Chair for Computer Science 10 (Media Computing and Human-Computer Interaction)

RWTH AACHEN UNIVERSITY

Exploring Midair Tangible Interaction over Tabletop Displays

Master's Thesis
submitted to the
Media Computing Group
Prof. Dr. Jan Borchers
Computer Science Department
RWTH Aachen University

by
Jonas Vogel

Thesis advisor:
Prof. Dr. Jan Borchers

Second examiner:
Prof. Dr. Ulrik Schroeder

Registration date: 14.02.2020
Submission date: 16.06.2020

Zentrales Prüfungsamt/Central Examination Office

RWTH AACHEN UNIVERSITY

Eidesstattliche Versicherung
Statutory Declaration in Lieu of an Oath

Vogel, Jonas

Name, Vorname/Last Name, First Name

Matrikelnummer (freiwillige Angabe)
Matriculation No. (optional)

Ich versichere hiermit an Eides Statt, dass ich die vorliegende ~~Arbeit/Bachelorarbeit/~~ Masterarbeit* mit dem Titel

I hereby declare in lieu of an oath that I have completed the present paper/Bachelor thesis/Master thesis* entitled

Exploring Midair Tangible Interaction over Tabletop Displays

selbstständig und ohne unzulässige fremde Hilfe (insbes. akademisches Ghostwriting) erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Für den Fall, dass die Arbeit zusätzlich auf einem Datenträger eingereicht wird, erkläre ich, dass die schriftliche und die elektronische Form vollständig übereinstimmen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

independently and without illegitimate assistance from third parties (such as academic ghostwriters). I have used no other than the specified sources and aids. In case that the thesis is additionally submitted in an electronic format, I declare that the written and electronic versions are fully identical. The thesis has not been submitted to any examination body in this, or similar, form.

Aachen, 16.06.2020

Ort, Datum/City, Date

Unterschrift/Signature

*Nichtzutreffendes bitte streichen

*Please delete as appropriate

Belehrung:
Official Notification:

§ 156 StGB: Falsche Versicherung an Eides Statt

Wer vor einer zur Abnahme einer Versicherung an Eides Statt zuständigen Behörde eine solche Versicherung falsch abgibt oder unter Berufung auf eine solche Versicherung falsch aussagt, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.

Para. 156 StGB (German Criminal Code): False Statutory Declarations

Whoever before a public authority competent to administer statutory declarations falsely makes such a declaration or falsely testifies while referring to such a declaration shall be liable to imprisonment not exceeding three years or a fine.

§ 161 StGB: Fahrlässiger Falscheid; fahrlässige falsche Versicherung an Eides Statt

(1) Wenn eine der in den §§ 154 bis 156 bezeichneten Handlungen aus Fahrlässigkeit begangen worden ist, so tritt Freiheitsstrafe bis zu einem Jahr oder Geldstrafe ein.

(2) Straflosigkeit tritt ein, wenn der Täter die falsche Angabe rechtzeitig berichtigt. Die Vorschriften des § 158 Abs. 2 und 3 gelten entsprechend.

Para. 161 StGB (German Criminal Code): False Statutory Declarations Due to Negligence

(1) If a person commits one of the offences listed in sections 154 through 156 negligently the penalty shall be imprisonment not exceeding one year or a fine.

(2) The offender shall be exempt from liability if he or she corrects their false testimony in time. The provisions of section 158 (2) and (3) shall apply accordingly.

Die vorstehende Belehrung habe ich zur Kenntnis genommen:

I have read and understood the above official notification:

Aachen, 16.06.2020

Ort, Datum/City, Date

Unterschrift/Signature

Contents

List of Figures

Abstract

Previous research has already established that users benefit from tangible interaction, for example in terms of precision and speed and because they allow eyes-free interaction. Usually, tangible interaction takes place directly on the tabletop surface, mostly because they can only be tracked when being in contact with it. We extend tangible interaction to the space above the tabletop, which allows a whole new set of interactions in 3D. Midair tangibles, for example, can be used to explore three-dimensional data, to easily inspect and manipulate a virtual 3D object, or to select targets that are out of reach for the user. To relate the position of midair devices with the tabletop, we use the world tracking and image recognition features of current augmented reality frameworks. Based on Apple's ARKit and iPhones used as tangibles, we implemented a reusable tracking framework which also allows the transmission of touch events between tabletop and the devices. Our tests and preliminary studies reveal that the general performance of our system is satisfactory, but highly depends on the display of suitable, feature-rich content on the tabletop. Based on the implemented framework, we developed five applications to explore different use-cases for midair tangible interaction, of which three have been evaluated in a user study to get qualitative feedback. The results from the study identify (1) the visualisation of the newly introduced third dimension as a key challenge in such scenarios, (2) that midair devices are suitable to solve the reachability problem and (3) that they can provide user-specific content in multi-user scenarios.

Überblick

Die bisherige Forschung hat bereits gezeigt, dass Menschen von der Interaktion mit Tangibles profitieren, beispielsweise in der Genauigkeit und Geschwindigkeit ihrer Eingaben und da sie ohne Hinzusehen verwendet werden können. Normalerweise findet die Interaktion mit Tangibles direkt auf der Oberfläche von Tischbildschirmen statt, in den meisten Fällen, da diese mit der Oberfläche in Kontakt stehen müssen, um vom System erkannt zu werden. Wir erweitern die Interaktion mit Tangibles in den Raum über den Tisch, was eine neue Reihe von Interaktionsmöglichkeit in 3D eröffnet. Zum Beispiel können Midair-Tangibles benutzt werden, um dreidimensionale Daten zu erforschen, um virtuelle 3D-Objekte zu untersuchen und zu manipulieren oder um Elemente auszuwählen, die außerhalb der Reichweite einer Person sind. Um die Position des Geräts in der Luft mit der des Tischs in Verbindung zu setzen, verwenden wir die World-Tracking- und Bilderkennungsmöglichkeiten von aktuellen Augmented-Reality-Frameworks. Basierend auf Apples ARKit und mehreren iPhones, die als Tangibles genutzt werden, haben wir ein wiederverwendbares Tracking-Framework implementiert, das außerdem den Austausch von Touch-Daten zwischen dem Tisch und den Geräten erlaubt. Unsere Tests und vorläufige Studien zeigen, dass die allgemeine Leistung unseres Systems zu unserer Zufriedenstellung ist, allerdings sehr von der Verwendung von passendem, markanten Inhalt auf dem Tischbildschirm abhängig ist. Basierend auf dem implementierten Framework haben wir fünf Anwendung entwickelt, um verschiedene Anwendungsfälle für Interaktionen mit Midair-Tangibles zu erforschen, von denen drei in einer Studie ausgewertet wurde, um qualitatives Feedback zu erhalten. Die Ergebnisse dieser Studie identifizieren die Darstellung der neu-eingeführten dritten Dimension als zentrale Herausforderung in solchen Szenarien, dass Midair-Tangibles es ermöglichen, das Problem der Erreichbarkeit zu lösen und dass sie personenspezifischen Inhalt auf den Geräten in Szenarien mit mehreren Personen ermöglichen.

Acknowledgements

With much appreciation, I'd like to thank my supervisor Christian Cherek for the support, his thoughts, ideas and the friendly collaboration throughout the creation of my thesis.

I enjoyed my time at RWTH Aachen's Media Computing Group very much and value the many things that I could learn from the lab's researchers, PhDs and students headed by Jan Borchers. Also, I want to thank Jan Borchers and Ulrik Schroeder for examining the thesis.

Third, thanks to all participants of the user study and the many early testers for their time and valuable input.

Finally, I am grateful for the support of my friends and family – during the time of my studies and beyond.

Thank you!

Conventions

The whole thesis is written in British English.

We use plurals in formulations with unidentified third persons.

Chapter 1

Introduction

Tangibles provide haptic feedback on interactive surfaces.

Interactive tabletop displays combine a large display with touch interaction. Many users are familiar with touch from their mobile devices. With the extent of a table instead of a small, portable display, interactive tabletops offer much visible space for users to work on and also afford to be used by multiple users at the same time. Such systems have been used in a variety of fields like museum exhibitions [Ma et al., 2015, Loparev et al., 2016], live music performance [Jordà et al., 2007] and – as they are becoming more affordable – even for fun in some fast food restaurants[1]. Initially, just as it is the case with touch interaction on smartphones or tablet computers, interactive tabletops provide no haptic feedback. Their display is usually covered by glass and does not allow users to feel any differences with their fingers, for example to get feedback if they tapped on the right button. For this reason, researchers put physical objects (tangibles) onto the tabletop and connected them with the display, so that they can be used to interact with the content on the screen. Research has shown that users benefit from interacting with tangible objects in terms of precision, speed and because they allow eyes-free interaction [Weiss et al., 2009, Voelker et al., 2015b]. In contrast to touch-only interaction, a tangible can be felt in the hand and even be heard moving over the surface [Cherek et al., 2018].

[1] https://thenextweb.com/shareables/2013/08/23/mcdonalds-happy-table-is-an-nfc-powered-virtual-playground-for-kids-and-adults-in-asia/

We extend tangible interaction to the space above the tabletop.

One domain for tangible user interfaces, which will be in the focus of interest for this thesis, is the use of the vertical space above the tabletop surface. Usually, tangible interaction takes place directly on the tabletop, mostly because tangibles can only be tracked when being in contact with the table. But tangibles also afford to be picked up, for example when being relocated from one end of the tabletop to the other, or when a user would want to inspect another side of a virtual, three-dimensional object that the tangible represents. Making use of the space above the tabletop introduces new interaction possibilities in 3D and adds a third dimension to the design space of tangible user interfaces. The literature research (see chapter 2 "Related work") shows that people started to make use of the vertical space above the tabletop display by stacking tangible blocks on top of each other [Baudisch et al., 2010, Chan et al., 2012]. Other examples include tangibles that can change their height or small items that can be moved inside a close range from the surface and are tracked by magnetic fields (cf. section 2.2 "On- and near-surface tangibles"). Interaction with full midair tangibles, i.e. tangibles being tracked freely in space are only covered by a few publications. One of these is the *PaperLens* prototype first published by Martin Spindler, Raimund Dachselt and others at ITS 2009 [Spindler et al., 2009]. Including other applications, they visualised different data layers above a tabletop display using paper-based midair tangibles. Section 2.4 contains the details of their research and other publications on midair tangibles allowing movement and rotation in all three dimensions.

Augmented reality technologies to track tangibles in midair

Now, more than ten years later, we started looking for new approaches to track tangibles freely over an interactive display using methods that allow a quicker setup and without requiring much technical load in terms of hardware and setup. Impressed by the improvements achieved in augmented reality (AR) these days, we came up with the idea to use state-of-the-art smartphones as tangible objects and use their capabilities in AR to detect the tabletop surface. Then, tabletop and smartphone could be connected and share their real-world positions, so that both know where the other is located. As a side effect, our midair tangibles would automatically come with a high-resolution display and multi-touch support, which can be used for even more elaborated scenarios. Having set the technical foundation, we will then further explore midair tangible interaction over tabletop displays. Based on the learnings from previous publications, we developed several ideas for applications that cover a broader range of scenarios and aspects that we would like to investigate. These include applications for 3D exploration, with a direct mapping between midair tangible and virtual object, for temporal data exploration, layer exploration and target selection (cf. chapter 4 "Applications"). Some of our applications were also designed for collaborative tasks, so that we can gain more insights about how midair tangibles can help users to work together. To evaluate our system and judge the general performance of the tracking approach, we conducted several preliminary user studies and gained early user feedback. Additionally, a remote user study for three of our applications helped us to get qualitative feedback for the different scenarios.

In conclusion, we provide the following contributions (in reading order of this document):

1. A comprehensive literature research on the use of midair tangibles and related fields
2. A reusable framework to track midair tangibles using augmented reality
3. Five applications to explore different use-cases for midair tangibles
4. An evaluation of three of the applications to get qualitative feedback

Chapter 2

Related work

As midair tangible, we understand physical objects that are trackable in the space above a tabletop display and allow movement and rotation in all dimensions. Being able to transmit their position and orientation to the system that runs the visualisation on the tabletop screen, midair tangibles can have displays themselves and be used as a peephole into the virtual world that the tabletop display belongs to or have a direct mapping to a virtual 3D object on the digital table. Some midair tangibles can also be placed and tracked on the surface and then are usable as regular tangibles.

Our understanding of tangible user interfaces with midair interaction

This chapter presents research, publications and related work on the way to midair tangible interaction. It will start with a short overview of general findings on tangible interaction and interactive tabletops. Putting the spot on midair interaction, research on on- and near-surface tangibles that extend to the third dimension is followed by prototypes that implement interaction on multiple layers in the space above a tabletop. The core of this chapter consists of a detailed summary of the existing work on full midair and off-surface tangibles. Next, important findings from related areas will be presented: the concept and examples of magic lenses, publications that combine an interactive tabletop display with augmented reality, applications with multiple devices and multi-user scenarios with tabletop displays. This selection of related papers and pro-

Focus of related work: Off-surface tangible interaction and learnings from related domains.

totypes is supposed to emphasise on the variety of application examples that can be used for midair interaction and help in the understanding of interaction design in this new domain.

2.1 A history of tangible user interaction research

From classic desktop computers to post-desktop user interfaces, that bridge the gap between the virtual and the physical world and centre on human activities.

Tangibles provide users with haptic feedback on a tabletop surface. Users benefit from interacting with tangible objects in terms of precision and speed and because they allow eyes-free interaction [Weiss et al., 2009, Voelker et al., 2015b]. First research on tangibles has been published now nearly two decades ago: At MIT Media Lab [Ishii and Ullmer, 1997], the authors envisioned so-called *tangible bits* to allow the user to manipulate virtual content in the centre of their interest. Users wouldn't have to sit behind their desktop computers any more, but could use everyday, tangible objects to to alter digital artefacts (see figure 2.1). In the publication, the term *tangible user interfaces* was shaped, which could succeed the commonly-used graphical user interfaces. MIT Media Lab also published some of the first application scenarios with tangibles: One publication [Underkoffler and Ishii, 1999] demonstrates the use of a system for urban planning where small physical building models can be placed on a tabletop surface. The system then cast shadows of these buildings onto the surface, which get

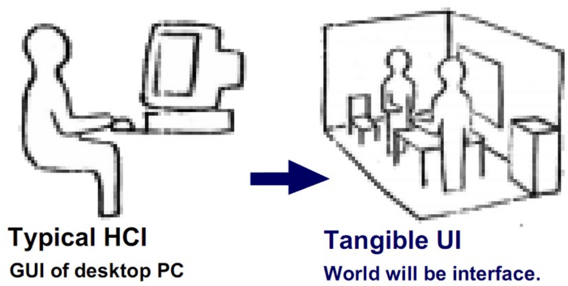


Figure 2.1: From graphical user interfaces to tangible user interfaces. [Ishii and Ullmer, 1997]

updated in real time when the building models are being moved. A second publication [Patten et al., 2001] presented electromagnetic tracking for the tangibles which allowed users to modify the tracked objects more accurately and with low-latency.

Twenty years of research

Since then, many other researchers contributed to the domain of tangible interaction. The research areas included technical findings, like tracking capabilities of tangible objects and the development of vision-based tabletop displays (which used a beamer either mounted on the ceiling or placed underneath the surface) to large LCD-based display panels with capacitive touch. Other research focussed on the implication of these new interaction techniques for user interaction design or new opportunities for multi-user scenarios and collaboration.

2.2 On- and near-surface tangibles

Classic tangibles have to stay in contact with the surface.

Most tangible interaction takes place directly on the tabletop, i.e. when a physical object rests on the table surface. In many cases, this is due to the technical limitation that the tangible's position has to be trackable. On capacitive touchscreens, tangibles close an electrical circle [Voelker et al., 2015a] and therefore have to be in physical contact with the surface. Multiple tangibles and the user's hands can be differentiated by the tabletop because of their respective capacitance.

Third dimension: Stack tangible blocks on top of each other.

One approach to make use of the third dimension with classic tangibles is to use building blocks, which are trackable cubes that can be stacked on top of each other. Researchers of the Hasso Plattner Institute in Potsdam, Germany, worked on this idea (see figure 2.2). For applications that require more complexity than the flat tangible user interfaces could provide, they present a solution of stackable building blocks on capacitive screens [Chan et al., 2012]: Inside a block, wires and connectors cause the capacitance of the bottom-most block to change when another block is stacked on top of it. The system then is able to recognise and differentiate the whole unit from single blocks or other

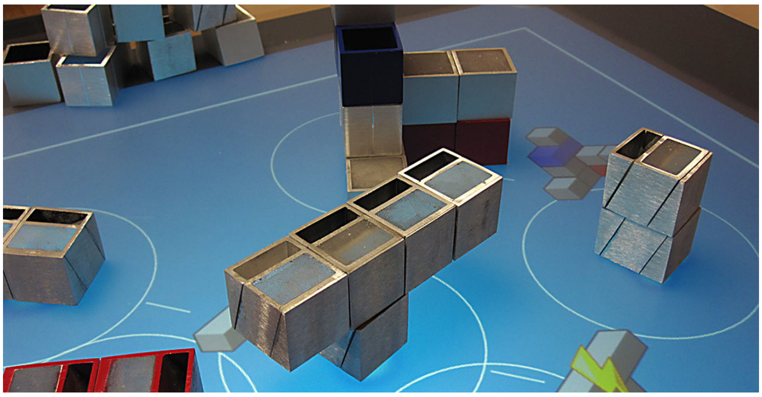


Figure 2.2: Tangible blocks can be stacked on top of each other. [Baudisch et al., 2010]

stacks. Application examples for building blocks include checkers, multi-dial knobs and a multi-purpose construction kit, which is founded on a system developed by the same group before [Baudisch et al., 2010].

Third dimension: Tangibles that can change their height.

Instead of stacking tangibles on top of each other, other approaches to extend tangible interaction to the third dimension change the height of the tangible itself. This sort of tangibles still rest on the tabletop and can be easily tracked, and include some kind of movable component in z-direction. The Interaction Technology Lab at University of Tokyo [Mi and Sugimoto, 2011] presents active height-adjustable tangibles with four degrees of freedom: movement in x- and y-direction on the tabletop, rotation, and movement in z-direction. The interaction is bi-directional, which means that the tangibles could move by themselves, or that users could manually move them to their liking. Following the urban planning example from Ishi et al. [1999], the prototype depicts building shadows which got longer when the height of the tangible is increased. Apparently inspired by nature, the prototype G-raff [Kim and Nam, 2015] follows a similar idea and is additionally able to elevate a smartphone (see figure 2.3). The phone is mounted on a device with a motorised, height-adjustable neck which is then placed on a tabletop display. With the device display showing only a part of the content that is displayed on the tabletop, and the possibility of augmenting that content with ad-

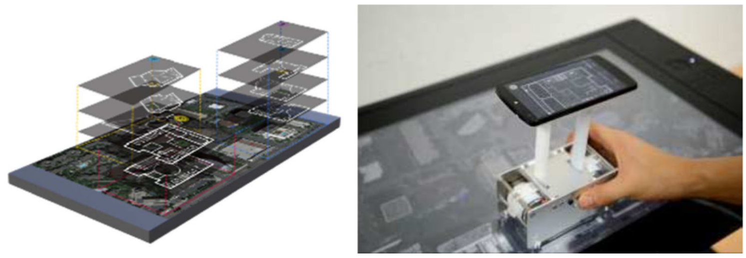


Figure 2.3: A tangible with an elevated smartphone (right). The device height represents different floors of the building plan (left). [Kim and Nam, 2015]

ditional data, the publication includes example use cases on 3D exploration: Users can search through floor and building plans by moving the device on the tabletop surface, and changing floors by prolonging the neck of the device. They also implemented an application were the device acts as a buffer for copied content from the tabletop. The device display shows the copied content and can be pasted again by gently pushing the smartphone downwards along the neck. A preliminary user study indicates that users easily recognised the connection between the height of the neck and the visualised data above the tabletop.

Tangibles tracked in near-surface space

Attempting to avoid the direct physical connection to the tabletop, GaussBits [Liang et al., 2013] extend the design space of tangible user interfaces to the near-surface space. Using a magnetic sensor grid that is attached to the back of the tabletop display, small magnetic items in close range to the surface become trackable. Non-ferrous materials such as the user's hand will not interfere and not be tracked by the sensor grid. It is to be noted however, that the system does not support free movement in all dimension (no six degrees of freedom) due to technical limitations. Also, the interaction space above the surface is limited to approximately 5cm. The implemented demo includes users controlling a small airplane (tilt, raising and lifting of the nose) and a widget to navigate on a map.

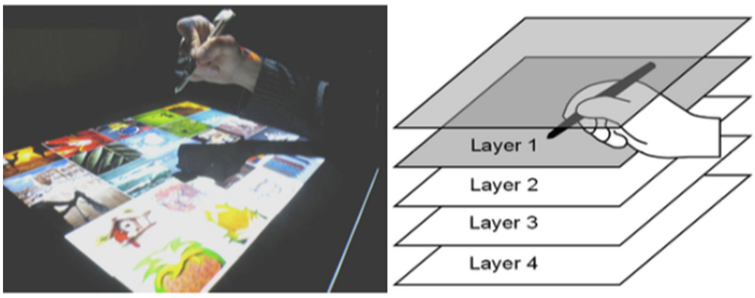


Figure 2.4: The space above the tabletop is divided into layers (right). Users can select elements from a layer in midair using a digital pen (left). [Subramanian et al., 2006]

2.3 Multi-Layer Interaction above Tabletops

Third dimension: Parallel layers over a surface where each distinct layer can be reached with a device to display or control data on the tabletop.

Another step towards midair tangibles that are able to move independently over a tabletop surface is the idea to divide the space above an interactive table into several layers parallel to its surface, with the lowest layer being the tabletop itself.

Soon after the first digital tables were sold, researchers [Subramanian et al., 2006] took the accompanying digital pencil and extended the interaction space above the table to several layers (see figure 2.4). The different layers represented different information of a map displayed on the digital table (bus lines, traffic information etc.). By moving the pencil to the respective layer, users could select or alter elements from that data group. A pilot experiment yields though that users soon get fatigued by holding the pencil up, especially higher above the surface. The authors suggest four layers at approximately 4 cm each for future designs.

The authors of the G-raff (see previous section) continued their work and mounted a transparent touch display on an arm structure, which they fixed on the frame of an interactive table [Kim and Nam, 2016]. With this construction, it is possible to move the transparent display in different positions and in different height above the tabletop, while the display stays parallel to the surface. Using a mechanical stand, this setup resolves the issue with fatigue that occurs when users have to hold a device with their own arms. The presented application scenarios include prototypes for 3D exploration and manipulation (scans of medical data) and the use of the device as secondary display to free the main screen from irrelevant information.

Spindler et al. investigated further into the amount and thickness of such interaction layers and did a comprehensive user study [2012c]. Their paper-based midair tangibles are held and moved in parallel to the surface by the user. The study results yield a similar layer thickness as was suggested by Subramanian et al. (4 cm, see above), but a higher overall interaction height of 44 cm. As a general rule, they suggest to use as few layers as possible and a "comfort zone" at a medium height for the most relevant information layers.

2.4 Midair Tangibles over Tabletop Displays

PaperLens, the lab's and other research on full midair tangibles.

This section presents related work on midair tangibles (as we defined them by our understanding) in conjunction with an interactive tabletop display. It starts with an summary of the extensive work of one research group on this subject and continues with publications on off-surface tangibles of other research groups. Next, the existing research in midair tangible interaction that has been done at our lab is presented. The section is concluded by a short attempt to summarise quantitative data on midair interaction. Later sections will cover topics from related areas or publications that do not fully meet our expectations of a tangible user interface with midair interaction.

