, M.; Olwal, A.; Barryte, Z.; Perlin, K.; Ishii, H. T(ether): Spatially-aware handhelds, gestures and proprioception for multi-user 3D modeling and animation. In *Proceedings of the 2nd ACM Symposium on Spatial User Interaction—SUI'14*; ACM Press: New York, NY, USA, 2014; pp. 90–93. [[CrossRef](#)]
155. Tsai, H.R.; Wu, T.Y.; Huang, D.Y.; Hsiu, M.C.; Hsiao, J.C.; Hung, Y.P.; Chen, M.Y.; Chen, B.Y. SegTouch: Enhancing Touch Input While Providing Touch Gestures on Screens Using Thumb-To-Index-Finger Gestures. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems—CHI EA'17*; ACM Press: New York, NY, USA, 2017; pp. 2164–2171. [[CrossRef](#)]
156. Zhu, F.; Grossman, T. BISHARE: Exploring Bidirectional Interactions Between Smartphones and Head-Mounted Augmented Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–14. [[CrossRef](#)]
157. Chen, Y.; Katsuragawa, K.; Lank, E. Understanding Viewport- and World-based Pointing with Everyday Smart Devices in Immersive Augmented Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems; CHI'20*, ACM Press: New York, NY, USA, 2020; pp. 1–13. [[CrossRef](#)]
158. Büschel, W.; Mitschick, A.; Meyer, T.; Dachselt, R. Investigating Smartphone-based Pan and Zoom in 3D Data Spaces in Augmented Reality. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services; MobileHCI'19*; ACM Press: New York, NY, USA, 2019; pp. 1–13. [[CrossRef](#)]
159. Rädle, R.; Jetter, H.C.; Marquardt, N.; Reiterer, H.; Rogers, Y. HuddleLamp: Spatially-Aware Mobile Displays for Ad-hoc Around-the-Table Collaboration. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces; ITS'14*; ACM Press: New York, NY, USA, 2014; pp. 45–54. [[CrossRef](#)]
160. Marquardt, N.; Hinckley, K.; Greenberg, S. Cross-device interaction via micro-mobility and f-formations. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology; UIST'12*; ACM Press: New York, NY, USA, 2012; pp. 13–22. [[CrossRef](#)]
161. Chen, X.A.; Grossman, T.; Wigdor, D.J.; Fitzmaurice, G. Duet: Exploring joint interactions on a smart phone and a smart watch. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems—CHI'14*; ACM Press: New York, NY, USA, 2014; pp. 159–168. [[CrossRef](#)]
162. Qian, J.; Cheung, A.; Young-Ng, M.; Yang, F.; Li, X.; Huang, J. Portalware: A Smartphone-Wearable Dual-Display System for Expanding the Free-Hand Interaction Region in Augmented Reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–8. [[CrossRef](#)]
163. Gombač, L.; Čopič Pucihar, K.; Kljun, M.; Coulton, P.; Grbac, J. 3D Virtual Tracing and Depth Perception Problem on Mobile AR. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems—CHI EA'16*; ACM Press: New York, NY, USA, 2016; pp. 1849–1856. [[CrossRef](#)]
164. Pfeuffer, K.; Hinckley, K.; Pahud, M.; Buxton, B. Thumb + Pen Interaction on Tablets. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems—CHI'17*; ACM Press: New York, NY, USA, 2017; pp. 3254–3266. [[CrossRef](#)]
165. Rekimoto, J. Pick-and-drop: A direct manipulation technique for multiple computer environments. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology; UIST'97*; ACM Press: New York, NY, USA, 1997; pp. 31–39. [[CrossRef](#)]
166. Wagner, D.; Langlotz, T.; Schmalstieg, D. Robust and unobtrusive marker tracking on mobile phones. In *Proceedings of the 2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, Cambridge, UK, 15–18 September 2008; pp. 121–124. [[CrossRef](#)]
167. Henrysson, A.; Billinghurst, M.; Ollila, M. Virtual object manipulation using a mobile phone. In *Proceedings of the 2005 International Conference on Augmented Tele-Existence; ICAT'05*; ACM Press: New York, NY, USA, 2005; pp. 164–171. [[CrossRef](#)]
168. Henrysson, A.; Billinghurst, M.; Ollila, M. Face to face collaborative AR on mobile phones. In *Proceedings of the Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)*, Vienna, Austria, 5–8 October 2005; pp. 80–89. [[CrossRef](#)]
169. Mohring, M.; Lessig, C.; Bimber, O. Video See-Through AR on Consumer Cell-Phones. In *Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality*, Arlington, VA, USA, 5 November 2004; pp. 252–253. [[CrossRef](#)]
