MIDAiR

Multimodal Interaction for Visual Data Analysis in Augmented Reality

Master's Thesis

by

Sebastian Hubenschmid
(01/794812)

at the

Universität Konstanz

Department of Computer and Information Science

First Reviewer:	Prof. Dr. Harald Reiterer
Second Reviewer:	Prof. Dr. Daniel A. Keim
Advisers:	Dr. Simon Butscher;
	Johannes Zagermann

Constance, 2019

Abstract

The research area of *immersive analytics* investigates how emergent technologies, such as mixed reality devices, can be helpful in analysing complex data. Prior work in this area shows how a multimodal interaction approach using a large touch-sensitive tabletop can be beneficial for interacting with an abstract 3D visualisation. This work improves upon prior work by extending the design space, utilising more input modalities to address several shortcomings.

Firstly, theoretical background for interacting with visualisations as well as several input modalities are investigated. This leads to functional requirements that guide the selection of used input modalities. Furthermore, related work is analysed in regard to the usage of input modalities in the realm of multimodal data analysis.

Secondly, the ***M**ultimodal **I**nteraction for Visual **D**ata Analysis **i**n Augmented **R**eality* (MIDAiR) system is presented. MIDAiR is a collaborative immersive analytics tool for the exploratory analysis of multidimensional abstract data. MIDAiR offers a novel combination of a spatially-aware tablet with an immersive augmented reality device that allows for multimodal interaction with a 3D visualisation. The used 3D visualisation consists of several linked 2D scatter plots, forming a 3D parallel coordinates visualisation. This enables users to easily filter and analyse data, allowing for the detection of clusters, trends, and outliers within the data set. MIDAiR thus expands the design space of prior work, offering a more flexible analysis workflow.

Thirdly, a usability study with eight participants was conducted to test the interaction concepts of MIDAiR. Although users needed some time to get used to the interaction concept, they fully employed most input modalities to accomplish the given tasks. Based on the results of this study, design recommendations and further research directions for immersive analytics tools are presented.

Publications

Parts of this work were previously published in:

> Sebastian Hubenschmid, Simon Butscher, Johannes Zagermann and Harald Reiterer. 'Employing Tangible Visualisations in Augmented Reality with Mobile Devices'.

as part of the workshop:

> Bongshin Lee, Arjun Srinivasan, John Stasko, Melanie Tory and Vidya Setlur. 'Multimodal Interaction for Data Visualization'. In: *Proceedings of the 2018 International Conference on Advanced Visual Interfaces – AVI '18*. ACM Press, 2018, pp. 1–3. DOI: 10.1145/3206505.3206602

Furthermore, parts of this work are also based on the documents:

> Sebastian Hubenschmid. 'Multimodal Interaction for Visual Data Analysis in Augmented Reality'. Seminar to the Master Project. 2018.

> Sebastian Hubenschmid. 'MIDAiR: Multimodal Interaction for Visual Data Analysis in Augmented Reality'. Master Project Report. 2019.

Contents

List of Figures

List of Tables

Abbreviations

AR . . . . . augmented reality

ART . . . . *__A__ugmented __R__eality above the __T__abletop*

HMD . . . head-mounted display

HUD . . . . head-up display

KPI . . . . . key performance indicator

MCV . . . . multiple coordinated views

MIDAiR . . *__M__ultimodal __I__nteraction for Visual __D__ata Analysis __i__n Augmented __R__eality*

NLI . . . . natural language interface

SSQ . . . . Simulation Sickness Questionnaire

TPI . . . . . Temple Presence Inventory

TUI . . . . tangible user interface

UEQ . . . . User Experience Questionnaire

VR . . . . . virtual reality

1

Introduction

The past few decades have seen a drastic increase in the amount of data being collected, making the analysis of such data ever more important. While recent advances in machine learning have helped in taming this flood of information, the visualisation of such information remains an integral part for gaining new insights, especially for large and complex data sets [31, 92, 113]. Much research has been dedicated to providing and choosing the correct visualisation, allowing us to employ our innate visual pattern detection capabilities [48].

However, it is not only essential to choose the correct visualisation type, but also to effectively interact with the visualisation that allows for a successful exploration and interpretation of the underlying data [92, 102, 113, 158]. Current visualisation frameworks such as Tableau [134] are hence feature-rich, offering both a large selection of different visualisations and many options to customise and explore the data set. Yet, visualisation frameworks for abstract data sets have traditionally focused on 2D visualisations, as there is a lot of scepticism towards 3D visualisations due to poorer understanding and inadequate interaction techniques [25, 92].

The research area of *immersive analytics* [87] aims to re-evaluate this scepticism through the use of new and emerging technologies such as immersive mixed reality environments. For example, research indicates that the use of egocentric navigation alongside an increased stereoscopic perception in immersive environments may contribute to a better understanding of abstract 3D visualisations [9, 21, 75, 135, 147, 157]. While current mixed reality devices can be beneficial for displaying abstract visualisations, their interaction is often still lacking: Current devices often rely on motion tracked controllers for virtual reality (VR) environments, or mid-air gestures for augmented reality (AR) environments. Although controllers and gestures are in many cases a natural fit for interacting with 3D objects, their usage quickly tires out the user's arms [4, 20, 50] and their lack of haptic feedback can

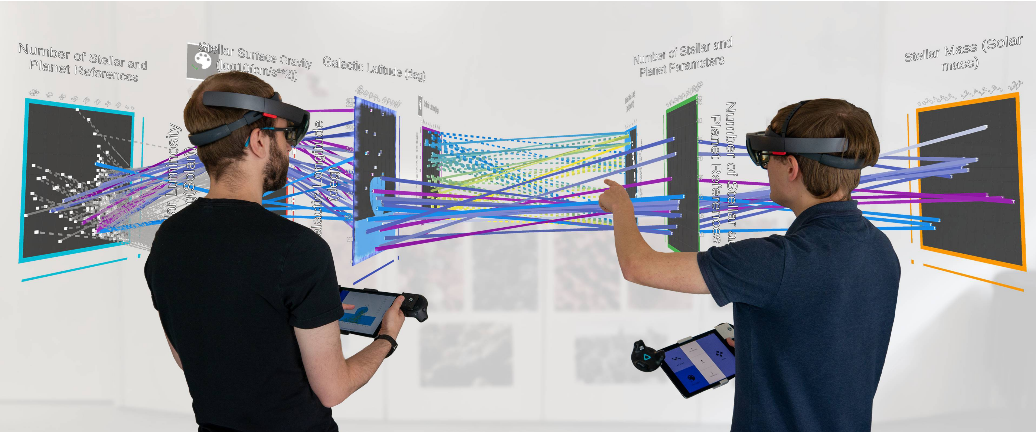


Figure 1.1: MIDAiR displays a 3D visualisation in an immersive AR environment. A spatially-aware tablet and an AR HMD offer multimodal input for interacting with the visualisation.

result in reduced accuracy [19, 24, 80]. Additionally, mid-air gestures, for example as used for the Microsoft HoloLens [55], often feel unnatural to the user when used for interaction [137].

A multimodal approach, on the other hand, can alleviate many of these issues; AR in particular is suited for a multimodal approach, since users can still fully use their physical environment (e.g. by using touch displays). Current AR products, such as the Microsoft HoloLens 2 [89], do already complement their use of mid-air gestures with both gaze and voice interaction (e.g. for text input), but still leave out many other modalities that can be useful, especially for interacting with information visualisations. While *immersive analytics* has already made some progress in exploring the benefits of mixed reality environments for abstract data visualisations, there is still little research on combining such visualisations with multimodal interaction [79], especially in combination with immersive AR environments.

This work thus aims to address this research gap with ***M**ultimodal **I**nteraction for Visual **D**ata Analysis **i**n Augmented **R**eality* (MIDAiR), an immersive analytics tool for interacting with 3D visualisations in an AR environment. MIDAiR combines a spatially-aware tablet with an immersive AR head-mounted display (HMD), resulting in a feature-rich system that uses multimodal input to facilitate interaction with 3D visualisations (see Figure 1.1). Users are able to see the entire visualisation in 3D space, or individual parts of it on their tablet in 2D, thus combining the advantages of both 2D and 3D visualisations.

1.1 Prior work

MIDAiR builds upon prior work in the area of immersive analytics by adressing the shortcomings of ***A**ugmented **R**eality above the **T**abletop* (ART) [21, 58]. ART was built for the exploratory analysis of high-dimensional data, combining an immersive AR environment with a large tabletop display for touch interaction (see Figure 1.2). The visualisation is composed of linked 2D scatter plots, creating a 3D parallel coordinates visualisation that is anchored to the tabletop. A control panel on the tabletop allows users to interact with the visualisation and individual scatter plots.

The ART system offers both touch input and spatial navigation in AR for performing different tasks. For plot arrangement (adding, reordering, removing scatter plots) and scatter plot configuration (assign dimensions, define filters) the touch input provides a familiar interface. Navigation is performed through a combination of touch input (scrolling on the table) and spatial movement around the tabletop.

A preliminary evaluation with ten domain experts revealed that the touch input for scatter plot configuration tasks was beneficial due to its preciseness when defining clusters. Although spatial movement allowed participants to navigate the visualisation intuitively, the fixed setup and large size of the tabletop made it difficult to move around the visualisation, forcing users to limit themselves to view the visualisation from the front. Still, spatial movement was appreciated for collaborative tasks, as users were able to view the visualisation from different angles. Because the visualisation is shared between users, changes made by one user also affect all other users – which may not always be intended. Users also had to switch their attention several times between the visualisation and tabletop when adding a filter, thus losing critical context. As a result, this work aims to address these limitations by extending the design space through multimodal interaction.

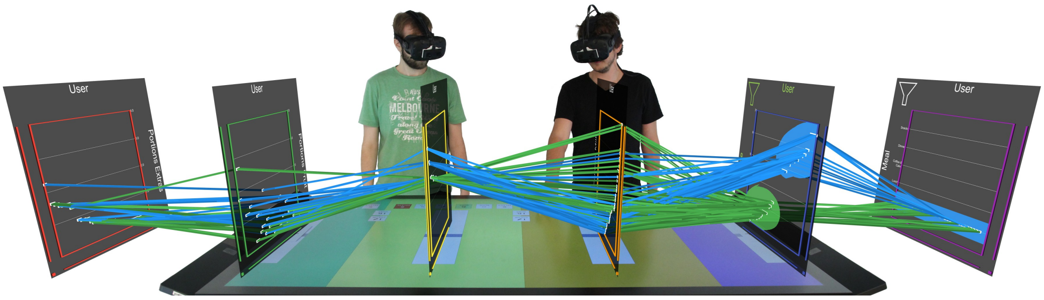


Figure 1.2: ART facilitates the collaborative analysis of multidimensional data. A 3D parallel coordinates visualisation is anchored to a touch-sensitive tabletop in augmented reality, allowing for familiar operation [21, 58].