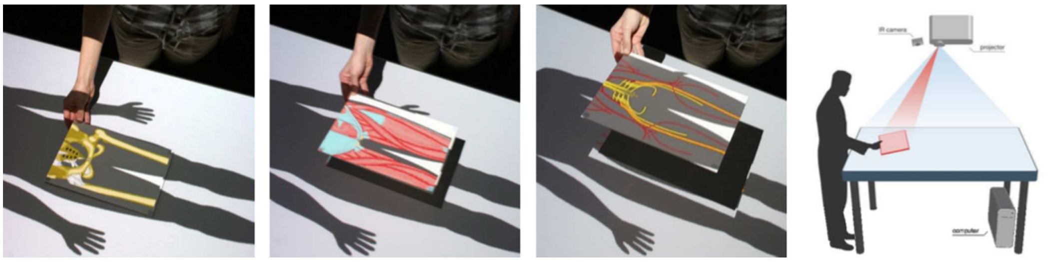


Figure 2.5: When raising the paper-based tangible over the tabletop, it shows different information layers (left). A beamer and an infrared camera is needed for the system to work. (right). [Spindler et al., 2009]

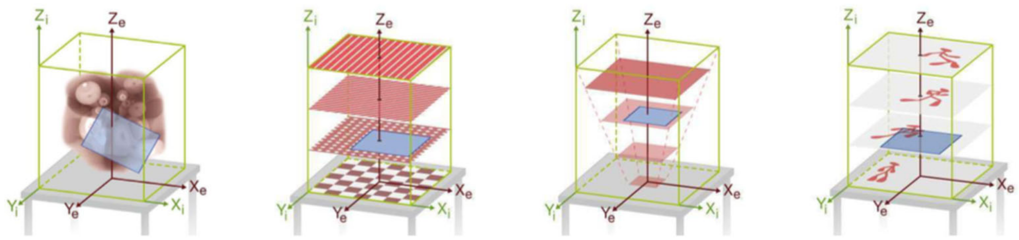
2.4.1 The PaperLens Prototype

Contributions by Spindler et al.

About ten years ago, Raimund Dachselt, Martin Spindler et al. from University of Magdeburg, Germany, did extensive research on using tangibles over an interactive tabletop [Spindler et al., 2009, 2010a,b, 2012a,b,c] [Spindler, 2012]. They contribute a tangible user interface with full midair support using paper-based tangibles, a vocabulary and classification for midair tangible interaction and various application examples. The following paragraphs will summarise their work, from technical findings over design considerations and study results to mentioned application scenarios.

A technical journey from simple, paper-based tangibles to active ones that support multi-touch.

Their first publication describes a system with paper-based tangibles called *PaperLens* over a vision-based tabletop display [Spindler et al., 2009]. The tangibles have roughly the size of a standard letter and are made of cardboard. An infrared camera is mounted on the ceiling above the tabletop and can track the tangibles using IR-reflecting markers attached to them. Next to the camera, a projector is used to display content on the tabletop and onto the tangible surface, as long as the tangible is not being moved outside the projection volume. The projected content on the tangible differs from the one on the tabletop surface, since the system can calculate the position of the tangible and adapt the projected image accordingly. This setup was later improved in favour of a back-projected tabletop display which used the top projector for the tangible only, because the content on the tabletop display got occluded by



(a) Volumetric Information Space (b) Layered Information Space (c) Zoomable Information Space (d) Temporal Information Space

Figure 2.6: By moving the tangible (blue rectangle) through the interaction space, users can study four types of exploration space. [Spindler et al., 2009]

the tangible (see figure 2.5). In the following years [2010b, 2012b], Spindler et al. extended the tracking capabilities to be more reliable and tried different forms of tangibles (circular, squared) and visual representations (colour, material). They also added gesture recognition (flip, shake and tilt), direct pointing capabilities for digital pens and touch input. Finally, the researchers added the possibility to use active displays (iPads) that support multi-touch, have a higher display resolution than the passive ones and are tracked using a Microsoft Kinect [2012a].

Information spaces and interaction vocabulary for midair tangibles

Next to the technical work, the authors contribute several categorisations and design considerations from their experience with tangible user interfaces that use midair tangibles. One distinction regards the space above the tabletop surface and identifies four *information spaces* (see figure 2.6) [Spindler et al., 2009]. In the volumetric information space, a tangible can be used to explore three-dimensional objects that are placed on the surface (and therefore extend to the space above it) without physical restrictions and with six degrees of freedom. The layered information space divides itself again into several, distinct layers, that can provide similar information as the volumetric space, but only has discrete values in the z-dimension. Multi-layer interaction has been presented with more detail in the previous section. The zoomable information space describes interactions with a tangible where its movement affects how much detail or scale is displayed on the tangible screen in comparison to the tabletop display. Finally, in the temporal information space, tangible movement is able to control time-based information, such as video playback. In figure

2.7 [2010b], Spindler et al. give an overview of the interaction vocabulary for their midair tangibles and summarise which interactions are possible with their system. Ranging from simple translation and rotation over gestures to common metaphors, the figure displays the variety of interactions that are becoming possible for tangible user interfaces combined with midair interaction.

User studies suggest intuitive use of the system.

To evaluate their system, Spindler et al. present a few usability studies, which tend to confirm a ease of use of the system. When studying the first implementation of their tangible user interface, the authors report that the given tasks were successfully accomplished by all twelve participants with almost no error, and agreement on statements about ease of use and ease on learning was high [2009]. After having demonstrated the system at conferences on human computer interaction and on their local university campus multiple times, the authors state that several hundred visitors have tried out their prototype, including average users, domain experts and children. The given feedback is stated to be very positive, and that "visitors appreciated how easy it was to learn the interaction techniques and to use the system" [2012]. The authors also did a quantitative study on multi-layer interaction [2012c], which was presented in the previous section 2.3 "Multi-Layer Interaction above Tabletops".

Presented application scenarios

In their publications, the researchers from the University of Magdeburg present a variety of application ideas for midair tangibles. Starting with advanced magic lenses that extend the information of a map displayed on the tabletop screen by additional information layers or by adding the possibility to zoom into certain regions of the map [2009], Spindler et al. also implemented prototypes for graph exploration, medical or geological volumetric data [2010b]. An in-depth prototype demonstrates the use of a tangible as a secondary display that shows the colour palette of an image processing software [2010a]. The publication on tangible windows [2012b] presents applications to copy and paste content using tangibles, to show alternative representations of virtual objects, for example wireframes, and that make use of gestures (flip the tangible to see the backside of a virtual object).

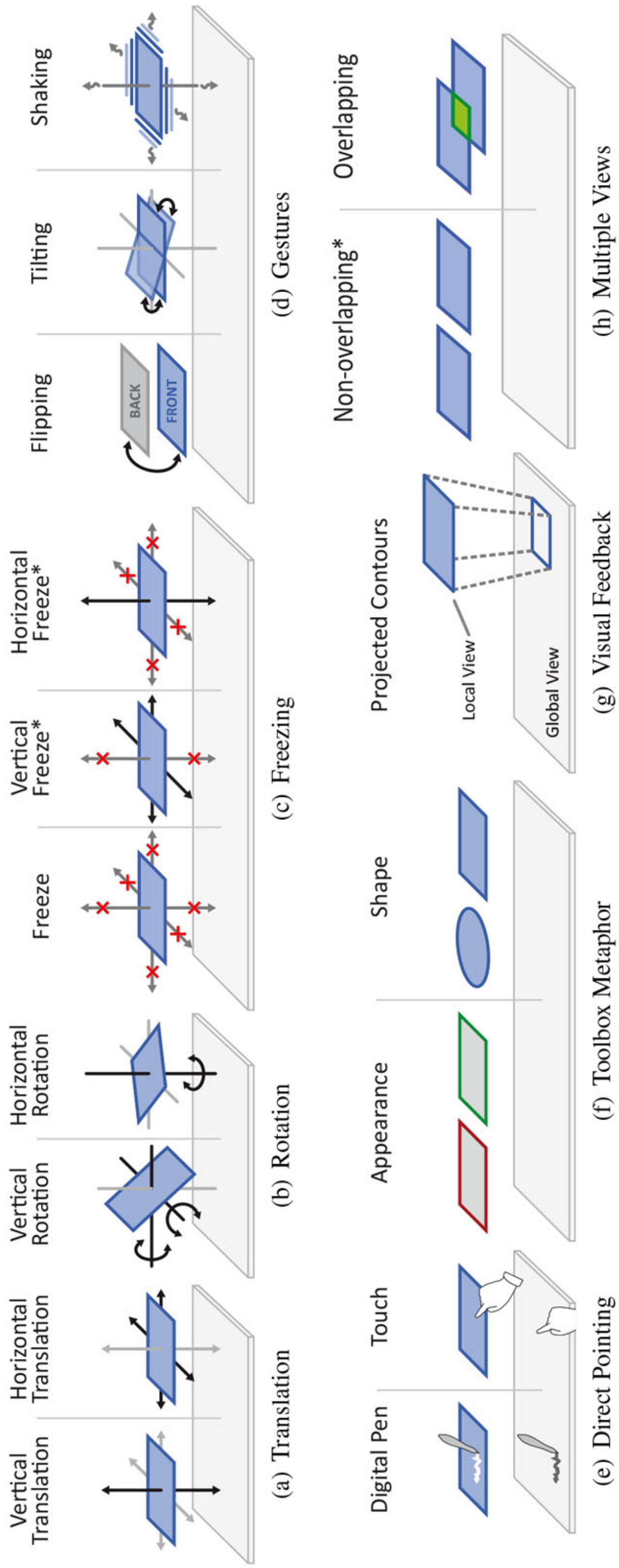


Figure 2.7: Summary of interactions and corresponding vocabulary for midair tangibles. [Spindler et al., 2010b]

2.4.2 Other Research on Midair Tangibles

The MIT Media Lab presents tracking using the camera of a smartphone and a marker on the main screen.

Newer research on off-surface tangibles in conjunction with an interactive tabletop display has been published by the MIT Media Lab [Leigh et al., 2015]. The presented prototype includes a smartphone that is used as a tangible to work, primarily, on a regular laptop screen, but they also show the system with a tabletop display. The camera of the phone is used to track a small marker that is displayed on the main screen, from which the position of the tangible can be determined. When the phone starts to move, the marker will grow and move along with the phone movement, so that tracking is possible at all times. Note that this design prohibits the display of the camera feed on the smartphone, since it will only show the marker, and not the content of the main screen. However, in some cases, the researchers were able to reproduce the content on the screen of the smartphone to keep the illusion of a see-through device. This vision-based technique is quite similar to the approach we will choose for our prototypes (see chapter 3 "Implementation"), which uses Apple's augmented reality framework, but was published two years before Apple even released its ARKit framework to the public. The authors also present a classification of basic interaction techniques that are possible with their system, as depicted in figure 2.8. These include the use of the tangible as boundary condition or spatial relation for virtual objects on the main screen, as magic lense (also see next chapter) and as secondary screen. To demonstrate their prototype, multiple example applications were developed. The smartphone is used as tool to cut-and-paste content, even across multiple devices. When hovering over hyperlinks on the main screen, the user can open the corresponding website on the smartphone browser by tapping on the screen. Another implementation uses the tangible as a controller for an image processing software and allows a user to change an image's blur by rotating the device.

Midair tangible views provide an immersive experience

Chan et al. introduced a tablet-based midair tangible that can be used to explore a three-dimensional version of a two-dimensional satellite image of a large city that is displayed on a tabletop surface [Chan et al., 2010]. Tracked via

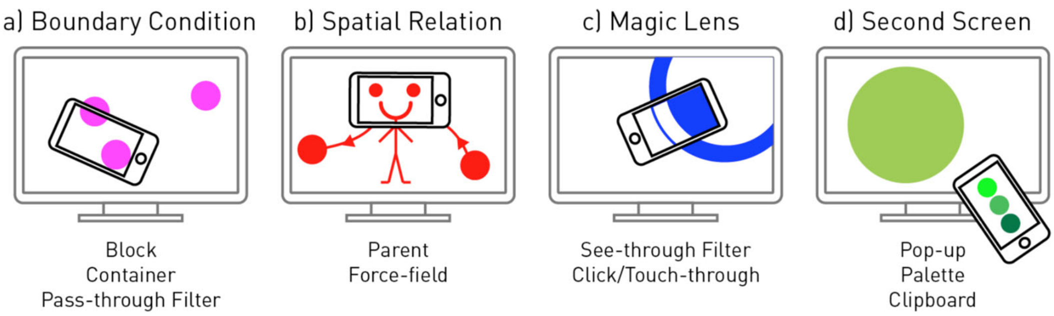


Figure 2.8: Interaction classification for the system published by [Leigh et al., 2015]. Each represents different modes of direct, near-surface interactions.

infrared, the tablet shows 3D silhouettes of the buildings in the city when moving around the tabletop, just as it is known from the 3D content of Google or Apple Maps. The authors found the tablet view to be immersive, to a degree that the it lets users loose awareness of their surrounding. They report the 3D view to isolate the users from others around the table and from the table itself.

Research on using midair tangibles for 3D object rotation

Our own lab at RWTH Aachen University has done various research on tangible interaction in the past [Weiss et al., 2009, Jansen et al., 2010, Voelker et al., 2013, 2015a,b, Cherek et al., 2018, 2019] and started investigating into off-surface tangibles and midair interaction three years ago. A master thesis [Asselborn, 2018] explored the use of tangibles for 3D object rotation and compared this new interaction technique with touch input and a 3D mouse. In particular, the transition design between on-surface and off-surface interaction was a point of interest, and how users perceive the analogy between physical object input and virtual object output. Technically, we developed an active tangible in form of a cube, which contains a gyroscope combined with an accelerometer for orientation detection and whose data is transmitted via bluetooth. Combined with the capabilities of on-surface tangibles, users then are able to move the cube on the tabletop surface and to lift the cube in midair and manipulate it with six degrees of freedom to control virtual objects. We implemented applications for 3D object rotation and one which used the tangible movement to navigate through a virtual, three-dimensional world that was displayed on the tabletop (see figure 2.9).

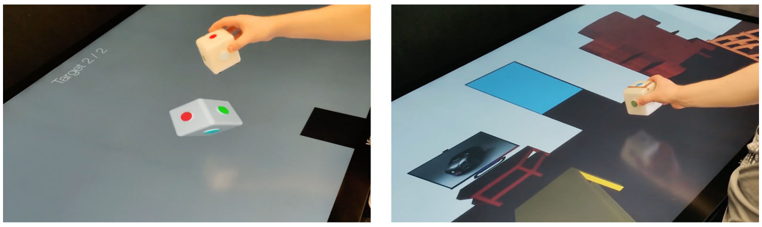


Figure 2.9: Manipulating content on the tabletop using a midair tangible. Left: The orientation of the physical cube affects the orientation of the virtual one. Right: The tangible is used to navigate in a virtual 3D world. [Asselborn, 2018]

A design space of on- and off-surface tangibles

Based on the knowledge gained in implementing the tangibles for 3D object rotations, we looked further into the design space of on-surface and off-surface tangible interaction [Cherek et al., 2019]. Figure 2.10 shows the new dimensions of midair tangible interactions. The figure is divided into on-surface interactions on the bottom and off-surface interactions on the top. X, Y and Z denote the respective dimension with linear positioning, and are followed by its rotational counterparts (i.e. roll, pitch and yaw). The design space is filled with existing input techniques, like a standard computer mouse, or on-surface tangibles, and completed by the novel interaction possibilities which include 3D manipulation or 6D midair manipulation. Studying the design space, it becomes possible to compare input techniques with each other and come up with new modalities by filling the blank regions in the design space.

2.4.3 Comparing 3D Manipulation Tasks

In comparison to other input techniques, midair tangibles tend to be intuitive to use for 3D rotation, but are not necessarily faster.

While qualitative research on midair tangible interaction has been done and yields positive feedback from users (as noted in one of the previous sections), only a few publications cover quantitative aspects. Researchers from TU Munich, Germany, compared interaction devices for translation and rotation of virtual objects in augmented reality [Reifinger et al., 2008]. They created a custom-made midair tangible with accelerometers and a gyroscope and

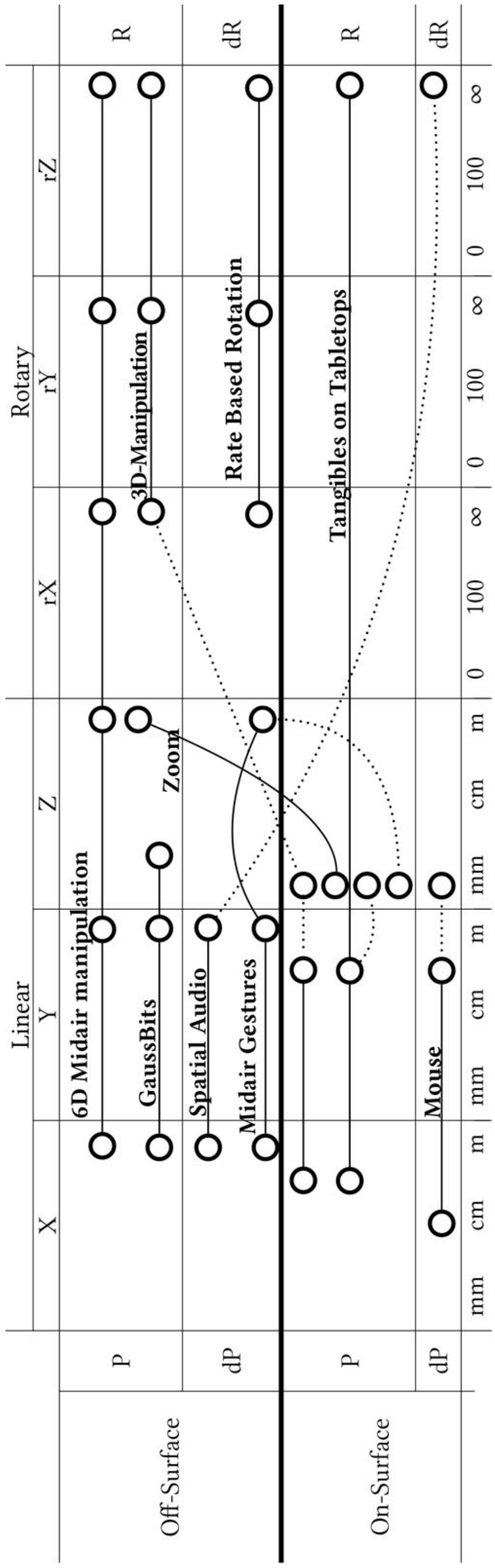


Figure 2.10: A design space of on- and off-surface tangibles. The top half represents interactions with tangibles when they are lifted off the table [Cherek et al., 2019].

used it for midair 3D manipulation. The study shows that this device performed as good as a combination of mouse and keyboard input and when using gestures. In our own study on 3D object rotation that compared midair tangibles with a similar design to touch input and a 3D mouse [Asselborn, 2018], many users stated that the tangible methods were more intuitive to use. Analysing task completion time, the tangible interaction was significantly faster than using touch, but the 3D mouse outperformed the other two input techniques. However, the 3D mouse was considered to have a steep learning curve by most participants.

2.5 Magic Lenses

Magic lense: see-through interface to provide additional information

The term magic lense describes a see-through interface between the main content of a user interface and the user. With such a tool, users can reveal an information layer in addition to the main content. Often, it is used to filter the content by some means or to offer supplementary data. Just as a magnifying glass that increases font when looking through it on a newspaper article, magic lenses use the same metaphor for virtual applications. Magic lenses are not limited to the physical domain: most modern operation systems provide a digital tool that increases the displayed content as accessibility feature and can be dragged around the screen using the mouse. Newer research introduced physical magic lenses that can be placed on top of a display and provide tangible feedback when being moved on the surface.

Close to today's understanding of augmented reality

The concept of magic lenses has been used in human computer interaction for a long time. Even today's popular developments in augmented reality base on the idea of a see-through interface that augment the content of the real world. One of the first publications that introduced magic lenses was presented by researchers at Xerox PARC and Bill Buxton [Bier et al., 1993]. They describe magic lenses as visual filters that "modify the presentation of application objects to reveal hidden information, to enhance data of interest, or to suppress distracting information." Buxton et al. complemented their concept later with an "eye-

Figure 2.11: An old handheld device used as magic lense over a physical map. The device screen shows details of the attraction when being moved close to its location on the map. [Reilly et al., 2005]

in-the-hand" metaphor where a user operated a palmtop to see additional information and which could even be used to navigate on a virtual screen [Fitzmaurice and Buxton, 1994]. This sounds close to our today's understanding of augmented reality, which will be covered in the next section. This section will summarise some publications and prototypes with magic lenses that could lead to further insights in this thesis' context of midair tangible interaction.

From palm computers to smartphones

With the increasing computational abilities and improvements in display technology, researchers started using handhelds or palm computers as magic lenses in the 2000s [Reilly et al., 2005, Olwal, 2006, Rohs et al., 2007, Morrison et al., 2009]. These predecessors of today's smartphones were often used in prototypes in conjunction with a regular, physical map and extended the static map with digital content (see figure 2.11). Tracking was implemented using RFID chips or via external tracking of build-in LEDs that the device was manufactured with. Newer prototypes made use of the device's camera and could match a specific location on a physical map using dedicated markers. With some devices, even more sophisticated image processing without visible markers was possible, just as it is the case with modern augmented reality frameworks. The presented applications include displaying attraction details on the device when hovering over the attraction on the physical map, or a detailed interactive road map of the area of interest. Rohs et al. [2007] additionally present a study to investigate map navigation. The results indicate that users benefit from magic lense interaction to explore the search space and that they finish searches faster in comparison to joystick interaction.

In the *MagPad* prototype, Xu et al. [2015] place regular paper-based documents (i.e. a printed research article) on a dedicated surface and use a smartphone to display additional, location-specific content. When browsing through the referenced literature, users can open a publication's abstract by holding the device close to the citation and performing a gesture. Another application example provides users with the possibility to translate the text of the printed document into another language, with the translation being visible on the screen of the phone. Also using a smartphone, other researchers implemented a magic lense to augment information displayed on a tablet computer [Strohmeier, 2015]. Similar to an implementation of the *PaperLens* prototype (see the section on related work on midair tangibles), the device could show another anatomic layer of the human body to complement the one that was displayed on the tablet. After having set a common ground between the two devices, the position of the device was tracked using the build-in sensors of the smartphone (accelerometer, gyroscope and magnetometer). The authors also investigated into cross-device interaction, which will be covered in one of the upcoming sections.

2.6 Interactive Tabletops and Augmented Reality

AR enhances the user's perception and interaction with the real world by integrating additional, virtual information.

Similar to the concept of a see-through interface known from magic lenses, augmented reality (AR) integrates additional information into the user's view of the real world. Following the characteristics of augmented reality systems in [Azuma, 1997], we understand augmented reality as a combination of real and virtual objects in a real environment, which aligns them with each other and runs interactively and in real-time. In distinction to virtual reality where the surrounding environment is virtual, augmented reality positions the user in the real world [Milgram et al., 1994] and enhances the user's perception and interaction with it. This section focusses on challenges that arise with the use of AR techniques and sets them in context to midair tangible interaction.

World tracking is based on the camera and image processing capabilities of modern smartphones.

First implementations realised augmented reality systems using head-mounted displays [Feiner et al., 1993]. Modern systems based on 3D glasses are commercially available and popular, but with the increasing computational power of today's smartphones, a more unobtrusive approach has become at hand for a broad range of users. Modern augmented reality frameworks for mobile devices (*ARKit* on iOS and iPadOS, *ARCore* on Android-based devices) process the device's camera feed to detect feature points that enables them to calculate movement of the device as well as distance and angle to each tracked point. In conjunction with the device's other sensors (accelerometer, gyroscope and magnetometer), the frameworks are able to localise the device in the surrounding world in real-time and with a precision of centimetres or less. Newest improvements make it possible to detect surfaces (plane detection), general objects like tables, custom objects that have been scanned before and even people in front of virtual content.

Interacting with augmented displays over a digital table introduces new challenges.

Hansen et al. [2005] present first design issues and possible solutions that occur with positioning using the device's camera and image processing. These can easily be transferred to midair tangible interaction. They describe a mixed interaction space above a detected surface that opens a new dimension to interact with the system and ranges from direct manipulation, rapid, incremental and reversible actions to the use of gestures. For example, one of the presented design issues concerns the mapping between the physical movement of the device in space and the action on the interface. The authors propose the use of natural mapping based on analogies or cultural meanings [Norman, 2002] with a relative or absolute conversion or semantic mapping (for example: move the device to the right to play the next song) as a solution.

Users are in risk to forget about their environment when interacting with augmented systems.