170. Wagner, D.; Schmalstieg, D. First steps towards handheld augmented reality. In *Proceedings of the Seventh IEEE International Symposium on Wearable Computers*, White Plains, NY, USA, 21–23 October 2003; pp. 127–135. [[CrossRef](#)]

171. Fitzmaurice, G.; Buxton, W. The Chameleon: Spatially aware palmtop computers. In *Proceedings of the Conference Companion on Human Factors in Computing Systems; CHI'94*; ACM Press: New York, NY, USA, 1994; pp. 451–452. [[CrossRef](#)]
172. Rekimoto, J. Tilting Operations for Small Screen Interfaces (Tech Note) 1996. In *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology (UIST '96)*; Association for Computing Machinery: New York, NY, USA, 1996; pp. 167–168. [[CrossRef](#)]
173. Wang, J.; Zhai, S.; Canny, J. Camera Phone Based Motion Sensing: Interaction Techniques, Applications and Performance Study 2006. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology (UIST '06)*; Association for Computing Machinery: New York, NY, USA, 2006; pp. 101–110. [[CrossRef](#)]
174. Agrawal, S.; Constandache, I.; Gaonkar, S.; Roy Choudhury, R.; Caves, K.; DeRuyter, F. Using mobile phones to write in air. In *Proceedings of the 9th International Conference on Mobile Systems, Applications, and Services—MobiSys'11*; ACM Press: New York, NY, USA, 2011; p. 15. [[CrossRef](#)]
175. Ruiz, J.; Li, Y.; Lank, E. User-defined motion gestures for mobile interaction. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems—CHI'11*; ACM Press: New York, NY, USA, 2011; p. 197. [[CrossRef](#)]
176. Teather, R.J.; MacKenzie, I.S. Position vs. Velocity Control for Tilt-Based Interaction 2014. In *Proceedings of the Graphics Interface 2014 (GI '14)*; Canadian Information Processing Society, CAN: Mississauga, ON, Canada, 2014; pp. 51–58.
177. Francone, J.; Nigay, L. Using the user's point of view for interaction on mobile devices. In *Proceedings of the 23rd French Speaking Conference on Human-Computer Interaction—IHM'11*; ACM Press: New York, NY, USA, 2012; p. 1. [[CrossRef](#)]
178. Kooima, R. Generalized Perspective Projection. *Acm Trans. Graph.* **2005**, *24*, 894–903.
179. Büschel, W.; Reipschläger, P.; Langner, R.; Dachsel, R. Investigating the Use of Spatial Interaction for 3D Data Visualization on Mobile Devices. In *Proceedings of the Interactive Surfaces and Spaces on ZZZ—ISS'17*; ACM Press: New York, NY, USA, 2017; pp. 62–71. [[CrossRef](#)]
180. Büschel, W.; Reipschläger, P.; Dachsel, R. Foldable3D: Interacting with 3D Content Using Dual-Display Devices. In *Proceedings of the 2016 ACM on Interactive Surfaces and Spaces—ISS'16*; ACM Press: New York, NY, USA, 2016; pp. 367–372. [[CrossRef](#)]
181. Hasan, K.; Ahlström, D.; Irani, P. SAMMI: A Spatially-Aware Multi-Mobile Interface for Analytic Map Navigation Tasks. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI'15*; ACM Press: New York, NY, USA, 2015; pp. 36–45. [[CrossRef](#)]
182. Spindler, M.; Schuessler, M.; Martsch, M.; Dachsel, R. Pinch-drag-flick vs. spatial input: Rethinking zoom & pan on mobile displays. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems—CHI'14*; ACM Press: New York, NY, USA, 2014; pp. 1113–1122. [[CrossRef](#)]
183. Pahud, M.; Hinckley, K.; Iqbal, S.; Sellen, A.; Buxton, B. Toward compound navigation tasks on mobiles via spatial manipulation. In *Proceedings of the 15th International Conference on Human-Computer Interaction with Mobile Devices and Services—MobileHCI'13*; ACM Press: New York, NY, USA, 2013; p. 113. [[CrossRef](#)]
184. Chen, X.A.; Marquardt, N.; Tang, A.; Boring, S.; Greenberg, S. Extending a mobile device's interaction space through body-centric interaction. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services—MobileHCI'12*; ACM Press: New York, NY, USA, 2012; p. 151. [[CrossRef](#)]
185. Rohs, M.; Oulasvirta, A. Target acquisition with camera phones when used as magic lenses. In *Proceedings of the Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems—CHI'08*; ACM Press: New York, NY, USA, 2008; p. 1409. [[CrossRef](#)]
186. Yee, K.P. Peephole Displays: Pen Interaction on Spatially Aware Handheld Computers. In *CHI '03: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; ACM Press: New York, NY, USA, 2003; p. 8.
