BISHARE: Exploring Bidirectional Interactions Between Smartphones and Head-Mounted Augmented Reality

Fengyuan Zhu and Tovi Grossman
University of Toronto, Toronto, CANADA
fyzhu@dgp.toronto.edu, tovi@dgp.toronto.edu

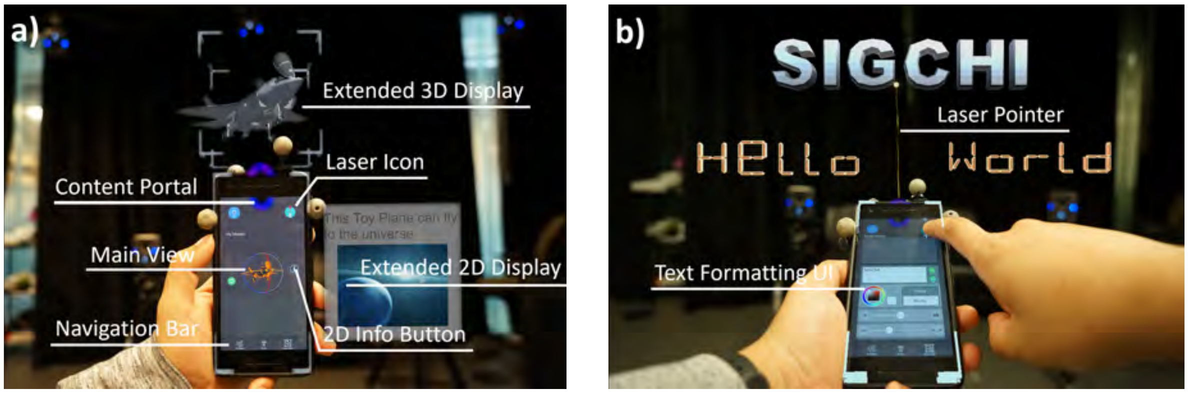

***Figure 1. We explore joint interactions between an Augmented Reality Head-Mounted Display and a smartphone. a) The HMD can be used to enhance smartphone interactions, for example, by displaying in-place previews of 3D content. b) The smartphone can be used to enhance AR centric task, for example, by supporting object selection or detailed UI configurations.**

ABSTRACT

In pursuit of a future where HMD devices can be used in tandem with smartphones and other smart devices, we present BISHARE, a design space of cross-device interactions between smartphones and ARHMDs. Our design space is unique in that it is bidirectional in nature, as it examines how both the HMD can be used to enhance smartphone tasks, and how the smartphone can be used to enhance HMD tasks. We then present an interactive prototype that enables cross-device interactions across the proposed design space. A 12-participant user study demonstrates the promise of the design space and provides insights, observations, and guidance for the future.

Author Keywords

Augmented Reality; Smartphones; Cross-Device Computing; Mixed-Reality Computing.

CSS Concepts

• **Human-centered computing~Human computer interaction (HCI)**; *Interaction Paradigms*; Mixed /augmented reality;

CHI '20, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6708-0/20/04...$15.00.
https://doi.org/10.1145/3313831.3376233

INTRODUCTION

Augmented Reality Head-Mounted Displays (ARHMDs), such as the Microsoft HoloLens [39] and Magic Leap [37], are becoming increasingly popular and sparking interesting discussion about the future of personal computing. Some have prophesized that as their form factor evolves, they could become the "ultimate" personal computing device [17].

However, in their current state, ARHMDs suffer from limitations, such as constrained fields of view [7], requiring unfamiliar freehand gestures, and lacking tactile feedback offered by true physical interactions. For these reasons, we do not expect in any near future that ARHMDs will replace our smartphones or other personal computing devices. To the contrary, we believe there are rich opportunities afforded by using ARHMD devices in tandem with smartphones, where their contrasting benefits can be jointly leveraged.

Along these lines, a number of research efforts have begun to explore the cross-device opportunities arising from using a mobile device and ARHMD together [3, 23, 40, 55]. For the most part, this prior work can either be classified as using HMDs to support phone-centric tasks, such as extending the display space of the phone [42], or alternatively, using the smartphone to support HMD-centric tasks, such as using the phone as a 6DOF controller for spatial content [40]. Little work has been dedicated to creating a coherent bidirectional platform where both of these interaction paradigms are enabled. Furthermore, prior research in this area is scattered across independent systems and research communities, creating somewhat of a fragmented research landscape.

*Note: Figures have been created by overlaying the rendering from the HoloLens App on top of a photo taken by a camera that was tracked in 3D.

Inspired by the vast work on cross-device interaction [10], this paper synthesizes prior work that has coupled the use of mobile devices and ARHMDs, into a holistic design space (*BISHARE*) of six classes of cross-device interactions that their joint use affords, for both phone-centric and HMD-centric tasks. This design space is grounded by both an analysis of the prior literature, and the contrasting affordances of the two interactive platforms. Our analysis of these prior works indicates that a critical gap exists in the literature: specifically, systems which allow for both phone-centric *and* HMD-centric tasks, as well as techniques to continuously transition between these two interaction paradigms.

To validate the concepts and elements of our design space, and to fill this important research gap, we developed an interactive prototype that supports joint interactions between an ARHMD and a smartphone. The prototype implements exemplary interaction techniques across the six elements of our design space, and demonstrates how each can be utilized for both phone-centric and HMD-centric tasks. To gather feedback on the design space and developed interaction techniques, we ran a user study that provides insights, observations, and guidance for future work.

This work has 2 main contributions. 1) We derive and implement a design space of joint interactions between smartphones and ARHMDs that each support both phone-centric and HMD-centric tasks and 2) We present findings and insights from a 12-person user study that enables an informed discussion regarding the challenges and opportunities associated with jointly using smartphones and ARHMDs.

RELATED WORK

Our related work covers a review of input techniques for smartphones, ARHMDs, and cross-device interactions. We note that a full treatment of these three topics is beyond the scope of this paper, and direct the reader to respective surveys for further details [9, 10, 30].

Smartphone Interaction

While touch may be the most common form of input on smartphones, researchers have explored many other input paradigms to increase their input vocabulary or expressivity. For instance, past research has explored how orientation [30], grip [22], tilting [47], and whacking [31], can all be used to interact with a phone. Relevant to our work is research that has explored ways to extend the input and output region of a phone beyond its physical boundaries. Many sensing techniques for such *Around-Device Interactions* have been developed, such as IR sensors [13], depth cameras [15], GSM signals [63], and external cameras [26]. We extend this work further by combining smartphones with AR devices, to allow users to leverage not only the space immediately surrounding the device, but the full surrounding spatial environment.

Head-Mounted Augmented Reality Interaction

A large variety of interaction techniques have also been explored for augmented reality environments. One of the most common forms of input is freehand spatial input [18, 34]. Another common form of input in AR is Gaze [35], which is commonly combined with freehand gestures [35, 44, 54], such as the air tap technique used for selection in the Microsoft HoloLens. Other forms of input for AR devices include speech and custom handheld controllers [9]. While the use of handheld controllers may detract for the real-world experience that AR offers, we believe that it can offer important advantages, which we will explore in detail.

Cross-Device Computing

Research in the area of cross-device computing and interactions has been rapidly growing, leading to an explosion of research topics, terminology, techniques, and systems. Brudy [10] conducted a recent analysis and presented a taxonomy based on a corpus of 510 papers in the *cross-device computing* domain, which serves as important groundwork for our own explorations. A large proportion of cross-device computing research that utilizes mobile devices are developed within the context of other 2D devices, such as tablets [25, 46], wearables [16], or large displays [50, 62].

Related to our efforts, cross-device computing is often used to extend the input and output areas of interactive devices [7, 8, 60]. Also related, a number of cross-device platforms have used a mobile phone to provide spatial input for a secondary display, like desktop [40, 49], mobile [33], tablet [61], large display [4, 45, 49 ,53, 61], tabletop [49, 50] and HMD [40, 41]. On the other hand, there is work combing see-through spatial displays to enhance desktop [36, 51, 57], mobile [51, 56, 58], and tabletop experiences [14, 24]. We extend these lines of work by exploring the design space afforded by coupling a smartphone with a spatial head-mounted display.