Other research [Yang and Maurer, 2010] summarises the existing publications specifically in the context of combining digital tables with augmented reality. Comparing the findings of their literature survey, the authors present challenges with the systems at time of publication, of which some are still applicable to newer systems: Since visual tracking requires the tracked surface to be visible in the camera view, awkward positioning may lead to arm fatigue of the users. Also, since the display size of mobile devices is limited, it may task the users to recognise smaller details on the screen, and users may have to walk further away from the table in order to fit it entirely into the device's viewfinder. Morrison et al. [2009] report corresponding observations in their on-field study using magic lenses where users were walking around the city and used a map and a device that augmented some of the map's content. For example, some users dropped their device on the ground while gesturing or organising their items, and a player walked into a lamppost while looking at the device. This indicates that the users' possibility to be aware of their immediate environment is challenged when interacting with devices that augment the reality around them and require their attention. Apart from challenges on the outside of a system, augmented reality interfaces require new thoughts on the design of software interfaces, too. For example, as virtual content blends seamlessly in within the real world, interfaces have to convey clearly with which objects a user can interact with [Aultman et al., 2018] .

A prototype that combines interactive touch displays with augmented reality.

A recent publication combines an interactive display with augmented reality to ease 3D modelling [Reipschläger and Dachselt, 2019]. Wearing 3D glasses, users are able to draw and alter three-dimensional content in midair using a digital pen or using the display's multi-touch capabilities (see figure 2.12). Users get immediate feedback of their changes through the augmented view. Combining these two input methods, the authors argue to mitigate issues of perception and ergonomics of which the former is commonly found when displaying three-dimensional content on a flat, 2D interface and the latter in midair interaction. For example, the touch surface adds natural, physical constraints and haptic feedback in at least one dimension, while midair gestures are performed without natural constraints in plain air.

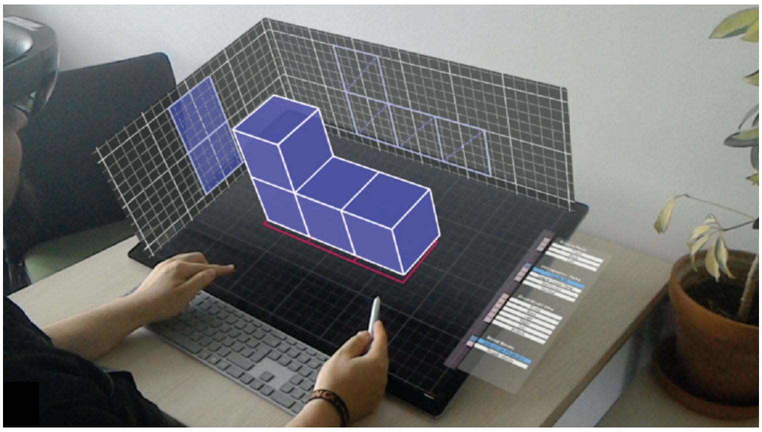


Figure 2.12: Combining augmented reality with interactive displays for 3D modelling. [Reipschläger and Dachselt, 2019]

2.7 Cross-Device Interaction

Combining multiple devices results in benefits from using different display sizes and input modalities.

Many of the research and prototypes that are presented in the previous sections combine interactions with more than one device. Midair tangible interaction relies on one tangible device in the users' hands, and on a second device, the interactive tabletop display, to present the main content. As an overview, figure 2.13 shows the many possibilities to combine multiple devices. The figure is based on a literature survey of over 500 papers in the cross-device computing domain [Brudy et al., 2019]. The top part of the figure shows sub disciplines of cross-device interaction, while the bottom part shows focus areas of the different research projects. The sub disciplines range from simple dual monitor systems over multi-mobile devices to tabletop and tangible interaction. Following this taxonomy, this thesis' research topic would be classified as cross-device interaction with a focus on interactions, collaboration and use-cases. The authors also note as on one of the key challenges of cross-device interaction: Communicating the possible actions to the user is crucial. This section will present research on cross-device or multi-device interaction (both terms are used synonymously) with a focus on what can be learned for midair tangible interaction.

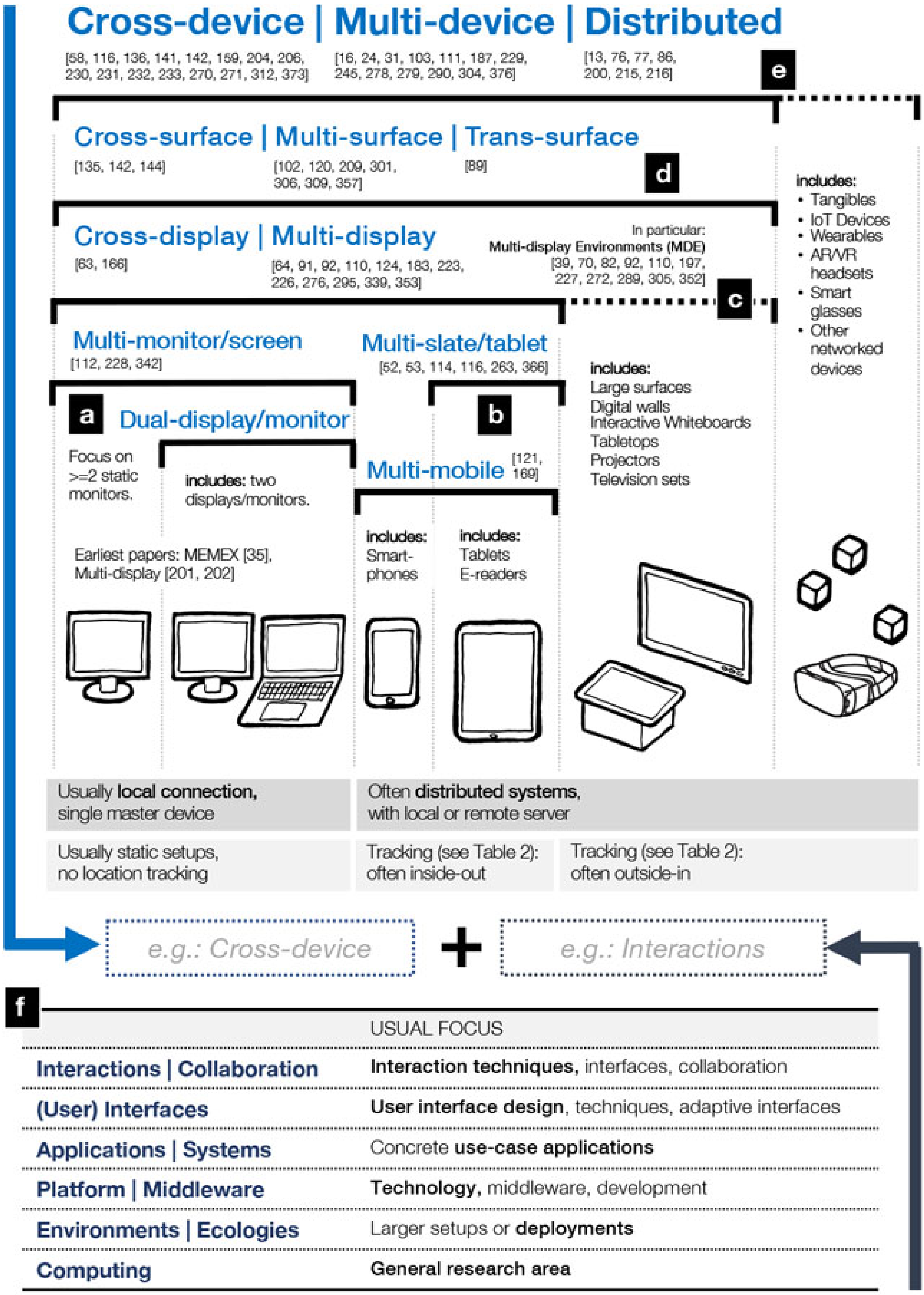


Figure 2.13: An overview of cross-device interaction research, with its sub disciplines at the top and focus areas at the bottom. *(Contains references to literature and tables of the original paper that are not included in this thesis.)* [Brudy et al., 2019]

Magic lense prototypes (see one of the previous sections) can be categorised into cross-device interaction, too. A more advanced example of a magic lense combines a tablet PC with a projected, wall-sized image [Sanneblad and Holmquist, 2006]. The prototype is used to demonstrate the interaction with very large computer graphics images, which forces users to choose between either a view of the whole image or a view of a detailed part of it, but not both at the same time. The presented solution used the tablet-sized device to show the details of the region covered by the device, while the whole image is projected onto the wall. The system was presented to several thousand people in a week-long public exhibit. Another earlier publication combines a vision-based tabletop display with phone-like devices [Wilson and Sarin, 2007]. Using bluetooth and computer vision, the system gains the ability to recognise phones on the surface of a table. The system can then be used to copy photos from the device to the tabletop, and sort or delete them on the bigger screen. It is also possible to copy the photos back to the original or to another device.

Early cross-device prototypes use small devices and vision-based, larger displays.

Systems that combine multiple devices benefit from the advantages in input modalities that each devices offer, or to mitigate issues that exist for one input device. For example, a regular, non-touch display of laptops or projectors can be combined with a second device that supports touch. This way, users can select elements on the large display using touch on the hand-sized device. The setup would also allow users to reach displays that are further away. Figure 2.14 shows a prototype that introduced such a system using live video images of surrounding displays on the mobile device [Boring et al., 2010]. Even though the authors report that their system does not offer enough stability and control to enable precise manipulation yet, the publication provides interesting scenarios for the use of midair tangibles in conjunction with an interactive tabletop display, for example to address the reachability problem or when other users around the tabletop would block a user from doing a certain action directly. In a later publication, the authors reverse their metaphor from *extending a large display with a device* to *extending the device content with a large screen* [Baur et al., 2012]. For this device-centric approach, the content of the device screen is projected to a large display. While

Using multiple devices together combines the benefits of each device and allow users to interact with devices that are further away.

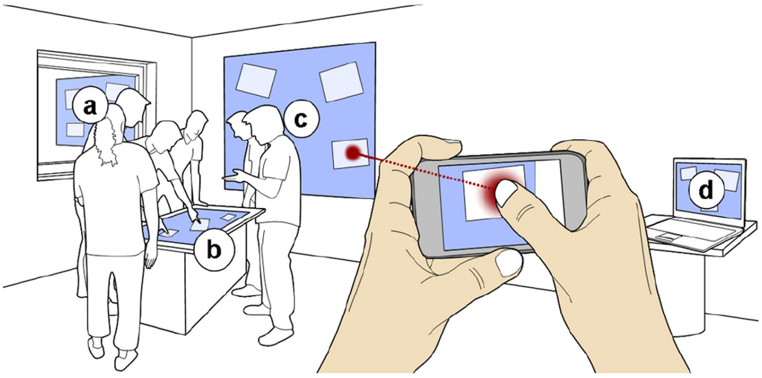


Figure 2.14: "[The device] allows users to manipulate content on distant displays that are unreachable, such as (a) displays outside a window, or (b) a tabletop system crowded with people. It allows users to manipulate devices that are incapable of touch interaction, such as (c) a wall projection or (d) a laptop. Users point the device at the respective display and manipulate its content by touching and dragging objects in live video. The device projects the touch input onto the target display, which acts as if it had occurred on itself." [Boring et al., 2010]

the device only shows a part of an image, the large screen could then show the whole image (similar to the magic lense metaphor used before), or the content could be mirrored one to one and viewed and edited on both displays collaboratively. In another use-case, a user starts a video on the device, which then gets displayed on the large display and the device turns into a remote control which can pause or forward the video.

Communicating the possible actions and making use of spatial awareness

Leaving the area of combining small devices with large screens, research on interaction between tablet- or smartphone-sized devices also gives interesting application examples and clues for midair tangible interaction. Even though its device is not tracked in space, a publication on a tangible bookmark that is used on a tablet provides two interesting insights and some very nice storyboards (see figure 2.15) [Bianchi et al., 2015]: First, the authors observed a rapid flow in the interaction between touch (browsing or scrolling through a document) and the tangible bookmark (making notes, . . .). Second, users reflected on poor affor-

Figure 2.15: One of the storyboards for the use of a bookmark-style tangible on a tablet from [Boring et al., 2010]. *"Liza reviews an academic paper. She uses her bookmark to look through the pages and check citations whilst reading. She stores a copy of an important figure so she can examine it throughout her read. She also makes notes to help organise her review."*

dances of gestural style interactions. When implementing document-focussed or gestural applications, this could be considered. Another aspect of our envisioned system using midair tangibles will be the spatial awareness between the tabletop and (possibly several) midair devices, so that each device can make use of the knowledge of its position in relation to the others. Spatial awareness between mobile devices has been a continued field of study with use-cases including collaborative visualisation, visual data exploration and a shared overview device for multiple users with tablets [Wozniak et al., 2016, Plank et al., 2017, Langner et al., 2018, Brudy et al., 2018]. Additionally, smaller mobile devices (i.e. smartphones) have been used in a gesture-like interaction to select cells or perform other actions in a spreadsheet processing application running on a tablet [Perelman et al., 2018], or to hand over content from one device to another using a "pouring" metaphor [Korzetz et al., 2019]

Tracking users' gaze when working with multiple devices allows new interaction methods.

A yet-to-be-published paper for The Conference on Human Factors in Computing Systems (CHI) 2020 from our lab presents a new approach of cross-device interaction using gaze [Voelker et al., 2020]. With the capability of gaze-tracking that is included in the recent, state-of-the-art smartphones, our prototype allows users to move content between multiple devices by performing a touch gesture and just looking at the other device. In a multi-user setup, tracking users' gaze and face authentication make it possible to show and hide content on different tablets with respect to which user is looking at it. Another prototype uses this ability in a card game to hide players' cards on their de-

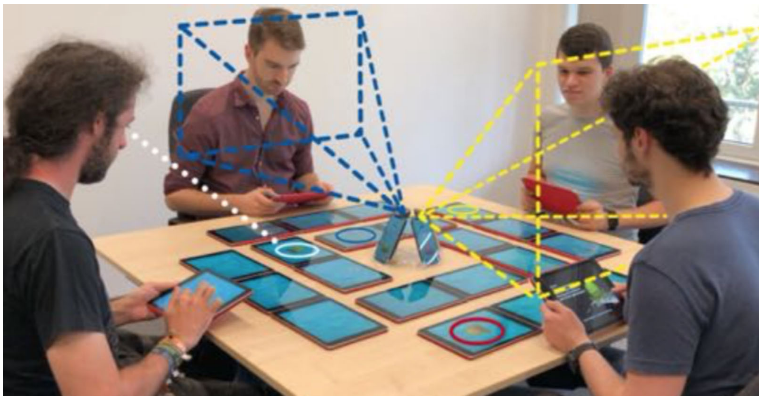


Figure 2.16: Users can interact with multiple devices using gaze. [Voelker et al., 2020]

vice when leaving the game or when another user peeks at a neighbours' device. A study in the setup shown in figure 2.16 with four users around the table and twenty tablets on the desk reveals that the system can reliably identify which tablet or collaborator a user is looking at.

2.8 Multiple Users and Collaboration

Tabletops allow users to work together on a large screen.

Large interactive tabletop displays, like the one we will be using for our implementation of midair tangibles, afford multiple users around them. Users benefit from the display that can show shared content large enough for everyone to see it and that can be used to visualise changes on which the users agreed upon, while being positioned in a circled way around the table instead of all facing a vertical display on a wall. Being able to look each other in the eyes, tabletops can facilitate communication. The surface of a table is metaphorically divided into three territorial regions to help coordinate users' interactions within the shared tabletop workspace, as has been shown by [Scott et al., 2004]: A personal territory for each user, a group territory as primary region for shared content and a storage territory for content that is not within the current scope of interest for the collaborators. Introducing midair tangible devices, it could be argued that the personal territory is extended to

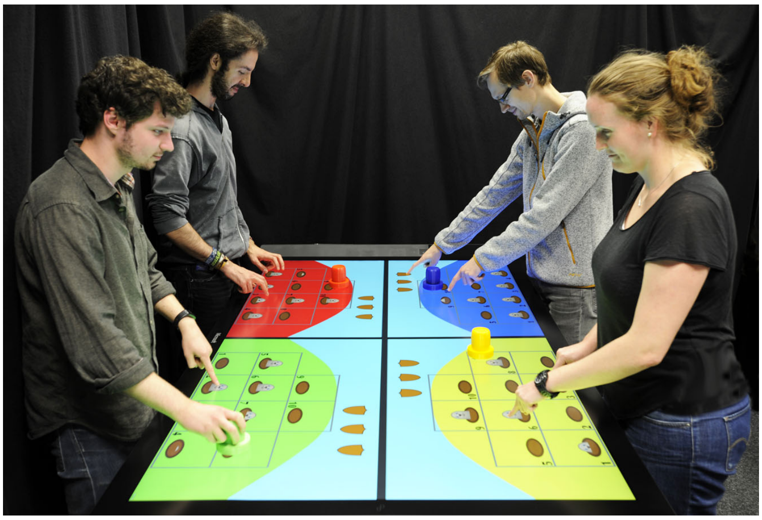


Figure 2.17: In a collaborative setting, tangibles make users aware of other users' actions. The image shows a game in which users could influence the opposing users' actions by placing a barrel. [Cherek et al., 2018]

the devices users hold in their hands and therefore by an additional level of privacy, since the devices usually are more protected from glances of other users.

In a multi-user scenario, users benefit from tangible interaction even more.

In a collaborative setting, benefits from using tangibles could emerge even more. For example, eyes-free interaction with the tabletop – which tangibles provide [Weiss et al., 2009, Voelker et al., 2015b] – is even more important when trying to communicate with a person on the opposite side of the table at the same time. Our lab has done more research on tangibles in a multi-user scenario [Cherek et al., 2018]. Our study reveals that in a collaborative setting, tangibles make users more aware of other users' secondary actions (see figure 2.17). Interestingly, some participants stated that not necessarily the physical object helped them to perceive a tangible, but also the gentle sound caused by its movement on the surface. Other research on tabletop collaboration focussed on the nature of the tasks at hand: Studies with the *Youtopia* system reveal that tasks that rely on each other, i.e. when users had to wait for another user to finish a task, lead to higher social, verbal and physical interaction [Antle et al., 2013, Fan et al., 2014].

Detecting the users around a tabletop

With multiple users interacting with a table, detecting the position of single users or even identifying single persons becomes of interest and would allow user-dependent views or content. On the *LumiSight Table*, four users around a vision-based tabletop can each see different versions of the same content, for example correctly rotated names on a map [Kakehi et al., 2005]. The *Medusa* system provides different views and interactions with respect to which side (or corner) of the table the user is located [Annett et al., 2011]. The authors also implemented a content control and do-not-disturb-mode for multi-user scenarios. With technological improvements, estimating the users' position before touching the table becomes possible and can be used by the system for intention guessing or to rotate virtual objects to the correct orientation from the users' view. [Suto et al., 2018]. As has been shown in the previous section, our lab successfully implemented a prototype that uses gaze to track individual users around a tabletop [Voelker et al., 2020].

Chapter 3

Implementation

This chapter summarises the implementation of midair tangibles over an interactive tabletop that are tracked using augmented reality techniques. After presenting basic design decisions and which devices we used specifically, the first section holds the first implementation step: recognising the tabletop display in the augmented reality framework. Now that the devices are aware of the position of the tabletop, they can transmit their position data to the system running on the interactive table. The following sections describe our findings for good marker creation, how we implemented device touch support and finally summarise the structure of the framework that we used to develop our different application scenarios (cf. chapter 4 "Applications").

Chapter summary: From relating the device and tabletop positions over device touch support to the final framework

Before starting with more detailed implementation steps, the idea to base our approach for midair tangibles on augmented reality techniques should be given some thoughts. From experiences from our other research projects [Wacker et al., 2019, 2020], we learned that the AR tracking technology has improved rapidly during the last years and could be suitable for our use-cases. By using an augmented reality framework, we hope to benefit in multiple ways: The device interaction would allow six degrees of freedom (free movement and rotation in all dimensions; cf. "design space" in section 2.4.2 "Other Research on Midair Tangibles"), and no special hardware would be needed, since ARKit runs on all recent smartphones and tablets by Ap-

Motivation why we choose the ARKit approach ple. We could use the devices' touch capabilities or other input and output modalities (display, speaker, front-camera, motion sensors, ...) that the devices provide. In our use-case of using the devices over an interactive tabletop display, we are able to control the markers that are used to track the devices' positions, since they are displayed on the screen of the table. In comparison to tracking static images on a regular table, controlling the markers would allow us for example to use dedicated markers for different levels of a game, or even to animate them. In general, our AR-powered midair tangibles support nearly every application scenario that has been summarised in the chapter on related work: magic lenses, 3D exploration, user-tracking, user specific-content on the device, user-centred content on the tabletop, spatial-awareness between devices and tabletop etc. With this thesis, we would like to find out if using augmented reality is suitable and stable enough for midair tangible interaction. Specifically, we would like to check if the tracking quality and the position updates tend to be good enough to deliver a natural experience to users.

Technical setup and pre-requisites

Technically, we developed our system in Swift[1] both for the iOS apps that run on the devices and the macOS app that is responsible for the content on the tabletop display. The core functionalities of the device apps will base on ARKit[2]. Apple's augmented reality framework integrates the device's camera and motion features to create a virtual representation of the surrounding world. As devices, we used an iPhone X and an iPhone 6s during implementation and tested our system with several other iPhones and a 6th-generation iPad, all running on iOS / iPadOS 13. Our interactive tabletop is a Microsoft Surface Hub 84 from 2016 with capacitive multi-touch and multi-pen support. The display has a size of 84 inch and allows a resolution up to 4K. Its physical dimensions including the framing are 220×117 centimetres. Figure 3.1 shows a sketch of the tabletop and a break down of the dimensions. The tabletop is connected to an Apple iMac Pro to control the displayed content and to process the touch data. Tabletop and devices transmit data using a local ethernet access point over wifi.

[1] https://developer.apple.com/swift/
[2] https://developer.apple.com/documentation/arkit/

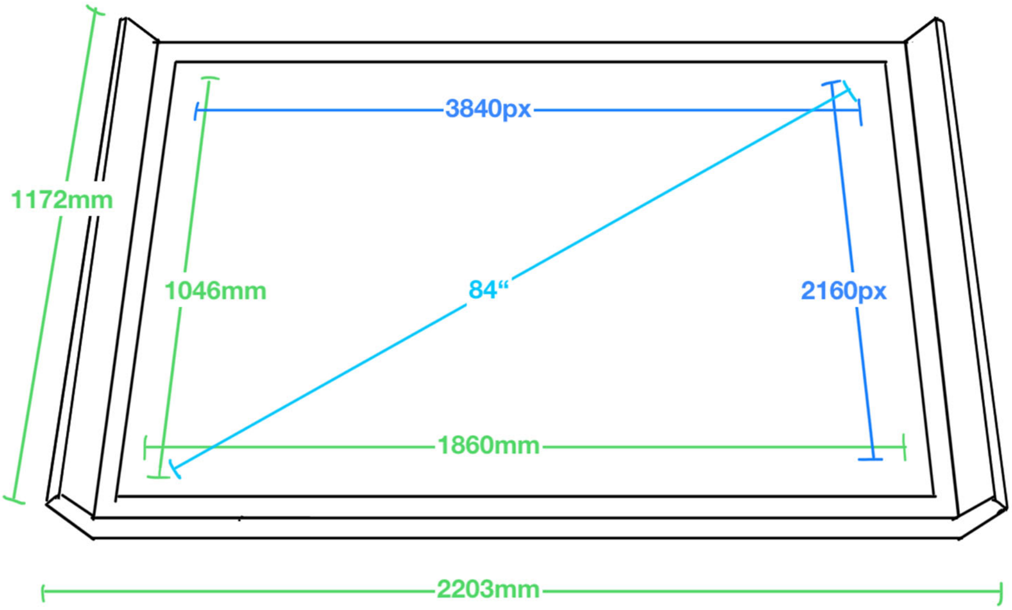


Figure 3.1: The physical and virtual dimensions of the Microsoft Surface Hub 84 that we use as interactive tabletop display. It comes with multi-touch support and allows a resolution up to 4K on its 84-inch-display.