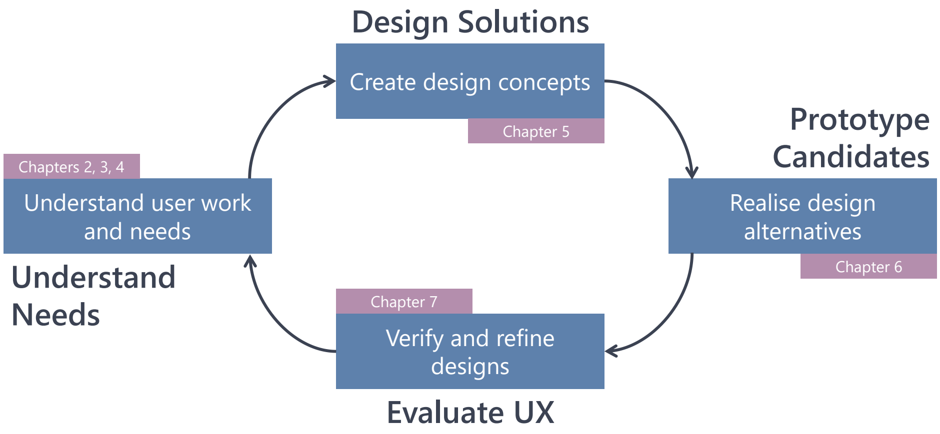


Figure 1.3: Four elemental activities of the UX design lifecycle. Adapted from Hartson and Pyla [47].

1.2 Outline

The creation of MIDAiR followed the UX design lifecycle as described by Hartson and Pyla [47] (see Figure 1.3). It describes four main activities for iteratively improving the usability of applications, which is reflected in this work:

The UX design lifecycle starts with *understanding the user's work and needs*, which is accomplished in three separate steps: The first step, covered in Chapter 2, was the analysis of current research on how to successfully interact with visualisations. This was achieved by performing a literature research of different visualisation interaction taxonomies and generating functional requirements that can be applied for multimodal interaction with 3D visualisations, which are also partially based on prior work. The second step, covered in Chapter 3, consisted of taking a look at available input modalities to evaluate their benefits, drawbacks, and possible use cases in regards to interacting with an immersive 3D visualisation. This chapter also examines how these input modalities can be combined for an effective multimodal interaction design. With Chapter 4, design considerations were extracted from several related projects which heavily influenced the design direction of MIDAiR.

The second step of the UX design lifecycle, *creating design concepts*, is covered by Chapter 5. This chapter describes the 3D visualisation in detail, explains the multimodal interaction concepts, and provides an overview over MIDAiR's features. Chapter 6 then describes how this design was implemented, thus corresponding to the *realisation of design alternatives* step of the UX design lifecycle.

Chapter 7 represents the last step of the lifecycle by *verifying and refining the design*. This chapter describes the setup and findings of a usability study and offers design improvements which may be implemented in another iteration of the lifecycle. Lastly, Chapter 8 concludes this work and presents opportunities for future work.

2

Visualisation Interaction Foundations

This chapter establishes the requirements for a multimodal visual data analysis tools in AR. Interaction with visualisation is well-studied, with several works analysing the necessary interaction features. Section 2.1 therefore reviews different interaction taxonomies for information visualisations. Section 2.2 then builds upon this foundation by elaborating specific requirements, based on both the ART system as well as the interaction taxonomies. Lastly, Section 2.3 provides a short overview over all established requirements. This topic was discussed in a preceding seminar [60], and has been updated and summarised for this work.

2.1 Interaction Taxonomies

Because interaction is an integral part of the data analysis process [31, 92, 113], several researchers provide guidelines on how to design an effective interaction model for information visualisations.

One of the earliest taxonomies is Shneiderman's *Visual Information Seeking Mantra* [122]: 'Overview first, zoom and filter, then details-on demand'. Shneiderman defines seven high-level tasks to illustrate this mantra: *overview*, *zoom*, *filter*, *details-on-demand*, *relate*, *history*, and *extract*.

Heer and Shneiderman [49] provide a taxonomy for interaction in information visualisations, divided into the three groups of *data & view specification*, *view manipulation*, and *process & provenance* (see Table 2.1). Each group consists of several task types: *Data & view specification* describes tasks related to choosing the correct *visualisation*, *filtering* and *sorting* the current data, and *deriving* values from the original data set. *View manipulation* describes *selecting* individual data points, *navigating* the visualisation, creating multiple *coordinated* views for multi-dimensional data,

Data & View Specifications	**Visualise** data by choosing visual encodings. **Filter** out data to focus on relevant items. **Sort** items to expose patterns. **Derive** values or models from source data.
View Manipulation	**Select** items to highlight, filter, or manipulate them. **Navigate** to examine high-level patterns and low-level details. **Coordinate** views for linked, multi-dimensional exploration. **Organise** multiple windows and workspaces.
Process & Provenance	**Record** analysis histories for revisitation, review and sharing. **Annotate** patterns to document findings. **Share** views and annotations to enable collaboration. **Guide** users through analysis tasks or stories.

Table 2.1: Taxonomy for visual data analysis, adapted from Heer and Shneiderman [49].

and *organising* these views and workspaces. Lastly, *process & provenance* comprises the ability to *record* the analysis workflow (e.g. for undoing actions), creating *annotations*, *sharing* views between multiple users, and provide user *guidance* (e.g. displaying tips for beginners).

Yi et al. [158] group general information visualisation tasks into seven categories: *select*, *explore*, *reconfigure*, *encode*, *abstract/elaborate*, *filter*, and *connect*. Ward et al. [146] extend this definition with the *hybrid* category (i.e. combining interaction techniques). Their groups do not, however, capture every interaction technique: for example, undo or redo actions are missing from these groups.

These taxonomies demonstrate that successful data analysis must support many different tasks, therefore requiring a feature-rich system. The question is thus: Which input modality is best suited for which task? For this, functional requirements are defined that can be assigned and examined with regard to the chosen modality later in this work.

2.2 Functional Requirements

This section establishes a set of requirements based upon both the interaction requirements from the previously discussed literature as well as a prior evaluation of the ART system [21, 58]. The purpose of these requirements is to address key limitations of the ART system and to make the visualisation system more flexible by allowing and encouraging the use of new input modalities. A short ID is placed next to each requirement for easier referral later on in this work. Given that the MIDAiR is based on the ART system, the requirements are grouped again by ART's interaction groups of *data manipulation*, *scatter plot arrangement*, *scatter plot configuration*, *navigation*, *inter-plot interaction*, and *collaboration*, with the addition of *visualisation interaction*.

2.2.1 Visualisation Interaction

(Annotate) **Create annotations.** Several works [49, 62, 113] allude to the importance of note taking, especially in a collaborative setting. This includes both text annotations as well as free-form drawings that are tethered to the visualisation or individual plots.

(History) **Record history (Undo / Redo).** Referred to as *history* [122] or *recording* history [49], undoing and redoing actions is an integral part of many applications, allowing users to retrace their steps or to correct mistakes.

(Details) **Display details.** *Details-on-demand* is an important part of both the *Visual Information Seeking Mantra* [122] and many desktop visualisations (e.g. mouse-over tooltips), but can be difficult to realise in AR. This can be extended to displaying other information, such as pictures attached to each data point.

2.2.2 Data Manipulation

(Filter) **Filter data.** Creating *filters* [49, 122] is an essential feature for any data analysis tool, allowing users to ignore irrelevant data.

(Brush) **Linking & brushing.** *Selecting* data [49, 158] by way of the linking and brushing technique [71] is useful for marking items of interest and gaining more insight into the selected data points. For example, by selecting data points in a multiple coordinated views (MCV) visualisation, the same data points are highlighted in all related views, thus allowing for easier comparison. This is also useful for highlighting clusters and outliers: Assigning a colour to each feature allows users to track these features through different scatter plots. To differentiate this from the Filter requirement, selections should only mark the related data points (e.g. by colour), but not filter out any data.

Derive new dimensions. Another way of manipulating data is to *derive* [49] or *extract* [122] from existing data. This can be implemented with common statistical methods (e.g. average, minimum, maximum), thus generating a dynamic data query. (Derive)

2.2.3 Scatter Plot Arrangement

Create & remove plots. An important part of the fluidity in ART's workflow was due to its simplicity in creating and deleting scatter plots. The new system should therefore make the creation and removal of new plots as simple and intuitive as possible to further encourage this fluidity. (Create)

Organise plot layout. Analysis goals can change quickly, so the system must be able to adapt. *Organising* [49] the plot layout to pursue a new analysis goal can therefore help towards this goal. ART's support for this is limited, as plots are automatically arranged in line. (Layout)

Connect plots. Also referred to as *relate* [122], *connect* [158], or *coordinate* [49], establishing connections is an important tool for features such as Filter and Brush. In addition, 3D visualisations can take advantage of this by visualising these links, thus revealing new patterns (e.g. see VisLink [26], ImAxes [29]). ART automatically links neighbouring scatter plots but allowing users to manually link plots – combined with a more flexible Layout – could be advantageous (e.g. see ImAxes [29], VisTiles [77]). (Connect)

Create workspaces. The combination of Connect and Layout naturally leads to the creation and management of workspaces. This can improve collaboration and allows for a non-linear workflow, where users can create a new workspace to follow different analysis goals. (Workspaces)

2.2.4 Scatter Plot Configuration

Change dimensions. To create new plots and to *explore* [158] the data set, users need to be able to change the dimensions of a visualisation. ART uses a scrollable list, allowing users to select a dimension through touch. Although this works well for a small number of dimensions, it quickly becomes cumbersome with many dimensions. (Dimensions)

(ZOOM) **Change zoom level.** *Zooming* [122] or *abstracting/elaborating* [158] allows users to focus on a particular subset within a dimension. This can also act as filter, only displaying visible data points in all connected plots (e.g. see VisTiles [77]).

(CONTEXT) **Avoid context switch.** Another shortcoming of ART is that users constantly have to switch between the tabletop and the AR visualisation. Especially when trying to configure a scatter plot, users first had to switch to the tabletop and select the appropriate dimension, which caused them to lose the relevant context.

2.2.5 Navigation

(NAVIGATE) **Navigate the visualisation.** *Navigation* [49] is essential for complex visualisations, allowing users to explore details or get an *overview* [122]. The egocentric navigation common in AR environments allows users to intuitively navigate the visualisation. However, a large tabletop as used in ART was a hindrance for navigation.