Combining 2D Mobile Devices and HMDs

Previous work that combines 2D mobile devices with HMDs can roughly be categorized as *mobile-centric*, where HMDs are used to enhance mobile interaction, and, *HMD-centric*, where mobile devices are used to enhance spatial interaction.

An inspirational example of mobile-centric joint interactions is seen with the Multifi system [23] which uses an ARHMD to enhance smartwatch or smartphone interactions, providing the users with a dynamically aligned head's up display that can be used for content previews, extended screens and interactive widgets. MultiFi is also one of the few systems that also demonstrates a set of HMD-centric interactions, where the phone can also be used as input for the HMD. Normand and McGuffin also propose a set of mobile-centric interactions, where the HMD extends the display space of a phone, and use the term VESAD to refer to a virtually extended screen-aligned display [42]. Gluey uses a head-worn display to support input and data transitions in distributed display environments [51]. For example, a user can move content from their desktop monitor to a nearby tablet, by shifting their view to the tablet while dragging content on the HMD with the mouse.

Conversely, researchers have also explored how 2D mobile devices can enhance interaction with spatial displays. Budhiraja et al. [11, 12] carried out an initial exploration of how a smartphone could be used to enable new types of mobile AR interactions. They proposed two ways the smartphone could be used for input – using its touch sensitive display and through spatial movements of the smartphone itself. Al-Sada et al. [1] explored how users could borrow different embedded inputs from smartwatches, smartphones and tablets to interact with AR content. Several researchers have also looked at how spatially tracked 2D devices, such as tablets [11, 19, 20, 55] and phones [40] can be used to enhance immersive design applications. For example, DualCAD is an immersive 3D design tool which allows a smartphone to be used as a complementary high-resolution display and input device [40]. TabletInVR uses a 3D-tracked multi-touch tablet in an immersive VR environment to augment and enhance 3D modelling operations [55]. To support spatial tracking of mobile devices in such systems, the device's IMU data can be used, but this may be prone to drift [28]. TrackCap instead uses the phone to "track a cap" that is mounted on the HMD [41].

The research described in this section primarily support a unidirectional relationship between the mobile and HMD devices – either the mobile device is used as an auxiliary device to enhance HMD-centric tasks, or the HMD is used to enhance mobile-centric tasks. Our work synthesizes this research into a cohesive design space, and builds upon these prior systems by supporting a bidirectional and continuous relationship between the mobile and HMD platforms within an interactive prototype.

Continuous Interaction between 2D and 3D Spaces

Prior research has explored how 2D and 3D systems can be utilized to create a continuous cross-device experience. Benko et al. introduced "cross-dimensional interaction techniques" to support the seamless transition of data between a 2D tabletop and 3D HMD [6]. Marquardt et al. propose the "continuous interaction space" which similarly explores gestures that begin with direct touch on a tabletop display and then continue with a freehand gesture [38]. Chen et al. expand on this work to support Air+Touch gestures that interweave touch and in-air gestures [15]. Serrano et al. used a projection setup to emulate and explore interactions for smartphone 3D display [52]. More recently, Roo and Hachet demonstrate a Mixed Reality ecosystem that allows users to incrementally transition from physical to virtual experiences across a variety of tabletop, handheld, and head-mounted displays [48]. XD-AR offers a similar framework for collaborative AR scenarios [54]. These research works all provide inspiration for our exploration, but none explicitly explore the continuous interaction space between a smartphone and ARHMD.

In summary, our work builds on the body of literature by 1) synthesizing prior work into a coherent bidirectional design space, revealing new interaction opportunities, and 2) implementing a set of interaction techniques that demonstrate the range of ways for which smartphones can enhance spatial-centric tasks, ways in which ARHMDs can enhance smartphone-centric tasks, and methods for transitioning between these two contexts of use.

CAPABILITIES OF SMARTPHONES AND ARHMDS

Our design space of joint interactions is largely grounded by the contrasting and complementing properties of smartphones and ARHMDs, which we first describe below (Table 1).

Capabilities and Limitations of Smart Phones

Smartphones have become ubiquitous within our society, and an essential part of many people's lives [27]. They are readily available and people are already familiar with their usage patterns.

Direct physical touch is the standard interaction method for today's smartphones supporting efficient, precise, and expressive input. A recent advance in smartphones is their ability to also track their 3D positions in space based on SLAM solutions like ARKit [4]. However, these solutions currently suffer from drifting and latency issues. Beyond providing a physical surface for touch input, the tangible nature of smartphones offer numerous interaction opportunities. Examples like holding orientations [30], grip gestures [22], haptic edges input [32] and physical collision [50, 62] show the potential of a phone's tangibility.

In terms of output, rendering high-resolution, full-color, 2D content is now the standard for smartphones. The display quality of smartphones makes them a device of choice for detailed viewing of text, images, and video.

One of the biggest limitations of smartphones is the physical size of their display and input areas. Furthermore, aside from handheld AR experiences [21], smartphones are unable to render stereoscopic imagery in a user's spatial environment. Finally, unlike HMDs, most smartphones cannot inherently track a user's head position, gaze, or freehand gestures.

Capabilities and Limitations of Head-Mounted AR

As with smartphones, ARHMDs have evolved extensively since their first development. Unlike smartphones, ARHMDs can vary greatly from one manufacturer to the next. In this paper, we consider modern ARHMDs such as the HoloLens, Meta, and Magic Leap.

The main benefit of ARHMDs is their ability to display stereoscopic imagery in the context of the user's spatial environment, which has resulted in an entire class of augmented reality systems [9]. Many ARHMDs also support spatial input via vison-based tracking of the user's hands. This allows users to interact with their surrounding environment using a combination of direct manipulation and freehand gestures. Another inherent benefit of ARHMDs is that they can track the position and orientation of the user's head. This offers a valuable input channel, as the user's area of attention can be inferred [14]. Head movements are also often used as an explicit input channel to control a cursor, combined with spatial or voice input for manipulations [35].

		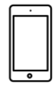	
Input	Precise Input	✓	x
	Spatial Input	Limited	✓
	Head Tracking	Limited	✓
	Tangibility	✓	x
	Familiar Interaction Patterns	✓	x
Output	High Resolution Display	✓	x
	Stereo Display	x	✓
	Spatial Display	Limited	✓
	Large Field of View	x	Limited

Table 1. A summary of some complementing properties of smartphones and augmented reality head-mounted displays.

One of the major limitations of ARHMDs is their rendering capabilities. Current hardware and rendering pipelines put constraints on the resolution, color spaces and field-of view. These constraints make it difficult to use ARHMDs for extensive reading or detailed information gathering tasks. Similarly, precise input can be difficult when using ARHMDs, as they typically require freehand gestures or custom 6DOF controllers to interact with spatial widgets. As demonstrated by Arora et al. [2], the absence of physical surfaces places a significant constraint on detailed input. Freehand gestures (such as the HoloLens "Bloom" gesture) can also be unfamiliar to users and are prone to recognition errors, making them difficult to learn, in comparison to well-known smartphone interactions.

DESIGN PRINCIPLES

An interesting observation from our above discussion is that many of the properties of smartphones and HMDs complement one another (Table 1). One's strength is the other's weakness. This provides additional motivation to design unified interactive platforms, which can jointly utilize their contrasting capabilities. Grounded by these contrasting capabilities, and our review of the related literature, we now describe a set of the design principles for creating joint interactions between smartphones and HMDs.

Supporting Continuous Bi-Directional Interaction

Most previous research developed joint-interaction between mobile devices and HMDs for either mobile-centric [23, 42, 51] or HMD-centric [3, 40, 55] interactions. Given the devices' contrasting capabilities, we believe there is an open opportunity to develop bidirectional frameworks which can support both 2D and spatial HMD centric tasks, as well as supporting a continuous transition between these two contexts of use.