3.1 Recognise the Tabletop in AR

Our first step towards midair tangibles using ARKit was to teach the device to recognise the tabletop display. We could then go ahead and display augmented 3D objects on the surface. The devices can the be used to explore these virtual objects in 3D by moving around the table (cf. section 2.4.1 "The PaperLens Prototype"). It would even be possible replace the whole content of the tabletop display when looking through the device and provide scenarios where the device is being used as a magic lense (cf. section 2.5 "Magic Lenses").

Plane detection works better than we anticipated.

We started with the idea to use ARKit's plane detection functionality which is able to detect plane surfaces in vertical and horizontal direction. With this approach, we could even avoid to use image or marker recognition and we wouldn't be forced to include trackable content on the tabletop display. We tested plane detection with Apple's

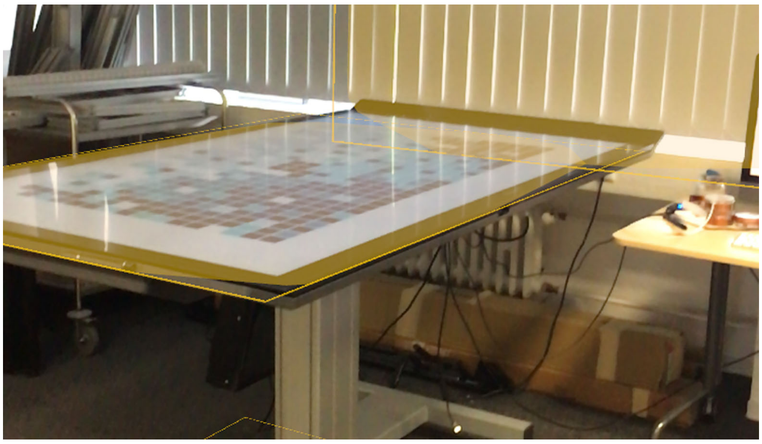


Figure 3.2: Trying out the plane detection feature. The yellow rectangles indicate planes that were recognised by ARKit.

sample app for Building Your First AR Experience[3] which visualises detected planes by convex polygons and estimates rectangular boundaries based on the detected features. The test was performed by moving around the table on three sides with varying distance, and the tabletop display showed a mosaic-like pattern. Apart from parts of the floor and the curtain that were recognised, too, the result is – by our expectations – surprisingly okay and are shown in figure 3.2. The system successfully identified the tabletop surface and drew a virtual rectangle that does not appear to deviate from the original tabletop by more than five to ten centimetres. However, the angled edges (that host cameras and speakers needed for the case the Microsoft Surface Hub is used as intended on a wall for video-conferences) seem to trouble the algorithm a little and resulted in bad recognition in these areas. Despite the rather good results, we decided to go on with image recognition because we wanted to see if the tabletop surfaces could be matched even closer.

Calculating the tabletop size by using image recognition

To detect the tabletop using image recognition, our plan was to show a dedicated image on the tabletop display of which we know the exact size and position on the screen. This image will act as a marker and is fed to ARKit's recog-

[3]https://developer.apple.com/documentation/arkit/building_your_first_ar_experience

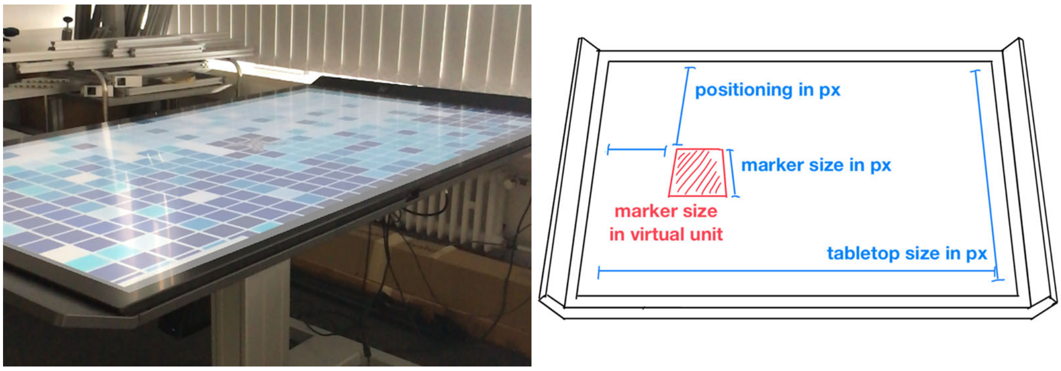


Figure 3.3: Overlaying the tabletop display with a virtual, semi-transparent plane (left). The dimensions can be translated into the virtual coordinate system (right).

nition algorithm. When it gets detected, we can use its measurements in the virtual coordinate system of the device to translate the tabletop's physical size to its virtual counterpart. The translation between the two coordinate systems then can be calculated by the values that we know: The virtual tabletop width equals to the tabletop's width divided by the marker width (both in pixels, see figure 3.3, right) multiplied by the marker width in the virtual coordinate system that is provided by ARKit. The height and the positioning can be computed accordingly. Using these values, we can place a virtual plane onto the table's surface, as depicted in figure 3.3, on the left. This virtual plane can then be filled with any content we like to replace or overlay parts of what is displayed on the tabletop display. Additionally, we have to calculate the physical size of the marker image when it will be displayed on the tabletop. ARKit expects these values when creating trackable reference objects to improve its recognition capability.

Our system is able to recognise the tabletop using marker tracking with a satisfying accuracy.

ARKit uses SceneKit[4] to position and render 3D content, which uses a node-based hierarchy as underlying structure. Each node contains information about position, rotation and scale with respect to its parent node. At first, designing and implementing in three-dimensional space was challenging. For example, when setting the virtual plane to the correct position, the nodes' coordinates have to be calculated in relation to their parent's node, and with the sup-

[4]https://developer.apple.com/documentation/scenekit/

port of multiple trackable markers (see below), complexity increased. For a test application, we implemented a temporal view of the arctic sea ice (see section 4.2 "Temporal Data Exploration") that shows aerial views of different years on the tabletop and on the device. We are satisfied with the accuracy of the virtual plane's position, especially when the device was held closer to the tabletop so that its frame was not visible any more. However, when the tracked marker is not visible in the camera's viewfinder, the tracking becomes more and more inaccurate. For this reason, we added the ability to our system to support multiple marker images by dividing the background image that is displayed on the tabletop into several, trackable regions. The creation of multiple markers is explained in one of the following sections. In the future, we would like to improve the tracking abilities even more by combining marker recognition with plane detection (see chapter 6.2 "Future Work"). For now, as we are able to recognise the tabletop display with a satisfying result, we continue on connecting the device and the tabletop so that both become spatially aware of each other.

3.2 Relate Position of Device and Tabletop using World Tracking

Using world tracking, the device is able to transmit its position and rotation in relation to the marker to the tabletop.

The core of the presented approach for tangible midair interaction is the position synchronisation between the tabletop and devices used around and above it. As has been explained in the previous section, the devices run on ARKit and are able to recognise the tabletop surface using image recognition of predefined marker regions. These marker images will be the shared common ground between the tabletop and a device. We implemented our system with ARKit's world tracking option[5] that is able to synthesise a virtual world from its surrounding by tracking feature points on surfaces and objects. Feature points are unique highlights in the camera feed that ARKit uses to create the virtual representation of the world around it, like regions with high-contrast or edges of physical objects. Now, when one of our markers on the tabletop is being recognised,

[5]https://developer.apple.com/documentation/arkit/world_tracking

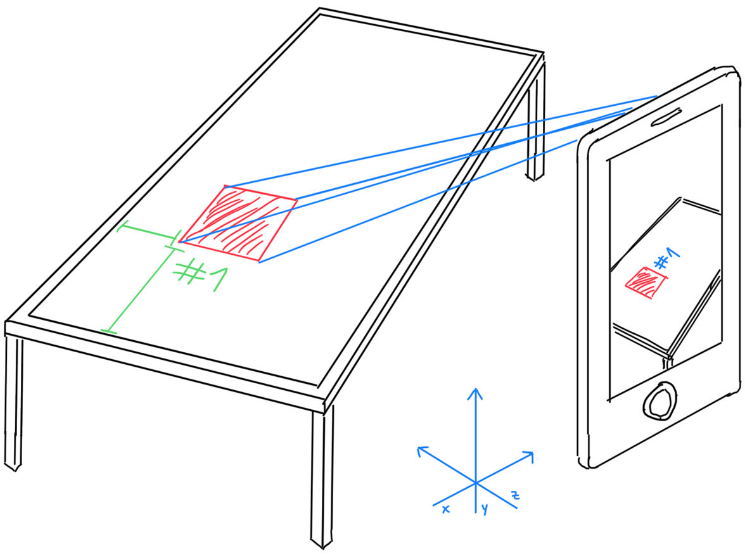


Figure 3.4: Relating the position of a device and the tabletop: The marker ID (red) and the vector between device camera and marker (blue) is transmitted to the tabletop.

ARKit places the reference image of the marker in its virtual world. We are then able to calculate the vector between the device and the marker in the unit of ARKit's coordinate system that includes the device's position and rotation in 3D. This location data is transmitted to the tabletop, together with an identifier of the tracked marker. Since the tabletop can retrieve the position of a marker on the surface by its identifier, it finally is able to recalculate the device's position from the location data with respect to itself. Figure 3.4 shows a conceptual visualisation of this approach.

To reduce the amount of computations on the device, our implementation actually transmits the position data of both the marker and the device to the tabletop in the device's virtual coordinate systems. The tabletop then mimics the device coordinate system, uses this to calculate the vector between device and marker and transfers the result to its own coordinate system. This approach puts more load on the data transmission and the computer running the tabletop visualisation, but proofed to be stable. To update the position in real-time, the position updates are transmitted

Position updates are transmitted 30 times per second using the websocket protocol.

30 times per second via wifi. As simple HTTP connections are not equipped for such a high rate, we chose to transmit the position data using websockets. The transmitted data itself is encoded as JSON strings based on Swift structs and can be extended to include further information from device to tabletop and vice-versa (cf. section 3.4 "Device Touch Support"). In the future, we could imagine using Apple's MultipeerConnectivity[6] feature to let the devices and the tabletop connect ad-hoc instead of using static IP addresses. Going even further, in a multi-device setup, devices could share their virtual map[7] of the world containing their model of the tabletop with each other to improve the overall tracking experience. Also, as all devices connect to the tabletop and transfer their position, it would easily be possible for the tabletop to transmit the positions of the other devices to each device, even translated into the corresponding virtual coordinate systems. We did not implement this, but imagine that it's possible to construct interesting application scenarios for such devices with spatial awareness. For example, it would be possible to allow or forbid certain actions when a device is close to another.

3.3 Marker Creation

Image detection relies heavily on the quality of the marker images.

Now that we have given the conceptual background and some details on exchanging the positions between devices and the tabletop, we'd like to provide some more insights into our findings in marker creation and how we achieved to split the tabletop display into multiple markers. Having good, trackable marker images displayed on the tabletop surface is of essence to our AR-based approach to track the device position, but at the same time introduces a limiting factor to the design of the visual content on the tabletop With a low tracking quality, ARKit is not able to provide us with a satisfying positioning. We experienced a rather small margin between a good tracking state and complete loss of tracking, which results in unforeseeable position updates.

[6]https://developer.apple.com/documentation/multipeerconnectivity

[7]https://developer.apple.com/documentation/arkit/creating_a_multiuser_ar_experience

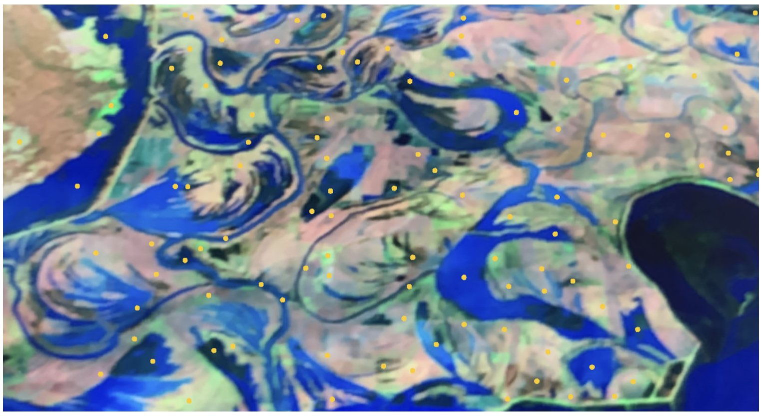


Figure 3.5: Example of a feature-rich image that we use in one of our applications. The yellow dots indicate feature points that were recognised by ARKit.

Feature-rich, high-resolution images are best.

Our first attempts as markers were based on AR stickers, which are based on geometrical forms and look a bit like QR codes. However, these did not yield good results with our AR framework. ARKit uses feature points (cf. section 3.2 "Relate Position of Device and Tabletop using World Tracking") to detect surrounding objects surfaces. Therefore, we choose large, feature- and contrast-rich images in a high resolution for our marker images (see figure 3.5). Furthermore, repetitive structures should be avoided, and marker images ideally have a large colour spectrum. When importing images as AR resources to Apple's IDE, Xcode, the images get analysed beforehand and feedback on their tracking capability is provided. To find good, high-resolution images (our tabletop runs in 4K) as background of which parts can be used as marker, we found the satellite images provided by NASA's earth observatory[8] to be a good resource.

The size of the markers depend on how far or close users will interact with the tabletop.

Another aspect of marker creation concerns the size of the markers. Just using the content of the whole tabletop screen as a marker is a possible solution for scenarios where the device (and the user holding it) is further away from the tabletop, so that all or most of the tabletop surface is visible in the device's viewfinder at once. With the focal lengths of

[8]https://earthobservatory.nasa.gov/images

the iPhones that we used, this results in a distance of about 1.5 m to 2 m from the tabletop, which we were not satisfied with for our use cases (see chapter 4 "Applications"). It is to be noted though that ARKit is able to recognise only parts of a marker and does not necessarily have to see the whole image, but the scope if this feature is rather limited. We then tried different marker sizes (12 cm, 27.5 cm, 42 cm, 90 cm) by splitting the tabletop background image into multiple squares of the given sizes. Our informal tests indicate that the bigger the size, the better the tracking. However, the 90 cm markers (resulting in two regions on the tabletop) are too big again and do not get recognised when holding the device closer to the screen. So by now the best marker size, with users standing close to the tabletop, seems to be around 40 cm, which we used for the implementation of our applications later. As we experimented with different marker sizes, we created a helper tool that can generate markers in the desired size based on a large image as input, and provides the marker images and their relative position on the screen in a format readable by Xcode, so that the result can easily be imported into a new AR project.

To improve tracking quality, the device will look for new markers after some seconds.

Now with the possibility to have multiple marker images on the tabletop surface, we improved our implementation that recognises the markers and sets them in relation to the tabletop surface. Before, the devices could only keep track of one marker, i.e. the first one that got recognised. If this first marker was located on the left side of the tabletop for example, the tracking quality decreased when moving to the right of the table, since the device had to calculate its position only by feature points and its sensors, without the ability to ground itself again by re-tracking the marker. At first, we thought of letting the device switch the positioning to a new marker as soon as it has been detected. But since it is possible to have multiple marker images in the device's viewfinder at the same time, this approach resulted in a very high rate of repositioning attempts, accompanied by flickering of the virtual tabletop representation and high computational need. To improve this, we instead added an expiry date of some seconds to a new-tracked marker. When this deadline has passed, the device starts looking if it finds a new marker that it can track and base its position on. In our preliminary tests, this approach showed to be a good trade-off between accurate positioning and power consumption of the device. Other marker-related challenges that occurred during implementation are listed within the evaluation chapter in section 5.1 "First Test and Early User Feedback" and include issues with lightning and sun that reflected on the tabletop surface and tracking in the near-surface areas.

3.4 Device Touch Support

To explore different use cases and interaction designs, we would like to make use of the device touch capabilities. By looking through the device screen onto the tabletop, users should be able to select elements on the tabletop surface by tapping on the device. This allows users to interact with the tabletop surface without actually reaching it. More on the reachability problem and benefits of combining input modalities from multiple devices has been summarised in the chapter on related work (section 2.7 "Cross-Device Interaction"). In order for our system to support device touch, the tabletop has to determine which part of its screen is currently visible in the device's viewfinder. From this, we can match the touch positions on the device screen into the tabletop coordinate system and perform a hit test on certain elements.

Combining device touch with the tabletop allows users to reach targets that are further away.

To calculate the current device viewport on the tabletop, we implemented a raycast from the camera position onto the surface. From the position updates (see one of the previous sections), we are able to retrieve a real-time, physical position of the device camera in the tabletop's coordinate system. Then, we project a line from at the camera position orthogonally to the device while targeting the surface. The intersection of this line with the surface determines the centre of the device viewfinder, which is shown in figure 3.6. We can transmit touch point positions relative to the centre from the device to the tabletop, which can then be translated into the right coordinate system using the intersection point. The raycast is performed at each incoming position update.

Touch locations are translated using a raycast of the device viewfinder onto the tabletop.

Figure 3.6: Raycast of the device's viewfinder onto the tabletop surface: The device representation can be seen at the bottom, with a blue circle at the camera position. A red line indicates the raycast from the camera to the tabletop surface, where a red rectangle highlights the part of the content that is currently in the focus of the device. When looking at the device screen, this can be verified by checking that the red rectangle indeed is in its centre. (The red outline indicates which marker is currently tracked by the device.)

Two touch modes where users can either focus their eyes on the device or on the tabletop

We implemented two different touch modes: direct touch and indirect touch using a cursor. In direct mode, tap locations on the device screen get exactly matched onto the tabletop using the viewfinder representation as described above. The system then performs a hit test to retrieve selected elements. In order to select elements using the direct touch mode, users have to focus on the device and usually look through the device screen on the tabletop. In indirect mode, the users control a cursor on the tabletop screen by moving the device. This cursor is always located at the intersection point described above, i.e. at the orthogonal projection of the device screen centre. Users can tap at any position on the device screen, which will trigger a hit test in the middle of the cursor. An implementation of the cursor can be seen in figure 4.7 located in the next chapter (section 4.4.2 "Collaborative Target Selection: Feeding Animals"). With this approach, they can focus their eyes on the tabletop display and move the cursor to and select elements of interest without looking at the device. The touch data is transmitted from the devices to the tabletop together with the position updates. The result of incoming touch transmissions can be send back to a device, for example to indicate that a certain element has been hit or to transmit any additional, action-specific data.

3.5 AMT Framework

The core functions of our implementation have been split into several Swift projects to improve reusability.

To facilitate the development of multiple applications using the described concepts and designs to track midair tangibles using augmented reality, we based the needed source code into a reusable framework. The AMT (AR-powered midair tangible) framework combines the functionality to track different devices and to translate the positions into the tabletop coordinate system. Figure 3.7 shows the structure of a project that uses the framework. Each project consists of an iOS app that runs the needed ARKit components and determines the device position using SceneKit. On the tabletop, a macOS app receives the position coordinates and translates the location data in its own, SceneKit-based coordinate system. To support touch, tangibles and simple 2D content on the tabletop, the macOS app includes a framework that was developed by our lab (see below) and is based on SpriteKit[9]. A project-based, common resource group can be used to store the marker images together with their positions on the tabletop and other shared assets. The main functions that handle the position updates are outsourced into three Swift packages that provide the same code base for all projects and have to be called by the respective delegate handlers in the iOS and macOS apps. The communication part depends on third-party plugins that implement the websocket protocol.

Integrating 3D content into Apple's 2D rendering framework

To ease the development of 2D user interfaces on the tabletop display and to support touch, we included the MultiTouchKit (MTK) that was developed by our lab specifically for the use with interactive tabletops like the Microsoft Surface Hub and other touch devices. [Linden, 2015,

[9]https://developer.apple.com/documentation/spritekit

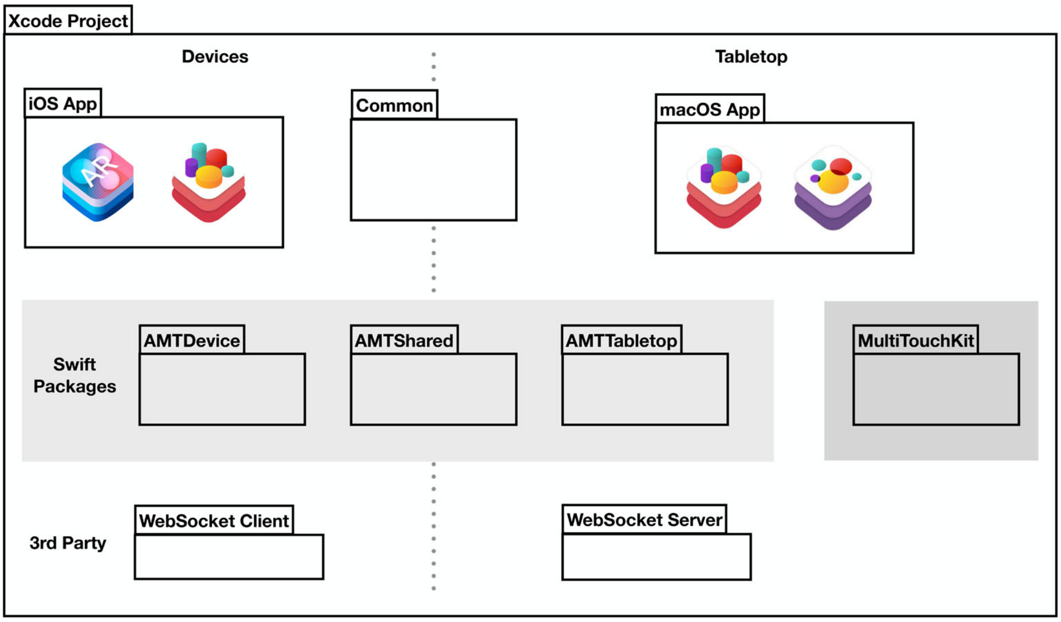


Figure 3.7: The project structure for our AMT (AR-powered midair tangible) framework consists of device-specific, tabletop-specific and shared resources and is based into Swift packages to improve reusability. The icons indicate the used technologies (from left to right): ARKit, SceneKit and SpriteKit.

Asselborn, 2018]. The framework directly uses the touch traces of the capacitive display and provides APIs to track multi-touch events as well as active and passive tangibles. Apple's 2D framework SpriteKit is used as a foundation for object placement and rendering. Fortunately, SpriteKit provides support to include 3D content that is based on SceneKit, which we need to handle the incoming position data of the devices. Unfortunately, Apple's implementation is not free of bugs and we experienced some issues with the positioning of elements in z-direction which have already been reported by [Asselborn, 2018] two years ago, haven't been fixed yet and still caused us some headaches in debugging, even though we were aware of them.

Chapter 4

Applications

Our literature review (see chapter 2) revealed several application scenarios for midair tangible interaction. Many of the presented prototypes could also be implemented with our approach to use augmented reality as tracking technology for off-surface tangibles. Based on the different types of interaction designs that we identified during the literature research, we decided to implement five applications to explore midair tangible interaction. Our applications cover use-cases for 3D exploration, temporal data exploration, a tangible midair controller with direct mapping to its virtual counterpart, layer exploration and target selection.

Contribution: Five prototypes that cover a range of use-cases for midair tangible interaction.

The main purpose of our first two apps is a proof-of-concept for our new technological approach and demonstrate the system's features for 3D and temporal data exploration. Apart from contributing prototypes with a range of different application scenarios, we added a focus on collaboration and multi-user interaction for our last, more advanced implementations. This chapter describes our applications and their use-cases more thoroughly, mentions specific challenges during implementation and relates each application to its evaluation part in the upcoming chapter.

Two applications focus on collaboration between multiple users.

Tip: All mentioned application examples for midair tangible interaction – including those of the related work, our own implementations and ideas from study participants – can be found using this document's index.

Figure 4.1: Mockup for 3D exploration using augmented reality: Users can move around the tabletop to observe all sides of the virtual cruiser through the device display.

4.1 3D Exploration

Augmented reality offers new ways to interact with 3D models.

Full midair support with six degrees of freedom opens new possibilities for interaction with three-dimensional content. In contrast to the common flat displays that need special interaction techniques to let users explore 3D models from all sides, the introduction of augmented reality frameworks made it possible for users to take a device in their hands and physically move around virtual objects to observe them from all angles. Figure 4.1 shows a mockup for a 3D exploration scenario: looking through the device and moving around the tabletop, users can explore all sides of a virtual cruiser and count the number of lifeboats on both side of the ship.