187. Mohr, P.; Tatzgern, M.; Grubert, J.; Schmalstieg, D.; Kalkofen, D. Adaptive user perspective rendering for Handheld Augmented Reality. In *Proceedings of the 2017 IEEE Symposium on 3D User Interfaces (3DUI)*, Los Angeles, CA, USA, 18–19 March 2017; pp. 176–181. [[CrossRef](#)]
188. Baricevic, D.; Hollerer, T.; Sen, P.; Turk, M. User-Perspective AR Magic Lens from Gradient-Based IBR and Semi-Dense Stereo. *IEEE Trans. Vis. Comput. Graph.* **2017**, *23*, 1838–1851. [[CrossRef](#)] [[PubMed](#)]
189. Čopič Pucihar, K.; Coulton, P.; Alexander, J. The use of surrounding visual context in handheld AR: Device vs. user perspective rendering. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems—CHI'14*; ACM Press: New York, NY, USA, 2014; pp. 197–206. [[CrossRef](#)]
190. Baričević, D.; Höllerer, T.; Sen, P.; Turk, M. User-perspective augmented reality magic lens from gradients. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology—VRST'14*; ACM Press: New York, NY, USA, 2014; pp. 87–96. [[CrossRef](#)]
191. Spindler, M.; Büschel, W.; Dachsel, R. Use your head: Tangible windows for 3D information spaces in a tabletop environment. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces—ITS'12*; ACM Press: New York, NY, USA, 2012; p. 245. [[CrossRef](#)]
192. Hill, J.S.A. Virtual Transparency: Introducing Parallax View into Video See-through AR. In *Proceedings of the 2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR '11)*, Basel, Switzerland, 26–29 October 2011; pp. 239–240. [[CrossRef](#)]

193. Kaufmann, B.; Ahlström, D. Studying spatial memory and map navigation performance on projector phones with peephole interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI'13*; ACM Press: New York, NY, USA, 2013; p. 3173. [\[CrossRef\]](#)
194. Winkler, C.; Pfeuffer, K.; Rukzio, E. Investigating mid-air pointing interaction for projector phones. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces; ITS'12*; ACM Press: New York, NY, USA, 2012; pp. 85–94. [\[CrossRef\]](#)
195. Willis, K.D.; Poupyrev, I.; Shiratori, T. Motionbeam: A metaphor for character interaction with handheld projectors. In *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems—CHI'11*; ACM Press: New York, NY, USA, 2011; p. 1031. [\[CrossRef\]](#)
196. Matsuda, Y.; Komuro, T. Dynamic layout optimization for multi-user interaction with a large display. In *Proceedings of the 25th International Conference on Intelligent User Interfaces; IUI'20*; ACM Press: New York, NY, USA, 2020; pp. 401–409. [\[CrossRef\]](#)
197. Markussen, A.; Boring, S.; Jakobsen, M.R.; Hornbæk, K. Off-Limits: Interacting Beyond the Boundaries of Large Displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*; ACM Press: New York, NY, USA, 2016; pp. 5862–5873. [\[CrossRef\]](#)
198. Lou, X.; Li, A.X.; Peng, R.; Hansen, P. Optimising Free Hand Selection in Large Displays by Adapting to User's Physical Movements. In *Proceedings of the 2016 Symposium on Spatial User Interaction; SUI'16*; ACM Press: New York, NY, USA, 2016; pp. 23–31. [\[CrossRef\]](#)
199. Walter, R.; Bailly, G.; Valkanova, N.; Müller, J. Cuenesics: Using mid-air gestures to select items on interactive public displays. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services; MobileHCI'14*; ACM Press: New York, NY, USA, 2014; pp. 299–308. [\[CrossRef\]](#)