Phone for Precise Interactions; AR for Spatial Interactions

Regardless of the task focus (AR centric vs. Phone centric), an important design principle is to utilize the phone for detailed and precise interaction, and utilize the ARHMD for coarse and spatial interaction [3]. For example, the phone may be used for tasks such as interacting with detailed widgets and rendering high-resolution imagery, while the HMD may be used for tasks such as visualizing 3D components and specifying large spatial regions of interest.

Preserving Known Interaction Patterns

When designing joint interactions between smartphones and ARHMDs, the legacy bias must be considered, where users resort to well-known interaction styles even when more effective and novel techniques are available [10]. Instead of developing completely new interaction patterns, it may be beneficial to leverage the interaction paradigms already familiar on each respective platform.

A Continuum of Display Spaces

Combining a smartphone with an ARHMD gives rise to a novel continuum of digital display spaces. This ranges from fully 2D (inherent to mobile devices), to fully spatial (inherent to ARHMDs). Furthermore, extensive prior research has shown the benefits of leveraging the space immediately surrounding an interactive device [13, 15, 26, 63]. To support continuous transitions along this continuum [38, 51, 52, 54, 55], we propose three semantically meaningful spaces that input, output, and content can transition between: within the phone, around the phone and within the spatial environment (Figure 2).

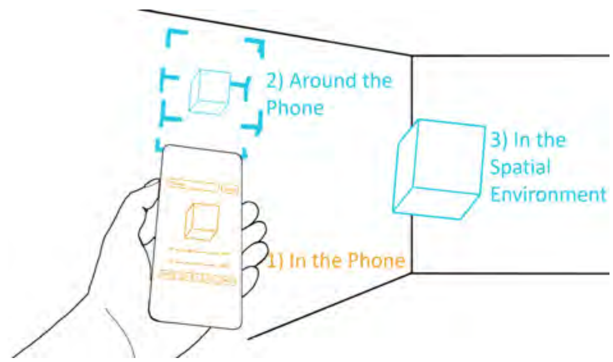

Figure 2. The continuum of display spaces arising from joint interaction between a smartphone and an ARHMD.

BISHARE: A DESIGN SPACE OF JOINT INTERACTIONS

Based on the above described principles, and our survey of the prior literature, we have identified six major categories of joint interactions between smartphones and ARHMDs. By considering each of these interactions through the lens of both phone-centric (P) and HMD-centric (H) tasks, we form BISHARE (**B**idirectional **I**nteractions between **S**martphones and **H**ead-Mounted **A**ugmented **Re**ality), a novel 2 x 6 bidirectional design space (Figure 3). Each cell of the design space in Figure 3 is populated with a single exemplary interaction technique, but actually represents an entire class of interaction opportunities. We note that some individual aspects of this design space have been explored in prior work, and include relevant citations where appropriate.

D1. Distributed Input

Distributed input relates to the class of techniques for which one platform is used to provide input for the other. For example, in a phone-centric task, a user may wish to use a spatial gesture to scroll a webpage or manipulate content, so that they do not occlude the content which they are working with [13, 59] (D1P). In a spatial-centric task, a user may wish to use the smartphone as a 6DOF device to perform ray casting for object selection of 3D components [41] (D1H). Using the touchscreen, recasting could be enhanced with gestures, for example, to cycle between overlapping objects, or define a subsequent manipulation.

D2. Distributed UI

Related to distributed input, distributed UI represents the concept of placing visual interactive widgets on one platform that are used to manipulate content or interactions on the other platform [10]. For example, the HMD can provide extra UI components surrounding the mobile device, to support a full viewing experience for when using mobile applications [42] (D2P). Conversely, the phone can be used to provide detailed UI elements for manipulating spatial objects [3] (D2H).

D3. Distributed Display

One of the most appealing interactions afforded by the combination of these platforms is the ability to extend and enhance display spaces. For example, the ARHMD can extend the mobile display area with 2D or 3D content, such as an extra screen for pictures, graphs and text [23, 52] (D3P). In spatial-centric tasks, the high-resolution display of the phone can be utilized to enhance the display space of the ARHMD, by extending its field of view [7], displaying detailed text or textures that would be difficult to perceive on the ARHMD, or acting as a focus + context lens [5, 21] (D3H).

D4. Content Transfer

An interesting opportunity afforded by the combination of a smartphone and ARHMD is the ability to transfer content between modalities [51]. In a phone-centric task, such as browsing a webpage, users could smoothly drag content from the phone to the area directly surrounding the phone, perhaps to create a miniature 3D representation of the content (D4P). Conversely, when interacting with spatial content, the 3D content could be brought into the spatial proximity of the phone for contextual viewing (D4H), or dragged into the phone itself for detailed manipulations [40].

D5. Cross-Platform Gestures

The use of free-hand gestures, local touch events, and controller-based input events can all be combined to create new forms of cross-platform gestures that incorporate both input platforms [16]. For example, to support content transfer in a phone-centric context (D4P), content on the phone can be dragged using a touch operation, and once reaching the edge, the movement can continue as a freehand pinch gesture [38] (D5P). In a HMD-centric setting, a user could combine spatial movements of the phone itself, with 2D touchscreen gestures on the phone, to perform manipulations of spatial content (D5H) .

D6. Tangible Interaction

Directly leveraging its tangibility, the phone can also be used as a spatial tangible input device [29]. In a 2D-centric context, different orientations of the phone could trigger different operations. For example, holding the phone flat (screen facing up) could create a virtual shelf to preview 3D content (D6P). When interacting with spatial content, the physical boundary of the phone could be used to collide with augmented objects as a mechanism to trigger object-specific events (D6H) [55].

Synthesis of Prior Work

Based on the above design space, we can synthesize prior work, comparing their coverage of each element of the design space (Table 2). Notably, our analysis of the prior work indicates that each prior system typically focuses on only one direction of the relationship between the two interactive platforms. The main exception is MultiFi, which considers the bidirectional relationship between a head-mounted display and mobile devices [23]. However, their implementation focused on 2D tasks and content, and thus did not exploit the unique realm of spatial 3D tasks and interactions that the combination of devices affords.

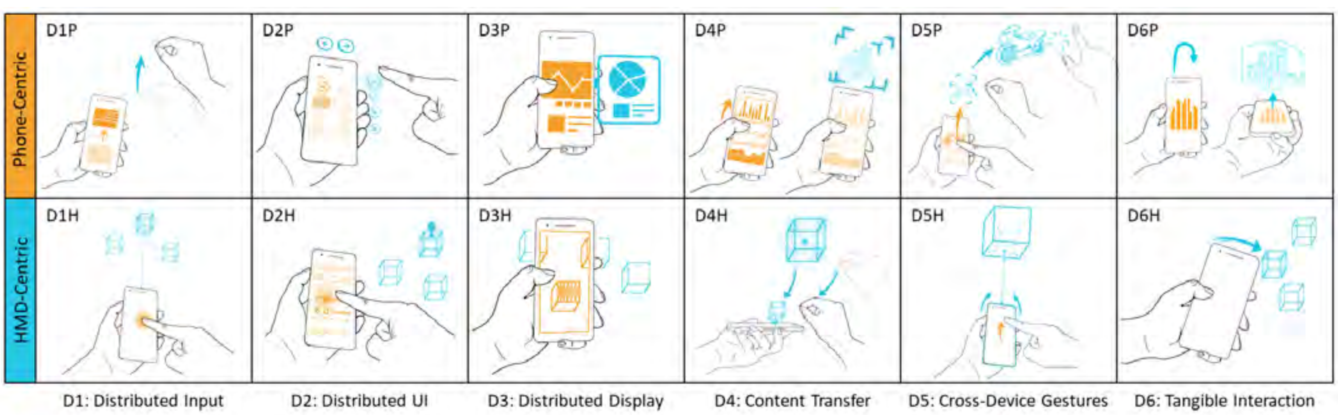

Figure 3. The BISHARE design space of joint interactions between a smartphone and augmented reality head-mounted display. Each cell contains a single example joint interaction, but represents a broader class of interaction techniques that may be possible.