Combing the tabletop with midair devices for 3D exploration

While the same is possible using 3D glasses, we argue that using an AR device in combination with an interactive tabletop display offers advantages in a number of cases: The large screen is able to show the most important information of a 3D data layer without the need for any device

Figure 4.2: Our simple demo application for 3D exploration: Users have to find Easter eggs that are "hidden" on the tabletop display by centring the "hiding places" in the device camera feed. The image shows a capture of the device screen after all eggs have been found.

(neither smartphone nor 3D glasses). Users then can discuss on the provided visualisation or manipulate the data layer by direct touch interaction to zoom in or to move the selection to another region of interest. Using a midair device, they can temporarily explore the data in the third dimension above the tabletop display. With a tablet-sized device, the observable information content even increases. The literature research (cf. chapter 2 "Related work") provides some application examples for this setup: While the tabletop displays a building's ground plot, the device can be used to search the floor plans of higher levels. Similarly, a city map could be displayed on the tabletop, with the device being used to explore each building's profile. The described system could be generally used for any volumetric data like 3D scans of medical data (e.g. MRIs) or geological content.

A simple application to get a first impression of our system's context.

At an early time in our development process, we implemented a simple 3D exploration application that tasks users to find hidden, virtual Easter eggs on the tabletop with the use of device. For this, certain spots on the displayed scenery (a house with a garden) were taken as markers for the image recognition algorithm of the device's AR framework (cf. section 3.3 "Marker Creation"). When the device's camera is moved to capture one of these regions, a virtual Easter egg appears at its position on the device display and a subtle audio feedback is given. Figure 4.2 shows the different spots with the found Easter eggs that we crafted by using the 3D software Blender[1]. The regions underneath the eggs (the bike, the stairs, the bed, . . .) have been used to train the AR algorithm. Our goal with this demo application was to get a first impression on how users interact with the tabletop and a device. The findings of a small, preliminary study are reported in section 5.2 "Preliminary Study: System Context and General Use". We did not implement more applications that focus on 3D exploration since these already are covered by some of the previous research projects in this field, for example within the publication of the *PaperLens* prototype by [Spindler et al., 2009].

4.2 Temporal Data Exploration

Compare the arctic sea ice development at different points of time on the tabletop and the device.

After having gained first experiences with marker recognition, we continued our implementation by creating an application that should replace the tabletop content when looking through a device. Comparable to the metaphor of magic lenses (cf. section 2.5 "Magic Lenses"), the system consists of a combination of a tabletop display with additional content provided by augmented reality on a handheld device. As scenario, we chose to visualise changes in earth climate and decided to use satellite images showing the development of the arctic sea ice. The high-resolution images have been made publicly available by the NASA Scientific Visualization Studio[2]. The tabletop displays shows the arctic and its ice regions from 1980. When looking through the device camera, the content of the whole tabletop display gets replaced by an image of the arctic's ice expansion of 2012. A first implementation of our app is presented in figure 4.3. It bases on ARKit's image recognition feature and uses parts of the image on

[1] https://www.blender.org
[2] https://svs.gsfc.nasa.gov/4750

Figure 4.3: The iOS app uses parts of the 1980s picture as marker for ARKit's image recognition and then calculates the correct size and position for the overlaying image from 2012.

the tabletop display as markers to provide context for the device. The implementational background has been presented previously in section 3.1 "Recognise the Tabletop in AR".

More ways to interact with temporal data

Conceptually, this application example makes use of the secondary display as comparison tool. Being able to compare images from two different points in time, the system enables users to make assumptions about temporal causality. We will reuse the idea of a comparison tool in another application that will be presented in section 4.4.1 "Collaborative Layer Exploration: Placing Windmills". In their research, [Spindler et al., 2009] identified the temporal information space as one exploration possibility in conjunction with a tabletop display (see section 2.4.1 "The PaperLens Prototype"). With their prototype, users were able to manipulate the playback of a video on the tabletop display by moving the tangible in a vertical direction. We followed this idea and extended our arctic sea ice application with another interaction possibility. To be able to compare the ice development more granularly, we added a slider to the device app, which can be seen in figure 4.4. Similar to video playback controls, users then are able to slide back and forth in time. The left-most position shows the same

Figure 4.4: Sliding through the years, users are able to observe the changes of arctic sea ice more granularly. The background shows the year 1984 with a much bigger ice region than in 2019 (on the device).

year as is been displayed on the tabletop, the right-most position the year 2019 and the positions in-between relate linearly to the corresponding year. When moving the slider slowly to the right, the changes in arctic sea ice are animated fluently. During the implementation of the Easter egg and the arctic applications, we successively improved our tracking framework (cf. section 3.5 “AMT Framework” which allowed us to continue to explore and implement more scenarios for tangible midair interaction. These are summarised in the next sections.

Credits of the used arctic satellite images:
NASA’s Scientific Visualization Studio
Quicklook Arctic Weekly EASE-Grid Sea Ice Age, Version 1: NSIDC/Mark Tschudi, Walter Meier, J Stewart
EASE-Grid Sea Ice Age, Version 4: NSIDC/Mark Tschudi, Walter Meier, J Stewart, Charles Fowler, Jim Maslanik

4.3 Tangible Midair Controller: Flying a Spaceship

Midair tangibles with direct mapping to their virtual counterpart and free movement in all dimensions.

A common use-case for on-surface tangibles is that they physically represent a virtual object within the digital world of the system displayed on the tabletop. Both the virtual object and the physical representation are fully linked to each other: when users move or rotate the tangible, its virtual counterpart behaves correspondingly. For example, the tangible could represent a figurine on a digital chess board [Voelker et al., 2014]. As visualised in the design space of on- and off-surface tangibles (see section 2.4.2 "Other Research on Midair Tangibles" and [Cherek et al., 2019]), regular tangibles only allow interactions with three degrees of freedom: They can be moved in x- and y-direction on the tabletop surface and rotated around their z-axis. With our concept of midair interaction, we would like to combine a direct natural mapping between the off-surface tangible and its virtual link with the increment of interaction space in midair to six degrees of freedom. Midair tangibles can be moved in x-, y- and z-direction and rotated simultaneously around any of these. Having this ability, users could for example lift the device to let their avatar overcome an obstacle.

With the device, users control a spaceship and have to avoid colliding with other spaceships.

Following this idea, we developed an application that lets users control a spaceship (see figure 4.5). The top of the device is linked to the spaceship's front and users can move the spaceship around over the tabletop surface, change its altitude or perform a roll manoeuvre by doing the respective movement with the device. Keeping up with the direct mapping between the two, the virtual spaceship mirrors the device positioning in real-time (cf. section 3.2 "Relate Position of Device and Tabletop using World Tracking"). On the tabletop, other moving spaceships which users have to avoid collisions with are being displayed. They can do so by moving away from the other spaceships or by flying above or underneath them. Furthermore, users are able to eliminate hostile spaceships by firing a laser cannon at them (not visible on the figure). The laser cannon is triggered by tapping on the device screen. To provide users with a challenge, we implemented a simple scoring system

Figure 4.5: We implemented a direct mapping between the device and the virtual spaceship (the smaller one on the right). When moving, tilting or rotating the device, the virtual object behaves accordingly and with delay.

that increments the users' score when destroying another spaceship and decreases when colliding with one.

Relating the altitude of different objects to each other

With the introduction of interaction in the third dimension, the tabletop scene has to provide depth information in order for users to determine the altitude of the displayed objects in relation to each other. We decided to implement shadows that vary in dependence to the object's altitude. A darker shadow close to the object indicates a lower altitude, while a lighter, distributed shadow should signify a larger distance to the ground. We also thought about changing the spaceship's size in relation to its altitude (the higher the spaceship, the bigger its size), but did not implement it. This application has been evaluated as part of a user study to gain, among others, first insights on the perception of depth information on a two-dimensional tabletop display. The study results are summarised in section 5.4 "Study: General System and Software Evaluation".

4.4 Midair Tangibles in Multi-User Scenarios

When collaborating with other users, midair tangibles enable user-specific content on the devices and user-aware content on the tabletop.

The extend and composition of interactive tabletop displays invites multiple users to work together on the table. The large display affords large, detail-rich content which users can discuss on and interact with using touch gestures, digital pens or other input devices (cf. section 2.8 "Multiple Users and Collaboration" of the chapter on related work). Collaboration scenarios can involve a task that is too complicated or too complex for one user and needs input from another pair of eyes. Similarly, users might not be able to immediately see the result of their actions and need a second person to provide feedback. Finally, collaboration enables tasks that need more than two hands to be mastered and make it possible to reach parts of the table that are not reachable by just one person. We implemented two applications that can also be used alone, but are primarily intended to be used by two users. With these applications, we want to gain more insights into user collaboration and how midair tangibles can provide help in this context.

4.4.1 Collaborative Layer Exploration: Placing Windmills

Different layers that influence the location for new windmills

In the first collaborative application that we implemented using midair tangibles, users have to compare different data layers with each other. These can be displayed and switched on the tabletop and independently on the device screens. As concrete use-case, we chose the application area of green energy and tasked users to place new windmills around a city's periphery. Users can create, move (and delete mistakenly placed) windmills using touch on the tabletop display. The windmill positioning depends on three conditions: the existence of enough wind to power the turbines, appropriate distance to housing areas and other buildings so that residents won't complain about the windmills and a required distance between the windmills. The different layer visualisations are shown in figure 4.6. At the initial state, the tabletop display shows a satellite

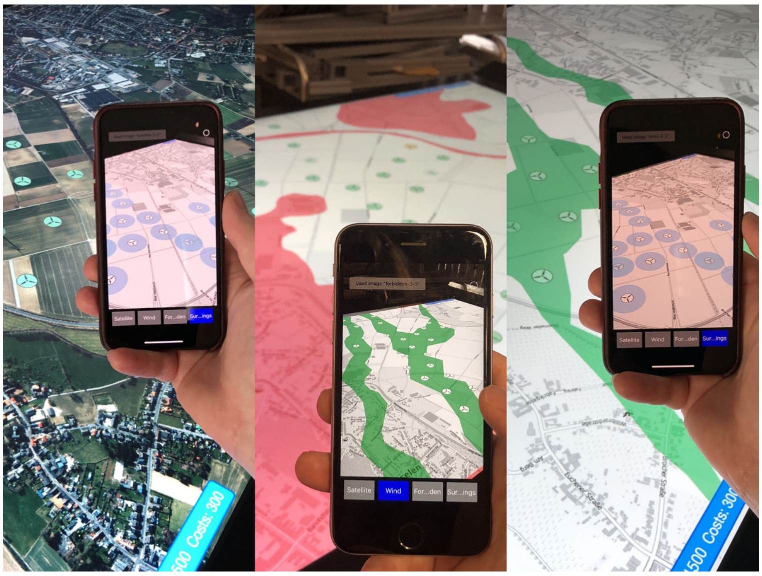


Figure 4.6: Users can explore multiple data layers to decide on windmill positioning by combining the tabletop layer (in the background) with different layers on their devices.

image of the area (background of left image). By tapping on corresponding buttons, users can change the tabletop to display wind regions (in green, background of right image) or forbidden regions (in red, background of centre image). The visualisation of the required surrounding of windmills is visible in blue on the device screen of the left image in figure 4.6, but can also be displayed as a layer on the tabletop.

Users can compare two layers by combining the tabletop view with another layer on the device screen.

The layer selection is mutual – it is not possible to combine multiple layer visualisations into one. To compare two visualisations at the same time, users have to pick up a device and can then overlay the tabletop layer with a second one when looking through the device screen (visible in all three images of figure 4.6). This interaction is similar to the metaphor of magic lenses which has been already described before (cf. section 2.5 “Magic Lenses”). Technically, the system follows the implementation steps to recognise the tabletop and replace the content using augmented reality as explained in section 3.1 “Recognise the Tabletop in AR”. Using the devices to compare different layers with the one displayed on the tabletop display, users have to work together to find spots for new windmills and sometimes could be required to negotiate a compromise which boundaries of a data layer will be crossed. Our system is able to compute a scoring for the chosen windmill positioning and gives feedback by changing the windmills' colour in a traffic lights' scheme and by rotating the propellors with different speeds. We evaluated the interaction design as part of a user study in section 5.4 "Study: General System and Software Evaluation".

Windmill positions are transmitted to the devices.

Side note to a challenge during implementation: After a first iteration, we soon realised that the created windmills are not visible when looking through the device and having chosen to replace the tabletop content with another layer. Therefore, we had to extend the transmission protocol between tabletop and devices to also include the positions of the windmills, so that the devices could display the windmills again on their visualisation layer.

4.4.2 Collaborative Target Selection: Feeding Animals

Addressing the reachability problem

Our second collaborative application focusses on target selection. Users are be able to use their devices as tool to collect elements that are displayed on the tabletop display without the need of direct physical interaction with the tabletop. To demonstrate how midair tangibles can be used to address the reachability problem, we implemented a game-like application that asks users to feed multiple animals. Each side of the tabletop display holds four animals that have to be fed with their particular food items. Underneath each animal, a progress bar indicates the food colour and how much food the animal has eaten (see figure 4.7). The progress bar will decrease over time, so that animals have to be given food within a constant interval. At the other end of the tabletop display, new food items will appear over time. The users' task is to move the food items from one end of the tabletop to the other and distribute them to the animal with the same colour denotation.

Figure 4.7: The device can be used to collect and drop off food items at the cursor position (in the centre of the device screen, next to the chicken). Users control the cursor by moving the device and without direct interaction with the tabletop.

Direct touch vs. device to select elements that are further away

We intentionally designed our system with a large distance between the location of new food items and the target zone with the animals. Users are given the choice to move the food items in a direct way by touching and dragging them over the surface. Since our tabletop provides multi-touch support, users can even handle food items with multiple fingers or hands at once. The second choice is to use the device: We implemented a selection mode that allows users to control a cursor that is being displayed on the tabletop by moving the device (cf. section 3.4 "Device Touch Support"). Users can collect food items in the centre of the cursor by tapping on the device screen. The collected items are stored within the device (visible in figure 4.7 at the bottom of the device screen) and can be dropped off again at the cursor position by a swipe-up gesture on the device. Collected items are stored last-in-first-out with an unlimited capacity. As with the other applications, we implemented a scoring system. Users are given points when animals are being fed and loose points when animals are getting hungry. In a planned user study (see section 5.3 "Intended Studies: User Collaboration and Midair Tangibles"), we'd like to observe how users will solve the task together and whether they will choose to use direct touch or the devices. In a semi-structured interview containing a video demonstration of our application (see section 5.4 "Study: General System and Software Evaluation"), we gathered first insights which interaction users could prefer.

Design iteration on cursor modes

During implementation, we iterated over different touch modes that have been previously described in section 3.4 "Device Touch Support". Our first approach was to use a direct mode where users could select elements by tapping on the corresponding location on the device screen. A small visualisation of the touch point was then shown on the tabletop. Early testers used this small visualisation to eliminate the need of looking through the device screen by constantly tapping on the device and then approaching the touch visualisation to a target by moving the device. Therefore, we implemented the indirect touch mode that is visible in figure 4.7 in a second iteration. The indirect touch mode always shows a prominent cursor on the tabletop display which is located at the intersection of the tabletop surface and the orthogonal projection of the device camera. Elements in the centre of the cursor are being selected when tapping on the device screen at any location. With this approach, users can focus their eyes on the tabletop display and select elements without looking at the device.

Chapter 5

Evaluation

Main contribution: Remote user study to get qualitative feedback for the applications

This chapter presents our different steps to get a first evaluation for the use of midair tangibles as we implemented them. We focus on qualitative feedback for our different use-cases, as we intended to explore the different scenarios and not to get quantitative measurements about their performance in the first step. Our main contribution is a remote user study in which we evaluated the spaceships, windmills and animals applications. The first section summarises early user feedback and interesting results from the first tests, especially regarding ARKit's tracking quality in the context of our implementation. Additionally, we conducted a preliminary study with the Easter egg application to gain first insights into the general use and context of our interactive system. Another focus of our research was the use of midair tangibles for collaboration between multiple users around a tabletop. We planned a corresponding study with sessions of two participants, but were not able to proceed with it on location. The experimental design and other thoughts on this study are archived here, so that the study can be conducted later. For all studies, associated documents are included in the appendix and the digital media library (cf. appendix A "Digital Media Library").

5.1 First Test and Early User Feedback

Early user feedback allowed us to cycle fast through implementation steps.

During implementation, we did some first tests and received early user feedback from students and staff from our lab. Our tabletop is placed in a larger project space that is publicly open to everyone in the lab. As they are walking by, people tend to come over and check out the latest prototypes. In our experience, the system seems to increase people's curiousity, possibly because humans are used to gather around a table, because of the large tabletop's visualisation capabilities, or because of the combination with the iPhones and iPad as midair tangibles. Additionally, we benefit from the low-tech setup that is possible by the use of ARKit which allows bypassers to easily try out a first demo and helps us to get fast feedback for an implementation state in the DIA (design, implement, analyse) cycle. We gathered early user feedback from around ten persons from our lab, all with HCI background, academic background (students, PhDs or higher), predominantly male and including this thesis' supervisor and examiner. Prototypes have also been shown to different groups visiting our lab as part of an introduction to the lab's research. These included the local macOS developer community, participants of a founder's workshop and single visitors from the international HCI community. The received feedback is summarised in the following paragraphs.

The tracking quality is to our satisfaction.

General performance: When coming up with the idea of using augmented reality techniques for midair device tracking, we wondered if the tracking quality would be good and stable enough (cf. chapter 3 "Implementation"). Our first test indicate that this is true, and also that the position updates are fast enough for a natural experience. The position, rotation and tilt of the device show to be sufficiently accurate, especially when holding it still. Users made statements like "That works surprisingly good", especially when being told that the tracking is done by using ARKit. While we observe a satisfying tracking accuracy in general, we experienced not much margin between a good tracking state and a complete loss of tracking, from which user can only recover by re-initialising the AR session. This happens mostly when moving the device with higher speed, or when rotating it by more than 180 degrees around one axis. Similar problems have been reported before by [Leigh et al., 2015]. One user noted to have observed a similar behaviour when using a popular AR app, and could only recover by force-quitting and restarting the app, too.

The device camera is not designed to capture images within close range.

Issues in near-surface area: Apart from fast acceleration, moving the device closer than around 15 cm to the tabletop display results in a decrease of the tracking quality. This has also been reported by [Leigh et al., 2015] and has its origin in the device camera's lack of near-focusing capability and its limited field of view resp. focal length: When being close to the display, the camera is not able to provide ARKit's algorithms for image recognition with sharp, feature-rich images and the tracking quality decreases. [Leigh et al., 2015] solved this by using customised colour patterns instead of feature-based image recognition. Our spaceships application was affected the most by near-surface issues, while the other applications did not demand users to hold the devices closer to the tabletop. As a result, our vision of a combined on- and off-surface tangibles that users can grab from the table to manipulate 3D objects in midair (cf. chapter 6.2 "Future Work") cannot be met with the current implementation. We came up with the idea of placing a trackable marker on the ceiling above the tabletop display. This could improve the tracking capabilities in the near-surface area, but we would have to turn the device around and hence loose the possibility to use the device screen (unfortunately, ARKit does not support world tracking using the front camera). Also, the technical overhead would increase significantly (the position and angle of the marker on the ceiling in relation to the tabletop display would have to be calculated thoroughly).

The tracking depends on sufficient marker images on the tabletop.

Dependence on tabletop background: The tracking quality highly depends on the content displayed on the tabletop display. ARKit is able to detect enough feature points for marker images with contrast-rich, colourful content (cf. section 3.3 "Marker Creation"). With this requirement, the freedom in designing content for the tabletop display is limited. When creating background assets for the tabletop, game designers either have to choose a proper, static background or find a clever way to provide a satisfying marker image for more dynamic scenarios.

Reflections on the tabletop surface affect the tracking.

Lightning and sun: In our tests, tracking quality is influenced by lightning and sun. The display of the tabletop is protected by a glass panel, which reflects ceiling lights and results in artefacts in the device camera feed that influence the tracking quality. We had to switch off the ceiling lights that are placed directly above the tabletop. Also, the tabletop is placed next to a window front, with inside office curtains that, when being closed, still allow light to come through at the edges. This negatively affects the tracking when it's a very sunny day or when the sun is pointing towards the window front, which regularly happens in the afternoons. For our user study and video recordings, we installed a black curtain around the tabletop environment and took some time to adjust two indirect, artificial light sources.

AR tracking causes devices to heat up.

Device issues: As we tried out different devices during our tests (iPhone 6s, iPhone X, iPhone Xs, 6th-generation iPad), we noticed that the newer devices perform better than the older ones. However, we would judge the tracking quality of the five years old iPhone 6s (running on iOS 13) still as ok. More of concern is the heat production: The devices tend to warm up quickly, caused by the high power consumption of the AR tracking. When using the device for a longer period, this could get discomfortable for users and certainly will affect battery status. As a rough estimate, we observed 20 percent of power consumption over an hour use of the device.

5.2 Preliminary Study: System Context and General Use

Study task: Finding Easter eggs

Next to our continuous, informal first tests that have been described in the previous section, we did a short study at a very early time in the development process. By observing users with an early prototype of our system, we hoped to gain basic insights into the system context and general use. We wanted to get a first impression how users behave when interacting with the tabletop using a device and identify interaction challenges that should be considered during implementation. For this preliminary study that took place a week before the Easter holidays, we asked participants to use our first 3D exploration application and find virtual Easter eggs that were "hidden" on the tabletop surface (cf. section 4.1 "3D Exploration"). Seven students from our lab (six male, one female) volunteered to help.

Arm posture and device orientation

General observations: Most participants performed well, some even noted that finding the Easter eggs was too easy. Some tapped onto the screen when an egg appeared, expecting that something more would happen, which it didn't. Most participants hold their arm static (with a 90 degree angle between forearm and upper arm) and moved along one side of the table to find the eggs. Only one participant stretched an arm to reach for the table's edges and moved less. Five of seven participants hold the phone in portrait mode, as it was handed over, and two in landscape mode, which fit better to the table's dimensions. All participants were using one hand to hold the phone, some tapped on the phone's screen with the other hand. Three participants moved the phone closer (around 30cm) to the table's screen to inspect an egg after it has been found. Two participants tried to fit the whole table onto the screen and hence moved away from the table for around 1.5m to try to catch all eggs at once. One of these two then hit a chair when moving backwards away from the table. It took the app longer to display the virtual egg on screen than to play the sound when an egg was found. Some participants were confused by this; two even moved the phone so fast that they did not understand what happened, because the egg was already out of scope and they only heart the sound.

AR decreases users' awareness of their surrounding.

Discussion: Although our preliminary study is not representative, we try to summarise some tendencies from our observations. It seems that users prefer having a comfortable arm position and move their body around instead of stretching their arm too much. More importantly, when focussing on the device, users tend to loose awareness of their surrounding, especially in augmented reality. Similar observations have been reported by [Chan et al., 2010].

For our setup, this underlines the importance to free the space around the table from any physical obstacles that users could stumble upon. In the future, we could also start thinking of other ways to help users to process their environment, and reduce the time that is needed to extract the information displayed on the device. A possible approach could be to provide a way of accessing the information again once it has been gathered using AR without the need of moving around, similar to the freezing mechanisms presented in the *PaperLens* prototype by [Spindler et al., 2009].

5.3 Intended Studies: User Collaboration and Midair Tangibles

A single-user and a multi-user study have been planned.

To evaluate our approach of midair tangible interactions and the different prototypes that we implemented, we planned to conduct two user studies. One study was to focus on the direct mapping between the device and a virtual object based on our implementation of the spaceships application. For the second study, we planned to expand our focus of interest to multi-user scenarios. Combining our midair tangibles with user collaboration, we hoped to get insights on how devices can help with the reachability problem on tabletops and how users handle specific content on their devices when working together. Consequently, we planned to invite two participants to each session. The second study based on the application scenarios for creating windmills and feeding animals that were presented in the previous chapter.

Switching to a remote study to eliminate physical contact

Unfortunately, the 2020 global COVID-19 pandemic prohibited us to conduct our studies on-location as planned. Looking for alternatives to evaluate our work, we decided instead to gather feedback from presenting video-recordings and conducting semi-structured interviews using a video-conference system. These will be discussed in the upcoming section 5.4 "Study: General System and Software Evaluation". While we are aware that this compromise will not yield the results that were to be expected from users that could try out the system with their own hands, the video study proved to able to provide us with decent general feedback on our interaction design and on the software that we implemented.