200. Nickel, K.; Stiefelhagen, R. Pointing gesture recognition based on 3D-tracking of face, hands and head orientation. In *Proceedings of the 5th International Conference on Multimodal Interfaces; ICMI'03*; ACM Press: New York, NY, USA, 2003; pp. 140–146. [\[CrossRef\]](#)
201. Vogel, D.; Balakrishnan, R. 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 Press: New York, NY, USA, 2004; pp. 137–146. [\[CrossRef\]](#)
202. Clarke, C.; Gellersen, H. MatchPoint: Spontaneous Spatial Coupling of Body Movement for Touchless Pointing. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology—UIST'17*; ACM Press: New York, NY, USA, 2017; pp. 179–192. [\[CrossRef\]](#)
203. Endo, Y.; Fujita, D.; Komuro, T. Distant Pointing User Interfaces based on 3D Hand Pointing Recognition. In *Proceedings of the Interactive Surfaces and Spaces on ZZZ—ISS'17*; ACM Press: New York, NY, USA, 2017; pp. 413–416. [\[CrossRef\]](#)
204. Azad, A.; Ruiz, J.; Vogel, D.; Hancock, M.; Lank, E. Territoriality and behaviour on and around large vertical publicly-shared displays. In *Proceedings of the Designing Interactive Systems Conference on—DIS'12*; ACM Press: New York, NY, USA, 2012; p. 468. [\[CrossRef\]](#)
205. Alt, F.; Bulling, A.; Gravanis, G.; Buschek, D. GravitySpot: Guiding Users in Front of Public Displays Using On-Screen Visual Cues. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology—UIST'15*; ACM Press: New York, NY, USA, 2015; pp. 47–56. [\[CrossRef\]](#)
206. Krueger, M.W.; Gionfriddo, T.; Hinrichsen, K. VIDEOPLACE—An artificial reality. *ACM SIGCHI Bull.* **1985**, *16*, 35–40. [\[CrossRef\]](#)
207. Shoemaker, G.; Tsukitani, T.; Kitamura, Y.; Booth, K.S. Body-centric interaction techniques for very large wall displays. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction Extending Boundaries—NordCHI'10*; ACM Press: New York, NY, USA, 2010; p. 463. [\[CrossRef\]](#)
208. Shoemaker, G.; Tang, A.; Booth, K.S. Shadow Reaching: A New Perspective on Interaction for Large Wall Displays 2007. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology (UIST '07)*; Association for Computing Machinery: New York, NY, USA, 2007; pp. 53–56. [\[CrossRef\]](#)
209. Walter, R.; Bailly, G.; Müller, J. StrikeAPose: Revealing mid-air gestures on public displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI'13*; ACM Press: New York, NY, USA, 2013; p. 841. [\[CrossRef\]](#)
210. Zhang, X.; Sugano, Y.; Bulling, A. Evaluation of Appearance-Based Methods and Implications for Gaze-Based Applications. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems—CHI'19*; ACM Press: New York, NY, USA, 2019; pp. 1–13. [\[CrossRef\]](#)
211. Sugano, Y.; Zhang, X.; Bulling, A. AggreGaze: Collective Estimation of Audience Attention on Public Displays. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology—UIST'16*; ACM Press: New York, NY, USA, 2016; pp. 821–831. [\[CrossRef\]](#)
212. Pfeuffer, K.; Vidal, M.; Turner, J.; Bulling, A.; Gellersen, H. Pursuit calibration: Making gaze calibration less tedious and more flexible. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology—UIST'13*; ACM Press: New York, NY, USA, 2013; pp. 261–270. [\[CrossRef\]](#)
213. Smith, B.A.; Yin, Q.; Feiner, S.K.; Nayar, S.K. Gaze locking: Passive eye contact detection for human-object interaction. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology—UIST'13*; ACM Press: New York, NY, USA, 2013; pp. 271–280. [\[CrossRef\]](#)