		D1		D2		D3		D4		D5		D6	
		Mobile	HMD	Mobile	HMD	Mobile	HMD	Mobile	HMD	Mobile	HMD	Mobile	HMD
Phone Centric	MultiFi [23]		✓	✓	✓	✓	✓	✓					
	VESAD [42]	✓			✓	✓		✓					
	Gluey [51]	✓				✓		✓		✓			
	Mobile True3D [52]		✓	✓	✓	✓						✓	
HMD Centric	HMD-HHD [11,12]		✓				✓		✓		✓		
	Input Forager [1]		✓		✓								
	Hyve3D [19,20]		✓				✓						
	SymbiosisSketch [3]		✓		✓								
	TabletInVR [55]	✓	✓	✓	✓		✓				✓		✓
	DualCAD [40]		✓		✓				✓		✓		✓
	TrackCap [41]		✓				✓				✓		
	BISHARE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

✓ Mobile-Centric ✓ HMD-Centric

Table 2. Summary of prior work within the context of our design space. Our design space covers the entire combination of elements that prior systems have explored in isolation.

INTERACTIVE PROTOTYPE

To address this gap in the research literature, we developed an interactive prototype, which supports joint interactions between a Smartphone and an ARHMD. Within this prototype we demonstrate both phone-centric and spatial HMD-centric interactions, as well as techniques to support continuous transitions between these two tasks contexts. The prototype allows users to manage and manipulate augmented 3D content in their surrounding environment. The interface has three main functionality modes: 3D Models, 3D Text, and Group Select (Figure 4).

We note that a subset of the techniques we present have been previously demonstrated in isolation, but we believe this is the first synthesis of such techniques across the entire spectrum of our design space, and more specifically, that broadly supports both phone centric and HMD centric tasks. When features are described, the associated area of the design space (D1-D6; P/H) will be noted in parenthesis.

Implementation

We used a Google Pixel 2 smartphone and Microsoft HoloLens as our primary hardware components. Both the smartphone and HoloLens ran a mobile app written in Unity 2017.4, with a Host/Client gaming logic to synchronize objects and events within the network. In order to support accurate tracking with low latency, we used a *Vicon system to track the Pixel 2 and HoloLens, and synchronized their relative coordinate systems through a local network. One of the user's finger was also tracked by the Vicon, to support phone edge swiping. Other freehand gestures like gaze, air-tap and dragging were enabled natively by the HoloLens OS. As vision algorithms improve, external tracking equipment may not be necessary. However, the Vicon system served our purpose of fully exploring the interaction design space. As such, our hardware setup should be considered an enabling technology, rather than our envisioned configuration.

*Vicon system is a commercial solution for Motion Capture purpose. Details can be found in https://www.vicon.com/

Figure 4. The 2D smartphone interface consists of three main views. a) 3D Models. b) 3D Text. C) Group Select

Extended Display Spaces

The prototype makes use of extended display spaces surrounding the smartphone, rendered by the ARHMD (D3) (Figure 1). At the top of the phone's display, there is a portal for the user to teleport objects between display spaces (D4). The bottom half of the portal is rendered on the phone in 2D, while the top half is rendered with the HoloLens in 3D. The space immediately above the portal is used for previewing 3D content, while the space to the side of the phone is used to display supplementary 2D content.

Object Selection

An icon on the phone interface is used to display a laser pointer in AR [41], which acts like a ray pointer for object selection (D1H) (Figure 5a). If the laser intersects an augmented object, the laser endpoint will be highlighted, and the phone interface will change to the view that matches the type of content being selected. When intersecting an object, the user can swipe downwards to transfer the object from AR to the phone (D5H) (Figure 5b).

Alternatively, objects in the spatial environment can be accessed using freehand gestures. An air tap and dragging gesture can be used to select and move an object in 3D space with a 1-1 mapping (Figure 6). The object can be transferred to the phone by dragging it towards the portal (D4H).

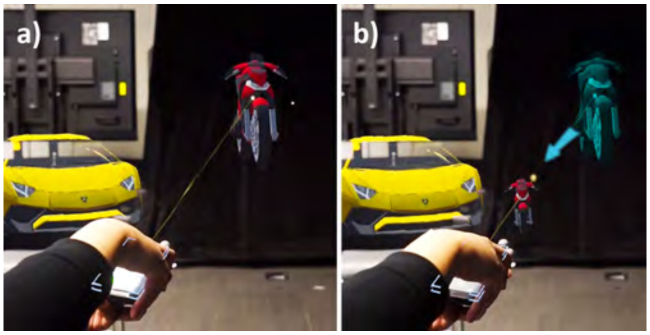

Figure 5. a) The phone can be used to control a laser. b) A swipe gesture brings the content into the 3D preview area.

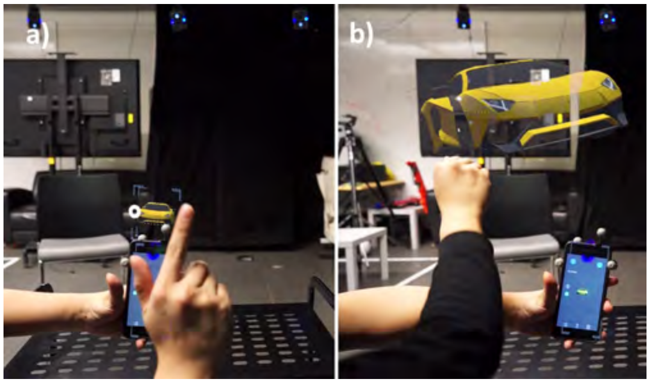

Figure 6. A freehand air tap and drag gesture can be used to move augmented content in the environment.

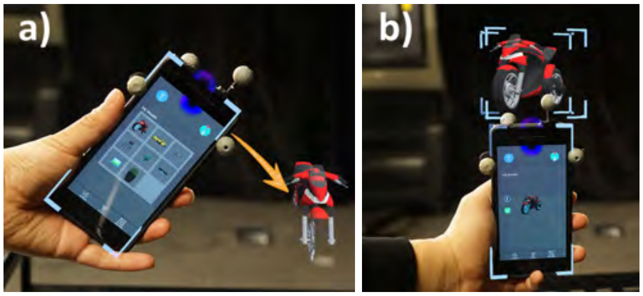

Figure 7. Knocking the phone against 3D augmented content will capture it into the 3D preview area of the phone.

A final method of selecting augmented content in 3D space is by knocking a corner of the phone against it with a spatial gesture (D6H). After a knocking gesture occurs, a preview copy of the 3D content is rendered on the extended display area of the phone (Figure 7). To provide feedback to the user, the phone will also vibrate once a collision is detected.

Application Modes

We now describe the three main modes of the prototype, accessed with icons at the bottom of the phone interface.

3D Model Operations

The 3D Models mode allows users to select, manipulate and place 3D models in the augmented environment. To start, the user can select which model they want to preview on the smartphone interface. They can tap a model to select it, and drag it to the portal to move it into the extended display space (D4P) (Figure 8a-b). This allows the user to see a miniature preview of the model before moving a full-scale version of the model into their spatial environment. The user can then use an air tap to continue the gesture (D5P) and move the model away from the phone, at which point it transitions to its actual 1-1 size [40] (Figure 6).

While viewing 3D models on the phone, users may want to get a quick 3D preview prior to moving it through the portal. To do so, we implemented a spring-loaded preview gesture, where the user flattens the phone (parallel to the ground), as if they were holding a shelf (D6P). A miniature version of the 3D model is displayed directly on top of this "shelf" (Figure 8c). During this temporary preview, the user can rotate the object by either rotating the phone, or swiping on its display (D1H). The preview is removed as soon as the phone is returned to its default orientation.