(REMOTE) **Remote control.** Users may utilise the AR environment to place visualisations far away from each other. Thus, the system should consider offering remote interaction, so that users do not need to walk to the visualisation they want to interact with.

(MOVE) **Complement AR movement.** Egocentric navigation in AR is intuitive, but also limited: for example, users cannot view the visualisation from above. Features such as ART's flip implementation can complement the AR movement, allowing users to overcome these limitations.

2.2.6 Inter-Plot Interaction

(DIFFERENCES) **Show relative differences.** When comparing data points of two different scatter plots, the lines can indicate trends and outliers. ART supports this comparison by allowing users to sort and colourise the lines based on relative differences between two plots.

(LINE-SELECT) **Select lines.** Once users find an outlier within the lines (e.g. different inclination than all other lines), they may want to further investigate this outlier. Similar to BRUSH, users should be able to quickly select or highlight lines, so that the line's behaviour can be investigated across different scatter plots.

2.2.7 Collaboration

Allow concurrent input. Studies indicate that collaborative data analysis is more effective when all users can work on the visualisation at the same time [6], and when they have control over different parts of the data [7]. (Concurrency)

Support separate workspaces. As extension to the Workspace requirement, enabling users to work simultaneously in different workspaces allows the users to work independently on different goals. (Separate)

Support workspace merging. Following the Separate workspaces requirement, users should also be able to *share* [49] their results. By allowing workspaces to merge together, findings can be compared more efficiently, and new findings can emerge. (Merge)

2.3 Summary

For a successful interaction with data visualisations, the system must support a large variety of tasks, as the previous taxonomies and functional requirements have demonstrated. To appropriately support these tasks, a multimodal interaction approach will be used, so that each task can be mapped to the most suitable input modality.

3

Input Modalities

This chapter examines viable input modalities for *immersive analytics* tools in AR environments. Although output modalities (e.g. olfactory [101] or auditory [109]) can be an effective complement to visual data analysis, the focus in this work lies solely on investigating input modalities for interaction with the data analysis tool. Furthermore, this work only concentrates on modalities suitable for an immersive AR environment; thus, modalities such as mouse and keyboard interaction are not covered. Similarly, although input modalities such as brain-computer interfaces show promising results – especially for mixed reality environments [27, 82] – their use is still restricted by the available technology and are therefore not covered in this work. As a result, this work examines the background of *Pen & Touch*, *Mid-Air Gestures*, *Tangible User Interfaces*, *Speech & Natural Language Interfaces*, *Gaze*, and *Proxemics* (see Figure 3.1). Each section also provides an overview over its modality's advantages and disadvantages, as well as its applicability for immersive analytics tools. Afterwards, Section 3.7 presents existing research on multimodal interaction, wwhile Section 3.8 provides a short summary over the discussed input modalities. This chapter has been adapted from a prior seminar work on this topic [60].

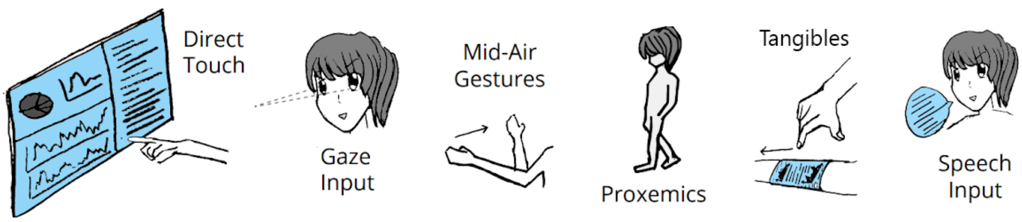


Figure 3.1: Suitable input modalities for AR include touch, gaze, mid-air gestures, tangibles, and speech input [91].

3.1 Pen & Touch

To reduce the scope of several decades worth of research into touch interaction, this section will focus on pen and touch interaction regarding recent information visualisation systems. Furthermore, because several major devices now natively supporting both touch and pen input (e.g. Apple iPad, Microsoft Surface) and because of the tight interplay between pen and touch [145], this work treats pen and touch as one input modality.

When designed with touch input in mind, a data analysis tool with touch interaction can offer a very natural [21, 51, 63] interface that is generally easy to learn and easy to use [21, 112, 125]. Unlike traditional desktop systems, touch interfaces work better without the use of nested menu entries [32] and are thus more accessible and simple to use, offering more fluidity [21, 112]. At the same time, this tendency to simplicity makes it harder to incorporate complex quantitative tools, which are essential for advanced users [32, 112].

However, unlike other modalities, touch interaction is strongly bound to its output display. Although the display provides somesthetic feedback [110], the action of directly touching a button also partially occludes the display [115]. Some projects [125, 126] address this through the use of indirect touch interfaces, decoupling the output from the actual input interface, which can be especially beneficial in an AR environment. The strong binding between input and output also makes touch appropriate for interaction with 2D visualisations due to its perceived directness but may not be appropriate for interacting with 3D visualisations, as finding intuitive 2D mappings for 3D content is challenging [63].

Our fingers are also often bigger than the targeted UI element, thus making it hard to interact precisely [124]. Pen interaction, on the other hand, prevents this *fat finger problem*, offering more accuracy [39, 145]. Research indicates that users see a clear division of labour for pen and touch interaction, with touch preferences for general interaction (e.g. manipulation) and pen for precise tasks (e.g. sketching, for Annotate) [51, 145]. Yet there is also a strong interplay between these two modalities, as users are faster and more accurate when using both [17], often preferring one-handed touch with pen interaction for many tasks [39].

While pen provides users with more accuracy, touch enables the use of 2D multi-touch gestures. Although some multi-touch gestures have now become commonplace (e.g. *pinch-to-zoom*), the use of such gestures has to be considered carefully, as contextual gestures can be hard to discover or lead to inconsistent behaviour [114]. Gestures should also be as simple as possible, as complex gestures can easily overwhelm users [12, 114, 153]. Designing a consistent set of gestures can therefore be challenging [116].

Overview: Pen & Touch

Advantages
- Precise input with somesthetic feedback
- Familiar interface for interacting with 2D screens

Disadvantages
- Limited to 2D plane, thus unsuitable for 3D interaction
- Hand (or other objects) may occlude input interface

AR data analysis use cases
- Create filters in scatter plot mapped to 2D screen (Brush, Filter)
- Create annotations with pen (Annotate)

3.2 Mid-Air Gestures

Mid-air gestures allow the user to naturally manipulate 3D objects – almost as if they were real objects – and provide the opportunity to trigger specific explicit interactions. These gestures can be categorised as *deictic gestures*, *manipulative gestures*, *semaphoric gestures*, *gesticulation*, or *language gestures* [70, 103].

Deictic gestures. Deictic gestures help to establish spatial context, usually by pointing at the object. This makes these gestures suitable for providing context to other modalities (e.g. speech input [15]). These gestures can also be used for explicit actions, for example for selecting objects by pointing at them.

Manipulative gestures. Manipulative gestures are 'those whose intended purpose is to control some entity by applying a tight relationship between the actual movements of the gesturing hand/arm with the entity being manipulated' [103]. These can range from simple drag-and-drop gestures on a 2D screen to full 3D gestures in AR or VR environments for grabbing and manipulating digital objects. They can be further restrained to match the digital object, for example by giving tactile feedback through tangible objects.

Semaphoric gestures. Semaphoric gestures are predefined templates that can be mapped to different commands (e.g. Microsoft HoloLens' *AirTap* gesture). These gestures can be either static hand positions (e.g. holding up two fingers) or dynamic hand movements (e.g. waving). However, given that these gestures do not usually appear in natural human communication, they can seem artificial [152].

Gesticulation. Gestures used during speech are classified as *gesticulation*. These gestures are often naturally used during communication and thus require no training.

Unlike semaphoric gestures, they can only rarely be matched against predefined templates. Thus, these gestures only gain meaning in combination with other modalities, such as speech input.

Language gestures. Sign languages form sentences by combining different semaphoric gestures. Due to their linguistical background, these gestures have a grammatical structure and hence only make sense in the right combination with other *language gestures*.

Especially *gesticulation* gestures are considered to be a very natural addition to speech input [36, 73, 74, 103, 151]. *Manipulative* gestures are also a natural fit for AR applications, since they allow users to grab and manipulate digital objects that appear in the users' surroundings.

However, usage of gestures should be considered carefully, since users can easily get tired, and mapping a large amount of actions to different gestures can quickly deplete the gesture vocabulary and overwhelm users [5, 20, 152]. Gestures can also be prone to false positives, whereby users inadvertently activate gestures (e.g. using *gesticulation* gestures during collaboration with other users which accidentally triggers system commands) [23]. In addition, gestures can be hard to combine, both for the tracking system classifier as well as for users [70].

This work also considers 3D UI elements as gestural interface: Although 3D UI elements do not necessarily require a specific gesture, they do share similar characteristics with mid-air gestures with a focus on the gesture's spatial position (e.g. a tap gesture must be performed in the position of the 3D button). In this regard, LaViola et al. [78] provide a comprehensive overview over 3D UIs in immersive environments, which would otherwise exceed the scope of this work.

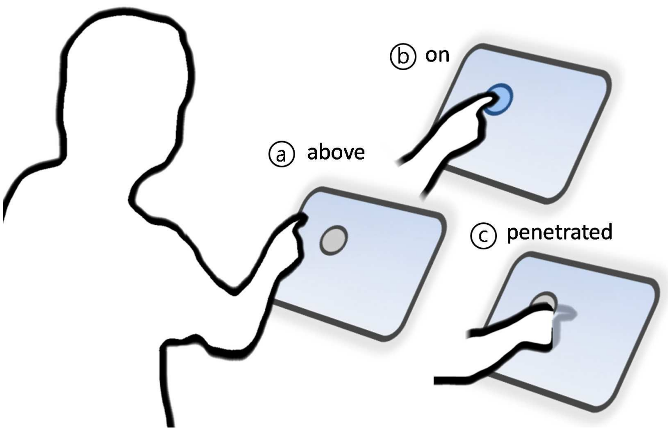


Figure 3.2: 3D UI elements support three states for mid-air touch interaction: (a) hovering above, (b) touching, and (c) penetrating the element [24].

