Research Article

Pascal Knierim*, Dimitri Hein, Albrecht Schmidt, and Thomas Kosch

The SmARtphone Controller

Leveraging Smartphones as Input and Output Modality for Improved Interaction within Mobile Augmented Reality Environments

Abstract: Current interaction modalities for mobile Augmented Reality (AR) are tedious and lack expressiveness. To overcome these prevalent limitations, we developed and evaluated a multimodal interaction concept by pairing a smartphone as an input and output modality for mobile AR. In a user study (n = 24), we investigated the effects on interaction speed, accuracy, and task load for (1) virtual object manipulation as well as (2) interaction with established graphical user interfaces (GUIs). Our findings show that a smartphone-based AR controller results in significantly faster and more accurate object manipulation with reduced task load than state-of-art mid-air gestures. Our results also indicate a significant enhancement for using the physical touchscreen as a modality compared to mid-air gestures for GUI interaction. We conclude that interaction in mobile AR environments can be improved by utilizing a smartphone as an omnipresent controller. Additionally, we discuss how future AR systems can benefit from tangible touchscreens as an additional and complementary interaction modality.

Keywords: Augmented Reality, Head-Mounted Display, Smartphone, Object Manipulation, Multi-Display Interaction

Figure 1: We conducted a user study comparing elaborated mid-air hand gestures (left) to hybrid interaction with a smartphone (right) as input and output controller in mobile AR environments.

1 Introduction

The next generation of head-mounted displays (sHMDs) for Augmented Reality (AR) is currently released with numerous technological advances. Enhanced wearing com-fort, a wider field of view (FoV), long-lasting battery life, and increased overall performance are some of the improved aspects. Today, consumers can already take advantage of AR experiences on their smartphones. Prominent use-cases are entertainment, education, navigation, or sightseeing. In the professional domain, wearable AR HMDs are deployed to support workers that perform demanding tasks, such as order-picking [14, 15], maintenance [41], repairs [27], or in the medical field to show vital information to improve surgical accuracy.

The expected generation of AR glasses incorporates many advances in displays and tracking technology. However, these devices will still offer only a restricted set of interaction capabilities that lack expressiveness. Voice commands are largely supported but can be unreliable under excessively loud conditions [40]. Further, using voice commands to interact with an AR system is not widely prevalent. Mid-air hand gestures are an alternative input method to interact with the AR system. Gestures are recognized by tracking either or both hands that are within the tracking space of the device. Depending on the device, the gesture-sensing space can vary. Furthermore, users need to approximate the boundaries they can interact in, to not break the gesture tacking space. Despite mid-air gestures are known to cause fatigue, AR glasses are still heavily relying on gestures as the primary interaction method. In

*Corresponding author: Pascal Knierim, Human-Centered Ubiquitous Media, LMU Munich, Munich, Germany, e-mail: pascal.knierim@ifi.lmu.de, ORCID: https://orcid.org/0000-0001-9578-9953
Dimitri Hein, Albrecht Schmidt, Human-Centered Ubiquitous Media, LMU Munich, Munich, Germany, e-mails: dimitri.hein@outlook.de, albrecht.schmidt@ifi.lmu.de
Thomas Kosch, Human-Computer Interaction, Technical University of Darmstadt, Darmstadt, Germany, e-mail: thomas.kosch@tk.informatik.tu-darmstadt.de

contrast, current Virtual Reality (VR) systems are provided with advanced controllers equipped with multiple tangible buttons, triggers, and precise tracking within their calibrated interaction space. Unfortunately, these accurate controllers rely on external tracking equipment that requires a calibration process. Hence, they are unsuited for spontaneous controller-based interaction in mobile AR environments.

Conversely, physical controllers provided with commercially available AR systems come with minimalistic input modalities. The controller's tracking is limited to rotations and is sometimes enhanced by simulated degrees-of-freedom (sDOFs) or interpolations to support extended interactions. In contrast to VR controllers, the number of buttons and input possibilities is noticeably reduced. In this article, we present our developed smartphone application that connects to AR glasses and acts as a universal controller to enhance the interaction with AR experiences and virtual objects. The smartphone serves as an input device by repurposing the multi-touch screen as a tangible handheld multi-touch input surface. Additionally, the smartphone display acts as a flexible high-resolution extension of the AR presentation space. Where previous work extended the smartphones' screen and interaction space using AR [17, 32], we use the input capabilities of smartphones to investigate the efficiency for direct interaction with content provided by the AR glasses. Hence, we combine the input and output modalities provided by smartphones' to interact with spatial available AR content.

In a user study with 24 participants, with particular emphasis on object manipulation and graphical user interface (GUI) interactions, we evaluated the speed, accuracy, and usability compared to elaborated gesture input. Instead of only utilizing multi-touch gestures, we leverage the smartphone screen as a visual extension for fine-grained input of the AR HMD. We found significant improvements in accuracy, interaction efficiency, and task load using our approach. This article introduces an approach to efficiently interact with AR environments by re-purposing our available smartphones. We contribute with our findings on the effects of integrating a multi-touch controller and high-resolution display in existing AR systems. With our combination of an AR HMD and a smartphone,[1] we enable users to interact fast and accurate with virtual objects and operate complex AR user interfaces with ease.

---

1 The software framework is available as a reference implementation to developers via GitHub: https://github.com/pknierim/The-SmARtphone-Controller

2 Research Questions

In past research, many systems have been proposed combining HMD and novel controllers. Smartphones were facilitated as secondary output screens or as input devices using touch or other built-in sensors. However, the potential of smartphones as an AR controller has not yet been fully explored and evaluated. None of the previous work specifically investigated the input capabilities of multi-touch gestures and performance compared to present mid-air gesture input. The visual split of AR content and user interfaces for interaction and manipulations was not yet explored in mobile AR environments. In this work, we specifically investigate these two interaction concepts.

RQ1 Does a smartphone-based input modality improve users' speed and accuracy in completing 3D manipulation tasks compared to mid-air gesture-based input?

RQ2 Does a smartphone-supported user-interface input modality offer improved usability and lower task load during interaction compared to mid-air gesture-based input?

3 Related Work

Our work builds on past research in AR and the recent development of novel interaction concepts for AR, VR, and mobile devices. We review research that motivated our development and work that explores the interaction space for mobile AR. Afterward, we summarize and discuss the interaction capabilities of selected commercially available AR and VR devices.

3.1 Interaction Concepts in AR and VR

The development of new AR and VR HMDs has ramped up over the last years. Major technology and entertainment companies have released the second to third iteration of HMDs to the mass market. Comfort, weight, FoV, visual, and audio quality have been continuously improved. Nevertheless, interaction concepts have not changed significantly. In the domain of mixed reality, interaction is mainly controller-driven, allowing to intuitively grab, throw, or precisely modify the virtual environment. Typically controllers require to be calibrated and are therefore unsuitable for mobile setups. New approaches using ultrasonic and magnetic sensors fused with gyroscope and accelerometer are promising but are still in their infancy. Unlike, commercial AR solutions offer a more fragmented

interaction space. Hand gestures, in combination with a head pose, are the most prominent ones.