214. Petford, J.; Nacenta, M.A.; Gutwin, C. Pointing All Around You: Selection Performance of Mouse and Ray-Cast Pointing in Full-Coverage Displays. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; CHI'18*; ACM Press: New York, NY, USA, 2018; pp. 1–14. [\[CrossRef\]](#)

215. Teather, R.J.; Stuerzlinger, W. Pointing at 3d target projections with one-eyed and stereo cursors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI'13*; ACM Press: New York, NY, USA, 2013; p. 159. [[CrossRef](#)]
216. Kela, J.; Korpipää, P.; Mäntyjärvi, J.; Kallio, S.; Savino, G.; Jozzo, L.; Marca, S.D. Accelerometer-based gesture control for a design environment. *Pers. Ubiquitous Comput.* **2006**, *10*, 285–299. [[CrossRef](#)]
217. Baudisch, P.; Sinclair, M.; Wilson, A. Soap: A pointing device that works in mid-air. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology—UIST'06*; ACM Press: New York, NY, USA, 2006; p. 43. [[CrossRef](#)]
218. Khamis, M.; Kienle, A.; Alt, F.; Bulling, A. GazeDrone: Mobile Eye-Based Interaction in Public Space Without Augmenting the User. In *Proceedings of the 4th ACM Workshop on Micro Aerial Vehicle Networks, Systems, and Applications—DroNet'18*; ACM Press: New York, NY, USA, 2018; pp. 66–71. [[CrossRef](#)]
219. Lander, C.; Gehring, S.; Krüger, A.; Boring, S.; Bulling, A. GazeProjector: Accurate Gaze Estimation and Seamless Gaze Interaction Across Multiple Displays. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology—UIST'15*; ACM Press: New York, NY, USA, 2015; pp. 395–404. [[CrossRef](#)]
220. Sibert, L.E.; Jacob, R.J.K. Evaluation of eye gaze interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI'00*; ACM Press: New York, NY, USA, 2000; pp. 281–288. [[CrossRef](#)]
221. Zhai, S.; Morimoto, C.; Ihde, S. Manual and gaze input cascaded (MAGIC) pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; CHI'99; ACM Press: New York, NY, USA, 1999; pp. 246–253. [[CrossRef](#)]
222. Santini, T.; Fuhl, W.; Kasneci, E. CalibMe: Fast and Unsupervised Eye Tracker Calibration for Gaze-Based Pervasive Human–Computer Interaction. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*; ACM Press: New York, NY, USA, 2017; pp. 2594–2605. [[CrossRef](#)]
223. Horak, T.; Badam, S.K.; Elmqvist, N.; Dachselt, R. When David Meets Goliath: Combining Smartwatches with a Large Vertical Display for Visual Data Exploration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems—CHI'18*; ACM Press: New York, NY, USA, 2018; pp. 1–13. [[CrossRef](#)]
224. Siddhpuria, S.; Katsuragawa, K.; Wallace, J.R.; Lank, E. Exploring At-Your-Side Gestural Interaction for Ubiquitous Environments. In *Proceedings of the 2017 Conference on Designing Interactive Systems—DIS'17*; ACM Press: New York, NY, USA, 2017; pp. 1111–1122. [[CrossRef](#)]
225. Katsuragawa, K.; Pietroszek, K.; Wallace, J.R.; Lank, E. Watchpoint: Freehand Pointing with a Smartwatch in a Ubiquitous Display Environment. In *Proceedings of the International Working Conference on Advanced Visual Interfaces—AVI'16*; ACM Press: New York, NY, USA, 2016; pp. 128–135. [[CrossRef](#)]
226. Houben, S.; Marquardt, N. WatchConnect: A Toolkit for Prototyping Smartwatch-Centric Cross-Device Applications. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*; CHI'15; ACM Press: New York, NY, USA, 2015; pp. 1247–1256. [[CrossRef](#)]