In terms of hardware, many hand tracking solutions (e.g. Microsoft HoloLens [55]) use computer vision to track the user's hands [70]. However, these cameras can be susceptible to lighting changes or interference. Furthermore, for cameras mounted on the AR headset, gestures can only be tracked if the user's hands are within view of the camera, meaning that users must partially look at their hands when interacting. Other solutions (e.g. gloves, exoskeletons, 3D controllers) offer more accuracy and off-camera interaction, but can be unwieldy and can restrict users from interacting with other physical objects, which is especially problematic for multimodal interaction in AR. Recent novel solutions use radar waves [83], offering better accuracy, latency, and off-camera interaction.

Another problem for mid-air gestures – and especially 3D UIs – is the *touch the void* issue [19, 24]. A study by Chan et al. [24] indicates that users have trouble judging the depth of 3D UI elements, and thus struggle to interact precisely with these UI elements. Indicators, such as when the user touches or penetrates the UI elements, or other visual position feedback may be necessary (see Figure 3.2). For this reason, recent work by Lopes et al. [85] adds force feedback to a user's actions through electric muscle stimulation, while Vogel and Balakrishnan [143] have used auditory and visual feedback to compensate for the lack of physical feedback.

Overview: Mid-Air Gestures

Advantages

- Natural interaction with 3D objects

Disadvantages

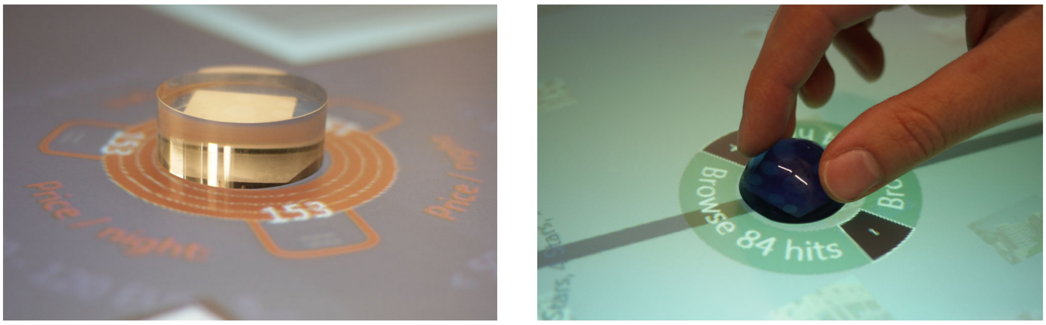
- Tiring
- Inaccurate & error-prone
- No tactile feedback

AR data analysis use cases

- Deictic gestures to establish spatial context for other modalities
- Moving scatter plots with manipulative gestures (Move)
- Mapping semaphoric gestures to common actions (e.g. History)

3.3 Tangible User Interfaces

Humans evolved by using their dexterity to turn physical objects into valuable tools. To this day, we have a strong relation to physical objects and tools in our daily life. The research field of tangible user interfaces (TUIs) aims to bring this object-based interaction to the digital world by coupling real objects with digital information (see Figure 3.3). Due to the steady growth since its origins as 'Graspable



(a) (b)

Figure 3.3: Facet-Streams binds digital filters to different physical tokens [68]. (a) Each token represents a filter that can be configured through touch. (b) A differently shaped token allows users to view the results.

Interface' [37], an in-depth discussion of this field would exceed the scope of this work. However, Ishii [64] and Shaer [121] provide an extensive overview over different definitions and taxonomies of TUIs. For brevity, this work summarises the main points relevant to information visualisation and interaction in AR. Furthermore, this work also considers mobile devices and wearables as TUIs, as their use within an AR environment can fit the characteristics of TUIs.

Ullmer and Ishii [141] define TUIs based on the relation between digital information and physical objects: tangibles 'couple physical representations (e.g. spatially manipulable physical objects) with digital representations (e.g. graphics and audio), yielding user interfaces that are computationally mediated but generally not identifiable as "computers" per se' [141]. They describe four characteristics for such TUIs:

1. Tangibles bind physical objects to underlying digital representations.
2. The physical interaction with these objects (i.e. their spatiality, relations, and connections) also embodies interaction with the digital realm.
3. Tangibles must, to some degree, represent their underlying digital information.
4. The physical objects are representative of certain key aspects of the underlying digital information.

Additionally, they describe four different approaches for TUIs: *spatial*, *relational*, *constructive*, and *associative* [141]. The *spatial* approach emphasises the tangible's location within a frame of reference (e.g. placing a tangible in a different position changes its digital state). The *relational* approach emphasises the arrangement of the tangibles (e.g. close tangibles form a connection). The *constructive* approach combines both previous approaches, allowing the construction of new tangibles by assembling modular elements. Lastly, the *associative* approach associates individual physical objects with digital information, but these objects do not derive meaning from referencing other objects.

Holmqvist et al. [53] categorise tangibles as either *container*, *token*, or *tool*: *Containers* are associated with any kind of digital information, without reflecting this information in its physical state. In contrast, *tokens* do reflect this information in their physical form while also being associated with digital information. Lastly, *tools* can be employed to change the underlying digital information.

Tangibles may be bound *dynamically* (i.e. the user can determine what a tangible is bound to) or *statically* (i.e. binding between tangible and digital information is fixed). According to Ullmer and Ishii [141], tangibles may be representationally bound to static and dynamic digital media, digital attributes, computational operations, remote entities, or simple or complex data structures. Given this work's focus on information visualisations, tangibles may therefore represent individual visualisations or tools for manipulating the underlying data or visualisations.

Due to their physical nature, users can draw upon their innate skills for organising and laying out information in the surrounding space [42]. This makes adding, sorting, or discarding information very natural (Create). Especially the spatial organisation of information may hold important implicit information for users (e.g. putting objects aside as a reminder to use this object later on) [77]. Although the physical state offers benefits (e.g. the ability to turn off the system and return to the same state later on), it also makes certain actions (e.g. undo, redo) more difficult (History). Furthermore, the physicality limits the amount of information that can be displayed simultaneously to the number of available tangibles.

Compared to more traditional interfaces where only one user at a time can interact with the system, tangibles naturally allow for every user to participate with the available artefacts [37]. This *space-multiplexing* therefore boosts co-located collaboration, enabling parallel interaction with the system (Concurrency). However, this also makes remote collaboration more difficult, as the physical state of the tangibles cannot be easily replicated. Some projects [18, 108, 150] try to alleviate this by using actuated tangibles that automatically synchronise their positions.

Overview: Tangible User Interfaces

Advantages

- Organisation of physical artefacts can hold implicit meaning
- Strong support for co-located parallel collaboration

Disadvantages
- Remote collaboration more difficult
- History feature harder to realise

AR data analysis use cases
- Bind visualisation position to individual tokens (Layout, Move)

3.4 Speech & Natural Language Interfaces

Thanks to recent improvements in machine learning and artificial intelligence, natural language interfaces (NLIs) and speech interaction have surged in popularity, as seen by the rise of voice-only home assistants (e.g. Amazon Echo, Apple HomePod, Google Home). While speech interaction without NLI can be useful on its own (e.g. as speech-to-text function), the combination with NLI allows the system to understand the user's intention. Similarly, NLI works well without voice (i.e. using a keyboard), but is inconvenient in an AR environment. This work therefore only focuses on the combination of NLIs with speech.

For natural dialogue and speech interaction alike, the current context is integral for correctly understanding a user's intention. Phrases like '*zoom in on this*' hold little meaning without knowing the previously established context of '*this*' [120]. Users may also want to name certain objects (e.g. '*call that . . . the calendar*') for easier reference later on [15]. Given the inherent ambiguity of certain statements (e.g. '*put that there*'), speech interaction is often used in conjunction with other modalities, such as touch or gesture interaction [15, 35, 120] – just as conversations are rich in both verbal and non-verbal communication [103]. The context can sometimes be

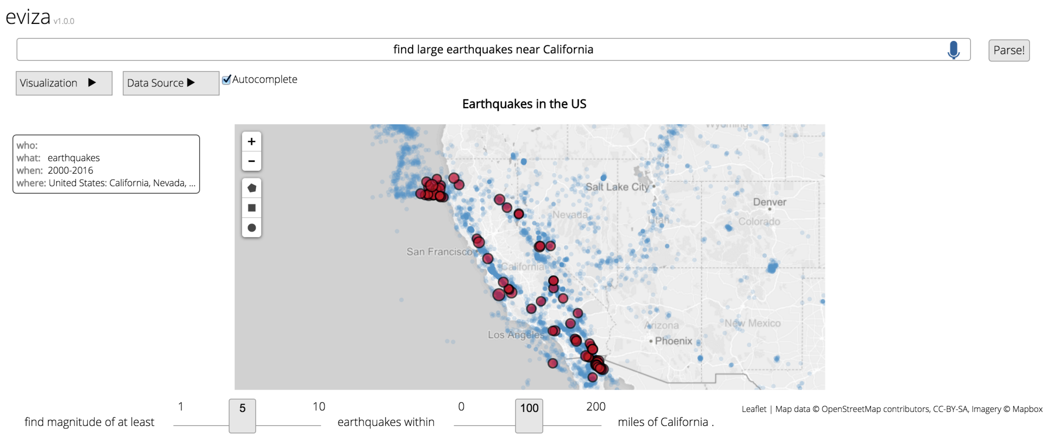


Figure 3.4: Eviza provides ambiguity widgets (bottom) for fuzzy queries (top), such as '*large*' and '*near*' [120].

established automatically through other sensors (e.g. '*show restaurants near me*' may use GPS) or by asking the user for clarification.

When interacting with data visualisations, Srinivasan and Stasko [129] differentiate between three query categories for interacting with data: *explicit*, *contextual and follow-up*, and *high-level* queries. For *explicit* queries, the system is provided with enough information to form a complete query on its own. *Contextual and follow-up* queries, in contrast, contain ambiguities that must be resolved through either contextual information from previous queries or input from other modalities. Lastly, *high-level* queries are akin to open-ended questions, as they can have multiple interpretations that can only be resolved by asking the user for clarification (e.g. '*show clusters*').

While this ambiguity and impreciseness is common during communication, it poses a big challenge for NLI. Although humans can make sense of fuzzy queries such as '*find large earthquakes near California*', the system must translate this into a specific strength in magnitude or range in metres. Some systems [41, 120, 129] therefore provide *ambiguity widgets*, which give the user control over resolving the fuzziness (see Figure 3.4).

The ambiguity can also cause problems when trying to determine whom the user is talking to, or when the user has completed the query. Especially in collaborative settings, users may want to seamlessly switch between talking to the system and talking to other collaborators. To distinguish between these two, some voice-based systems use keywords or other explicit actions (e.g. button press) to indicate that the system should be listening [129].

Pauses are also common in natural language (e.g. '*Create a blue square … there*') but can be difficult to handle [15]. On the one hand, the system should be responsive and react quickly to the user's query. On the other hand, additions to the query may drastically change the intended action, or the system may no longer be listening and thus ignore the addition.