Today, we use smartphones as a ubiquitous computing device to interact with our environment [5]. We control our home appliances, buy tickets, navigate, or engage with location-based games. Current smartphones comprise numerous sensors that facilitate a good understanding of the environment. Further, devices are becoming more connected than ever and act as a remote interface for current cameras,[2] speakers,[3] or toys.[4] There is a large body of work in which various input techniques have been proposed to interact with virtual elements displayed on an HMD in a mobile augmented environment. In smartphone-enabled handheld mobile AR experiences, direct selection and manipulation of objects give a natural and convenient user experience. However, maintaining visual tracking while holding the smartphone in one hand and interact via touch-gestures with the other hand is challenging [4, 23].

Selection and manipulation in mobile handheld AR got improved by freezing the input [4] or providing additional devices or modified pens [39]. With the availability of optical see-through smart glasses, ubiquitous interaction techniques were investigated. Wearable input is often facilitated through touch surfaces encircling users' clothes [11, 36] or even fully garment-integrated sensors [25] that are extended by AR. A flexible wearable alternative are sensor-enabled smartwatches. They can provide natural input for short interaction with virtual objects [22, 24]. Further interaction techniques include the use of the user's body. From mid-air hand gestures to foot-tapping [31], viable solutions for interacting with virtual elements exist.

A different approach was presented by Normand and McGuffin [32]. Instead of augmenting the environment or the user's body, they augment a smartphone to display additional content around it. A user study highlights their approach's feasibility, hence promoting to display context-aware information around the smartphones' display. With hybrid systems that are including AR, VR, or large displays as primary screen technology and an additional handled secondary display or smartphone as a controller enable seamless advanced interaction in mobile context [7, 18]. With smartphones following a "bring-your-own-device" approach, they have been proposed to support spontaneous interaction with large public displays via the user's

smartphone [33]. Practical examples include gaming, were large displays promote multiplayer gaming in social settings [28]. Similarly, Grubert et al. [17] proposed how smart artifacts and the user's body can be augmented with virtual objects using smartphones. Previous work dealt with the use of the smartphone as an input proxy for other devices. Al-Sada et al. [1] presented the two interaction concepts *Input Borrowing* and *Interaction-Event Mappings* which combine smartphone input with other smart devices or AR applications. They show the feasibility of their approach in a user study. However, their user study was employed in a seated scenario.

Detecting the interaction space and virtual objects of interest for manipulation was investigated by previous research. Prominent interaction metaphors for object selection and manipulation are image plane-based or ray casting techniques. Motions of the controller are translated to a spatial ray or are projected onto a plane to support interaction. The inertial measurement unit (IMU) integrated in tangible interfaces [10] or smartphones [18] senses the orientation that is directly translated into the interaction space [21]. Sophisticated and highly specialized handheld controllers were built. Incorporated with a touchscreen, six-degree-of-freedom (6DOF) tracking, and tactile buttons, interaction with immersive applications for VR and AR are viable [29, 35]. Recently, Mohr et al. [30] developed an application that turns a regular smartphone into a 6DOF pointing and selection device to retrofit AR or VR HMDs. They confirmed the feasibility of re-purposing smartphones as an input controller without any hardware modifications. Within this context, Babic et al. presented *Pocket6*, a smartphone application that uses the integrated tracking capabilities of a smartphone to enable subtle interaction via gestures [3]. Their work shows the technical feasibility of using explicit and implicit gestures for smartphones that can be used to interact with content on the smartphone or virtual content integrated into the environment. The applicability and usability of smartphone gestures have to be considered as well. For example, Serrano et al. [37] found that phone rotations can lead to unpleasant wrist postures over time. Thereby, the usability of gestures has to be evaluated prior to developing and implementing the novel user interface. In this context, Zhu et al. [42] presented design recommendations for the interaction between AR and smartphones. They present *BISHARE*, an application that enables interaction in AR using a smartphone controller. However, the design recommendations do not evaluate the efficiency, perceived usability, and task load in a user study. A selection of AR and VR devices supported interaction modalities are listed in Table 1 for reference.

---

2 GoPro – gopro.com/en/us/shop/quik-app-video-photo-editor

3 Spotify – spotify.com/us/connect

4 Sphero – sphero.com/collections/all/products/sphero-mini – all links last access 2021-03-13.

Table 1: Overview of out-of-the-box supported interaction modalities of selected AR and VR systems separated into embodied and peripheral (Controller (C)) interaction. ● fully supported, ◐ partially supported, ○ not supported.

	Embodied				Peripheral			
	Headpose	Eye Gaze	Gesture	Voice	C Motion	C Buttons	C Touch	Smartphone
HoloLens	●	○	●	●	◐	◐	○	○
HoloLens 2	●	●	●	●	○	○	○	○
Magic Leap One	●	●	●	●	●	●	◐	◐
Varjo VR-3	●	●	●	○	○	○	○	○
Google Glass	○	○	○	●	○	◐	●	◐
Aryzon [9]	●	○	○	○	○	○	○	○
Daydream	●	○	○	○	◐	●	◐	○
Vive Pro EYE	○	●	○	○	●	●	◐	○
Valve Index	○	○	○	○	●	●	●	○
Oculus Rift S	○	○	○	○	●	●	◐	○
Oculus Quest 2	○	○	○	○	●	●	◐	○

Previous work has extensively researched how AR and VR environments can benefit from smartphones as an additional controller for interaction. However, previous work is either limited to the evaluation of gestures or the development of design recommendations, leaving a gap in the evaluation of the overall efficiency and perceived usability of applications when using smartphones in tandem with virtual environments. We close this gap by presenting a user study investigating the smartphones' feasibility as a second screen and interaction controller for an AR application. We select suitable interaction concepts, provide details about the implementation, present the design and the results of a user study in the following.

4 Interaction Concept

Current interaction with AR applications can be divided into pointing or selection and point manipulations in space. Interaction with free-floating graphical user interfaces (sGUIs) or menus can be reduced as a combination of pointing and selection. The combination of these two interactions is the fundamental requirement for any interaction with AR environments. Our approach targets an easy to understand system that utilizes known gestures, paradigms, and adaptations of current solutions [26, 38]. For object manipulation, we focus on an eyes-free input to not distract the user from the AR elements. This concept aims to keep the task load low while maintaining

high flexibility and functionality. Users can manipulate elements with high fidelity via swipes and touch gestures, buttons, and more sophisticated user interfaces that can be displayed on the smartphones' multi-touch screen for fast and intuitive touch interactions. In both cases, kinesthetic as well as tactile feedback is provided through the smartphone. Auditive feedback is either provided via the smartphone or the HMDs speakers to guide the user's attention.