Once a 3D model is placed in the spatial environment, a full screen view on the phone can be used for further detailed transformations. A rotation widget allows the user to adjust the orientation of the model, and an icon can be tapped to return the model to its default orientation (Figure 9) (D2H).

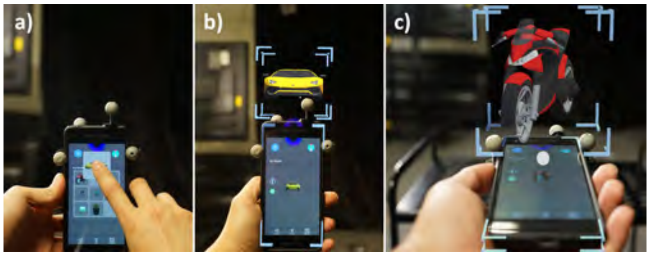

Figure 8. a-b) Dragging a model into the portal displays a 3D preview. c) Flattening the phone parallel to the ground will display a 3D preview, as if it were resting on a shelf.

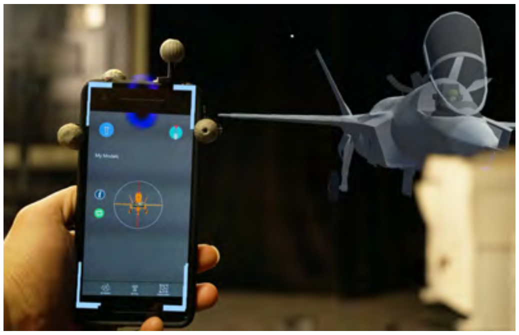

Figure 9. A full screen rotation widget can be used to make small adjustments to the model's orientation.

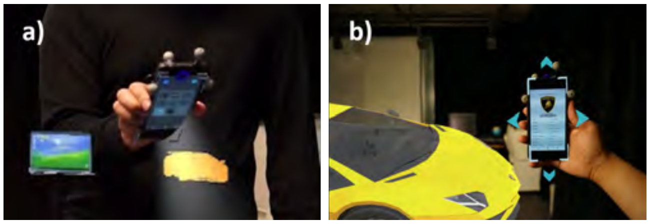

Figure 10. a) When the phone display is held away from the user, a flashlight is emitted, causing objects to be displayed in wireframe. b) When the phone intersects a 3D model, a detailed webpage about it is displayed on the phone.

If the user wishes to view an alternative rendering of the augmented content, they can hold the phone facing away from them, as if they were using the screen of the phone to illuminate a dark room (D6H). A virtual flashlight is projected, and intersected objects are rendered as a wireframe, as if being exposed to an x-ray (Figure 10a).

In this full screen view, an info icon can also be tapped to display a cross-device tooltip [42] presenting basic meta-data related to the model next to the phone (Figure 1) (D3P). The user can swipe the info icon left or right, to display the cross-device tooltip on either the left or right side of the phone.

The user can also retrieve a webpage associated with a 3D model, by placing the phone within its boundaries (D3H) (Figure 10b). Because the content is displayed on the phone, it can include extensive information about the model, such as a Wiki article, which would otherwise be difficult to read in AR. When the webpage is displayed on the phone, widgets in AR are displayed around the phone, which can be used to scroll up and down, or navigate forwards and backwards between pages (D2P). Alternatively, a pinch gesture can be used to scroll the pages up a down [13] (D1P). These distributed input techniques avoid the occlusion typically present when interacting by direct touch.

3D Text Operations

To further explore the use of distributed UI (D2H), we implemented a 3D text tool. Users can type words or short sentences, apply formatting, and place the text in the 3D environment. Text entry, in particular, is trivial to perform on the phone, but quite challenging in AR [23].

Tapping on a text entry box displays a native soft keyboard on the phone. A set of detailed widgets and sliders can then be used to change the formatting of the text (Figure 4b). Text can be moved to the 3D preview area using the portal (Figure 11a) (D4P). Similarly, text can be selected from the AR space using the provided selection operations, and then edited remotely on the phone (Figure 11b) (D2H).

The laser can be used to select 3D text in the AR environment, and swiping the laser icon downwards brings the text onto the phone. Additionally, the user can swipe upwards to copy the text format currently configured on the phone's UI and apply it to the 3D text being intersected by the laser. This allows the user to apply the same format to multiple pieces of content quickly, demonstrating a benefit of using the phone as an enhanced 6DOF input device (D1H).

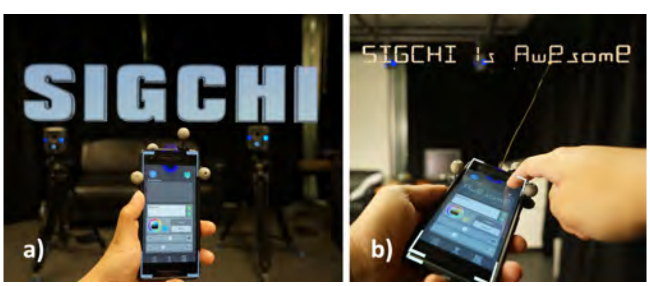

Figure 11. a) 3D text can be displayed in the 3D preview area. b) Text can be remotely edited on the phone.

Figure 12. In Group Select mode, a swipe along the top of the phone performs a top alignment of the selected objects.

Group Select Operations

The final mode of the prototype is used for group operations. Group operations are carried out in a landscape orientation, leveraging the affordance of capturing a photo or video with a mobile phone (D6H). The mode can be entered by either re-orienting the phone or tapping the icon. Instead of rendering a direct see-through video on the phone's display, individual objects that are in the phone's current field of view are rendered (D3H) (Figure 12).

Users can toggle the selection of multiple objects by tapping them (D2H). Once selected, groups of objects can be repositioned, aligned, or deleted. The user can also lock/unlock groups of objects with a lock icon. Once locked, subsequent 3D movement of the phone repositions them as a group (D1H). Icons are provided to align multiple objects to the left, right, top, bottom, front and back. Additionally, the user can tap or swipe with the finger along the physical edges of the phone to perform alignment or distribution in the corresponding direction (D6H) (Figure 12).

USER EVALUATION

Our interactive prototype was designed to explore a variety of techniques across our design space. As such, the goal of our evaluation was not to formally validate the design space, but to instead gather initial feedback on the joint interactions.

Participants and Procedure

We invited twelve users to participate in a 30-minute session. Participants were aged 21–35 (7 male, 5 female). Participants had a range of experience with VR/AR systems: Three had extensive experience, three had light familiarity, and six had little or no experience.

The evaluation began with an introduction of the prototype. Users were first shown the phone interface, and its main features. Users were then given a walkthrough of the prototype, lasting approximately 20 minutes. During this walkthrough, users were shown individual features of the prototype and asked to accomplish simple atomic tasks with those features.

After the guided walkthrough, users were asked to perform a simple high-level task independently, lasting approximately 10 minutes. The task consisted of building a scene, which required participants to place 2 models on a physical desk, and add 2 text labels of different styles to mark-up both physical and augmented objects. The detailed walk-through procedure and task description can be found in the Appendix.

After the study, participants completed a questionnaire together with a short interview. The questionnaire asked participants to rate each individual feature on a Likert scale from 1-7. In the interview, participants were asked about their most and least favorite aspects of their experience. Participants were then asked to share any other comments they had about the prototype and its features.

Results

The results of the study were positive and encouraging. Users enjoyed using the prototype, found the user experience compelling (and even magical), and were impressed with the variety of techniques that it offered.