The following pages summarise the research questions, hypotheses and experimental design for our planned studies. Together with the original documents (questionnaires, forms, observation protocols, . . .) in the appendix, all necessary information should be given to conduct the studies at a later point in time.

The original study can be conducted later.

5.3.1 User Study on Direct Midair Interaction

With our first user study, we planned to retrieve insights on user perception of midair tangibles with a direct mapping to its virtual counterpart. We expect participants to adapt quickly to the system since the approach follows a natural mapping between the physical and the virtual object. Within the study, participants are asked to control a spaceship using the device which can move in all dimension. Section 4.3 "Tangible Midair Controller: Flying a Spaceship" presents the application and also contains more information on the conception of the interaction design.

Study focus: User perception of midair tangibles with direct mapping to a virtual object

After being introduced to the system, the participants' task will be to gain as many points as possible in a given timeframe. They can win points by shooting at other spaceships and loose points if they collide with them. Collisions can be avoided by changing the spaceship's course as well as by changing its altitude, which is easily achievable by lifting the device up or down above the tabletop surface. Over time, it gets more difficult to avoid collisions because more and more spaceships will appear. We expect the whole experiment to occupy 30 minutes of the participants' time.

Study task: Flying a spaceship

Research questions:
How will users perceive a direct mapping between a virtual 3D object and a virtual tangible with six degrees of freedom? How will users interpret height visualisations on the 2D tabletop display?

Hypotheses:
Users quickly learn the conceptual mapping between the midair device and the linked virtual object. The chosen height visualisation using shadows will help users to differentiate altitude between the displayed spaceships.

Hardware setup and surrounding:
The main visualisation is displayed on our Microsoft Surface Hub 84, whose technical specifications are presented in section 3 "Implementation". There is no touch interaction with the tabletop display in this scenario. The users control a spaceship that is displayed on the tabletop display with an iPhone X that is connected to wifi. Both the tabletop and the device run the spaceships application based on our swift implementation of the ARKit-powered midair tangible framework. The area around the tabletop has been cut off from the public space by curtains to prevent distractions from outer factors. The space around the tabletop has been freed of obstacles and allows participants to move around three sides of the tabletop (cf. section 5.2 "Preliminary Study: System Context and General Use"). The lightning in the study area has been adapted to prevent reflections on the glass surface of the tabletop (cf. section 5.1 "First Test and Early User Feedback"). Additionally, a wide-angle camera (GoPro Hero 3) has been installed on a stand to record the tabletop and the user.

Measure:
Qualitative feedback will be gained from observation notes from the study conductor during the study, by analysing video and audio from the study recordings and from a questionnaire that users will be asked to fill out at the end of their participation.

Experimental procedure:
Before starting the session, the video recording system has to be checked. Also, the applications on both tabletop and the device have to be ready. The device has to be charged and restarted; the tabletop surface and the devices have to be cleaned. After welcoming the participants, they can be given an overview of the study by following the consent form. After having signed the consent form, participants have to answer some questions on demographics and be

asked for approval of the video recordings.
Then, they should be introduced to the system and be shown the capabilities of the device interaction. After having familiarised themselves with the system for some minutes, the participants should try to gain as many points as possible within five minutes. During that time, the study conductor can take notes of interesting observations. At the end, participants can be asked for any general feedback on the system or the study itself. Finally and after being expressed gratitude, they are free to go.

5.3.2 Study on User Collaboration

Study focus: Midair tangibles in multi-user scenarios

We planned a second user study with two of our applications that focusses on midair tangibles and multi-user scenarios. As a tabletop affords multiple users around it, we want to observe how users will collaborate with the use of midair tangibles to solve a complex layer exploration task and how midair tangibles can be used to select elements on the tabletop that are out of their physical reach. More conceptual background for the scenarios are given in the introduction of section 4.4.

We gather qualitative feedback.

Both parts of the study are done in one session and should take 45 to 60 minutes in total. The experiment is split into two rounds, one with the windmills and the other with the animals application. In contrast to the study using the spaceships application, we explicitly will not pressure participants to solve their tasks as fast as possible. Instead, we want to give them time to try out different aspects of the systems and find out which strategies and interactions work best for them. As before, we focus on user observations and attempt to gather qualitative feedback from our participants. The following paragraphs summarise the study focus and experimental design for the whole experiment, starting with common parts and then continuing with the specifics for each use-case. The appendix contains the original forms, questionnaires, documents and protocol templates that are mentioned throughout the following descriptions.

Hardware setup and surrounding:
The setup is the same as explained before for the spaceships study – except that two iPhone X are available (one for each participant).

Measure:
During both experiments, qualitative feedback will be gathered primarily from user observations. The study conductor is tasked to write down notes during the study or to mark the time of interesting actions. By analysing video and audio from the study recordings, detailed observations can be summarised and compared to each other. Additional feedback will be provided by a questionnaire that users will be asked to fill out at the end of their participation. Also, participants are asked to give individual comments after each round.

Experimental procedure (start and end of session):
Before starting the session, some preparations have to be done. The video recording system has to be checked if the camera is charged, has enough storage capacity and if the angle is still in its correct position. Also, the applications on both tabletop and devices have to be ready. The devices have to be charged and restarted; the tabletop surface and the devices have to be cleaned. After welcoming the participants, they can be given an overview of the study by following the consent form. After having signed the consent form, participants have to answer some questions on demographics and how close they are to their teammates. Being asked for approval of the video recordings, the participants then should stand on the opposite sides of the tabletop. It could be a good idea to emphasise that their should not feel to be examined and are free to talk to each other.
Now, the actual experiments can start (see below). After the two rounds, participants can be asked for any general feedback on the system or the study itself. Finally and after being expressed gratitude, they are free to go.

Layer Exploration: Windmills

In the first round, participants will work on the windmill application (see section 4.4.1 "Collaborative Layer Exploration: Placing Windmills") and are given the task to find spots for new windmills that are limited by wind, housing areas and the windmills' surroundings while achieving the best possible scoring.

Research questions:
How can midair tangibles help users to explore different data layers? How do users combine different data layer views on midair tangibles and a tabletop display? How do users share user-specific content with each other?

Observations goals:
How do they use the system: Each one using the device? One with the tabletop and one with a device? Both with the tabletop? Comparing the tabletop layer with a device layer?
How do they work together: Both on their own in their own region? Discussing about a positioning when looking on a tabletop layer? Discussing about a positioning when one person refers to something visible on the device (so that the other person does not see the referenced area)? Showing the device to the other person or handing it over?

Hypothesis:
Users will collaborate and choose to distribute the three layers needed to make decisions on the three screens (tabletop and both devices).

Experimental procedure:
After the welcome part (see above), the application on the tabletop should be started and the participants should be given an introduction. During the introduction, the system and the tasks have to be explained. After being asked if they are ready, the participants should start their task. During their performance, first observations notes can be written down. After around fifteen minutes, they should be asked to stop. Then, the participants should fill out the questionnaire and should being asked if they have any-

thing else to add or discuss. Next, the experiment continues with the animals application.

Target Selection: Animals

The second round bases on the animal application (see section 4.4.2 “Collaborative Target Selection: Feeding Animals”). The participants’ task will be to feed animals either by making use of the tabletop’s multi-touch capabilities or by using midair devices to collect and drop off food items. Again, they are encouraged to get a score as high as possible.

Research question:
How can midair tangibles help users to reach targets that are further away?

Observations goals:
How do they use the system: Primarily using touch or using the devices? How difficult is it for users to control the cursor with the device? How much do they move when using touch and when using the devices?
How do they work together: Which strategy do they apply to solve the task? Will one person use touch and the other the device? Will they each possess one device and move along to feed the animals? Will one person be responsible to move the food items close the animals and the other person to distribute them to the animals?

Hypothesis:
Users will learn that the device interaction is faster, will choose to both use their devices and split the task by the sides of the table.

Experimental procedure:
After the welcome part and the first round of the study with the windmills scenario (see above), participants are presented with the second application. The procedure is basically the same as before: after being introduced and explained the tasks, the participants have about fifteen minutes to try out the system and feed the animals as good as

possible, while first observations can be noted down. After their performance, the participants should be asked to fill out the questionnaire and encouraged to give other feedback. Then, the study is concluded as described above.

5.4 Study: General System and Software Evaluation

Participants were asked to share qualitative feedback remotely.

To get qualitative feedback for the spaceships, windmills and animals application, we conducted a remote user study. Eleven participants took part from their homes. For each application, we showed a short video and asked the participants to describe what they see, followed by a semi-structured interview to learn how they perceived the current application. Each part took 10 to 15 minutes, so we ended up with 45 minutes altogether. After the study, we analysed the recordings of each session to find similar statements and summarise general feedback. The following paragraphs describe the study in detail. First, more information on the participants is given, followed by the experimental design for all rounds. Then, divided into the applications, research questions, hypothesis and results are presented. Each part concludes with a short, application-specific discussion, and the section closes with a general discussion of the study. The observation goals and questions defined in the previous sections for the study on location served as a guideline for the interviews. Additional documents (study protocol and playbook) are included in the appendix (cf. B “User Study Documents”). The videos and codes from the analysis are archived in the digital media library (cf. A “Digital Media Library”).

Participants:
Eleven participants (six male, five female) with a mean age of 26 (21-29) took part in the study. To decouple personal data from study data, each participant was given a random number that was used to identify participants during analysis and in the summaries below. Four participants were studying computer science, three were students in other fields, four participants were employees. Seven of

the eleven participants had experience in computer science or related fields. Being asked for previous experience with an interactive tabletop system, three responded positively. Due to the remote setup, participants were located in different cities in Germany.

5.4.1 Experimental Design

Hardware setup and surrounding:
The study was conducted remotely using the video-conference software Zoom, which allowed participants to take part in the study from home. Consequently, participation was only possible with a computer and a connected microphone (sharing video was voluntary) as well as a stable internet connection. Two interviews had to be rescheduled due to current server overload of the Zoom software. At the beginning of the interview, participants were asked to mute notifications on their devices, so that the session would not be interrupted.

Experimental procedure:
Each session consisted of an introduction, three rounds of video and semi-structured interview for the spaceships, animals and windmills application and the possibility to give general feedback at the end. All interviews were conducted in German. In the introduction, participants were asked permission to record their voice and – voluntarily – their image. They were given an overview of the session and purpose, procedure, risks and confidentiality of their statements were explained to them. After that, participants were asked for demographical data. Before starting with the first video, participants were told that the videos were recorded in the users' perspective and we encouraged the participants to speak aloud and describe what they see in the videos. Then, the rounds with each application in the order described above started. For few participants, we paused the video to give them time to finish a longer statement, or encouraged them to speak more. For each application, we prepared some questions in alignment with our observation goals beforehand that were used as a base for the semi-structured interviews after the videos. Generally, we

tried to keep the interviews as fluid as possible and to retain a natural conversation. After the three rounds, participants were asked if they wanted to provide any other comments or feedback. After stopping the recording, they were thanked for their participation.

Measure:
To analyse qualitative feedback from the participants, we recorded their voice when they described what they saw in the videos and during the semi-structured interviews and took notes during the session. For each application, we created coding tables that were filled with a few expectable observations beforehand and (primarily) by listening to the recordings. The German-conducted interviews were translated directly into English language in that process. We used the observation notes as first measure during the coding. For about the first third, the code creation was done shortly after the interview. For the following two-thirds, we extracted the codes from the recordings after all sessions were conducted on a per-application basis. As comparison measure, we also gathered demographic data from our participants.

5.4.2 Direct Mapping: Spaceships

Based on the spaceships application that has been presented in section 4.3 "Tangible Midair Controller: Flying a Spaceship", the video shows – from the users' point of view – how a spaceship is controlled by the movement of a device. The virtual spaceship has a direct mapping to the tangible. The video is about two and half minutes long and contains scenes in which the user controls the spaceship to fly above and underneath other spaceships and fires a laser cannon at them.

Research questions:
How will users perceive a direct mapping between a virtual 3D object and a virtual tangible with six degrees of freedom? How will users interpret height visualisations on the 2D tabletop display?

Hypotheses:
Users understand from the video that the spaceship is controlled by moving the device. Users understand from the video that they can fly above and underneath virtual objects. Users wonder that the controlled spaceship flew backwards or sidewards.

Results

The following paragraphs summarise the results from the coding process and contain analysis of what participants stated during video playback as well as during the semi-structured interview afterwards. The summarise is divided into paragraphs on direct midair control, altitude perception and physics perception and finishes with user-contributed suggestions on how to use the device display.

The direct mapping was always recognised.

About Direct Midair Tangible Control:
All participants stated that the spaceship is controlled by the device. Some (4 of 11) explicitly noted that the spaceship can move in all dimensions that the device offers. A few (2 of 11) participants compared the device control to touch (without being asked) and stated that controlling the spaceship using touch the same way as with the device would not be possible.

Altitude perception was challenging.

On Altitude Perception and the Third Dimension:
Many participants had difficulties with the altitude perception and how the spaceship's altitude is controlled. All recognised that the controlled spaceship can fly above and underneath another one, but not all of them (10 of 11) mapped this to the raising and lowering of the device. About half of the participants (5 of 11) stated that they have difficulties estimating the controlled spaceships altitude in comparison with the other spaceships. A few (3 of 11) wondered why the spaceship's size did not change with relation to the altitude (the higher the spaceship, the bigger its size). However, the shadow of the controlled spaceship seems to help participants (5 of 11) to get a sense for the third dimension. One participant even suggested to explicitly improve the shadow rendering.

On In-Game Physics and Actual Device Movement:
When controlling the spaceship with the device, the movement is not limited to in-game physics, but only depends on how the user actually moves the device. This allows for rapid acceleration, stopping or abrupt direction changes which would not be possible with real-world physics. Surprisingly, only 5 of 11 participants reported they saw backwards or sidewards movement when being asked about if something were off with the spaceship movement. In comparison with a controlled airplane or a car, which physics users are expected to know pretty well, the system seems to benefit from the fictional origin of the spaceship and that users do not seem to expect a realistic movement from it. Two participants gave a hint for this by saying: "I do not know the physics of a spaceship too much - it can possibly move backwards or sidewards anyways." Other participants (3 of 11) noted the difference between the game world and the real world and expect the device - if a more realistic spaceship movement is being wished for - to be moved accordingly. On the other hand, two participants stated the spaceship does not appear as an object inside the game, since it does not respond the game's physics. Also, four of eleven participants were confused that nothing happened when spaceships collide, or expected something to happen. One participant had a suggestion to circumvent this issue: instead of directly controlling the spaceship, the device could control a virtual target to which the spaceship flies, so that the spaceship responds to in-game physics on its way to the target.

Violation of in-world physics was tolerated.

About the content on the device display:
Some minor statements and ideas came up for the content of the device display, in which we did not put much effort (the display only showed the live camera feed). One participant suggested to show a view from the spaceship's cockpit, i.e. what one could see when actually sitting inside the cockpit and controlling the spaceship. This would then allow to check the altitude of the other spaceships. The display could also be used for collision detection and show a hint when approaching another spaceship on a collision course.

Suggestion: Altitude visualisation on device display

Discussion

Results suggest a strong cognitive mapping between tangible and virtual mapping.

The study results indicate that our hypotheses, that users understand that the spaceship is controlled by moving the device and that they can fly above and underneath virtual objects can be confirmed. This is backed up by the corresponding statements from all participants. On the hypothesis on the direct mapping between device and virtual representation (that users wonder that the controlled spaceship flew backwards or sidewards), the results indicate similar findings to regular tangibles that can be moved on the surface. Users seem to build a strong mapping between the tangible and the virtual object it represents. Some participants tend to apply the rules of the virtual world to the tangible, even though the physical world behaves differently and would allow other movement. However, our results show that this is not true for all participants.

More work has to be done on altitude visualisation.

As a main finding, we identified the visualisation of the newly introduced third dimension – in our application the altitude of the spaceships – as a key challenge for midair tangible interaction with a direct mapping between tangible and virtual object. Our first approach using shadows seems to lead in a good direction, but we suggest experimenting with more ideas to map the z-axis onto the flat display in future prototypes.

5.4.3 Target Selection: Animals

The video contains a scenario with animals that have to be fed, which as been described before in section 4.4.2 "Collaborative Target Selection: Feeding Animals". In the first part of the video, this is done by using direct touch, in the second part, with the help of a device. The device can be used for interaction with objects that are further away. The video is about three minutes long and shows – from the users' point of view – how animals are fed with the right food item, what happens when selecting the wrong colour and that animals can only be given a certain amount of food until they are hungry again.

Research questions:
How can midair tangibles help users to reach targets that are further away? Would users initially prefer touch or device interaction for target selection?

Hypothesis:
Users understand from the video that they can use the device to reach targets that are further away.

Results

The following paragraphs summarise the results from the coding process and contain an analysis of what participants stated during video playback as well as during the semi-structured interview afterwards. The results are structured into three categories, starting with game logic and software evaluation, continuing with the comparison of touch and device interaction and finishing with findings on strategies for collaboration.

The order of the food items on the device screen was unclear.

On the Game Logic and Software Evaluation:
Evaluating the statements concerning the game logic and other software aspects, most participants seem to understand what is going on pretty well. All participants observed that animals only do eat food in their own colour, and that they do not eat more food when the indicator bar is full. About the indicator bars, most participants explicitly stated that it increases when feeding an animal (7 of 11), and decreases over time (8 of 11). P2 suggested that the amount of food that an animal ate could be visualised with circles instead of the bar, so that it would be easier to guess how many food items the animals still needs until it is satisfied. Most (9 of 11) participants emphasised on the multi-touch capabilities of the tabletop ("It is possible to feed multiple animals and with multiple food items using touch.") and all participants identified the device to be able to collect and drop off food items. Seven of eleven participants suspected the device to have an unlimited capacity to store food items, which it does. Uncertainty arose on the display of the food items on the device screen. Five participants recognised that the last food item that was selected is

the first one that will be dropped off. P0 and P6 suggested that on the device screen, the collected food items should be displayed in a vertical manner instead of horizontally, so that new items will be pushed on top of the stack, and the older ones are at the bottom. This mapping would also be better supported by the swipe-up gesture, as P0 added. P2 wished that the selected food item would be displayed more prominently in the centre of the device screen.

Initial preference for touch, but device is more capable.

Comparing Touch and Indirect Pointing Using the Device:
When being asked to compare the touch input with the indirect pointing using the device to move food items to the animals, the participants gave a variety of arguments for both interaction techniques. When using the device, most participants (9 of 11) stated that they would not have to move that much than with touch input. Some participants (4 of 11) added that they would to have reach further with touch than with the device, or that they would use the device to select targets that are further away. P1 suggested that the device could help people with shorter arms or height to reach the targets. Another point that immediately was contributed by nine of the eleven participants was that the device is able to hold more food items than it would be possible using their fingers. Interestingly just by looking at the videos (that only show one user), a few participants recognised that when working with multiple users on the tabletop, another user's hand could be in the way when selecting food items using touch. Three participants, one of them without an HCI background, stated that "The device prevents that users interfere with each other when tapping on items close to each other". When working with devices and touch in a mixed setting, two participants stated that "Another user's hands will occlude the device cursor." When being asked which interaction the participants suspect to be faster, the response is divided into half: Six of eleven participants suspected to be faster with the device, five participants suspected to be (initially) faster with touch. Four of the latter explained their preference for touch by being more familiar with this interaction technique. Again four of the eleven participants (some that suspected to be faster with touch, some that did not) stated to feel a more direct connection to the game with touch than with the device. The lesser movement with the

device is not judged as an advantage by all participants: interestingly, P4, P8 and P9 stated that they would move more using touch in a positive way. Other arguments for touch interaction included that they suspected direct touch to be more precise than the device cursor, since it seemed more difficult to hold the device still (4 of 11). Two participants preferred touch because they stated to need two steps (select and drop) to move a food item to an animal with the device vs. one step for direct touch. One participant would have used touch for close food items, only.

Strategies for Collaboration:
Besides the participant's preference for touch or device input, they were asked for ideas when feeding the animals together with another users. Half of the participants (6 of 11) thought about dividing the task and let each user feed the animals on their side. Two participants came up with the idea that one user would collect food items with the device and drop them off next to the animals, and the other user would then distribute the food items to the animals using touch. A few participants (3 of 11) wondered if it would be best to try to pick up the food items in a specific colour order to be faster when feeding the animals or to first select the food items randomly as fast as possible and then check the colours later. They came to no conclusion and wished to actually try out the system on-location for this. P2 noted that the device cursors would be difficult to separate from each other in a multi-user scenario, since the cursors both would have the same colour, and suggested the possibility to personalise them. P0 suggested that, when working together, the collected items could be synced between the device to be even faster.

Discussion

For this application, we started with the hypothesis that users understand from the video that they can use the device to reach targets that are further away. Based on our results, we can confirm this expectation, since most users seem to understand that they do not have to move that much with the device. When focussing more on the com-

Device pointing proofed to be useful to reach targets that are out of reach.

parison between touch and indirect device interaction however, some participants stated to initially prefer touch. This suggests that device pointing could be a good use when touch is not possible, or when many users are working on the tabletop.

5.4.4 Layer Exploration: Windmills

The video shows the application to position windmills that as been introduced in section 4.4.1 “Collaborative Layer Exploration: Placing Windmills”. With the help of the system, users are enabled to find good spots for new windmills, since possible locations are bound by wind, housing and proximity to other windmills. The video is about four minutes long and shows the creation process of new windmills, how they can be moved around and how their position can be evaluated by the system. Switching the current layer visualisation is shown, as well as using the device to explore different layers than the one currently visible no the tabletop. Then, the video demonstrates how the devices and the tabletop can be used in conjunction to correct bad placements for windmills.

Research questions:
How can midair tangibles help users to explore different data layers? What strategies are possible to combine tabletop and midair interaction?

Hypothesis:
Users understand from the video that they can use the device as second screen to find good windmill positions.

Results

The results from the coding process and the analysis of what participants stated during video playback as well as during the semi-structured interview are summarised in the following paragraphs. The summary starts with participants’ statements on the combination of tabletop and device for layer exploration, continues with feedback on the application task and finishes with statements regarding collaboration and more ideas for use-cases provided by the participants.

On Layer Exploration Using the Device and the Tabletop:
When viewing the video, all participants noted that the device can show another layer than the one displayed on the tabletop, and that with the use of the device, it is possible to compare a windmill position on two layers with the tabletop screen and the device simultaneously. A few participants wished to merge the layers together: P2 and P6 would have liked the different maps merged into one layer on the tabletop; P2 alternatively suggested to have only the surroundings visible on all maps; P6 alternatively suggested that the device could display a merged layer that consists of the one displayed on the tabletop and the one selected on the device. On the other hand, P8 stated as a benefit that the device offered a clear view of the selected layer, in case the tabletop layers were more complicated or difficult to perceive. P10 did a similar statement by saying to have a smaller, focussed area on the device. Two participants suspected merging all layers into one on the tabletop would be too challenging to perceive. Due to technical limitations when replacing the tabletop content with virtual content on the device screen, the user's hand is not visible when looking through the device. Three participants noted this when looking at the videos.

Device and tabletop can be used to compare different layers.

Feedback on the Application Tasks:
On the task itself (finding locations for windmills by comparing the different layers), all participants realised (some sooner, some later) that the current placement gets evaluated when tapping a button on the tabletop surface, and what the conditions (wind layer in green, forbidden layer in red, surroundings in blue) are. One participant suggested to update the scoring when moving the windmills. Three participants valued the device as beneficial when trying to change the layers, because they wouldn't have to reach to the buttons on the tabletop and could change them on the device instead, since the buttons were located in the corner of the tabletop screen.

The scenario inspired users to come up with other examples.