4.1 Object Manipulation

Object manipulation can be separated into translation, rotation, and scaling along the three different axes (i. e., the x-, y-, and z-axis). For object *translation*, users can move the focused object horizontally (along the x- and z-axes) by touching the screen and swiping horizontally or vertically. Users can adjust the height (y-axis) of an object by double-tapping with and subsequent swipe. For an intuitive translation, a reference coordinate system is created every time the user begins a translation. The coordinate system is created according to the head position concerning the object. Early user tests showed that the decoupling of head and smartphone rotation and position while translating objects is very intuitive and coherent to the user.

For the *rotation* around the y-axis, we selected a common approach known from map interactions with smartphones. Rotation is initiated by touching the screen with two fingers and then continuously rotate them around each other. To adjust the *scale*, we adopted the pinch gesture done with two fingers. The space between the fingers is translated directly to the size of the object that is modified. During object manipulation, no information is visible on the display. The interaction design is represented in Figure 2. For comparison, we depicted the elaborated mid-air gestures supported by commercial AR HMDs in Figure 3.

4.2 Secondary Screen Support

Supplementary to object modifications, AR applications often require input on free-floating or space anchored 2D GUIs. Our smartphone supported approach leverages two different possibilities to interact with these kinds of interfaces. First, by implementing a simple remote-like controller similar to the object manipulation. Users can swipe and tap to interact with the AR display space presented user interface elements like buttons, slides, or checkboxes.
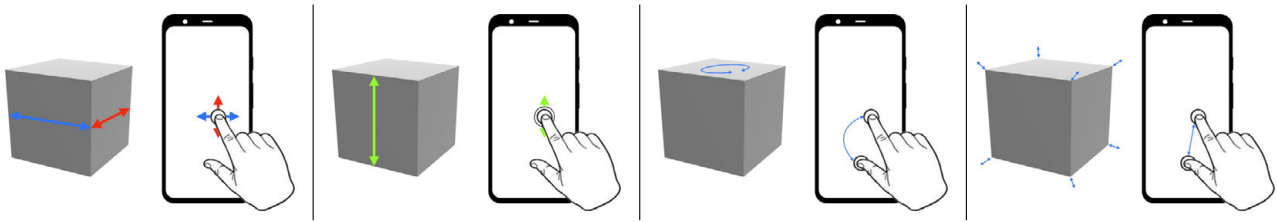
Figure 2: Interaction concept for object manipulation with a smartphone. From left to right: 1) single tap + swipe: translation along x- and z-axes (left/right, back/forth) 2) double tap + swipe: translation along y-axis (up/down) 3) two finger rotation: rotation 4) pinch: scale.
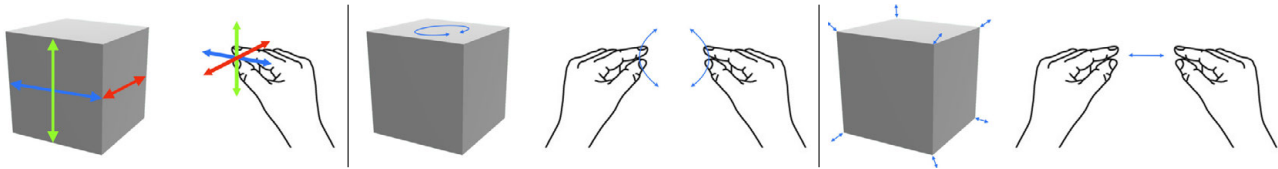

Figure 3: Default mid-air gesture interaction concept for object manipulation. From left to right: 1) Air tap and hold + move hand along the corresponding axis: translation along with hand movement (left/right, back/forth, up/down) 2) two hand air tap and hold + counter hand movement: rotation 3) two hand air tap and hold + move apart: scale.

In the second approach, the entire user interface is transferred on demand onto the smartphones' display and allows seamless and direct touch and swipe interaction with the represented elements.

5 Implementation

To investigate the unique features of using a smartphone as ubiquitous input and output controller in AR environments, we implemented our apparatus incorporating a Microsoft HoloLens and a Google Pixel 2 XL Smartphone.

5.1 Smartphone Application

We developed a native Android application to provide seamless input and output. Touch input, swipes, and gestures are sensed and directly sent to the AR HMD. User interface elements can be displayed at the request of the AR HMD. Any GUI input is forwarded to the HMD accordingly. For bidirectional communication between both devices, a Bluetooth radio frequency communication channel is established. Touches, swipes, and gestures are transferred via this channel to manipulate visible objects in AR. Bidirectional status messages are transferred for UI input, states of the application, or haptic feedback control commands. We used Google's Protocol Buffer because it provides a fast and platform-neutral serialization protocol of the structured data.

5.2 HMD Application

The HoloLens displays the AR environment and processes any incoming interaction commands from the Smartphone application. For a smooth and enhanced user experience, any continuum user-input is low pass filtered to remove any jitter. If a selected virtual object contains a context menu, the presentation of this user interfaces is triggered on the connected smartphone. Both the HoloLens application and the AR environment are implemented using the Unity game engine 2018.4.7 and the Mixed Reality Toolkit v2.

6 Method

Our approach enables immersed users to manipulate virtual objects in AR environments with a smartphone. This study aims to evaluate the effect of a touch screen as input and output modality in contrast to state-of-the-art mid-air gestures on the object manipulation performance. We further investigate the effect of task complexity on the overall task completion time and task load.

Two different tasks were elaborated to understand the qualities of a touchscreen as an input and output controller regarding performance, user experience, and task load. First, participants rearranged virtual objects in 3D space, followed by a set of modification tasks requiring to interact with a free-floating context menu.

We used a repeated-measures within-subject design with a within-subject independent variable (IV) INPUT. The IV INPUT has two levels in the first experiment: *mid-air gesture* as a baseline and *multi-touch* since we focus only on the input capabilities of the smartphone. We chose *mid-air gesture* as a baseline since it is well supported across several state-of-the-art AR HMDs and provides the necessary input quality and adaptability in contrast to voice or dedicated input solutions like the HoloLens clicker. In the second experiment, the IV has an additional level of *multi-touch display* since we are also using the smartphone screen as output.

6.1 Subjects

We recruited 24 participants (12 female, 12 male) via our universities' mailing lists. Participants received either 10 EUR or course credits as compensation for their participation. Four had previous AR experience, one of them used AR glasses for professional purposes. The study received ethics clearance according to the ethics and privacy regulations of our institution.