Despite these problems, voice commands can be a powerful way to interact with a system. Given the vast amount of analysis options, data analysis systems usually rely on complicated menus or plain console commands. While neither is suitable for an AR environment, voice commands can offer similar expressiveness. However, like a console, the user has to be familiar with the available commands [111]. Proactive measures [129], such as autocompletion, can facilitate the discovery of available commands when typing [120], but is not suited for speech interaction, making exploration difficult.

Overview: Speech & Natural Language Interfaces

Advantages

- Hands-free interaction
- Ability to express queries, offering instant access to any function

Disadvantages

- Ambiguity
- Difficult to communicate what user can (not) do

AR data analysis use cases

- Input for search- or other text-fields (i.e. speech-to-text)
- Access statistical measures (e.g. show average, min, max) (Derive, Sort)
- Formulate data queries (e.g. SQL-like queries)

3.5 Gaze

Since the early 1980's [15], eye tracking has been used not only for studies, but also as imprecise input modality. Holmqvist et al. [54] hereby differentiate between *(inter)active gaze contingency* and *passive gaze contingency*: In *(inter)active gaze contingency*, users deliberately control the interface and issue actions by using their gaze as input (e.g. pressing a button by resting their eyes on the button for a few seconds). In contrast, with *passive gaze contingency* users are unaware that their gaze is used as input (e.g. during foveated rendering[1]).

Although eye tracking technology has been steadily improving over the last few decades, eye tracking still suffers from several issues. Jacob [65, 66] as well as Stellmach and Dachselt [131] group these issues into eight categories: *Midas touch*, *unconscious eye movements*, *inaccurate targeting*, *eye behaviour*, *synchronisation of multimodal inputs*, *double-role*, *reliability of tracking*, and *unfamiliarity*.

Midas touch. Users often need to look at a target before deciding if they want to trigger an interaction. Gaze interaction may therefore trigger unwanted interactions, if the system cannot interpret the user's intention correctly. One solution is to use gaze dwelling, which only triggers the interaction after users have rested their eyes on the target for a while. Another solution is to use a multimodal approach, only using gaze input to establish context while relying on other modalities (e.g. touch [130, 132, 139, 156] or gestures [159]) to trigger the action.

[1]Foveated rendering (also known as gaze-contingent rendering [33]) uses gaze to render only the small area visible to the fovea in high resolution, while the rest is rendered in a much lower resolution, thus reducing the necessary computing power.

Unconscious eye movements. Not all our eye movements are conscious and deliberate; changes in the environment may cause our eyes to dart unconsciously towards the change, thus distorting interaction.

Inaccurate targeting. Eye tracking technology is still rather inaccurate and is thus inappropriate for actions that require precision. Although coarse gaze direction can already be useful on its own, several approaches try to overcome this inaccuracy (e.g. local magnification [130, 160], avoiding closely positioned targets [65], or using areas of interest [54, 69]).

Eye behaviour. Eye movements consist mainly of saccades[2], which are highly erratic and thus hard to track accurately. Even when fixated on an object, our eyes dart around involuntarily due to microsaccades, which are imperceptible to humans, but cause jitter and inaccuracy in eye trackers [33, 54]. These saccades also make gaze-based interaction unsuitable for use cases such as drawing or moving a slider. Although our eyes are capable of *smooth pursuit* of an object, this often involves tracking a moving object (e.g. a bird taking flight) and can not be used deliberately for interaction [33, 54].

Synchronisation of multimodal inputs. When gaze is used in combination with other modalities (e.g. selecting an object with gaze, confirming with touch), the user's eyes may already target the next object before the action from another modality is registered. Thus, the interface must anticipate a delay between these modalities.

Double-role. UI changes caused by gaze input can be distracting, thus capturing the attention of the user's eyes which, in turn, distorts the captured gaze input. This can be resolved by tracking the direction of users' head instead of their gaze, and by ensuring that gaze-activated controls are non-distracting.

Reliability of tracking. Despite their technological improvements, eye trackers still do not offer perfect reliability. This is further complicated by different lighting conditions, glasses, makeup, or other obstructions. Additionally, the eye tracker may suffer from drift, thus becoming more inaccurate over time [54].

Unfamiliarity. Users are generally unfamiliar with *(inter)active gaze contingency*, and therefore do not expect to trigger any action with their eyes. Consequently, gaze input should be used carefully with other modalities (i.e. *gaze-supported interaction* [130, 132]) or without the user's awareness (i.e. *passive gaze contingency*).

[2] 'Saccades are rapid eye movements used in repositioning the fovea to a new location in the visual environment' [33].

Still, gaze input can be a highly useful input modality, if employed correctly: Due to their quickness, saccades do make for a beneficial, yet imprecise pointing modality, especially when compared to other input modalities [13, 130, 148]. Because users often need to look at the interaction target anyway, gaze can be inherently faster than other pointing modalities. Furthermore, tracking a user's head instead of gaze can solve many of the discussed issues, making the input more steady and predictable, albeit more tiring [131].

Overview: Gaze

Advantages

- Hands-free interaction
- Fast target acquisition

Disadvantages

- Inherently inaccurate & error-prone
- Prone to unwanted interaction

AR data analysis use cases

- Display additional information through head gaze (Details)
- Establish context for other modalities

3.6 Proxemics

The term *proxemics* was first used in a theory by Hall [46] to describe different interpersonal spatial relationships, both in humans and in animals. For this, Hall defines four discrete *proxemic zones* (see Figure 3.6a): *intimate* (0–50 cm), *personal* (50 cm–1 m), *social* (1–4 m), and *public* (>4 m). The exact ranges also depend on other characteristics, such as the current location, cultural factors, age, gender, or personal relationship between individuals [1]. In addition, orientation is also an important factor for classifying these relationships [45]. Hall further differentiates between *fixed* and *semi-fixed features* [46], which can influence interpersonal be-

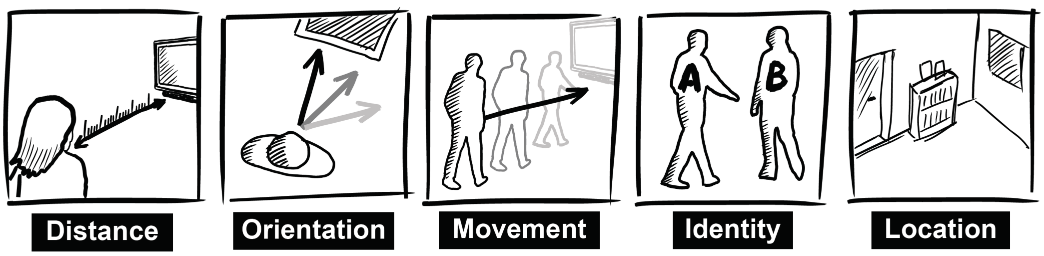


Figure 3.5: Five dimensions of proxemic interaction [44].

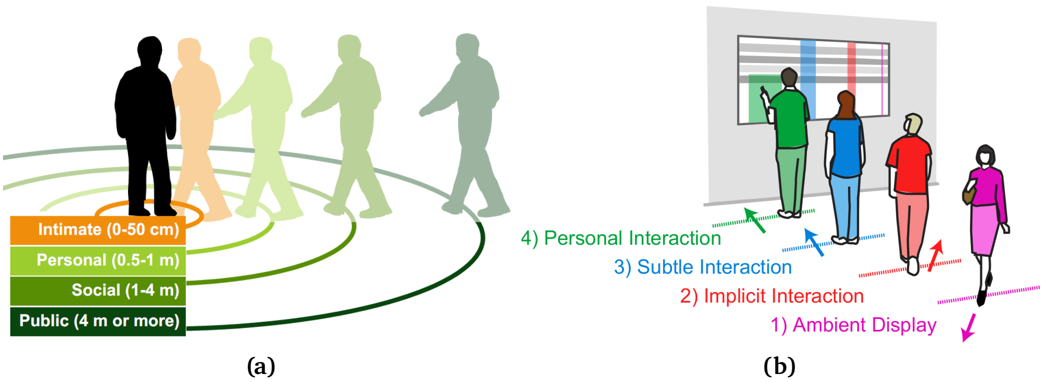


Figure 3.6: (a) Four proxemic zones to determine interpersonal spatial relationships [86]. (b) Four interaction phases, in which the system offers different amounts of interactivity and information [144].

haviour: *Fixed features* describe the immobile features of the environment, such as doors, windows, or the room layout. In contrast, *semi-fixed features* describe the moveable features, such as chairs and other furniture.

Using concepts from Vogel and Balakrishnan's *interaction phases* [144] (see Figure 3.6b), Greenberg extended Hall's definition of proxemics [44] in the context of ubiquitous computing [149], thus introducing five different dimensions for measuring spatial relations (see Figure 3.5): *distance*, *orientation*, *movement*, *identity*, and *location*. Whereas Hall's definition of the four *proxemic zones* aims to translate physical distance between entities into social distance, Greenberg et al. use inter-entity relations, measuring for example the distance between a user and a mobile device. By observing these dimensions, a system can anticipate the user's intention, thus allowing the user to interact implicitly or offering relevant interaction possibilities.

Distance. Although the objective *distance* between two entities can be measured as a continuous value, it is much more useful to think of the measure in terms of discrete values (e.g. Hall's proxemic zones [46], Vogel and Balakrishnan's interaction phases [144], or a binary measure, such as if two entities are within the same room).

Orientation. Similar to the *distance* dimension, the *orientation* can be measured as continuous value, but may be more useful as a discrete value (e.g. facing towards or away from another entity) – provided that the tracked entity has a defined front face.

Movement. The *movement* dimension measures the change of distance over time, describing both the speed and direction of the movement.

Identity. The *identity* dimension describes different properties about the entities themselves. This can range to exact information about their attributes (e.g. size, colour, form), over a general description (e.g. person, mobile device, display), to general associations (e.g. employee, guest). This is especially useful for personalisation and privacy, as the application can adjust the amount of information and interactivity based on a person's *identity* [8].

Location. The *location* dimension describes the context of an entity's physical location, as defined by its *fixed* and *semi-fixed* features. An application can trigger certain actions based on a user's current *location* (e.g. turn on monitors when entering the room).