Multi-User and More Applications Ideas:
Most participants picked up the thought of using the system together with another person and a second device, and came up with the strategy to distribute the three layers to the two devices and the tabletop display, so that each layer is visible at the same time (9 of 11). Inspired by the video, four participants explicitly stated that in a multi-user scenario, users could have their own specific content on the device while the tabletop would show information relevant to all users. P8 extended this idea when stating that the device could help to input user-based content on the tabletop without reaching to the specific area, which could possibly be occupied by another user. P2 and P9 had a similar thought and suggested that they would have liked to be able to move the windmills not just by tapping on the tabletop display, but by using the device display in the same manner, too. In comparison to the other apps and previous experiences (participants were shown the spaceships and animals applications before), six participants stated that this application scenario were a good motivation to use the tabletop and a device. Some shared more application ideas: P0 and P9 could imagine the system for urban or agricultural planning. P8 suggested the use for planning a building construction site, where the tabletop shows the structural design and representatives from each department can use a device to check the critical parts of the building plans for their domain, i.e. electrical wiring, water pipes etc.

Discussion

Study results suggest the use of midair devices for user-specific content.

This application scenario received the most positive feedback for the combination of a tabletop display and device interaction. Our hypothesis that users understand from the video that they can use the device as second screen to find good windmill positions can be confirmed by the study results. While some participants suggested to merge all visible layers, others judged this approach to be too challenging to perceive. We think that the possibility to provide user-specific content in addition to the shared content on the tabletop can be especially helpful for multi-user scenarios or complex tasks.

5.4.5 General Discussion

Key findings of the study

Based on the studies of the spaceships, animals and windmills applications, we could gather various qualitative feedback on the use of midair tangibles. The results are of a more general nature and cannot reveal the same profoundness as a study on-location, where participants could hold the devices in their hands, but proofed to give a good first impression of the designed interactions. The previous paragraphs summarised the results in details and contained short discussions on each application scenario. The key findings of our experiments are:

1. Users have a strong direct mapping between tangible and virtual object, but visualisation of the newly introduced third dimension is a challenge on the flat tabletop display.

2. Midair tangibles show to be suitable to select for targets that are out of reach. However, direct touch interaction tends to be more intuitive.

3. Especially in multi-user scenarios, the use of midair devices for user-specific content gets high approval by our participants. The magic lense metaphor (cf. section 2.5 "Magic Lenses") shows to be suitable.

Lessons learned from conducting a study remotely

Additionally, we'd like to share some lessons learned from doing a remote study instead of one on location. We think that a remote study can be a good approach to get fast, initial feedback on an interactive system from users outside the development group. A remote study can be faster (in our case 10-15 minutes per scenario instead of 30 minutes per application on location) and participants do not have to come to the lab. This can make it easier to find participants, since they do not have to be from the same city or country. Also, feedback from participants with a different cultural background becomes possible. Is to be noted though that user feedback is not that profound as on location, but can give a good sense of the general direction and usability of a system.

Chapter 6

Summary and Future Work

This thesis presented how midair tangibles can be used to extend the interaction space of tangibles into the third dimension above an interactive tabletop. By exploring and evaluating different use-cases of midair tangible interaction, we showed multiple solutions, challenges and opportunities for this novel interaction setup using augmented reality technologies as basis for the device tracking. The following paragraphs summarise the presented findings on related work, the implemented framework and applications and their evaluation. It finishes with a collection of future work in this research area.

We extended tangible interaction to the space above the tabletop.

6.1 Summary and Contributions

Not many publications include midair tangibles that are tracked in the space above the tabletop and allow movement and rotations in all dimensions. One of these is the *PaperLens* prototype [Spindler et al., 2009] which we presented in section 2.4.1 “The PaperLens Prototype”. However, many related fields cover parts of midair tangible interaction. Our literature research includes on- and near-surface tangibles and multi-layer interaction above inter-

The literature research revealed few publications but many related fields.

active tabletops. Further learnings can be found in publications on magic lenses, cross-device interaction and augmented reality and have been summarised in chapter 2.

We implemented a reusable framework that uses AR to track midair tangibles.

The main contribution consists in the development of a cross-platform framework to track midair tangibles using augmented reality technologies (cf. chapter 3 "Implementation"). To establish physical awareness between an interactive tabletop and midair devices, we use the world tracking and image recognition features of Apple's augmented reality framework. ARKit provides us with a location vector between the device and a marker of predefined size and position that is displayed on the tabletop screen. Additionally, it uses the device sensors to re-create a virtual image of its surrounding. We can then transmit this location data to the system running the content on the tabletop, and are able to compute the exact device location with the knowledge of the marker position. Concerning tracking quality, our tests show that the general performance of our system is good and even exceeded our initial expectation. ARKit's tracking capabilities though highly depend on suitable, feature-rich content on the tabletop.

Five applications to explore different use-cases for midair tangible interaction

Based on the implemented framework, we created five applications to explore different use-cases for midair tangible interaction. The first two applications served as proof of concept to recognise the tabletop in AR. Users can search for Easter eggs that were "hidden" within the scenery on the tabletop display or use the arctic app to replace an image of arctic sea ice displayed on the tabletop with one from another point in time, just like a magic lense. The other three applications provided more sophisticated use-cases and were evaluated in a remote user study (see next paragraph). With the spaceships application (section 4.3 "Tangible Midair Controller: Flying a Spaceship"), we implemented a direct mapping between the tangible and its virtual representation just as it is a common use-case for regular tangibles, but extended it into the third dimension. We used shadows to visualise the altitude of the spaceships. The windmills application (section 4.4.1 "Collaborative Layer Exploration: Placing Windmills") makes use of midair tangibles to allow users to compare different restrictions to build new windmills around the city of Würselen.

With this scenario, we focus on midair tangibles as secondary screens that enable users to make decisions and can provide user-specific content in multi-user settings. Finally, in the animals application (section 4.4.2 "Collaborative Target Selection: Feeding Animals", the device is being used to select targets that out of physical reach for the user. For this, we implemented a simple game where users have to feed animals with the correct food either by using touch on the tabletop or the devices.

We conducted a remote user study to get qualitative feedback

To get qualitative feedback for the spaceships, windmills and animals applications, we conducted a remote user study. Eleven participants took part from their home. For each application, we showed them a video and asked the participants to describe aloud what they see, followed by a semi-structured interview to learn how they perceived the system. After the study, we analysed the recordings of the sessions to find similar statements and summarise general feedback. Our study results indicate that the visualisation of the newly introduced dimension on the tabletop (for example: the altitude of the spaceships) is a key challenge in such scenarios. For the windmills applications, many participants favoured this combination of tabletop and device interaction. The difference between the touch and device interaction in the animals application revealed that even though many participants stated to not have to move that much when using the device, some initially preferred the touch interaction.

Study design for midair tangibles with user collaboration

Initially, we also wanted to focus more on how midair tangibles can help users to collaborate which each other. We contribute a study design for sessions with two users on location using the windmills and the animals application, which can be conducted in the future.

6.2 Future Work

Apart from study on location, which will hopefully be conducted by our lab in the near future, we present future work and ideas from our implementation and first tests, but also from participants from our study. Some improvements for different implementation aspects have already been presented directly in the corresponding sections and can be reviewed their for further context.

Tracking loss in the near-surface areas

Our initial vision with a floating experience from on-surface to midair tangibles that can be picked up from the tabletop and then used for 3D interaction has not fully been met. Especially in the near-surface area, the tracking quality of our AR-based system decreases because the iPhones' cameras are not able to focus on such close targets. Hence, ARKit is not able to recognise enough feature points to provide a satisfying tracking. To improve the situation here, we came up with two ideas: First, it could be worth a try to combine the AR tracking with the tracking of regular tangibles using the capacitive display of the tabletop. For this, capacitive pads would have to be attached to the lower sides of the devices, so that they can be tracked on the surface in a passive way. For the spaceships application, this would mean that users would be able to reliably land the controlled spaceship on the surface. Second, as the display is not needed for the spaceships scenario with a direct mapping between virtual and physical object, we thought of flipping the iPhone around and attaching a marker on the ceiling. With a marker of suitable size, we expect the tracking to improve in the near surface area, and it would even allow to land the spaceship. This setup would though come with the cost of a more complicated hardware installation, since size, position and angles of the marker at the ceiling would have to be determined with precision to compute the devices' positions in relation to the tabletop. Of course, a combination of both approaches would be possible, too.

Further improvements for the tracking framework

As has been presented in section 3.1 "Recognise the Tabletop in AR", we tried out plane detection and image recognition to recognise the tabletop and relate the position of the devices with it. An interesting step would be to combine both approaches to improve the tracking, since the tabletop remains static and in parallel to the ground all the time. Especially when the marker recognition is difficult, for example when being further away from the tabletop, or when the content of the display does not allow for good marker placements at all positions, the results from the plane detection could help the system to recover from losses in the image recognition quality. Another technical idea was presented by a participant of the local Apple developer community when being told about the communication implementation using websockets: Currently, our implementation relies on communicating over the lab's general wifi access points including fixed IP addresses. Apple's ad-hoc network capabilities could provide a more sophisticated way of establishing real-time communication between the tabletop and multiple devices, and would reduce the risk of package lost if the lab's network is occupied by other traffic. However, we did not experience any network delay that let us start measuring the network performance.

Future work for aspects of cross-device interaction and AR

On the other hand, using Apple's ad-hoc network would allow us to easily share ARKit's world maps (the virtual image of feature points that it used for world tracking) between devices, as this feature is easily provided by Apple's frameworks. When using multiple devices, a new device could receive an existing world map of another device and use it to determine its position much quicker and without tracking a marker on its own. Additionally, this would again allow users take their devices further away from the tabletop without loosing the ability to provide satisfying tracking and position results. Another idea in the field of cross-device interaction that our implemented framework easily enables, is to make use of the spatial knowledge between the devices. As has been presented in section 3.2 "Relate Position of Device and Tabletop using World Tracking", it would be possible the share the location of other devices with a single device, so that the system could forbid certain actions when two devices are close to each other.

Feedback from user studies

One issue that came up during our preliminary user study was that augmented reality creates interfaces with an immersive nature. The danger for users of loosing track of their surrounding is of importance, and more thoughts have to been done how interfaces, also in our context in combination with an interactive tabletop, can ease users' perception of their surrounding (cf. section 5.2 "Preliminary Study: System Context and General Use"). For example, interfaces could provide freeze views of the device display or reduce the complexity of the displayed data to improve users' awareness of the space around them. From our remote user study (section 5.4 “Study: General System and Software Evaluation”, we learned that improvements of the device UI of the animals application would help users to understand the order of the select items.

Midair tangibles over tabletop displays open a new research field.

In general, our approach to track midair tangibles using augmented reality has proven to open a new research field for tangible interaction in midair. Our different applications show that various use-cases for combining midair devices and a tabletop display exist and were met with positive reception by the participants in our studies. With the feedback from the participants and the previous publications, future work to develop prototypes of new application scenarios can be worthwhile.

Appendix A

Digital Media Library

All digital files with relation to this thesis are made available in the thesis' digital media library. It contains not only the files and images included in thesis, but also more images and videos from the applications and a demo of the implemented framework, sketches that were done during implementation, diagrams and all documents from the study preparation.

The digital media library is available at the lab's server "Oliver" at `/Public/Research Projects/TABULA/Data/Midair Tangibles - Thesis Jonas Vogel`.

The repositories of the framework's source code and the applications are available on the lab's server at `/Public/Research Projects/TABULA/Software/Sources/ARKit Midair Tangibles`.

Appendix B

User Study Documents

The following pages include documents and questionnaires that were created for the user studies and which had to be filled out by the participants.

Contents:

B.1 Intended User Study: User Collaboration and Midair Tangibles

With the help of the following documents and the experimental design described in section 5.3 “Intended Studies: User Collaboration and Midair Tangibles”, all information to conduct the studies should be given, so that the study can take place at a later point of time. The grey areas in some of the documents should help to separate them from each other during the study.

Contents:

1. Study Consent Form
2. Questionnaire on demographics, previous experience and teammate
3. Questionnaire on layer exploration
4. Questionnaire on target selection
5. Playbook for the study conduction

Informed Consent Form

Exploring Collaborative Midair Tangible Interaction

PRINCIPAL INVESTIGATOR Jonas Vogel, Media Computing Group, RWTH Aachen University
Phone: +49 (0)241 80-21051, Email: jonas.vogel@rwth-aachen.de

Purpose of the study: The goal of this study is to understand how users will interact with midair tangibles in a collaborative setting. Participants will be asked to position virtual objects on an interactive tabletop display, either with the help of a midair tangible or by direct touch. To analyse the interaction, users will be video-recorded and fill out a questionnaire.

Procedure: You will work together with another participant on two applications that involve midair tangibles. In the first application, you will try to find positions for virtual objects based on different information layers. You can switch the layers either by direct touch, or by using a midair tangible device as magic lens. In the second application, you will be asked to move objects into different target zones and can do this either by direct touch or by selecting and dropping the objects with the use of a midair tangible. This study should take about 40 minutes to complete.

After each application, we will ask you to fill out a questionnaire about the tested system.

Risks/Discomfort: You may become fatigued during the course of your participation in the study. You will be given some opportunities to rest, and additional breaks are also possible. There are no other risks associated with participation in the study. Should completion of either the task or the questionnaire become distressing to you, it will be terminated immediately.

Benefits: The results of this study will be useful for further understanding of midair tangible interaction.

Alternatives to Participation: Participation in this study is voluntary. You are free to withdraw or discontinue the participation.

Cost and Compensation: Participation in this study will involve no cost to you. There will be snacks and drinks for you during and after the participation.

Confidentiality: All information collected during the study period will be kept strictly confidential. You will be identified through identification numbers. No publications or reports from this project will include identifying information on any participant. If you agree to join this study, please sign your name below.

Video recordings: You will be video-recorded (image and sound) during the study. The recordings will not be made public and only be used for study analysis. You are free to withdraw or discontinue the participation at any time, if you don't want to be recorded.

O (Optional): I agree to a publication of a short video clip or a still photography in the written thesis, thesis presentation (usually made public on the lab's website), scientific paper or article.

O I have read and understood the information on this form.
O I have had the information on this form explained to me.

Participant's Name	Participant's Signature	Date
	Principal Investigator	Date

If you have any questions regarding this study, please contact the PI (contact data above).

Collaborative Study: Demographics

Participant ID: ___

Gender

◯ female ◯ male ◯ non-binary

Age

Dominant hand

◯ left ◯ right

Do you have previous experience with interactive tabletops?

◯ yes ◯ no

How close are you to your teammate?

we have never met before ◯ ◯ ◯ ◯ ◯ ◯ ◯ very close

User Study: Collaborative Layer Exploration

Participant ID: ___

1. It was easy for me to solve the task.

strongly disagree ◯ ◯ ◯ ◯ ◯ ◯ ◯ strongly agree

2. It helped me having another person solve the task with me.

strongly disagree ◯ ◯ ◯ ◯ ◯ ◯ ◯ strongly agree

3. We worked together to find the positions for the windmills.

strongly disagree ◯ ◯ ◯ ◯ ◯ ◯ ◯ strongly agree

4. The device helped me comparing the windmill's position on two layers.

strongly disagree ◯ ◯ ◯ ◯ ◯ ◯ ◯ strongly agree

5. In a few words, explain your strategy to solve the task.

6. Do you have any other comments that you would like to share?

User Study: Collaborative Target Selection

Participant ID: ___

1. It was easy for me to solve the task.

strongly disagree ◯ ◯ ◯ ◯ ◯ ◯ ◯ strongly agree

2. It helped me having another person solve the task with me.

strongly disagree ◯ ◯ ◯ ◯ ◯ ◯ ◯ strongly agree

3. The device helped me to solve the task.

strongly disagree ◯ ◯ ◯ ◯ ◯ ◯ ◯ strongly agree

4. The device helped me to reach targets that were further away.

strongly disagree ◯ ◯ ◯ ◯ ◯ ◯ ◯ strongly agree

5. I am satisfied with the system and its usability.

strongly disagree ◯ ◯ ◯ ◯ ◯ ◯ ◯ strongly agree

6. In a few words, explain your strategy to fulfil the task.

7. Do you have any other comments that you would like to share?

Study Playbook

0. Preparation
Video recording ready to go, charge and restart phones;
Notes ready.

1. Welcome

2. Following the Consent Form: What are we going to do?
Sign the Consent Form. Ask if they wish to obtain a copy.
Clean the table and devices.

3. Start the video recording and the timer.
Ask both to stand on the opposite sides of the tabletop.
No pressure. Feel free to talk to each other.

4. Windmill app.

Start the app.

Introduction

1. "Your task will be to find new spots for windmills around the city of Würselen."
2. "Here you can see three existing windmills, and you can create new windmills by pressing the button and move them around. You can delete windmills by moving them over the left and right edges."
3. "There are some rules that influence your choice, and you can get the result of your positioning by tapping the button." *Show the results live.*
    a. "Of course windmills cannot be placed onto housing areas or on top of each other. Also, placing them too close to housing areas or to each other results in higher building costs."
    b. "Windmills have to be placed in or next to regions with wind to have the best power generation. Placing them too close to each other will reduce the amount of power that they can harvest from the wind."
4. "To check the housing areas, the wind regions and the windmills surroundings, you can either use the buttons on the tabletop, or you can use the device to show and check the layers."
5. "Your task is to maximize the power and minimize the building costs. You have about 15min."

Reset the tabletop and the phones.

Ask if they are ready, then let them work and check the clock.

Ask to fill out the questionnaire.
Prepare the next app.

Anything else to add or discuss?

5. Animals app.

Start the app and the devices.

Introduction

- “Your task is to feed the animals. Each animal eats food with the color specified by the indicator bar.”
- “You can move the food by direct interaction on the tabletop, or you can collect food with your device and drop it off again.”
- “The amount of food an animal has decreases over time. New food will appear over time, and there will be at most four food entities per type present (either on the table or in a device).”
- “You get points by feeding animals, and lose points when an animal is hungry.”
- “Your task is to get 1000 points, or as much as possible in about 10min.”

Ask if they are ready, then let users work and observe what happens.
Fill out the questionnaire, anything else to add?

6. General Feedback

Anything else?

7. Thank you! Snacks.

Timing estimate:
5min Welcome
5 min Windmill explanation
15 min Windmill task
5 min Windmill QA
5 min Animals explanation
10 min Animals task
5 min Animals QA

B.2 Remote User Study: General System and Software Evaluation

Here, some documents for the remote study in section 5.4 “Study: General System and Software Evaluation” are listed. The analysis coding tables can be found in the digital media library.

Contents:

1. Playbook for the study conduction
2. Protocol for notes during the study conduction

Video Study Playbook

Invitation

Hi, hast du Lust, an der neuen Studie für meine Abschlussarbeit teilzunehmen? Statt vor Ort nun per Video-Konferenz.
In dieser Studie interessiere ich mich dafür, wie Nutzer*innen mit Eingabegeräten über einem interaktiven Tisch umgehen. Sie wird ca. 45min dauern und per Zoom-Videokonferenz stattfinden. Du kannst also von Zuhause mit Laptop und stabiler Internetverbindung teilnehmen. Instruktionen/Starthilfe dazu folgen. Ich werde dir drei Videos mit Anwendungsbeispielen des von mir entwickelten Systems zeigen. Nach jedem Video führen wir ein offenes Interview durch. Zur späteren Analyse wird die Videokonferenz aufgezeichnet.

Schreib' mir einfach einen Terminvorschlag.

Confirmation
Wir starten am X um Y. Die Studie wird ca. 45min dauern und per Zoom-Videokonferenz stattfinden. Du brauchst dazu einen Computer / Tablet mit Mikrofon und eine stabile Internetverbindung. Die Zoom-Software funktioniert im Browser, ich empfehle aber, den Client (kostenlos, keine Registrierung notwendig) zu installieren, da mit diesem die Übertragungsqualität besser ist. Die Videokonferenz findet unter folgender URL statt: ~. Wenn du den Zoom-Client noch nicht installiert hast, wirst du ebenfalls unter diesem Link dazu aufgefordert.

Introduction

1. Dauer: ca. 45min.
2. Bitte stelle dein Handy und den Computer für die Zeit auf “nicht stören”.
3. Ich werde die Session aufzeichnen, um später einzelne Aspekte nochmals untersuchen zu können. Du kannst entscheiden, ob du deine Kamera nun dazu an- oder ausschalten möchtest. Der Ton wird in jedem Fall aufgezeichnet.
   Ist das für dich in Ordnung? Dann starte ich jetzt die Aufzeichnung.
4. Ich erzähle nun kurz, was passiert:
   Purpose/Benefits: Ich interessiere mich dafür, wie Nutzer*innen mit Eingabegeräten über einem interaktiven Tisch umgehen und möchte dazu neue Erkenntnisse gewinnen.
   Procedure: Du wirst dazu drei Videos sehen, in denen Anwendungen dazu gezeigt werden. Nach jedem Video führen wir dann ein offenes Interview.
   Risks/Discomfort/Alternatives: Deine Teilnahme ist freiwillig, du kannst jederzeit aufhören.
   Costs/Compensation: Es entstehen dir keine Kosten, du bekommst keine Vergütung.
   Confidentiality: Die gesammelten Informationen, insbesondere die Video- und Tonaufzeichnungen, werden vertraulich und nur im Rahmen meiner Abschlussarbeit verwendet.
   Bist du damit einverstanden?

Stop/start recording

Zur Vergleichbarkeit mit anderen Studien notiere ich noch folgende Infos:
Geschlecht und Alter
Hast du schon Erfahrung an einem interaktiven Tisch?

Alle Videos zeigen mich, wie ich mit einem Eingabegerät über dem interaktiven Tisch arbeite. Ich hatte die Kamera auf dem Kopf, du siehst also alles wie “aus meinen Augen”. Versuche während das Video läuft, zu beschreiben, was du siehst - mit einem Fokus darauf, was ich im Video tue. Du kannst mich jederzeit bitten, das Video zu stoppen oder kurz zurückzuspulen.

Stop/start recording

Tangible Midair Controller

Start screen sharing, turn off own camera.

Introduction:

1. “You will see some Star Wars spaceships flying around, for example the Millenium Falcon, a Star Destroyer and some TIE Fighters.”
2. “The video is about two and half minutes long.”

Play video

Observation notes (~codes) and questions:

1. User seems to understand that spaceships is moved by device.
   Q: “How does the spaceship move?”

2. User seems to understand that the displayed content is represented in the space above the tabletop / that the third dimension, i.e. the virtual objects’ height, is covered by the space above the tabletop.
   Q: “What happened when the Falcon crossed the Star Destroyer?”

3. User wondered why the spaceship flew backwards or sidewards.
   Q: “Did the spaceship movement felt natural to you?”

4. (Added during study) What would you expect if the Falcon collides with another spaceship?

Questions after video playback:

1. Do you have any comments that you would like to share?

Stop/start recording

Target Selection

Start screen sharing, turn off own camera.

Introduction:

1. “You will see a scenario with some cute animals that have to be fed. In the first part of the video, this is done by using direct touch, in the second part, with the help of a device.”
2. “The video is about three minutes long.”

Start video

Observation notes (~codes):

1. The indicator color determines the food color.
2. Animals only eat food in their own color.
3. Animals have enough food after being fed some times. / Animals do not eat more food when the indicator bar is full.
4. Food can be selected and dropped off using the device.
5. The last food item that was selected is the first one that will be dropped off.
6. Users tend to move more when using direct touch in comparison to using the device.

Questions after video playback:

1. What advantages and disadvantages can you think of in using the device instead of direct touch?
2. Imagine using the presented system with a second person: What strategies would you think of to keep the animals fed? Which strategy would you choose to use?
3. Do you have any other comments that you would like to share?

Stop/start recording

Layer Exploration

Start screen sharing, turn off own camera.

Introduction:

1. “You will see a map of the city of Würselen. To fulfill the global climate goals, we want to build new windmills around it.”
2. “With the help of the system, we are able to find good spots for new windmills, since possible locations are bound by wind, housing and proximity to other windmills.”
3. “The video is about four minutes long.”

Start video

Observation notes (~codes):

1. Placing a windmill in a wind region is good.
2. Placing a windmill in a forbidden region is bad.
3. Placing windmills too close to each other is bad.
4. The initial, semi-transparent windmills cannot be moved.
5. The device can show another layer than the one displayed on the tabletop.

Questions after video playback:

1. What advantages and disadvantages can you think of in using the device instead of changing the layers on the tabletop display itself?
2. Imagine using the presented system with a second person: What strategies would you think of to find the best spots for new windmills? Which strategy would you choose to use?
3. Do you have any other comments that you would like to share?

At the end

Stop/start recording

Any other comments?