6.2 Apparatus

The apparatus for this study comprised a Microsoft HoloLens and the Google Pixel 2 XL running Android 9. The smartphone has a bright, high-resolution display (538 PPI) with a presentation and interaction area of 136×68 mm. Both devices are connected via Bluetooth and run our developed applications presenting the different stimuli and logging the data. The developed smartphone application serves as an input controller as well as a secondary display. For the baseline, only the HoloLens with the build-in mid-air hand gesture support was facilitated. Our experiment was conducted in a room with controlled light conditions for consistent visibility of holograms and a free interaction area of approximately 3×3 m.

6.3 Procedure

After welcoming the participants, we asked them to sign the consent form. We explained the course of the study, all devices, and the interaction concepts to the participants. Afterward, we adjusted the AR glasses to the participant's head and ensured that the participant can comfortably perceive the entire display area. In the last preparation step, we ask the participants to start our application, which guides them through the study. The application starts with a tutorial and aids the participant in getting used to the different interaction concepts. Participants could freely practice the first input modality until they understand them and feel comfortable during the tutorial. Specific questions regarding the input and study were answered, and the main task was explained. Participants were requested to finish the tasks as fast and as accurately as possible. Then participants started with the object manipulation task. After manipulating all 12 objects (three sets of four objects), they had to fill out the NASA-TLX [19] questionnaire on a dedicated laptop. Participants could take a break at this point since all tasks were performed in a standing position to minimize fatigue. The participants continued with the second task using the same input modality. Again, a tutorial was presented introducing the new task and allow the participant to practice and familiarize themselves with a different function. After accomplishing the new tasks, participants filled out the NASA-TLX questionnaire. Finally, they repeated both of the experiments with the other input modality.

The input modality and the tasks were presented in a counterbalanced order using a full Latin square to prevent any sequence or learning effects. Throughout the study, we logged all interactions with the system for subsequent offline analyses. After successfully finishing both experiments, we asked for comments about the user experience and which input modality they ultimately prefer. Participants completed the study, including the debriefing, within 40 to 55 minutes.

6.4 Experiment 1: Object Manipulation

For the first experiment, we designed an object manipulation task in which participants had to select, translate, rotate, and scale different colored cubes. Four virtual cubes with an edge length of 50 cm were placed in front of the participant. After the selection of a cube, a white line guides the participant to the ghost representation of the cube representing the target state. Participants were asked to perform the manipulations using the given input modality as precise and fast as possible. Both conditions are visualized in Figure 4. Task complexity was increased through three sets of four cubes. The first set only includes a translation to achieve the target transformation. The translation requires movements of the cube between 25 cm and 125 cm along all three axes in a predefined direction. The second set of four cubes additionally includes rotation. The rotation offset was between +/−30 degree and +/−150 degree, whereat participants could rotate the cube either way in one seamless interaction. Lastly, participants
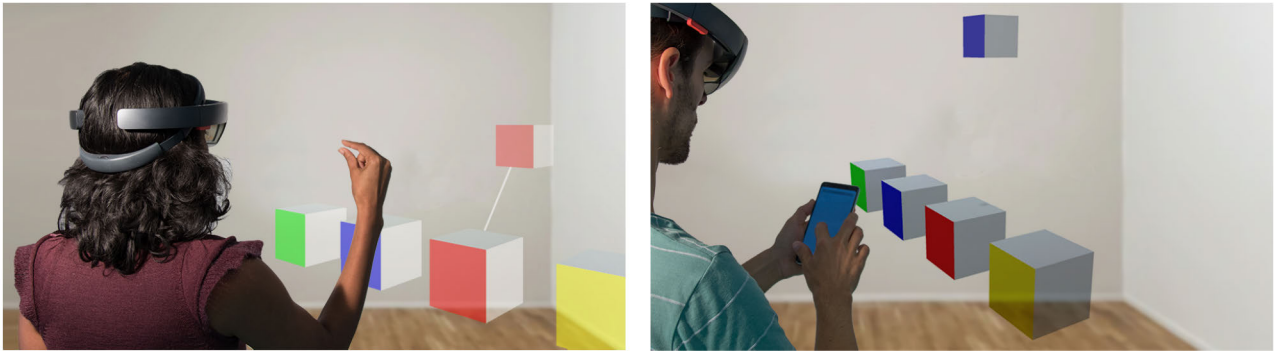
Figure 4: The first experiment includes manipulation of the position, rotation, and size of colored cubes. Arrows and lines help to compensate for the limited field of view and enhance orientation in space. The first condition includes manipulation via mid-air gestures (left). The second condition of the first experiment involves multi-touch as an input modality (right). Participants can manipulate colored cubes by swiping and multi-touch gestures on the smartphone.
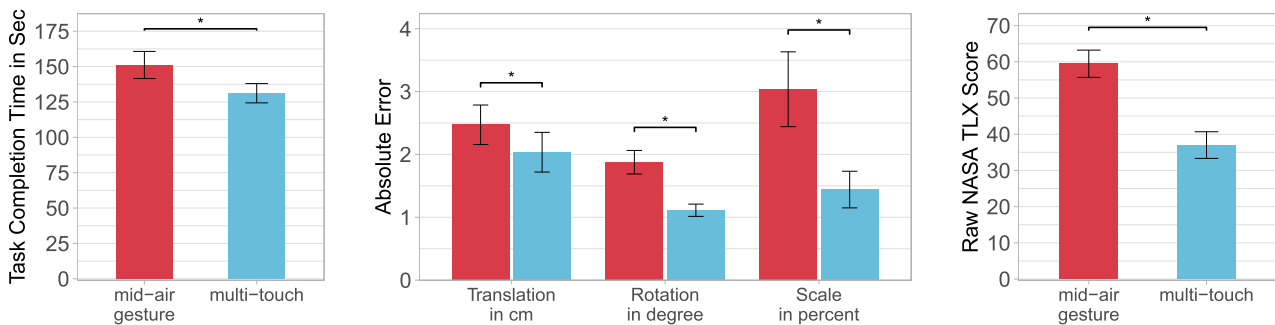

Figure 5: Mean values for TCT, relative error, and raw TLX Score for each condition of the first experiment. Error bars show standard error of the mean (SE). Asterisk indicate statistically significant differences between conditions.

had to translate, rotate, and scale each of the cubes, respectively. Matching the size of the target cube required the participants to set the scaling factor between 0.6 and 1.8, creating cubes between 30 cm, and 90 cm edge length.

We measured the accumulated task completion time (TCT) starting from the first modification till the last modification of each cube since we were only interested in the object manipulation performance. Further, we recorded the accuracy by calculating the euclidean distance in cm, absolute rotation in degree, and scale offset in percent between the ghost representation of the target cube and the cube placed by the participant. Finally, we assessed the task load through the raw NASA-Task Load Index (raw TLX). In the first experiment, we recorded a total of 578 object manipulations.