Overview: Proxemics

Advantages

- Contextual cues for other modalities

Disadvantages

- Interaction opportunities may not be obvious to user, thus requiring strong indicators [86]

AR data analysis use cases

- Show interaction opportunities and more information when user is near a visualisation (Details)
- Establish links based on proximity between visualisations (Connect)

3.7 Multimodal Interaction

Human interaction is inherently multimodal, as people use a combination of gaze, gestures, speech, and other modalities to communicate with each other. A multimodal interface, therefore, seeks to 'leverage natural human capabilities to communicate [...], bringing more sophisticated pattern recognition and classification methods to human-computer interaction' [138]. Although users tend to switch between unimodal and multimodal interaction depending on the task [97], they still strongly prefer multimodal interaction [96, 99].

Existing research [14, 34, 78, 97, 98, 100, 111, 154, 155] points to several benefits of multimodal interaction over unimodal interaction, such as increased flexibility, increased user satisfaction, better task performance (in some tasks), and better adaptability to changes in the environment. Furthermore, multimodal interaction allows for more reliability in the face of both error handling and error avoidance: For example, users may want to issue unambiguous commands through the faster speech interface and fall back to pen input for ambiguous commands (e.g. searching for a hard-to-pronounce name). The resulting redundancy can be useful for offering greater accessibility for users with permanent or temporary disabilities. This redundancy also gives users respite from more tiring modalities (e.g. gestures). Lastly, multimodal interaction provides the opportunity to choose the most suitable modality for any given task, or to combine different modalities to make the interface as natural and efficient as possible.

To formalise the various ways of combining modalities, Nigay and Coutaz [94] establish a classification for multimodal interactions (see Table 3.1). Their classification differentiates between how modalities are combined temporally (*use of modalities*), and if data from different modalities is combined (*fusion*). For *use of modalities*, they discern between systems that only allow one single modality at a time (*sequential*), or systems that can process input from multiple modalities simultaneously (*parallel*). For *fusion*, they differentiate between combining input from multiple modalities for one action (*combined*) or treating these modalities as separate actions (*independent*). There are thus four different ways to use multimodal interaction: *Alternate*, *synergistic*, *exclusive*, and *concurrent*.

The *parallel* use of modalities in particular poses a difficult challenge. Although the user's intent is often a *parallel* use of modalities, their actions are executed sequentially [99], as modalities can have different temporal constraints [138]. Some modalities may trigger an action at a distinct point in time (e.g. *deictic* gestures), whereas other modalities provide continuous output (e.g. issuing a command by voice) [34, 97, 138]. Thus, when matching a pointing gesture to the correct verbal signal (e.g. 'delete *that*'), the system must anticipate some delay to correctly interpret the action.

The *combined* fusion of modalities can also cause difficulties for implementing a multimodal system, as applications can integrate multimodal signals *early* or *late* [34, 138]. *Early* integration of multimodal signals means combining data from all input channels first, before allowing the system to classify the action, whereas *late* integration classifies each individual modality first, before attempting to combine these signals into an action. Although a *late* integration can miss important multimodal context, the implementation is generally much simpler. Often, fusion engines [76] can be used to handle the input for the right context, task, user, and time.

		Use of Modalities	
		Sequential	Parallel
Fusion	Independent	*Exclusive*	*Concurrent*
	Combined	*Alternate*	*Synergistic*

Table 3.1: Classification of multimodal interfaces [94].

Lastly, Reeves et al. [105] provide guidelines that should be kept in mind when designing a multimodal user interface. For brevity, the following list summarises the main points applicable to this work's topic:

- **Designing Multimodal Input and Output:** Exploit the strengths of each modality, taking advantage of user's cognitive and physical abilities.
- **Adaptivity:** Adapt modalities to the current usage context (e.g. switch from speech to pen interaction in a more public setting).
- **Consistency:** Always use a consistent output modality, regardless of what input modality was used for an action.
- **Feedback:** Clearly communicate to users what input opportunities are available (e.g. when a speech interface is actively listening), without overloading the user with information.
- **Error Prevention/Handling:** Increase robustness through complementary modalities and allow switching to another modality in case an error does occur (see also Grasso et al. [43] or Jacob and Sibert [67]).

3.8 Summary

The previous sections have shown the strengths and weaknesses of several input modalities and discussed guidelines how these modalities can be successfully merged for a multimodal interaction approach. While some tasks (e.g. creating sketch annotation) may fit well for only one modality (i.e. pen input), other tasks may need to use multiple input modalities in parallel (e.g. placing scatter plots may combine gestures and gaze), and still others can be mapped to several alternate different input modalities (e.g. voice commands as alternative to buttons or text input). Yet the shortcomings of each modality must be kept in mind – and, if possible, circumvented with a multimodal approach – when applying these modalities to the each task.

4

Related Work

This chapter presents related work in the area of multimodal data analysis. Section 4.1 therefore examines several systems that offer multimodal input for interacting with data analysis systems, while Section 4.2 provides an overview concerning the used modalities. Although visualisations play an integral role for immersive analytics systems, the discussion thereof has already been extensively covered in ART [21, 58] (notably concerning ImAxes [29] and VisLink [26]) and is thus not present in this work. Recent works presented several frameworks for developing immersive analytics applications (e.g. DXR [123], IATK [28]), but were not considered for this work as they either do not perform well, or were published after the MIDAiR prototype was completed. Furthermore, Kraus et al. [75], Batch et al. [9], and Tadeja et al. [135] have since further investigated the benefits of 3D visualisations in immersive mixed reality environments. Parts of this chapter were already presented in the preceding seminar [60], and have been updated where appropriate.

4.1 Multimodal Data Analysis

This section explores the practical application of multimodal input for visual data analysis systems. Since there is still little research on this topic, the following sections will describe and analyse several research prototypes in detail: *Tangible Data Analysis* [40], *VisTiles* [77], *Proxemic Lens* [5], *When David Meets Goliath* [56], *Orko* [128], *DebugAR* [107] & *DesignAR* [106], *Smartphone-Based Pan and Zoom* [20], and *Immersive Insights* [23]. Each section provides an overview over the used input modalities and a simplified multimodal interaction classification [94] and discusses how their ideas may be applicable to this work.

Figure 4.1: *Tangible Data Analysis* uses smart tangibles to display and interact with visualisations [40].

4.1.1 Tangible Data Analysis

Fuchs et al. [40] used Sifteo Cubes for a public, collaborative data visualisation system, combining touch interaction with tangible cubes (see Figure 4.1). Each cube is equipped with a low-resolution display for showing a simple data visualisation, as well as sensors to detect touch, proximity to other devices, movement, and rotation (e.g. allowing detection of shake and tilt). These cubes are used to represent either a single dimension (*data mode*), or an auxiliary display for exploring a connected cube in more detail (*exploration mode*). Users can switch between these two modes by shaking the cubes, while rotation switches between different dimensions in *data mode*. Cubes in *data mode* show a glyph visualisation of the selected data point. Once two or more cubes in *data mode* are placed next to each other, the cubes switch to a bar chart visualisation, showing the differences between each connected neighbour. If a cube in *data mode* is placed next to a cube in *exploration mode*, the latter cube will display more details about the data cube, thus offloading additional information to a new screen. Lastly, users can sort, group, and filter dimensions by physically rearranging or discarding cubes.

This allows for a very natural way of interacting and exploring a data set. However, the limited display size severely restricts the data visualisation of each cube, and the physicality of the cubes prohibits features such as automatic layouting [40]. Both challenges may be solved with AR: The display can be offloaded to a virtual display in AR, thus increasing the resolution significantly for more complex visualisations. Additionally, AR can provide optimal layout suggestions by placing virtual cubes in the intended positions, giving users the option to rearrange the cubes in the suggested layout (Layout). Similarly, these suggestions can also be used for a history feature by hinting at the previous layout in AR (History).

Summary: Tangible Data Analysis

Overview

Tangibles to display and organise individual visualisations; *touch* for basic interaction.

Classification

Concurrent, because modalities can be used in *parallel*, but *independent* from each other for different actions.

Lessons Learned

- Assign visualisations to physical objects for intuitive organisation
- Natural collaboration support by sharing cubes between users
- Physicality of the cubes prohibits certain features
- Limited display space

Applicability

- Use physical items for more natural organisation of visualisations (Layout, Workspaces)

4.1.2 VisTiles

VisTiles [77] is a visualisation framework using mobile devices, mimicking a multiple coordinated views (MCV) visualisation (see Figure 4.2). Unlike desktop MCV visualisations where different visualisations and controls are present on one screen, VisTiles distributes individual elements onto different mobile devices. The system has both *data tiles*, which display a single interactive data visualisation (e.g. scatter plot, bar chart), as well as *control tiles*, which offer control elements to configure a *data tile* (e.g. changing dimensions).

Figure 4.2: VisTiles uses mobile devices to display and control visualisations. Devices can be combined in different ways, depending on their orientation [77].

Due to the physical nature of these tiles the visualisation becomes tangible, allowing users to assign additional meaning through the organisation of the different tiles. This implicit meaning is used to link tiles together, based on their distance and orientation (e.g. placing them side-by-side). Linking a *control tile* and a *data tile* allows the *control tile* to change parameters of the connected *data tile*. Connecting two *data tiles*, on the other hand, provides the user with several different options, such as extending the visualisation across both tiles or synchronising filters and encodings. A notable linking option is *filter-by-viewport*, where the linked visualisations create filters based on the visible data. For example, zooming on a scatter plot automatically removes any data points that are no longer visible in the scatter plot on all linked tiles.

Similar to other MCV visualisations, VisTiles synchronises filters, brushes, and selections between all tiles within the same *workspace*. *Workspaces* are established based on the proxemic distance between devices, allowing for multi-user scenarios.

Multiple tiles within the same *workspace* and with the same configuration can also act in an overview and detail mode, where one tile displays the whole visualisation and a connected tile shows a smaller, more detailed viewport. For some visualisations (e.g. parallel coordinates), the detail tile can then be moved around to scroll through the visualisation, acting as tangible slider.

Although the spatial arrangement was generally well-received, and the side-by-side arrangements were rated intuitive and logical, results from a preliminary user study indicated that using these spatial arrangements for connecting data tiles and *workspaces* caused too many problems. Users had no feedback at what distance devices were linked together, or when a *workspace* was formed. Furthermore, users could accidentally activate different functions (e.g. switch *workspaces*), especially when space was scarce due to too many devices. Some users also wanted to take a closer look at some visualisations, thereby picking up the mobile device and causing existing connections to automatically disappear. Similarly, users may not want to link two devices just because two visualisations are placed side-by-side. Consequently, actions should not activate changes based on device movement alone; instead, the system should offer suggestions to the user.

Dedicating one device to one task was also beneficial, especially for offloading widgets onto a separate device. However, users also wanted to use a dedicated device to control tiles from a distance, only falling back to the physical connection as a shortcut. To address these issues, the final prototype removes the notion of *workspaces* and allows *control tiles* to select visualisations from a distance. Placing devices side-by-side no longer automatically triggers a connection but instead offers a pop-up menu with suggestions. Likewise, the prototype asks the user if the link should be broken, once two linked tiles are no longer physically side-by-side.