Stop recording

Thank you!

Video Study Protocol

Participant ID:

Start time:

Introduction notes:

Falcon

Video start time:

Time	Observation
	User seems to understand that spaceships is moved by device.
	User seems to understand that the displayed content is represented in the space above the tabletop.
	User wondered why the spaceship flew backwards or sidewards.

Do you have any comments that you would like to share?

Target Selection

Video start time:

Time	Observation
	The indicator color determines the food color.
	Animals only eat food in their own color.
	Animals have enough food after being fed some times. / Animals do not eat more food when the indicator bar is full.
	Food can be selected and dropped off using the device.
	The last food item that was selected is the first one that will be dropped off.
	Users tend to move more when using direct touch in comparison to using the device.

What advantages and disadvantages can you think of in using the device instead of direct touch?

Imagine using the presented system with a second person: What strategies would you think of to keep the animals fed? Which strategy would you choose to use?

Do you have any other comments that you would like to share?

Layer Exploration

Video start time:

Time	Observation
	Placing a windmill in a wind region is good.
	Placing a windmill in a forbidden region is bad.
	Placing windmills too close to each other is bad.
	The initial, semi-transparent windmills cannot be moved.
	The device can show another layer than the one displayed on the tabletop.

What advantages and disadvantages can you think of in using the device instead of changing the layers on the tabletop display itself?

Imagine using the presented system with a second person: What strategies would you think of to find the best spots for new windmills? Which strategy would you choose to use?

Do you have any other comments that you would like to share?

Bibliography

Michelle Annett, Tovi Grossman, Daniel Wigdor, and George Fitzmaurice. Medusa: A proximity-aware multi-touch tabletop. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, UIST '11, pages 337–346, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450307161. doi: 10.1145/2047196.2047240. URL `https://doi.org/10.1145/2047196.2047240`.

Alissa N. Antle, Alyssa F. Wise, Amanda Hall, Saba Nowroozi, Perry Tan, Jillian Warren, Rachael Eckersley, and Michelle Fan. Youtopia: A collaborative, tangible, multi-touch, sustainability learning activity. In *Proceedings of the 12th International Conference on Interaction Design and Children*, IDC '13, pages 565–568, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450319188. doi: 10.1145/2485760.2485866. URL `https://doi.org/10.1145/2485760.2485866`.

David Asselborn. Reaching a new dimension: Using tangibles for 3d object rotation. Master's thesis, RWTH Aachen University, Aachen, September 2018.

Andrew Aultman, Spencer Dowie, and Nelly Ann Hamid. Design heuristics for mobile augmented reality game user interfaces. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356213. doi: 10.1145/3170427.3188580. URL `https://doi.org/10.1145/3170427.3188580`.

Ronald T. Azuma. A survey of augmented reality. *Presence: Teleoper. Virtual Environ.*, 6(4):355–385, August 1997. ISSN

1054-7460. doi: 10.1162/pres.1997.6.4.355. URL https://doi.org/10.1162/pres.1997.6.4.355.

Patrick Baudisch, Torsten Becker, and Frederik Rudeck. Lumino: Tangible blocks for tabletop computers based on glass fiber bundles. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1165–1174, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781605589299. doi: 10.1145/1753326.1753500. URL https://doi.org/10.1145/1753326.1753500.

Dominikus Baur, Sebastian Boring, and Steven Feiner. Virtual projection: Exploring optical projection as a metaphor for multi-device interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1693–1702, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450310154. doi: 10.1145/2207676.2208297. URL https://doi.org/10.1145/2207676.2208297.

Andrea Bianchi, So-Ryang Ban, and Ian Oakley. Designing a physical aid to support active reading on tablets. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 699–708, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450331456. doi: 10.1145/2702123.2702303. URL https://doi.org/10.1145/2702123.2702303.

Eric A. Bier, Maureen C. Stone, Ken Pier, William Buxton, and Tony D. DeRose. Toolglass and magic lenses: The see-through interface. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, pages 73–80, New York, NY, USA, 1993. Association for Computing Machinery. ISBN 0897916018. doi: 10.1145/166117.166126. URL https://doi.org/10.1145/166117.166126.

Sebastian Boring, Dominikus Baur, Andreas Butz, Sean Gustafson, and Patrick Baudisch. Touch projector: Mobile interaction through video. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2287–2296, New York, NY, USA, 2010. Association for Computing Machinery. ISBN

9781605589299. doi: 10.1145/1753326.1753671. URL `https://doi.org/10.1145/1753326.1753671`.

Frederik Brudy, Joshua Kevin Budiman, Steven Houben, and Nicolai Marquardt. Investigating the role of an overview device in multi-device collaboration. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356206. doi: 10.1145/3173574.3173874. URL `https://doi.org/10.1145/3173574.3173874`.

Frederik Brudy, Christian Holz, Roman Rädle, Chi-Jui Wu, Steven Houben, Clemens Nylandsted Klokmose, and Nicolai Marquardt. 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, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702. doi: 10.1145/3290605.3300792. URL `https://doi.org/10.1145/3290605.3300792`.

Li-Wei Chan, Hsiang-Tao Wu, Hui-Shan Kao, Ju-Chun Ko, Home-Ru Lin, Mike Y. Chen, Jane Hsu, and Yi-Ping Hung. Enabling beyond-surface interactions for interactive surface with an invisible projection. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology*, UIST '10, pages 263–272, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781450302715. doi: 10.1145/1866029.1866072. URL `https://doi.org/10.1145/1866029.1866072`.

Liwei Chan, Stefanie Müller, Anne Roudaut, and Patrick Baudisch. Capstones and zebrawidgets: Sensing stacks of building blocks, dials and sliders on capacitive touch screens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 2189–2192, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450310154. doi: 10.1145/2207676.2208371. URL `https://doi.org/10.1145/2207676.2208371`.

Christian Cherek, Anke Brocker, Simon Voelker, and Jan

Borchers. Tangible awareness: How tangibles on tabletops influence awareness of each other's actions. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356206. doi: 10.1145/3173574.3173872. URL `https://doi.org/10.1145/3173574.3173872`.

Christian Cherek, David Asselborn, Simon Voelker, and Jan Borchers. Off-surface tangibles: Exploring the design space of midair tangible interaction. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359719. doi: 10.1145/3290607.3312966. URL `https://doi.org/10.1145/3290607.3312966`.

Min Fan, Alissa N. Antle, Carman Neustaedter, and Alyssa F. Wise. Exploring how a co-dependent tangible tool design supports collaboration in a tabletop activity. In *Proceedings of the 18th International Conference on Supporting Group Work*, GROUP '14, pages 81–90, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450330435. doi: 10.1145/2660398.2660402. URL `https://doi.org/10.1145/2660398.2660402`.

Steven Feiner, Blair MacIntyre, Marcus Haupt, and Eliot Solomon. Windows on the world: 2d windows for 3d augmented reality. In *Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology*, UIST '93, pages 145–155, New York, NY, USA, 1993. Association for Computing Machinery. ISBN 089791628X. doi: 10.1145/168642.168657. URL `https://doi.org/10.1145/168642.168657`.

George Fitzmaurice and William Buxton. The chameleon: Spatially aware palmtop computers. In *Conference Companion on Human Factors in Computing Systems*, CHI '94, pages 451–452, New York, NY, USA, 1994. Association for Computing Machinery. ISBN 0897916514. doi: 10.1145/259963.260460. URL `https://doi.org/10.1145/259963.260460`.

Thomas Riisgaard Hansen, Eva Eriksson, and Andreas

Lykke-Olesen. Mixed interaction space: Designing for camera based interaction with mobile devices. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1933–1936, New York, NY, USA, 2005. Association for Computing Machinery. ISBN 1595930027. doi: 10.1145/1056808.1057060. URL `https://doi.org/10.1145/1056808.1057060`.

Hiroshi Ishii and Brygg Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, CHI '97, pages 234–241, New York, NY, USA, 1997. Association for Computing Machinery. ISBN 0897918029. doi: 10.1145/258549.258715. URL `https://doi.org/10.1145/258549.258715`.

Yvonne Jansen, Thorsten Karrer, and Jan Borchers. Mudpad: Tactile feedback and haptic texture overlay for touch surfaces. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, pages 11–14, New York, NY, USA, 2010. Association for Computing Machinery. ISBN 9781450303996. doi: 10.1145/1936652.1936655. URL `https://doi.org/10.1145/1936652.1936655`.

Sergi Jordà, Günter Geiger, Marcos Alonso, and Martin Kaltenbrunner. The reactable: Exploring the synergy between live music performance and tabletop tangible interfaces. In *Proceedings of the 1st International Conference on Tangible and Embedded Interaction*, TEI '07, pages 139–146, New York, NY, USA, 2007. Association for Computing Machinery. ISBN 9781595936196. doi: 10.1145/1226969.1226998. URL `https://doi.org/10.1145/1226969.1226998`.

Y. Kakehi, M. Iida, T. Naemura, Y. Shirai, M. Matsushita, and T. Ohguro. Lumisight table: an interactive view-dependent tabletop display. *IEEE Computer Graphics and Applications*, 25(1):48–53, 2005.

Chang-Min Kim and Tek-Jin Nam. G-raff: An elevating tangible block for spatial tabletop interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 4161–4164, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450331456. doi: 10.1145/2702123.2702444. URL https://doi.org/10.1145/2702123.2702444.

Chang Min Kim and Tek-Jin Nam. Exploring the layered use of transparent display on a large tabletop display. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, CHI EA '16, pages 2555–2562, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450340823. doi: 10.1145/2851581.2892297. URL https://doi.org/10.1145/2851581.2892297.

Mandy Korzetz, Romina Kühn, and Thomas Schlegel. Turn it, pour it, twist it: A model for designing mobile device-based interactions. In *Proceedings of the 5th International ACM In-Cooperation HCI and UX Conference*, CHIuXiD'19, pages 20–23, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450361873. doi: 10.1145/3328243.3328246. URL https://doi.org/10.1145/3328243.3328246.

Ricardo Langner, Tom Horak, and Raimund Dachselt. Demonstrating vistiles: Visual data exploration using mobile devices. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, AVI '18, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356169. doi: 10.1145/3206505.3206583. URL https://doi.org/10.1145/3206505.3206583.

Sang-won Leigh, Philipp Schoessler, Felix Heibeck, Pattie Maes, and Hiroshi Ishii. Thaw: Tangible interaction with see-through augmentation for smartphones on computer screens. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '15, pages 89–96, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450333054. doi: 10.1145/2677199.2680584. URL https://doi.org/10.1145/2677199.2680584.

Rong-Hao Liang, Kai-Yin Cheng, Liwei Chan, Chuan-Xhyuan Peng, Mike Y. Chen, Rung-Huei Liang, De-Nian Yang, and Bing-Yu Chen. Gaussbits: Magnetic tangible bits for portable and occlusion-free near-surface interactions. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1391–1400, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450318990. doi: 10.1145/2470654.2466185. URL `https://doi.org/10.1145/2470654.2466185`.

Rene Linden. Multitouchkit: A software framework for touch input and tangibles on tabletops and. Master's thesis, RWTH Aachen University, Aachen, September 2015.

Anna Loparev, Lauren Westendorf, Margaret Flemings, Jennifer Cho, Romie Littrell, Anja Scholze, and Orit Shaer. Bacpack for new frontiers: A tangible tabletop museum exhibit exploring synthetic biology. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*, ISS '16, pages 481–484, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450342483. doi: 10.1145/2992154.2996878. URL `https://doi.org/10.1145/2992154.2996878`.

Joyce Ma, Lisa Sindorf, Isaac Liao, and Jennifer Frazier. Using a tangible versus a multi-touch graphical user interface to support data exploration at a museum exhibit. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '15, pages 33–40, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450333054. doi: 10.1145/2677199.2680555. URL `https://doi.org/10.1145/2677199.2680555`.

Haipeng Mi and Masanori Sugimoto. Hats: Interact using height-adjustable tangibles in tabletop interfaces. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '11, pages 71–74, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450308717. doi: 10.1145/2076354.2076368. URL `https://doi.org/10.1145/2076354.2076368`.

Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *SPIE Vol. 2351, Telemanipulator and Telepresence Technologies (1994)*, pages 282–292, 1994.

Ann Morrison, Antti Oulasvirta, Peter Peltonen, Saija Lemmela, Giulio Jacucci, Gerhard Reitmayr, Jaana Näsänen, and Antti Juustila. Like bees around the hive: A comparative study of a mobile augmented reality map. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 1889–1898, New York, NY, USA, 2009. Association for Computing Machinery. ISBN 9781605582467. doi: 10.1145/1518701.1518991. URL `https://doi.org/10.1145/1518701.1518991`.

Donald A. Norman. *The Design of Everyday Things*. Basic Books, Inc., USA, 2002. ISBN 9780465067107.

Alex Olwal. Lightsense: Enabling spatially aware handheld interaction devices. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, ISMAR '06, pages 119–122, USA, 2006. IEEE Computer Society. ISBN 1424406501. doi: 10.1109/ISMAR.2006.297802. URL `https://doi.org/10.1109/ISMAR.2006.297802`.

James Patten, Hiroshi Ishii, Jim Hines, and Gian Pangaro. Sensetable: A wireless object tracking platform for tangible user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '01, pages 253–260, New York, NY, USA, 2001. Association for Computing Machinery. ISBN 1581133278. doi: 10.1145/365024.365112. URL `https://doi.org/10.1145/365024.365112`.

Gary Perelman, Marcos Serrano, Celia Picard, Christophe Bortolaso, Mustapha Derras, and Emmanuel Dubois. Cell selection for spreadsheets on tablets: Stacking-based interaction. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450356213. doi: 10.1145/3170427.3186546. URL `https://doi.org/10.1145/3170427.3186546`.

Thomas Plank, Hans-Christian Jetter, Roman Rädle, Clemens N. Klokmose, Thomas Luger, and Harald Reiterer. Is two enough?! studying benefits, barriers, and biases of multi-tablet use for collaborative visualization. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, pages 4548–4560, New York, NY, USA, 2017. Association for Computing Machinery. ISBN 9781450346559. doi: 10.1145/3025453.3025537. URL `https://doi.org/10.1145/3025453.3025537`.

S. Reifinger, F. Laquai, and G. Rigoll. Translation and rotation of virtual objects in augmented reality: A comparison of interaction devices. In *2008 IEEE International Conference on Systems, Man and Cybernetics*, pages 2448–2453, Oct 2008. doi: 10.1109/ICSMC.2008.4811662.

Derek Reilly, Michael Welsman-Dinelle, Colin Bate, and Kori Inkpen. Just point and click? using handhelds to interact with paper maps. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services*, MobileHCI '05, pages 239–242, New York, NY, USA, 2005. Association for Computing Machinery. ISBN 1595930892. doi: 10.1145/1085777.1085820. URL `https://doi.org/10.1145/1085777.1085820`.

Patrick Reipschläger and Raimund Dachselt. Designar: Immersive 3d-modeling combining augmented reality with interactive displays. In *Proceedings of the 2019 ACM International Conference on Interactive Surfaces and Spaces*, ISS '19, pages 29–41, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450368919. doi: 10.1145/3343055.3359718. URL `https://doi.org/10.1145/3343055.3359718`.

Michael Rohs, Johannes Schöning, Martin Raubal, Georg Essl, and Antonio Krüger. Map navigation with mobile devices: Virtual versus physical movement with and without visual context. In *Proceedings of the 9th International Conference on Multimodal Interfaces*, ICMI '07, pages 146–153, New York, NY, USA, 2007. Association for Computing Machinery. ISBN 9781595938176. doi: 10.1145/1322192.1322219. URL `https://doi.org/10.1145/1322192.1322219`.

Johan Sanneblad and Lars Erik Holmquist. Ubiquitous graphics: Combining hand-held and wall-size displays to interact with large images. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, AVI '06, pages 373–377, New York, NY, USA, 2006. Association for Computing Machinery. ISBN 1595933530. doi: 10.1145/1133265.1133343. URL https://doi.org/10.1145/1133265.1133343.

Stacey D. Scott, M. Sheelagh T. Carpendale, and Kori M. Inkpen. Territoriality in collaborative tabletop workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, CSCW '04, pages 294–303, New York, NY, USA, 2004. Association for Computing Machinery. ISBN 1581138105. doi: 10.1145/1031607.1031655. URL https://doi.org/10.1145/1031607.1031655.

Martin Spindler. Spatially aware tangible display interaction in a tabletop environment. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '12, pages 277–282, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450312097. doi: 10.1145/2396636.2396679. URL https://doi.org/10.1145/2396636.2396679.

Martin Spindler, Sophie Stellmach, and Raimund Dachselt. Paperlens: Advanced magic lens interaction above the tabletop. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, ITS '09, pages 69–76, New York, NY, USA, 2009. Association for Computing Machinery. ISBN 9781605587332. doi: 10.1145/1731903.1731920. URL https://doi.org/10.1145/1731903.1731920.

Martin Spindler, Michel Hauschild, and Raimund Dachselt. Towards making graphical user interface palettes tangible. In *ACM International Conference on Interactive Tabletops and Surfaces*, ITS '10, pages 291–292, New York, NY, USA, 2010a. Association for Computing Machinery. ISBN 9781450303996. doi: 10.1145/1936652.1936721. URL https://doi.org/10.1145/1936652.1936721.

Martin Spindler, Christian Tominski, Heidrun Schumann, and Raimund Dachselt. Tangible views for information visualization. In *ACM International Conference on*

Interactive Tabletops and Surfaces, ITS '10, pages 157–166, New York, NY, USA, 2010b. Association for Computing Machinery. ISBN 9781450303996. doi: 10.1145/1936652.1936684. URL `https://doi.org/10.1145/1936652.1936684`.

Martin Spindler, Wolfgang Büschel, and Raimund Dachselt. Towards spatially aware tangible displays for the masses. In *Proceedings of the 2012 AVI Workshop on Designing Collaborative Interactive Spaces for e-Creativity, e-Science and e-Learning*, 5 2012a.

Martin Spindler, Wolfgang Büschel, and Raimund Dachselt. Use your head: Tangible windows for 3d information spaces in a tabletop environment. In *Proceedings of the 2012 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '12, pages 245–254, New York, NY, USA, 2012b. Association for Computing Machinery. ISBN 9781450312097. doi: 10.1145/2396636.2396674. URL `https://doi.org/10.1145/2396636.2396674`.

Martin Spindler, Marcel Martsch, and Raimund Dachselt. Going beyond the surface: Studying multi-layer interaction above the tabletop. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1277–1286, New York, NY, USA, 2012c. Association for Computing Machinery. ISBN 9781450310154. doi: 10.1145/2207676.2208583. URL `https://doi.org/10.1145/2207676.2208583`.

Paul Strohmeier. Displaypointers: Seamless cross-device interactions. In *Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology*, ACE '15, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450338523. doi: 10.1145/2832932.2832958. URL `https://doi.org/10.1145/2832932.2832958`.

Sriram Subramanian, Dzimitry Aliakseyeu, and Andrés Lucero. Multi-layer interaction for digital tables. In *Proceedings of the 19th Annual ACM Symposium on User Interface Software and Technology*, UIST '06, pages 269–272, New York, NY, USA, 2006. Association for Computing Machinery. ISBN 1595933131. doi: 10.1145/

1166253.1166295. URL `https://doi.org/10.1145/1166253.1166295`.

Shota Suto, Toshiya Watanabe, Susumu Shibusawa, and Masaru Kamada. Multi-touch tabletop system using infrared image recognition for user position identification. *Sensors*, 18(5):1559, May 2018. ISSN 1424-8220. doi: 10.3390/s18051559. URL `http://dx.doi.org/10.3390/s18051559`.

John Underkoffler and Hiroshi Ishii. Urp: A luminous-tangible workbench for urban planning and design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, pages 386–393, New York, NY, USA, 1999. Association for Computing Machinery. ISBN 0201485591. doi: 10.1145/302979.303114. URL `https://doi.org/10.1145/302979.303114`.

Simon Voelker, Kosuke Nakajima, Christian Thoresen, Yuichi Itoh, Kjell Ivar Øvergård, and Jan Borchers. Pucs: Detecting transparent, passive untouched capacitive widgets on unmodified multi-touch displays. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces*, ITS '13, pages 101–104, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450322713. doi: 10.1145/2512349.2512791. URL `https://doi.org/10.1145/2512349.2512791`.

Simon Voelker, Christian Corsten, Nur Al-huda Hamdan, Kjell Ivar Øvergård, and Jan Borchers. An interaction model for grasp-aware tangibles on interactive surfaces. In *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, ITS '14, pages 279–282, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450325875. doi: 10.1145/2669485.2669494. URL `https://doi.org/10.1145/2669485.2669494`.

Simon Voelker, Christian Cherek, Jan Thar, Thorsten Karrer, Christian Thoresen, Kjell Ivar Øvergård, and Jan Borchers. Percs: Persistently trackable tangibles on capacitive multi-touch displays. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15, pages 351–356, New York, NY, USA, 2015a. Association for Computing Machinery. ISBN 9781450337793. doi: 10.1145/2807442.2807466. URL `https://doi.org/10.1145/2807442.2807466`.

Simon Voelker, Kjell Ivar Øvergård, Chat Wacharamanotham, and Jan Borchers. Knobology revisited: A comparison of user performance between tangible and virtual rotary knobs. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces*, ITS '15, pages 35–38, New York, NY, USA, 2015b. Association for Computing Machinery. ISBN 9781450338998. doi: 10.1145/2817721.2817725. URL `https://doi.org/10.1145/2817721.2817725`.

Simon Voelker, Sebastian Hueber, Christian Holz, Christian Remy, and Nicolai Marquardt. Gazeconduits: Calibration-free cross-device collaboration through gaze and touch. In *To appear in the Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pages 451:1–451:10, New York, NY, USA, April 2020. ACM. ISBN 978-1-4503-5970-2/19/05. doi: 10.1145/3313831.3376578. URL `https://doi.org/10.1145/3313831.3376578`.

Philipp Wacker, Oliver Nowak, Simon Voelker, and Jan Borchers. 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, New York, NY, USA, 2019. Association for Computing Machinery. ISBN 9781450359702. doi: 10.1145/3290605.3300849. URL `https://doi.org/10.1145/3290605.3300849`.

Philipp Wacker, Adrian Wagner, Simon Voelker, and Jan Borchers. Heatmaps, shadows, bubbles, rays: Comparing mid-air pen position visualizations in handheld ar. In *To appear in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, pages 719:1–719:11, New York, NY, USA, April 2020. ACM. ISBN 978-1-4503-6708-0/20/04. doi: 10.1145/3313831.3376848. URL `https://doi.org/10.1145/3313831.3376848`.

Malte Weiss, Julie Wagner, Yvonne Jansen, Roger Jennings, Ramsin Khoshabeh, James D. Hollan, and Jan

Borchers. Slap widgets: Bridging the gap between virtual and physical controls on tabletops. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 481–490, New York, NY, USA, 2009. Association for Computing Machinery. ISBN 9781605582467. doi: 10.1145/1518701.1518779. URL https://doi.org/10.1145/1518701.1518779.

Andrew D. Wilson and Raman Sarin. Bluetable: Connecting wireless mobile devices on interactive surfaces using vision-based handshaking. In *Proceedings of Graphics Interface 2007*, GI '07, pages 119–125, New York, NY, USA, 2007. Association for Computing Machinery. ISBN 9781568813370. doi: 10.1145/1268517.1268539. URL https://doi.org/10.1145/1268517.1268539.

Paweł Wozniak, Nitesh Goyal, Przemysław Kucharski, Lars Lischke, Sven Mayer, and Morten Fjeld. Ramparts: Supporting sensemaking with spatially-aware mobile interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 2447–2460, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450333627. doi: 10.1145/2858036.2858491. URL https://doi.org/10.1145/2858036.2858491.

Ding Xu, Ali Momeni, and Eric Brockmeyer. Magpad: A near surface augmented reading system for physical paper and smartphone coupling. In *Adjunct Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, UIST '15 Adjunct, pages 103–104, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450337809. doi: 10.1145/2815585.2815740. URL https://doi.org/10.1145/2815585.2815740.

Jishuo Yang and Frank Maurer. Literature survey on combining digital tables and augmented reality for interacting with a model of the human body. technical report, University of Calgary, 2010.

Index

Typeset June 11, 2020