Overall, users found the combination of the phone and ARHMD to be engaging:

"The usage of combining both a phone and AR screens turned out to be very smooth, and the designs provide the potential of incorporating multiple modalities into a user's view" (P1)

"The combination of mobile + AR as an entire system is very good, intuitive." (P6)

An interesting set of comments was that mobile phones offered a familiar interaction paradigm for an otherwise unfamiliar platform, supporting our design principle of preserving known interaction patterns:

"I have very limited experience with AR, using mobile device easily transfer my previous knowledge into the AR world." (P6)

"Distributed input from the mobile device to AR content is very useful, and I can easily transfer my previous experience on touch screen to spatial operation." (P11)

During the initial feature walkthrough, users were able to perform each of the individual interactions successfully. For users with less experience, it took a short amount of time to get used to the general interaction paradigm, and in particular the mid-air gestures. Overall, we were happy to see that even participants without extensive AR experience we eventually able to understand the prototype and its gestures.

All participants successfully completed the final task (see Appendix for final scenes), and were able to do so independently with minimal assistance. Aside from Vicon tracking errors that occurred for two participants, no major errors were made during the final task. This task did not require the use of every feature in the prototype, but it was promising to see that participants understood how to use the features of the prototype together, continuously transitioning between mobile and spatial interactions. The general method participants followed were to first create content on the phone, transfer it into the 3D environment, and then use the phone to fine tune its positions. Participants enjoyed the experience of anchoring their 3D virtual content to existing physical objects. However, it could have been further aided with snapping and/or alignment mechanisms that considered the physical environment.

Overall, the individual features of the interactive prototype were rated positively. Figure 13 shows the responses to each feature on a Likert scale from 1-7, grouped by their respective area of the design space. Across the design space, all features were rated highly, with medians of 6 in all cases except for D1H (5.5) and D4P (7). Interesting to note is that there was little difference in the responses between phone-centric and HMD-centric features. This further justifies the promise of a bidirectional framework which allows users to transition between both task contexts.

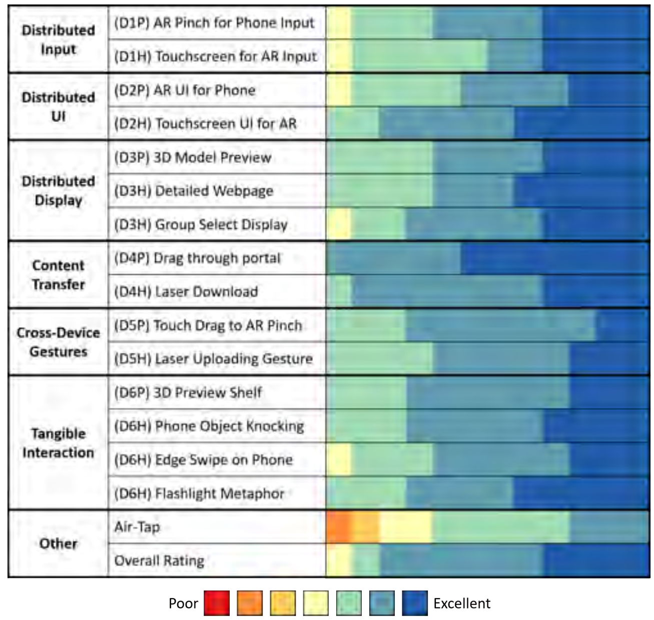

Figure 13. Responses to individual features, grouped by their associated area of the design space.

Of the individual features, the most popular set of techniques across participants was content transfer, which participants appreciated for moving content in both directions:

"Transferring 2D content into 3D from phone to spatial is like a magic for me." (P6)

"The portal and laser pointer metaphors as moving data between the phone and environment, it's a cool mental model. And the laser downloading is amazing." (P9)

Ironically, the one feature that was not well received was the air tap gesture (Figure 13), which happens to be a default technique for controlling spatial content on many AR platforms. A number of the participants mentioned that the gesture was difficult or awkward to use:

"Air tap is very difficult to use! (p3)

"The air tap AR interactions for direct manipulation is the worst part in the system." (p12)

This finding supports our motivation that freehand gestures in AR can be difficult to use, whereas tactile interaction with smartphone screens is familiar and efficient, and less prone to recognition errors.

Participants also noticed the limitation of the HoloLens field of view, which is something that we hope will improve with future evolutions of ARHMDs:

"The limitation of AR FOV limited the performance of working with real sized models". (P1)

"FOV is small, and that actually strength the combination usage of the phone and AR, as a physical prop, you can fully control where you want object to be spawned" (P9)

Overall, the observation sessions provided encouraging feedback on the prototype, motivating further developments and future evaluations.

DISCUSSION

This paper has introduced a design space of joint interactions between ARHMDs and smartphones. Our implementation demonstrates the feasibility of the associated interaction techniques, and our initial observation sessions demonstrate the promise of its potential. In this section, we further reflect on our experience and findings, and discuss limitations and generalizations.

Our design space was grounded on prior work, and also driven by our discussion of the complementing capabilities of ARHMDs and smartphones. It should be noted that we omitted from this discussion some of the capabilities that the platforms share. For example, voice-based input and audio output are possible using smartphones and ARHMDs. As such, our design space should not be considered as an exhaustive list of all joint interactions that are possible.

Similarly, the design space contributes six classes of interaction techniques, which could each lead to new interactions that have yet to be explored. For example, our prototype leverages the phone as a tangible device (D6) through physical edge gestures, orientation detection, and knocking gestures. Prior research has shown that there are many more potential uses of spatial tangible devices, such as being used as physical props [29] or for detecting grip [22].

One interesting observation about our prototype is that it supported more HMD-centric interaction interactions than phone-centric interactions. This was not an intentional decision, but perhaps a reflection of a similar balance seen in prior literature (e.g. Table 2). However, since this was not an intentional decision, we would be cautious about drawing any strong conclusions from this observation.

The task domain we chose for our evaluation was simple and generic – consisting of the manipulation and placement of 3D content within an augmented reality environment. While this application domain was a useful testbed for our explorations, generalizations to specific usage domains (gaming, education, training, etc.), could reveal new interaction opportunities that fall within our design space.

It is important to note that our implementation did not exploit the proximity of real-world objects in the physical environment beyond the phone. For instance, many promising techniques for snapping augmented objects to the physical world have been explored [43], and would be interesting to incorporate into our design space.

While the results of our work are encouraging, several limitations should be pointed out. First, our implementation required the use of external motion tracking cameras. Future implementations should consider fully contained systems. Second, the limited FOV of the HMD constrained the overall user experience – an issue we hope will improve with new generations of AR devices. Third, we did not perform a technical evaluation of our gesture set. During the observation sessions, there were some instances of gesture interference – such as knocking being recognized as flattening the phone. We believe such issues could be addressed with additional efforts on gesture recognition.

FUTURE WORK

In terms of technical implementations, we believe that in the near future it will be possible to implement a more robust system without the need for any external tracking equipment. In particular, recent research has shown promising ways to track the smartphone in 3D space [4, 41].

Regarding our evaluation, we believe there are numerous interesting studies that could be conducted in the future. For example, it would be interesting to specifically study the learnability of cross-device gestures, and examine techniques that could provide dynamic guidance or feedback. It may also be interesting to do formal comparisons of joint smartphone and HMD platforms to equivalent standalone platforms, using HMDs or handheld augmented reality.

Finally, one of the main concerns of our work, which could be explored further, may be the necessity to have a handheld device. Many prior AR systems are untethered and standalone, to support as "natural" of an experience as possible. However, we believe that in many scenarios, the improved experience that results from having a tactile and high precision input device, will outweigh the desire for a freehand experience. Indeed, it was interesting to note that the main gesture that users struggled with in our study was air tap, a native HoloLens freehand gesture.

CONCLUSION

We have presented our exploration that enables cross-device interactions between a smartphone and ARHMD. Our exploration covered six main classes of joint interaction techniques, under both phone-centric and HMD-centric task contexts. Our implementation of interactive prototype shows the feasibility of the proposed techniques, while a user evaluation indicated that users found the joint interaction compelling and were able to use the prototype without extensive training, by leveraging known interaction patterns. We believe that as smartphones and ARHMDs continue to evolve and improve, so too will the opportunities to combine them into unified interaction platforms. We hope our contributions will serve as important groundwork for this future, where AR does not replace our other personal computing devices, but instead can be used together with them in complementing ways.