Summary: VisTiles

Overview

Tangible mobile devices to display individual visualisations; *touch* for interacting with visualisations; *proxemics* for linking visualisations.

Classification

Concurrent, due to *parallel* use of modalities which are used *independently*.

Lessons Learned

- Mapping visualisation to tablets is 'like working with paper' [77]
- Separation of concerns by distributing UI across different devices, leading to a more focused UI
- Proxemic distance for linking tablets together not obvious
- Unintended changes when picking up a linked device
- Number of visualisations restricted by available mobile devices

Applicability

- Use dedicated interaction devices for separation of concerns
- Control tiles as a remote tool for controlling visualisations from a distance (Remote)

4.1.3 Proxemic Lens

Badam et al. [5] created an interactive data visualisation for large display environments using lenses to explore several explicit and implicit interaction techniques (see Figure 4.3). The displays show an overview over different line charts where users can create or delete lenses to view visualisations in more detail, pan and zoom through the data in the lens, and move the lens around. Lenses can be combined which facilitates comparisons between line charts, thus enabling collaboration.

To determine which actions are suitable for implicit or explicit input, they first created two systems where each action was mapped to either an implicit input (proxemics or gaze), or an explicit input (mid-air gesture). A halo around the user's feet indicated the view direction of the user as interpreted by the system, thus making the proxemic view direction more obvious. Similarly, a line connecting two users indicated the proxemic distance between users, with static lines indicating different proxemic zones for display interaction. An initial evaluation study with 12 participants revealed that:

1. Gaze interaction (head-dwell on a line chart) for lens creation was prone to errors and led to many false negatives and false positives. This can be attributed to the *Midas touch* problem: Users need to look at the visualisation

Figure 4.3: Proxemic Lens allows users to create lenses in a large display environment and compare data from these lenses collaboratively [5].

to decide where to create a lens, but the act of looking may already trigger an interaction. Similarly, lens positioning and data zooming and panning based on head direction led to involuntary movements of the lens and was generally inaccurate. In contrast, more explicit interaction (i.e. hand gesture for lens creation) worked well.

2. Some implicit interactions (e.g. moving and scaling a lens based on body position) were rated as intuitive. Furthermore, the lines indicating different proxemic zones greatly helped to communicate the different thresholds to users. However, while the users liked merging lenses based on proximity, many did not like the automatic splitting. This mirrors the results from VisTiles [77], where users generally liked the proxemic interaction, but wanted more control over when to link visualisations.

3. With gestures, users generally felt more in control of the application. However, gestures that required prolonged interaction (e.g. moving or scaling a lens) quickly exhausted users. Although gestures were generally more accurate than their implicit counterpart, certain gesture combinations (e.g. rolling motion while pointing to create a lens) caused inaccurate targeting.

Summary: Proxemic Lens

Overview

Gestures to create and control lenses; *proxemics* for different interaction zones; *gaze* was considered but discarded for final prototype.

Classification

Synergistic, because certain actions required the *combined* and *parallel* use of both proxemics and gestures.

Lessons Learned

- Indicators for zones makes proxemic interaction more effective
- Avoid gaze interaction, as it can lead to accidental interaction
- Gesture combinations can lead to inaccurate targeting

Applicability

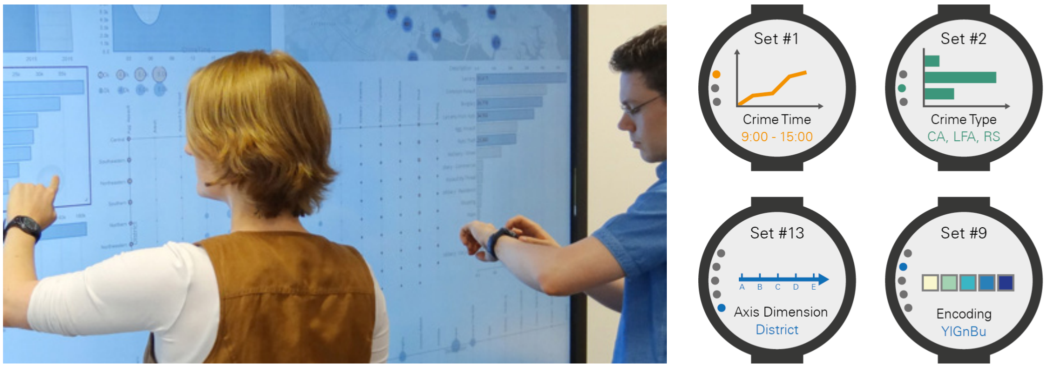
- Indicators for proxemics can be displayed in AR environment

4.1.4 When David Meets Goliath

Horak et al. [56] extend a MCV visualisation on a large interactive display with a smartwatch (see Figure 4.4). Although large displays offer several benefits for MCV visualisations, the sheer size of the display can make interaction tiring and configuration widgets may obscure the visualisations. Smartwatches can solve these problems by offloading configuration widgets to a wearable device, which allows users to interact with visualisations from a distance.

Their conceptual framework gives smartwatches the ability to push different configuration aspects (e.g. colour encoding, axis dimensions, filters) from the smartwatch to a visualisation on the large display and vice versa. The smartwatch thus acts as a personal storage device for graph configurations. The following interactions are supported: exchanging content between watch and display through horizontal swipes along the *proximodistal axis* (i.e. along the arm); scrolling through content stored on the smartwatch through either vertical swipes along the *axial axis* (i.e. orthogonal to the arm) or by utilising a physical control (e.g. rotatable bezel); touch gestures to manipulate content stored on the smartwatch (e.g. combine data, create filters) or view additional information; and arm movements for pointing gestures to select visualisations from a distance.

To establish the context of what should be manipulated, users must either touch the relevant element (when standing in front of the display), double-tap to permanently select it, or point at it (when the user is away from the display). In addition, they define four *connective areas* to further contextualise which component



(a) **(b)**

Figure 4.4: When David Meets Goliath combines smartwatches with a large display [56]. (a) Users can select visualisation through touch on the display and interact with the visualisation on the smartwatch. (b) Content can be stored on the smartwatch, such as filters, dimensions, or encodings.

of the visualisation a user wants to manipulate: the *marks* (e.g. lines and points of a visualisation), *canvas*, *axes*, and the axis *origin*. Each *connective area* offers specialised customisation options on the smartwatch. For example, selecting the *canvas* allows users to switch between different data sets stored on the smartwatch, whereas touching the *axes* allows users to change the dimension. Furthermore, by scrolling through the content on the smartwatch, the changes are instantly previewed on the selected visualisation; the changes are applied only once the user performs a horizontal swipe gesture.

Because users like to take a step back to gain an overview, Horak et al. defined three interaction zones: *direct touch*, *close proximity*, and *far distance*. The smartwatch helps to bridge the gap between these zones, as users can switch between a *distant interaction* mode and a close interaction mode (e.g. by enabling pointing gestures). Although the authors note that this switch could be enhanced with proxemics, the current system uses a double-tap gesture on the smartwatch to switch between these modes.

A user study revealed that many participants preferred this remote interaction over the close interaction mode. Most interactions with the smartwatch also occurred while looking at the large display, thus indicating that the wearable interaction is often eyes-free. Participants also used the rotatable bezel for scrolling through content, acting akin to a TUI. In general, the study indicates that the addition of a smartwatch allowed for a more flexible workflow, reduced attention switches, and was rated more seamless and intuitive than just using the large display. The smartwatch was, however, associated with an increased interaction cost.

Summary: When David Meets Goliath

Overview

Touch and *deictic gestures* for establishing context; *tangible* smartwatch for interaction; *proxemic* zones as possible extension.

Classification

Alternate, due to *sequential* use of touch for context, then smartwatch for triggering a *combined* interaction.

Lessons Learned

- Wearable device allows for *eyes-free interaction*
- For large visualisations, users prefer to step back to get overview
- Possible increased interaction cost due to wearable

Applicability

- Context-sensitive watch for remote configuration of plots (Remote)
- Avoid context switch with eyes-free interaction (Context)

4.1.5 Orko

Orko [128] uses a network visualisation to study multimodal interaction, offering both a NLI as well as touch input (see Figure 4.5). Users can use touch to interact with individual nodes directly or to pan and zoom. Additionally, users can issue data queries either through an input box at the top of the visualisation with a keyboard, or through voice commands that are activated by a keyword or by pressing a button.

Considering the inherent ambiguity of NLIs, Orko adds ambiguity widgets similar to other systems [41, 120]. The system further suggests different actions to the user (e.g. *find connections*), if no action was specified. The current query is also displayed beneath the input box, allowing users to adjust their queries.

The system supports both *sequential* as well as *parallel* use of modalities, differentiating between them based on a short time gap. Despite the support for both, study participants never used the modalities in parallel, and only 20% of the time sequentially. While this may be in part due to unsuitable study tasks for parallel use of modalities, this pattern concurs with findings from Oviatt [97].

Further findings reveal that, for the most part, users interacted unimodally, with about 30% of interactions happening through the touch interaction and about 50% through speech. Touch was used for highlighting connections and changing graphical encodings, whereas speech interaction was used mainly for searching, filtering, and topology-based tasks with multiple nodes. However, the voice recognition still failed roughly 16% of the time, causing frustration among users.