6.5 Results 1: Object Manipulation

We analyzed the TCT and accuracy concerning the position, rotation, and scale deviation. We further assessed

the perceived task load using the NASA-TLX questionnaire. We used a repeated-measures t-test for statistical comparison. The significance level for all comparisons was set to $\alpha = .05$. The results of the first experiment are graphically depicted in Figure 5.

6.5.1 Task Completion Time

The aggregated TCT for all manipulations of *mid-air gesture* input ($M = 453.47$, $SE = 34.84$) was significantly higher than *multi-touch* input ($M = 393.53$, $SE = 25.01$), $t(23) = 2.341$, $p = .028$, $r = 0.438$. The effect size estimate indicates that the difference in TCT created by the input modality was a medium effect. To understand the strength of the smartphone input we subsequently analysed each of the three task complexities individually using t-tests. We found significant differences between *mid-air gesture* input ($M = 173.47$, $SE = 15.74$) and *smartphone input* ($M = 135.42$, $SE = 10.52$), $t(23) = 2.799$, $p = .010$, $r = 0.571$, for the medium complex task including transla-

Figure 6: The conditions of the second experiment. Participants are either controlling the GUI via mid-air gestures (left), touch gestures on the smartphone or the GUI is displayed on the smartphone itself (center/right) while the task and elements are displayed in AR.

tion and rotation but no significant differences for the others tasks (all $p > .05$).

6.5.2 Accuracy

For statistical analysis, we split the accuracy measure into translation, rotation, and scaling error. The differences between the conditions of all metrics were not normally distributed, as assessed by the Shapiro-Wilk test (all $p < .001$). Therefore, we analyzed the data using Wilcoxon signed-rank tests. The accuracy using *multi-touch* input was significantly improved compared to *mid-air gesture* input in all subcategories (all $p < .001$). We measured the largest effect size for the scaling factor ($Z = -3.848662$, $r = 0.467$), followed by the translation error ($Z = -5.32343$, $r = 0.365$) and rotation error ($Z = -4.114372$, $r = 0.350$).

6.5.3 Task Load

We used the raw TLX as a subjective, multidimensional assessment tool to rate the perceived task load of each input condition [19]. The raw TLX scores were not normally distributed ($p = .010$). The Wilcoxon signed-rank test showed that using *mid-air gesture* input ($M = 59.46$, $IQR = 27.25$) elicit a statistically significant change in the perceived task load in comparison to our *multi-touch* input ($M = 37.00$, $IQR = 24.75$, $Z = -4.273$, $p < 0.001$). Indeed, the effect size $r = 0.617$ suggests a large practical significance.

We summarize that the utilized input modality has a significant effect on the object manipulation performance measured using the relative error concerning translation, rotation, and scale. Further, the task completion time can, dependent on object manipulation complexity, significantly be reduced. The raw TLX score was the lowest when using the multi-touch display of the smartphone.

6.6 Experiment 2: GUI Interaction

For the second experiment, we designed a menu with a list of modification options for a virtual 3D object. The participant had to interact with the menu and change the settings according to the presented information. Therefore, the target state was textually shown next to the object to modify. The menu comprises drop-down items, radio-buttons, slider, and regular buttons.

In this experiment, INPUT has three levels. Additional to *mid-air gesture* and *multi-touch*, we introduce *multi-touch display*. In that condition, the menu was directly presented on the smartphones' display, and the participant could directly interact via touch. In the other conditions, the menu was anchored in space, facing the user, and had to be operated via mid-air or multi-touch gestures. Both visualizations are represented in Figure 6. In total, participants had to complete 36 tasks, 12 in each condition.

Again, we measured the TCT and the raw TLX score. Since the correct input for all parameters was required to complete the task, we did not measure an error rate. TCT was measured from first to last input for each sub-task individually. In total, we recorded 864 menu interaction.

6.7 Results 2: GUI Interaction

We statistically compared the TCT, and the raw TLX, using one-way repeated measures analysis of variance (ANOVA). The results of the second experiment are displayed in Figure 7.

6.7.1 Task Completion Time

We analyzed the average time the participants needed to adjust the 3D model according to the specifications without any error. We measured the highest TCT in the mid-
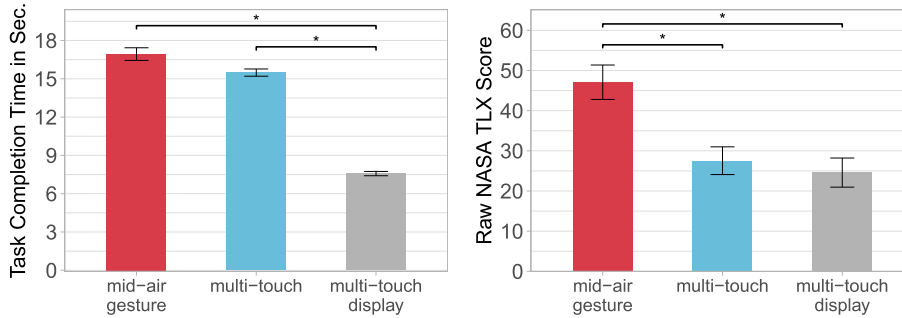
Figure 7: Mean values for the TCT (left) and raw TLX Score (right) for all three conditions of the second experiment. Error bars show standard error of the mean (SE). Asterisk indicate statistically significant differences between conditions.

air gesture condition (M = 16.93, SD = 8.43). Time decreased in the multi-touch input condition (M = 15.48, SD = 4.81), and was lowest in the multi-touch display condition (M = 7.57, SD = 2.85).

A repeated measures ANOVA with a Greenhouse-Geisser correction revealed a statistically significant difference between TCT measurements, $F(1.51, 428.06)$ = 240.25, p < .001, partial η^2 = .46. Bonferroni-adjusted post-hoc analysis revealed a significant difference (p < .001) in completion times of *mid-air gesture* input and *multi-touch display* input (9.38, 95 %-CI[8.17,10.59]) and *multi-touch display* input and *multi-touch* input (7.89, 95 %-CI[7.16,8.24]).

6.7.2 Task Load

We conducted further analyses to assess how the participants subjectively perceived the task load while interacting in AR. Using the smartphones' display for input led to a lower perceived task load (M = 24.58, SD = 17.75) compared to the *multi-touch* input (M = 27.54, SD = 16.92) and *mid-air gesture* input (M = 47.08, SD = 21.00).

A repeated measures ANOVA revealed a significant difference between the input modalities for the GUI interaction task, $F(2, 46)$ = 60.56, p < .001 with a large estimated effect size (partial η^2 = .73). Bonferroni-adjusted post-hoc analysis revealed a significant difference in perceived task load between *mid-air gesture* input (both p < .001) and *multi-touch* (19.54, 95 %-CI[13.43, 25.66]) as well as *multi-touch display* input (22.50, 95 %-CI[16.43, 28.56]).