227. Harrison, C.; Benko, H.; Wilson, A.D. OmniTouch: Wearable multitouch interaction everywhere. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology—UIST'11*; ACM Press: New York, NY, USA, 2011; p. 441. [[CrossRef](#)]
228. Liu, M.; Nancel, M.; Vogel, D. Gunslinger: Subtle Arms-down Mid-air Interaction. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology—UIST'15*; ACM Press: New York, NY, USA, 2015; pp. 63–71. [[CrossRef](#)]
229. Haque, F.; Nancel, M.; Vogel, D. Myopoint: Pointing and Clicking Using Forearm Mounted Electromyography and Inertial Motion Sensors. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems—CHI'15*; ACM Press: New York, NY, USA, 2015; pp. 3653–3656. [[CrossRef](#)]
230. von Zadow, U.; Büschel, W.; Langner, R.; Dachselt, R.I.M.L. SleeD: Using a Sleeve Display to Interact with Touch-sensitive Display Walls. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces—ITS'14*; ACM Press: New York, NY, USA, 2014; pp. 129–138. [[CrossRef](#)]
231. Parzer, P.; Probst, K.; Babic, T.; Rendl, C.; Vogl, A.; Olwal, A.; Haller, M. FlexTiles: A Flexible, Stretchable, Formable, Pressure-Sensitive, Tactile Input Sensor. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*; CHI EA'16; ACM Press: New York, NY, USA, 2016; pp. 3754–3757. [[CrossRef](#)]
232. Saponas, T.S.; Harrison, C.; Benko, H. PocketTouch: Through-fabric capacitive touch input. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology—UIST'11*; ACM Press: New York, NY, USA, 2011; p. 303. [[CrossRef](#)]
233. Bailly, G.; Müller, J.; Rohs, M.; Wigdor, D.; Kratz, S. ShoeSense: A new perspective on gestural interaction and wearable applications. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems—CHI'12*; ACM Press: New York, NY, USA, 2012; p. 1239. [[CrossRef](#)]
234. Yeo, H.S.; Lee, J.; Kim, H.i.; Gupta, A.; Bianchi, A.; Vogel, D.; Koike, H.; Woo, W.; Quigley, A. WRIST: Watch-Ring Interaction and Sensing Technique for Wrist Gestures and Macro-Micro Pointing. In *Proceedings of the 21st International Conference on Human–Computer Interaction with Mobile Devices and Services*; MobileHCI'19; ACM Press: New York, NY, USA, 2020; pp. 1–15. [[CrossRef](#)]
235. Barrera Machuca, M.D.; Stuerzlinger, W. The Effect of Stereo Display Deficiencies on Virtual Hand Pointing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems—CHI'19*; ACM Press: New York, NY, USA, 2019; pp. 1–14. [[CrossRef](#)]
236. McArthur, V.; Castellucci, S.J.; MacKenzie, I.S. An empirical comparison of “wiimote” gun attachments for pointing tasks. In *Proceedings of the 1st ACM SIGCHI Symposium on Engineering Interactive Computing Systems*; EICS'09; ACM Press: New York, NY, USA, 2009; pp. 203–208. [[CrossRef](#)]

237. Jiang, H.; Ofek, E.; Moraveji, N.; Shi, Y. Direct pointer: Direct manipulation for large-display interaction using handheld cameras. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI'06*; ACM Press: New York, NY, USA, 2006; p. 1107. [[CrossRef](#)]
238. Wilson, A.; Shafer, S. XWand: UI for Intelligent Spaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*; Association for Computing Machinery: New York, NY, USA, 2003; pp. 545–552. [[CrossRef](#)]
239. Silfverberg, M.; MacKenzie, I.S.; Kauppinen, T. An Isometric Joystick as a Pointing Device for Handheld Information Terminals 2001. In *Proceedings of Graphics Interface 2001 (GI '01)*; Canadian Information Processing Society, CAN: Mississauga, ON, Canada, 2001; pp. 119–126.