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* (MUM '16), 115-122. DOI: https://doi.org/10.1145/3012709.3012719

[2] Rahul Arora, Rubaiat Habib Kazi, Fraser Anderson, Tovi Grossman, Karan Singh, and George Fitzmaurice. 2017. Experimental Evaluation of Sketching on Surfaces in VR. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17), 5643-5654. DOI: https://doi.org/10.1145/3025453.3025474

[3] Rahul Arora, Rubaiat Habib Kazi, Tovi Grossman, George Fitzmaurice, and Karan Singh. 2018. SymbiosisSketch: Combining 2D & 3D Sketching for Designing Detailed 3D Objects in Situ. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI '18), Paper 185, 15 pages. DOI: https://doi.org/10.1145/3173574.3173759

[4] 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* (SUI '18), 2-10. DOI: https://doi.org/10.1145/3267782.3267785

[5] Patrick Baudisch, Nathaniel Good, and Paul Stewart. 2001. Focus Plus Context Screens: Combining Display Technology with Visualization Techniques. In Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology (UIST '01), 31–40. DOI: https://doi.org/10.1145/502348.502354

[6] Hrvoje Benko and Edward W. Ishak. 2005. Cross-Dimensional Gestural Interaction Techniques for Hybrid Imrnersive Environments. In *Proceedings of the 2005 IEEE Conference 2005 on Virtual Reality* (VR '05), 209-216, 327. DOI: https://doi.org/10.1109/VR.2005.1492776

[7] Hrvoje Benko, Eyal Ofek, Feng Zheng, and Andrew D. Wilson. 2015. FoveAR: Combining an Optically See-Through Near-Eye Display with Projector-Based Spatial Augmented Reality. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (UIST '15), 129-135. DOI: https://doi.org/10.1145/2807442.2807493

[8] Xiaojun Bi, Tovi Grossman, Justin Matejka, and George Fitzmaurice. 2011. Magic Desk: Bringing Multi-touch Surfaces into Desktop Work. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '11), 2511–2520. DOI: https://doi.org/10.1145/1978942.1979309

[9] Mark Billinghurst, Adrian Clark and Gun Lee. 2015. A Survey of Augmented Reality. *Foundations and Trends® in Human–Computer Interaction*: Vol. 8: No. 2-3, pp 73-272. http://doi.org/10.1561/1100000049

[10] Frederik Brudy, Christian Holz, Roman Rädle, Chi-Jui Wu, Steven Houben, Clemens Nylandsted Klokmose, and Nicolai Marquardt. 2019. 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* (CHI '19), Paper 562, 28 pages. DOI: https://doi.org/10.1145/3290605.3300792

[11] 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. https://doi.org/10.1109/ISMAR.2013.6671837.

[12] Rahul Budhiraja, Gun A. Lee, and Mark Billinghurst. 2013. Interaction techniques for HMD-HHD hybrid AR systems. In 2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 243–244. DOI: https://doi.org/10.1109/ISMAR.2013.6671786

[13] Alex Butler, Shahram Izadi, and Steve Hodges. 2008. SideSight: Multi-"Touch" Interaction Around Small Devices. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology* (UIST '08), 201–204. DOI: https://doi.org/10.1145/1449715.1449746

[14] Simon Butscher, Sebastian Hubenschmid, Jens Müller, Johannes Fuchs, and Harald Reiterer. 2018. Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multidimensional Data. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI '18), Paper 90, 12 pages. DOI: https://doi.org/10.1145/3173574.3173664

[15] Xiang "Anthony" Chen, Julia Schwarz, Chris Harrison, Jennifer Mankoff, and Scott E. Hudson. 2014. Air+Touch: Interweaving Touch & In-air Gestures. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (UIST '14), 519–525. DOI: https://doi.org/10.1145/2642918.2647392

[16] Xiang 'Anthony' Chen, Tovi Grossman, Daniel J. Wigdor, and George Fitzmaurice. 2014. Duet: Exploring Joint Interactions on a Smart Phone and a Smart Watch. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14), 159–168. DOI: https://doi.org/10.1145/2556288.2556955

[17] Jessica Conditt. (2017). Worlds collide: VR and AR in 2018. Retrieved on April 4th, 2019 from https://www.engadget.com/2017/12/20/vr-and-ar-in-2018/.

[18] Laura Dipietro, Angelo M. Sabatini, and Paolo Dario. 2008. A Survey of Glove-Based Systems and Their

Applications. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 38, 4 (July 2008): 461–482. DOI: https://doi.org/10.1109/TSMCC.2008.923862

[19] Tomas Dorta, Gokce Kinayoglu, and Michael Hoffmann. 2014. Hyve-3D: A New Embodied Interface for Immersive Collaborative 3D Sketching. *In ACM SIGGRAPH 2014 Studio* (SIGGRAPH '14), 37:1–37:1. DOI: https://doi.org/10.1145/2619195.2656325

[20] Tomás Dorta, Gokce Kinayoglu, and Michael Hoffmann. 2016. Hyve-3D and the 3D Cursor: Architectural co-design with freedom in Virtual Reality. *International Journal of Architectural Computing* 14, 2: 87–102. https://doi.org/10.1177/1478077116638921

[21] George W. Fitzmaurice. 1993. Situated information spaces and spatially aware palmtop computers. *Commun. ACM* 36, 7 (July 1993), 39-49. DOI: http://doi.org/10.1145/159544.159566

[22] Mayank Goel, Jacob Wobbrock, and Shwetak Patel. 2012. GripSense: Using Built-in Sensors to Detect Hand Posture and Pressure on Commodity Mobile Phones. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (UIST '12), 545–554. DOI: https://doi.org/10.1145/2380116.2380184

[23] 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* (CHI '15), 3933-3942. DOI: https://doi.org/10.1145/2702123.2702331

[24] Martin Hachet, Benoit Bossavit, Aurélie Cohé, and Jean-Baptiste de la Rivière. 2011. Toucheo: Multitouch and Stereo Combined in a Seamless Workspace. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (UIST '11), 587–592. DOI: https://doi.org/10.1145/2047196.2047273

[25] Peter Hamilton and Daniel J. Wigdor. 2014. Conductor: Enabling and Understanding Cross-device Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14), 2773–2782. DOI: https://doi.org/10.1145/2556288.2557170

[26] Khalad Hasan, David Ahlström, Junhyeok Kim, and Pourang Irani. 2017. AirPanes: Two-Handed Around-Device Interaction for Pane Switching on Smartphones. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17), 679-691. DOI: https://doi.org/10.1145/3025453.3026029

[27] Mikiyasu Hakoama, and Shotaro Hakoyama. 2011. The impact of cell phone use on social networking and development among college students. *The American Association of Behavioral and Social Sciences Journal*(AABSS), 15.1 (2011).