We encapsulate that using the smartphone as a controller significantly reduced the perceived task load independent of the utilized input method. However, only the utilization of both capabilities, the multi-touch screen as input, and the display as output is significantly reducing the required interaction time.

6.7.3 Overall Preference

In the overall ranking, 16 out of 24 participants preferred the *multi-touch* input for object manipulation. Participants clarify that interactions felt more accurate and less frustrating, especially during rotation and scaling of objects. Only one participant was in favor of *mid-air gesture* input, while seven participants had no explicit preference. For the GUI interaction experiment overwhelming 22 participants preferred the *multi-touch display* interaction. Exact modifications of the GUI slider subjectively result in the most exasperation using *mid-air gestures*. Hence, none preferred this input modality for GUI interactions.

7 Discussion

Our results show that with the smartphone as a tangible multi-touch input controller, users can modify the virtual environment significantly faster compared to state-of-the-art mid-air gesture input. Due to tangible direct multi-touch interaction, the perceived task load is significantly reduced. Our approach benefits, in particular, from using multi-touch interactions that are not overloaded. Results indicate that using the two-finger movement for rotation in parallel to two-finger pinching slows down the overall interaction speed. We examine several factors that are potentially responsible for the overall advances in interaction.

The systems provide visual and auditory feedback in all conditions. However, direct kinesthetic and tactile feedback is only provided in the smartphone-based conditions. Users benefit in general from the well-known device and can perform gestures easier. Contrary, there is no haptic feedback while performing mid-air gestures since

users are less trained to perform these particular gestures. Added physical fatigue during long-lasting interaction further impairs interaction. This gorilla arm effect is also confirmed by several works evaluating mid-air interaction [6, 20].

Interestingly we could not observe any adverse effects for the multi-touch display condition. Typically switching attention to a secondary display causes overhead for the user [13, 16, 34]. However, the participants mentioned that switching and focusing on the AR display and the smartphone display feel unnatural in the first place but quickly became intuitive. Data supports this through the lowest perceived task demand. Contrasting to Gabbard et al. [2, 16] we facilitated a stereo HMD where the number of attention switches was low. We further assume that the potential overhead is minimized since virtual information was approximately displayed at an optical distance of 2.0 m away from the user. Thus, virtual distance and focal distance match and perceptual conflicts, such as space misperception or the vergence-accommodation conflict, were mitigated. In a different scene, where multiple virtual objects are widespread at different distances, perceptual issues such as depth distortion or size misperception could occur [12]. These effects could negatively influence the overhead when attention switching.

Setting users' performance, task load, and accuracy into contrast, our results suggest deploying controller-based multi-touch interaction for simple and complex modifications and interaction in AR environments. Based on the quantitative results and overall preference, we conclude that the smartphone-based input modality provides improved usability. We assume that users benefit through a relaxed body posture, in particular during extended interaction sessions. For future AR systems that require modification of the virtual elements, our findings imply that the support of a smartphone comprising a multi-touch screen offers the best results.

7.1 Limitations and Future Work

Our smartphone controller approach currently supports only a subset of interaction paradigms necessary to fully interact in generic AR applications. However, we demonstrated the feasibility and potential of multi-touch and a secondary touch display in AR environments. Extending our approach with already existing positional tracking solutions [8, 30] would enable fluent, precise, and convenient input in mobile or spontaneous interaction scenarios. Additional sensors like accelerometer, gyroscope, or

proximity could be incorporated to further increase the interaction space.

Our approach uses a well-developed high-precision touch interface for interaction. In contrast, mobile mid-air gesture tracking is more complex and less accurate through different factors that are given by the spatial tracking and temporal camera resolution. Our participants are used to smartphone interaction and gestures on multi-touch displays. Mid-air gestures are less common in everyday interaction with computing systems. Hence, our results may not be generalizable for users with extensive hand gesture experience that may be less prone to fatigue. In our study scenario, the relationship between the secondary handheld display and augmented space is evident. In more complex scenarios, methods need to be implemented to keep it comprehensible to the user when and how information is displayed on the secondary screen. For future research, we envision an investigation on how our smartphone-based controller performs outside the lab and whether other modalities outperform our interaction concept, e. g., during micro-interactions on the move.

8 Conclusion

Consumer Augmented Reality (AR) headsets still lack expressive controllers and interaction concepts. Current implementations of mid-air gestures are slow and physically demanding. This article shows the potential of transforming unmodified smartphones into ubiquitous controllers and extensions for mobile augmented reality experiences. Our approach comprises a smartphone paired with an AR headset. Thus, users can effortlessly interact with the virtual content via multi-touch gestures and further extend the interaction space through a handheld high-resolution touchscreen. We conducted two experiments with 24 participants and investigated the interaction performance with the smartphone as a controller compared to mid-air gestures as a baseline. We found significant differences in task completion time, accuracy, and perceived task load in both experiments. Users benefited from the smartphone as an input controller for fine-grained manipulations, haptic feedback, and the tangible display. Our work contributes to an empirical study showcasing smartphones' viability as a multimodal and ubiquitous input device in mobile AR experiences. The combination of AR glasses and smartphones enables us to build upon an existing interaction space. In the future, our smartphones can be a supportive device in our pockets for extended and improved interaction within mobile AR environments.

References

[1] Mohammed Al-Sada, Fumiko Ishizawa, Junichi Tsurukawa, and Tatsuo Nakajima. 2016. Input Forager: A User-Driven Interaction Adaptation Approach for Head Worn Displays. In *Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia* (Rovaniemi, Finland) (*MUM '16*). Association for Computing Machinery, New York, NY, USA, 115–122. https://doi.org/10.1145/3012709.3012719.

[2] M. S. Arefin, N. Phillips, A. Plopski, J. L. Gabbard, and J. E. Swan. 2020. Impact of AR Display Context Switching and Focal Distance Switching on Human Performance: Replication on an AR Haploscope. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 571–572.

[3] Teo Babic, Harald Reiterer, and Michael Haller. 2018. Pocket6: A 6DoF Controller Based On A Simple Smartphone Application. In *Proceedings of the Symposium on Spatial User Interaction* (Berlin, Germany) (*SUI '18*), Association for Computing Machinery, New York, NY, USA, 2–10. https://doi.org/10.1145/3267782.3267785.

[4] Huidong Bai, Gun A. Lee, and Mark Billinghurst. 2012. Freeze View Touch and Finger Gesture Based Interaction Methods for Handheld Augmented Reality Interfaces. In *Proceedings of the 27th Conference on Image and Vision Computing New Zealand* (Dunedin, New Zealand) (*IVCNZ '12*). Association for Computing Machinery, New York, NY, USA, 126–131. https://doi.org/10.1145/2425836.2425864.