240. König, W.A.; Gerken, J.; Dierdorf, S.; Reiterer, H. Adaptive pointing: Implicit gain adaptation for absolute pointing devices. In *Proceedings of the CHI'09 Extended Abstracts on Human Factors in Computing Systems*; CHI EA'09; ACM Press: New York, NY, USA, 2009; pp. 4171–4176. [[CrossRef](#)]
241. Vogt, F.; Wong, J.; Po, B.; Argue, R.; Fels, S.; Booth, K. Exploring collaboration with group pointer interaction. In *Proceedings of the Proceedings Computer Graphics International, Crete, Greece, 19 June 2004*; pp. 636–639. [[CrossRef](#)]
242. Olsen, D.R.; Nielsen, T. Laser pointer interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*; CHI'01; ACM Press: New York, NY, USA, 2001; pp. 17–22. [[CrossRef](#)]
243. Forlines, C.; Vogel, D.; Balakrishnan, R. HybridPointing: Fluid switching between absolute and relative pointing with a direct input device. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology—UIST'06*; ACM Press: New York, NY, USA, 2006; p. 211. [[CrossRef](#)]
244. Khan, A.; Fitzmaurice, G.; Almeida, D.; Burtnyk, N.; Kurtenbach, G. A remote control interface for large displays. In *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology—UIST'04*; ACM Press: New York, NY, USA, 2004; p. 127. [[CrossRef](#)]
245. Wehbe, R.R.; Dickson, T.; Kuzminykh, A.; Nacke, L.E.; Lank, E. Personal Space in Play: Physical and Digital Boundaries in Large-Display Cooperative and Competitive Games. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–14. [[CrossRef](#)]
246. Jakobsen, M.R.; Jansen, Y.; Boring, S.; Hornbæk, K. Should I Stay or Should I Go? Selecting Between Touch and Mid-Air Gestures for Large-Display Interaction. In *Proceedings of the 15th IFIP TC 13 International Conference on Human-Computer Interaction—INTERACT 2015*; Springer: Berlin/Heidelberg, Germany, 2015; Volume 9298, pp. 455–473. [[CrossRef](#)]
247. von Zadow, U.; Reipschläger, P.; Bösel, D.; Sellent, A.; Dachselt, R. YouTouch! Low-Cost User Identification at an Interactive Display Wall. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*; AVI'16; ACM Press: New York, NY, USA, 2016; pp. 144–151. [[CrossRef](#)]
248. Leigh, S.w.; Schoessler, P.; Heibeck, F.; Maes, P.; Ishii, H. THAW: Tangible Interaction with See-Through Augmentation for Smartphones on Computer Screens. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction—TEI'14*; ACM Press: New York, NY, USA, 2014; pp. 89–96. [[CrossRef](#)]
249. Simeone, A.L.; Seifert, J.; Schmidt, D.; Holleis, P.; Rukzio, E.; Gellersen, H. A cross-device drag-and-drop technique. In *Proceedings of the 12th International Conference on Mobile and Ubiquitous Multimedia—MUM'13*; ACM Press: New York, NY, USA, 2013; pp. 1–4. [[CrossRef](#)]
250. Bolt, R.A. “Put-that-there”: Voice and gesture at the graphics interface. In *Proceedings of the 7th Annual Conference on Computer Graphics and Interactive Techniques*; SIGGRAPH'80; ACM Press: New York, NY, USA, 1980; pp. 262–270. [[CrossRef](#)]
251. Stellmach, S.; Dachselt, R. Looking at 3D User Interfaces. 2012. In *CHI 2012 Workshop on The 3rd Dimension of CHI (3DCHI): Touching and Designing 3D User Interfaces, CHI '12, Austin, TX, USA*; ACM Press: New York, NY, USA, 2012; pp. 95–98.
252. Zhai, S.; Buxton, W.; Milgram, P. The partial-occlusion effect: Utilizing semitransparency in 3D human-computer interaction. *ACM Trans.-Comput.-Hum. Interact.* **1996**, *3*, 254–284. [[CrossRef](#)]
253. Schuchardt, P.; Bowman, D.A. The benefits of immersion for spatial understanding of complex underground cave systems. In *Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology—VRST'07*; ACM Press: New York, NY, USA, 2007; p. 121. [[CrossRef](#)]
254. Rekimoto, J. A vision-based head tracker for fish tank virtual reality-VR without head gear. In *Proceedings of the Proceedings Virtual Reality Annual International Symposium'95, Research Triangle Park, NC, USA, 11–15 March 1995*; IEEE: Hasselt and Diepenbeek, Belgium, 1995; pp. 94–100. [[CrossRef](#)]
255. Cruz-Neira, C.; Sandin, D.J.; DeFanti, T.A.; Kenyon, R.V.; Hart, J.C. The CAVE: Audio visual experience automatic virtual environment. *Commun. ACM* **1992**, *35*, 64–72. [[CrossRef](#)]