[28] 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* (SUI '15), 89-98. DOI: https://doi.org/10.1145/2788940.2788947

[29] Ken Hinckley, Randy Pausch, John C. Goble, and Neal F. Kassell. 1994. Passive Real-world Interface Props for Neurosurgical Visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '94), 452–458. DOI: https://doi.org/10.1145/191666.191821

[30] Ken Hinckley, Jeff Pierce, Mike Sinclair, and Eric Horvitz. 2000. Sensing Techniques for Mobile Interaction. In *Proceedings of the 13th annual ACM symposium on User interface software and technology* (UIST '00), 91-100. DOI: http://doi.org/10.1145/354401.354417

[31] Scott E. Hudson, Chris Harrison, Beverly L. Harrison, and Anthony LaMarca. 2010. Whack Gestures: Inexact and Inattentive Interaction with Mobile Devices. In *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction* (TEI '10), 109–112. DOI: https://doi.org/10.1145/1709886.1709906

[32] Sungjune Jang, Lawrence H. Kim, Kesler Tanner, Hiroshi Ishii, and Sean Follmer. 2016. Haptic Edge Display for Mobile Tactile Interaction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16), 3706-3716. DOI: https://doi.org/10.1145/2858036.2858264

[33] Haojian Jin, Christian Holz, and Kasper Hornbæk. 2015. Tracko: Ad-hoc Mobile 3D Tracking Using Bluetooth Low Energy and Inaudible Signals for Cross-Device Interaction. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (UIST '15), 147-156. DOI: https://doi.org/10.1145/2807442.2807475

[34] David Kim, Otmar Hilliges, Shahram Izadi, Alex D. Butler, Jiawen Chen, Iason Oikonomidis, and Patrick Olivier. 2012. Digits: Freehand 3D Interactions Anywhere Using a Wrist-worn Gloveless Sensor. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (UIST '12), 167–176. DOI: https://doi.org/10.1145/2380116.2380139

[35] Mikko Kytö, Barrett Ens, Thammathip Piumsomboon, Gun A. Lee, and Mark Billinghurst. 2018. 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), Paper 81, 14 pages. DOI: https://doi.org/10.1145/3173574.3173655

[36] Jinha Lee, Alex Olwal, Hiroshi Ishii, and Cati Boulanger. 2013. SpaceTop: Integrating 2D and Spatial 3D Interactions in a See-through Desktop Environment. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13), 189–192. DOI: https://doi.org/10.1145/2470654.2470680

[37] Magic Leap. Retrieved April 4th, 2019 from https://www.magicleap.com/ .

[38] Nicolai Marquardt, Ricardo Jota, Saul Greenberg, and Joaquim A. Jorge. 2011. The Continuous Interaction Space: Interaction Techniques Unifying Touch and Gesture on and Above a Digital Surface. In *Proceedings of the 13th IFIP TC 13 International Conference on Human-computer Interaction - Volume Part III* (INTERACT'11), 461–476. DOI: https://doi.org/10.1007/978-3-642-23765-2_32

[39] Microsoft HoloLens. Retrieved April 4th, 2019 from https://hololens.com

[40] A. Millette and M. J. McGuffin. 2016. DualCAD: Integrating Augmented Reality with a Desktop GUI and Smartphone Interaction. In *2016 IEEE International Symposium on Mixed and Augmented Reality* (ISMAR-Adjunct), 21–26. DOI: https://doi.org/10.1109/ISMAR-Adjunct.2016.0030

[41] 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* (CHI '19), Paper 585, 11 pages. DOI: https://doi.org/10.1145/3290605.3300815

[42] Erwan Normand and Michael 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. DOI: https://doi.org/10.1109/ISMAR.2018.00043

[43] Benjamin Nuernberger, Eyal Ofek, Hrvoje Benko, and Andrew D. Wilson. 2016. SnapToReality: Aligning Augmented Reality to the Real World. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI '16), 1233-1244. DOI: https://doi.org/10.1145/2858036.2858250

[44] Ken Pfeuffer, Benedikt Mayer, Diako Mardanbegi, and Hans Gellersen. 2017. Gaze + Pinch Interaction in Virtual Reality. In *Proceedings of the 5th Symposium on Spatial User Interaction* (SUI '17), 99–108. DOI: https://doi.org/10.1145/3131277.3132180

[45] 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* (UIST '15), 57–62. DOI: https://doi.org/10.1145/2807442.2807471

[46] Roman Rädle, Hans-Christian Jetter, Nicolai Marquardt, Harald Reiterer, and Yvonne Rogers. 2014. 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), 45-54. DOI: https://doi.org/10.1145/2669485.2669500

[47] Jun Rekimoto. 1996. Tilting Operations for Small Screen Interfaces. In *Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology* (UIST '96), 167–168. DOI: https://doi.org/10.1145/237091.237115

[48] Joan Sol Roo and Martin Hachet. 2017. One Reality: Augmenting How the Physical World is Experienced by Combining Multiple Mixed Reality Modalities. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (UIST '17), 787-795. DOI: https://doi.org/10.1145/3126594.3126638

[49] 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* (CHI '17), 5892-5904. DOI: https://doi.org/10.1145/3025453.3025661

[50] Dominik Schmidt, Julian Seifert, Enrico Rukzio, and Hans Gellersen. 2012. A Cross-device Interaction Style for Mobiles and Surfaces. In *Proceedings of the Designing Interactive Systems Conference* (DIS '12), 318–327. DOI: https://doi.org/10.1145/2317956.2318005

[51] Marcos Serrano, Barrett Ens, Xing-Dong Yang, and Pourang Irani. 2015. Gluey: Developing a Head-Worn Display Interface to Unify the Interaction Experience in Distributed Display Environments. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (MobileHCI '15), 161-171. DOI: https://doi.org/10.1145/2785830.2785838

[52] 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 & services* (MobileHCI '14). ACM, New York, NY, USA, 135-143. DOI: https://doi.org/10.1145/2628363.2628375

[53] Shaishav Siddhpuria, Sylvain Malacria, Mathieu Nancel, and Edward Lank. 2018. Pointing at a Distance with Everyday Smart Devices. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (CHI '18), Paper 173, 11 pages. DOI: https://doi.org/10.1145/3173574.3173747

[54] Maximilian Speicher, Brian D. Hall, Ao Yu, Bowen Zhang, Haihua Zhang, Janet Nebeling, and Michael Nebeling. 2018. XD-AR: Challenges and Opportunities in Cross-Device Augmented Reality Application Development. *Proc. ACM Hum.-Comput. Interact. 2, EICS* (June 2018), 7:1–7:24. DOI: https://doi.org/10.1145/3229089

[55] 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* (CHI '19), Paper 13, 13 pages. DOI: https://doi.org/10.1145/3290605.3300243

[56] 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* (CHI '19), Paper 619, 12 pages. DOI: https://doi.org/10.1145/3290605.3300849

[57] Christian Weichel, Manfred Lau, David Kim, Nicolas Villar, and Hans W. Gellersen. 2014. MixFab: A Mixed-reality Environment for Personal Fabrication. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14), 3855–3864. DOI: https://doi.org/10.1145/2556288.2557090

[58] Dirk Wenig, Johannes Schöning, Alex Olwal, Mathias Oben, and Rainer Malaka. 2017. WatchThru: Expanding Smartwatch Displays with Mid-air Visuals and Wrist-worn Augmented Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (CHI '17), 716-721. DOI: https://doi.org/10.1145/3025453.3025852

[59] Daniel Wigdor, Clifton Forlines, Patrick Baudisch, John Barnwell, and Chia Shen. 2007. Lucid Touch: A See-through Mobile Device. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology* (UIST '07), 269–278. DOI: https://doi.org/10.1145/1294211.1294259

[60] Christian Winkler, Markus Löchtefeld, David Dobbelstein, Antonio Krüger, and Enrico Rukzio. 2014. SurfacePhone: A Mobile Projection Device for Single- and Multiuser Everywhere Tabletop Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI '14), 3513–3522. DOI: https://doi.org/10.1145/2556288.2557075

[61] Paweł Woźniak, Lars Lischke, Benjamin Schmidt, Shengdong Zhao, and Morten Fjeld. 2014. Thaddeus: A Dual Device Interaction Space for Exploring Information Visualisation. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (NordiCHI '14), 41–50. DOI: https://doi.org/10.1145/2639189.2639237

[62] Robert Xiao, Scott Hudson, and Chris Harrison. 2016. CapCam: Enabling Rapid, Ad-Hoc, Position-Tracked Interactions Between Devices. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces* (ISS '16), 169-178. DOI: https://doi.org/10.1145/2992154.2992182

[63] Chen Zhao, Ke-Yu Chen, Md Tanvir Islam Aumi, Shwetak Patel, and Matthew S. Reynolds. 2014. SideSwipe: Detecting In-air Gestures Around Mobile Devices Using Actual GSM Signal. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (UIST '14), 527–534. DOI: https://doi.org/10.1145/2642918.2647380