[5] Rafael Ballagas, Jan O. Borchers, Michael Rohs, and Jennifer G. Sheridan. The Smart Phone – A Ubiquitous Input Device. *IEEE Pervasive Computing* (2006).

[6] Sebastian Boring, Marko Jurmu, and Andreas Butz. 2009. 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*, (Melbourne, Australia) (*OZCHI '09*). Association for Computing Machinery, New York, NY, USA, 161–168. https://doi.org/10.1145/1738826.1738853.

[7] Rahul Budhiraja, Gun A. Lee, and Mark Billinghurst. 2013. Interaction techniques for HMD-HHD hybrid AR systems. In *ISMAR*. 243–244.

[8] Rahul Budhiraja, Gun A. Lee, and Mark Billinghurst. 2013. Using a HHD with a HMD for mobile AR interaction. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 1–6.

[9] Aryzon Cardboard. 2019. Aryzon: Augmented Reality powered by your smartphone. https://www.aryzon.com.

[10] Gerhard Reitmayr, Chris Chiu, Er Kusternig, Michael Kusternig, and Hannes Witzmann. [n. d.] iOrb – Unifying Command and 3D Input for Mobile Augmented Reality. In *Proc. IEEE VR Workshop on New Directions in 3D User Interfaces*. 7–10.

[11] David Dobbelstein, Christian Winkler, Gabriel Haas, and Enrico Rukzio. PocketThumb: A Wearable Dual-Sided Touch Interface for Cursor-Based Control of Smart-Eyewear. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 2, Article 9 (June 2017) 17 pages. https://doi.org/10.1145/3090055.

[12] David Drascic and Paul Milgram. 1996. Perceptual issues in augmented reality. In *Stereoscopic displays and virtual reality systems III*, Vol. 2653. International Society for Optics and Photonics, 123–134.

[13] Anna Eiberger, Per Ola Kristensson, Susanne Mayr, Matthias Kranz, and Jens Grubert. 2019. Effects of Depth Layer Switching between an Optical See-Through Head-Mounted Display and a Body-Proximate Display. In *Symposium on Spatial User Interaction* (New Orleans, LA, USA) (*SUI '19*). Association for Computing Machinery, New York, NY, USA, Article 15, 9 pages. https://doi.org/10.1145/3357251.3357588.

[14] Markus Funk, Andreas Bächler, Liane Bächler, Thomas Kosch, Thomas Heidenreich, and Albrecht Schmidt. 2017. Working with Augmented Reality? A Long-Term Analysis of In-Situ Instructions at the Assembly Workplace. In *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments* (Island of Rhodes, Greece) (*PETRA '17*). Association for Computing Machinery, New York, NY, USA, 222–229. https://doi.org/10.1145/3056540.3056548.

[15] Markus Funk, Thomas Kosch, and Albrecht Schmidt. 2016. Interactive Worker Assistance: Comparing the Effects of Head-Mounted Displays, In-Situ Projection, Tablet, and Paper Instructions. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. https://doi.org/10.1145/2971648.2971706.

[16] J. L. Gabbard, D. G. Mehra, and J. E. Swan. Effects of AR Display Context Switching and Focal Distance Switching on Human Performance. *IEEE Transactions on Visualization and Computer Graphics* 25, 6 (June 2019), 2228–2241. https://doi.org/10.1109/TVCG.2018.2832633.

[17] Jens Grubert, Matthias Heinisch, Aaron Quigley, and Dieter Schmalstieg. 2015. MultiFi: Multi Fidelity Interaction with Displays On and Around the Body. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (*CHI '15*). Association for Computing Machinery, New York, NY, USA, 3933–3942. https://doi.org/10.1145/2702123.2702331.

[18] Taejin Ha and Woontack Woo. 2011. ARWand: Phone-Based 3D Object Manipulation in Augmented Reality Environment. In *2011 International Symposium on Ubiquitous Virtual Reality (ISUVR)*. IEEE, 44–47.

[19] Sandra G. Hart. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (Nov. 2016), 904–908.

[20] Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed Endurance: A Metric to Quantify Arm Fatigue of Mid-Air Interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (*CHI '14*). Association for Computing Machinery, New York, NY, USA, 1063–1072. https://doi.org/10.1145/2556288.2557130.

[21] Juan David Hincapié-Ramos, Kasim Ozacar, Pourang P. Irani, and Yoshifumi Kitamura. 2015. GyroWand: IMU-Based Raycasting for Augmented Reality Head-Mounted Displays.

In *Proceedings of the 3rd ACM Symposium on Spatial User Interaction* (Los Angeles, California, USA) (*SUI '15*). Association for Computing Machinery, New York, NY, USA, 89–98. https://doi.org/10.1145/2788940.2788947.

[22] Teresa Hirzle, Jan Rixen, Jan Gugenheimer, and Enrico Rukzio. 2018. WatchVR: Exploring the Usage of a Smartwatch for Interaction in Mobile Virtual Reality. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI EA '18*). Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188629.

[23] Jinki Jung, Jihye Hong, Sungheon Park, and Hyun S. Yang. 2012. Smartphone as an Augmented Reality Authoring Tool via Multi-Touch Based 3D Interaction Method. In *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry* (Singapore, Singapore) (*VRCAI '12*). Association for Computing Machinery, New York, NY, USA, 17–20. https://doi.org/10.1145/2407516.2407520.

[24] Daniel Kharlamov, Brandon Woodard, Liudmila Tahai, and Krzysztof Pietroszek. 2016. TickTockRay: Smartwatch-Based 3D Pointing for Smartphone-Based Virtual Reality. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology* (Munich, Germany) (*VRST '16*). Association for Computing Machinery, New York, NY, USA, 365–366. https://doi.org/10.1145/2993369.2996311.

[25] Konstantin Klamka and Raimund Dachselt. 2018. ARCord: Visually Augmented Interactive Cords for Mobile Interaction. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI EA '18*). Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188456.

[26] Jingbo Liu, Oscar Kin-Chung Au, Hongbo Fu, and Chiew-Lan Tai. Two-Finger Gestures for 6DOF Manipulation of 3D Objects. *Computer Graphics Forum* 31, 7 (2012), 2047–2055. https://doi.org/10.1111/j.1467-8659.2012.03197.x. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8659.2012.03197.x.

[27] Tariq Masood and Johannes Egger. Augmented reality in support of Industry 4.0 – Implementation challenges and success factors. *Robotics and Computer-Integrated Manufacturing* 58 (2019), 181–195.

[28] Sven Mayer, Lars Lischke, Jens Emil Grønbæk, Zhanna Sarsenbayeva, Jonas Vogelsang, Paweł W. Woundefinedniak, Niels Henze, and Giulio Jacucci. 2018. 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* (Montreal QC, Canada) (*CHI '18*). Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3173574.3174113.

[29] Mark Mine, Arun Yoganandan, and Dane Coffey. 2014. Making VR Work: Building a Real-World Immersive Modeling Application in the Virtual World. In *Proceedings of the 2nd ACM Symposium on Spatial User Interaction* (Honolulu, Hawaii, USA) (*SUI '14*). Association for Computing Machinery, New York, NY, USA, 80–89. https://doi.org/10.1145/2659766.2659780.

[30] Peter Mohr, Markus Tatzgern, Tobias Langlotz, Andreas Lang, Dieter Schmalstieg, and Denis Kalkofen. 2019.

TrackCap: Enabling Smartphones for 3D Interaction on Mobile Head-Mounted Displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (*CHI '19*). Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3290605.3300815.

[31] Florian Müller, Joshua McManus, Sebastian Günther, Martin Schmitz, Max Mühlhäuser, and Markus Funk. 2019. Mind the Tap: Assessing Foot-Taps for Interacting with Head-Mounted Displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland UK) (*CHI '19*). Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300707.

[32] E. Normand and M. J. McGuffin. 2018. Enlarging a Smartphone with AR to Create a Handheld VESAD (Virtually Extended Screen-Aligned Display). In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 123–133. https://doi.org/10.1109/ISMAR.2018.00043.

[33] Krzysztof Pietroszek, James R. Wallace, and Edward Lank. 2015. Tiltcasting: 3D Interaction on Large Displays Using a Mobile Device. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) (*UIST '15*). Association for Computing Machinery, New York, NY, USA, 57–62. https://doi.org/10.1145/2807442.2807471.

[34] Umar Rashid, Miguel A. Nacenta, and Aaron Quigley. 2012. The Cost of Display Switching: A Comparison of Mobile, Large Display and Hybrid UI Configurations. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (Capri Island, Italy) (*AVI '12*). Association for Computing Machinery, New York, NY, USA, 99–106. https://doi.org/10.1145/2254556.2254577.

[35] Houssem Saidi, Marcos Serrano, Pourang Irani, and Emmanuel Dubois. 2017. TDome: A Touch-Enabled 6DOF Interactive Device for Multi-Display Environments. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (*CHI '17*). Association for Computing Machinery, New York, NY, USA, 5892–5904. https://doi.org/10.1145/3025453.3025661.

[36] Stefan Schneegass and Alexandra Voit. 2016. GestureSleeve: Using Touch Sensitive Fabrics for Gestural Input on the Forearm for Controlling Smartwatches. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers* (Heidelberg, Germany) (*ISWC '16*). Association for Computing Machinery, New York, NY, USA, 108–115. https://doi.org/10.1145/2971763.2971797.

[37] Marcos Serrano, Dale Hildebrandt, Sriram Subramanian, and Pourang Irani. 2014. Identifying Suitable Projection Parameters and Display Configurations for Mobile True-3D Displays. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Toronto, ON, Canada) (*MobileHCI '14*). Association for Computing Machinery, New York, NY, USA, 135–143. https://doi.org/10.1145/2628363.2628375.

[38] Hemant Bhaskar Surale, Aakar Gupta, Mark Hancock, and Daniel Vogel. 2019. 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* (Glasgow, Scotland UK) (*CHI '19*). Association for Computing Machinery, New York, NY, USA, Article 13, 13 pages.

https://doi.org/10.1145/3290605.3300243.

[39] Philipp Wacker, Oliver Nowak, Simon Voelker, and Jan Borchers. 2019. 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* (Glasgow, Scotland UK) (*CHI '19*). Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300849.

[40] W. Zhao and V. Madhavan. 2005. Integration of voice commands into a virtual reality environment for assembly design. In *Proceedings of the 10th annual international conference on industrial engineering theory, applications & practice* (Clearwater Beach, FL, USA).

[41] Xianjun Sam Zheng, Cedric Foucault, Patrik Matos da Silva, Siddharth Dasari, Tao Yang, and Stuart Goose. 2015. Eye-Wearable Technology for Machine Maintenance: Effects of Display Position and Hands-Free Operation. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) (*CHI '15*). Association for Computing Machinery, New York, NY, USA, 2125–2134. https://doi.org/10.1145/2702123.2702305.

[42] Fengyuan Zhu and Tovi Grossman. 2020. BISHARE: Exploring Bidirectional Interactions Between Smartphones and Head-Mounted Augmented Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (*CHI '20*). Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3313831.3376233.

Bionotes

Pascal Knierim
Human-Centered Ubiquitous Media, LMU Munich, Munich, Germany
pascal.knierim@ifi.lmu.de

Pascal Knierim is a postdoctoral researcher at the Human-Centered Ubiquitous Media group at the LMU Munich, Germany. Prior, he was at the Human-Computer Interaction and Cognitive Systems group at the University of Stuttgart and spent some time as an intern at Microsoft Research in Cambridge. His research focus includes augmented and virtual reality as ubiquitous computing technology. He investigates how to enhance input and output modalities for mixed reality applications and how they will impact our everyday lives. He developed advanced skills in ideation, conceptual design, development, and evaluation of novel mixed reality applications based on many research projects.

Dimitri Hein
Human-Centered Ubiquitous Media, LMU Munich, Munich, Germany
dimitri.hein@outlook.de

Dimitri Hein is a software engineer at the MaibornWolff GmbH in Munich. His focus lies in the development of virtual and augmented reality applications and the conceptualization and ideation of application features. Prior, he studied computer science at the LMU Munich and received his Master's in 2019.

Albrecht Schmidt
Human-Centered Ubiquitous Media, LMU Munich, Munich, Germany
albrecht.schmidt@ifi.lmu.de

Albrecht Schmidt is a professor at the LMU Munich and has a chair for Human-Centered Ubiquitous Media at the Department of Informatics. Prior, he directed the Human-Computer Interaction Research group at the Institut for Visualizations and Interactive Systems at the University of Stuttgart. His research encompasses the development and evaluation of augmented and virtual reality applications to enhance the user's life quality through technology. This includes the amplification of human cognition and physiology through novel technology.

Thomas Kosch
Human-Computer Interaction, Technical University of Darmstadt, Darmstadt, Germany
thomas.kosch@tk.informatik.tu-darmstadt.de

Thomas Kosch is a postdoc and group leader of the Human-Computer Interaction Group at the Technical University of Darmstadt (TU Darmstadt). He obtained his Ph. D. in 2020 at the Ludwig Maximilian University of Munich (LMU Munich). His research focuses on the analysis of physiological and behavioral data for real-time interaction in augmented and virtual reality. This includes the implementation, design, and creation of adaptive user interfaces and intelligent virtual environments that adjust to the user's perceived workload, memory capacities, and attention.