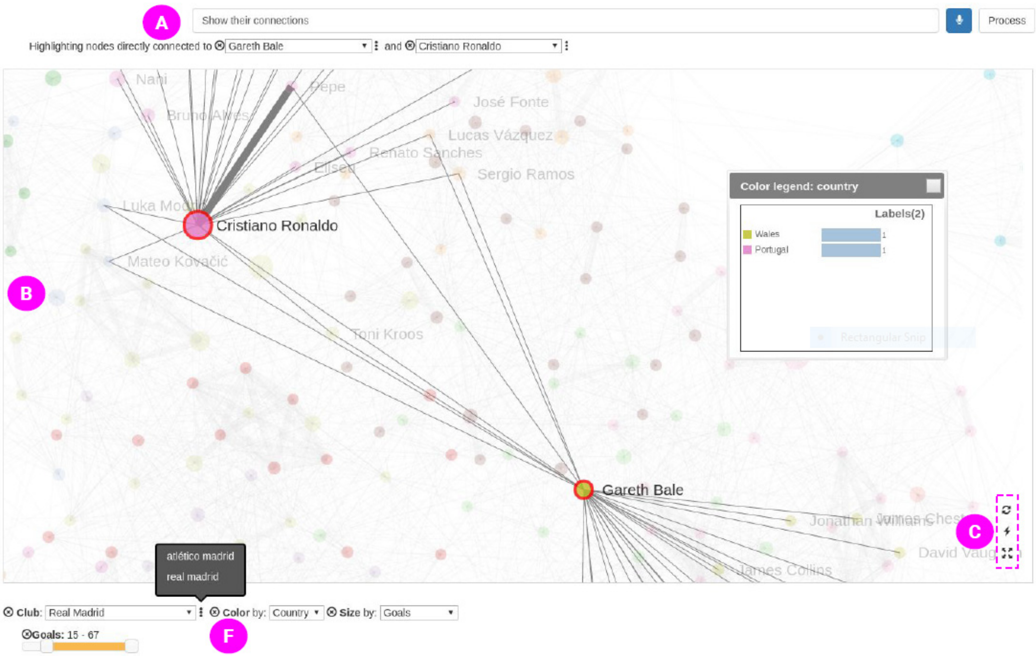


Figure 4.5: Orko uses (a) natural language queries to (b) explore a network data visualisation, while (f) widgets allow users to resolve ambiguities (adapted from Srinivasan and Stasko [128]).

Summary: Orko

Overview

Speech for data queries; *touch* for resolving ambiguities and navigation.

Classification

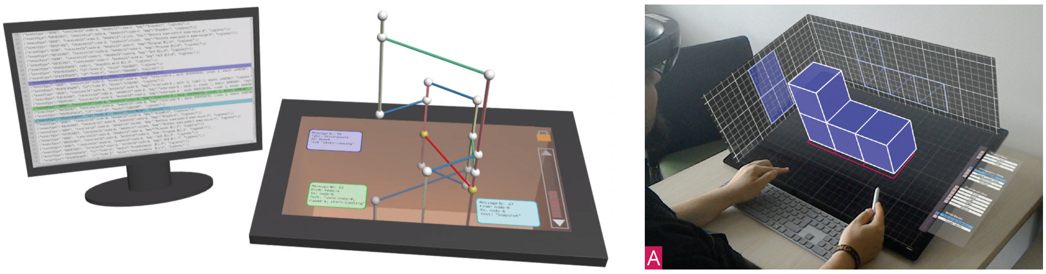
Synergistic, due to support for *parallel* use of touch and speech for *combined* fusion, allowing *contextual and follow-up* data queries. However, because users mostly used *independent* fusion with *sequential* use of modalities, system can also be considered *exclusive*.

Lessons Learned

- Users felt in control due to ambiguity widgets
- Little use of *alternate* interface (i.e. *contextual and follow-up* queries)
- No use of *synergistic* interface
- Voice recognition often still unreliable

Applicability

- Offer speech interface for data queries (Derive, Sort)
- Add widgets for adjusting current queries



(a) (b)

Figure 4.6: (a) DebugAR shows a 3D visualisation above an interactive display for debugging distributed systems [107]. (b) DesignAR allows users to create 3D models by seamlessly combining the 3D model with an interactive display [106].

4.1.6 DebugAR & DesignAR

Reipschläger et al. developed DebugAR [107] to facilitate the debugging process of distributed systems (see Figure 4.6a). DebugAR combines two displays (one touch, one conventional display) with an immersive AR environment, and a node-based 3D visualisation that sits on the touch display. Users can interact with the 3D visualisation via several widgets on the touch screen, or select a log entry from the conventional display for linking and brushing. However, no user studies were conducted to evaluate this prototype with regard to their interaction concept.

Although the similar DesignAR by Reipschläger and Dachselt [106] is not related to data analysis per se, many of the proposed interaction concepts can be also applied to data visualisations, as it represents a more elaborate interaction design than DebugAR. DesignAR also combines an interactive display with an immersive AR environment, offering multimodal input through pen and touch, mid-air gestures, and egocentric navigation (see Figure 4.6b). Thus, users can create and manipulate 3D objects by interacting with the system, and place these objects within the room. For this system, Reipschläger and Dachselt propose three different levels of spatial proximity between the AR content and the display: (L1) The centre of the display, where content is perfectly aligned to the display; (L2) the edges of the display, where content is still partially aligned to the display; and (L3) where content no longer has a clear spatial connection to the display, for example when placing objects away from the display.

By using these three levels for different tasks and modalities, the system can offer a seamless transition between AR and on-screen content. The centre of the screen (L1) therefore uses pen and touch input for creating new 3D models, either through an object browser, for sketching 2D shapes, or by tracing the outline of real-world objects and extrapolating a 3D shape. The use of a pen increases accuracy when compared to touch input, and evokes a strong association of creating 3D content; the touch input, on the other hand, allows for natural gestures for translating, rotating, and scaling content. The screen edges (L2) are mainly used for displaying widgets, such as buttons. To further take advantage of the AR environment, these widgets can be offloaded from the screen to the AR environment. When offloaded, small handles on the touch display still allow users to interact with these menus through familiar touch operation. Lastly, the AR environment (L3) is used for placing the created 3D content within the room. This task does not require precise input and thus uses mid-air gestures, which are more suited for interacting with 3D content in a 3D environment.

Summary: DebugAR & DesignAR

Overview
Pen for precise interactions with 3D content; *touch* for widgets and gestures; *mid-air gestures* for interacting with 3D content in a 3D scene; *egocentric navigation* for viewing 3D content.

Classification
Concurrent, because modalities can be used in *parallel*, but are *independent* from each other for different actions.

Lessons Learned
Not applicable: No user study.

Applicability

- Offloading menus to AR (e.g. History)
- Pen input for precise 2D input (Annotate, Filter)
- Different interaction zones for employing different tasks and modalities

4.1.7 Smartphone-Based Pan and Zoom

Büschel et al. [20] investigate different smartphone-based pan and zoom techniques for interacting with 3D data spaces. Each technique employs unimanual input (thus only one hand is preoccupied), eyes-free interaction (to focus on the AR visualisation), a high degree of compatibility (matching the physical actions to the digital response), as well as robustness and conciseness (e.g. avoid accidental interaction due to fatigue). They define three design dimensions: D1: Degree of spatiality (‘*How many DoF [degrees of freedom] are controlled through spatial input?*’); D2: degree of simultaneity (‘*How many DoF can be controlled in parallel/simultaneously?*’); and D3: degree of guidance (‘*How many DoF are controlled through gestures with some sort of alignment to give guidance to the user?*’).

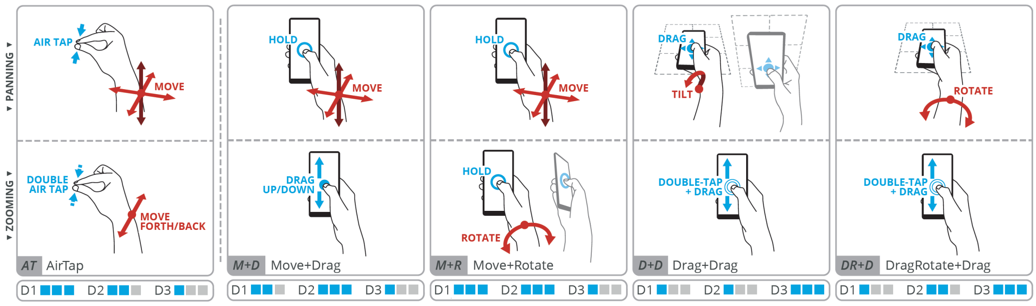


Figure 4.7: Büschel et al. compare five smartphone-based pan and zoom techniques for interacting with 3D data spaces. Indicators underneath each technique show the degree of each design dimension [20].

In total, five techniques were compared (see Figure 4.7): (1) *AirTap* uses mid-air gestures as a baseline; (2) *Move+Drag* combines touch gestures for zoom and the device's spatial position for panning; (3) *Move+Rotate* uses the device's spatial position for both pan and zoom; (4) *Drag+Drag* uses touch gestures for both zoom and pan, where the 3D panning direction is based on the device's spatial orientation (a double-tap switches between zoom and pan); (5) *DragRotate+Drag* again uses touch-gestures for both zoom and pan, except that a physical rotation gesture moves the data space up and down.

A study with 25 participants indicates that techniques with high *degree of spatiality* (i.e. *Move+Drag*, *Move+Rotate*) are well-suited for 3D interactions, feel intuitive, and allow for performing pan and zoom at the same time. A high *degree of guidance* (i.e. *Drag+Drag*, *DragRotate+Drag*) seems beneficial for 2D tasks, although the *DragRotate+Drag* technique does invoke a high mental demand. Mid-air gestures (i.e. *AirTap*) showed problems in terms of fatigue and unreliable gesture recognition. Lastly, although egocentric navigation was available, the participants did not move around, as the tested techniques were sufficient for navigation.

Summary: Smartphone-Based Pan and Zoom

Overview

Comparison between *mid-air gestures*, *touch*, and *tangibility* (i.e. device position and rotation) in immersive AR with *egocentric navigation*.

Classification

Not applicable: Comparison between different techniques.

Lessons Learned

- Use actions with high *degree of spatiality* for 3D interaction
- Use actions with high *degree of guidance* for 2D interaction
- Mid-air gestures tire out users

Applicability

- Use device's position and rotation for 3D position (Layout)
- Map touch gestures to 2D interaction tasks (Zoom)
- Use touch gestures for eyes-free interaction (Context)

4.1.8 Immersive Insights

Cavallo et al. [23] use the hybrid reality environment Dataspace [22] for the collaborative exploratory data analysis tool Immersive Insights (see Figure 4.8). The application combines several large displays to display different 2D data views with a shared interactive table in the centre and optional AR devices to further enrich the data views (e.g. with more information, or 3D visualisations). Users can assign data views to different screens, allowing for separate workspaces, while the central tabletop allows for coordination between collaborators. Without AR devices, users can either interact via touch on the displays or central table, issue complex voice commands (e.g. '*apply agglomerative clustering with 4 clusters to solution 1*'), or use physical keyboards (e.g. for filtering tasks). The AR devices further add mid-air gestures and gaze interaction, allowing users to interact remotely with the displays by emulating touch input. In addition, several sensors enable proxemics between the users and data views (e.g. showing details when a user is near a display).

Two separate studies with 12 data scientists revealed that a pure mixed reality approach is still inferior to a hybrid reality environment in terms of task duration. Furthermore, the use of mid-air gestures was problematic due to a steep learning curve, several false-positives, and, in combination with gaze for context, poor accuracy – especially when compared to touch.

Figure 4.8: Immersive Insights leverages a hybrid reality environment to display several data views, combining large displays, projectors, and immersive AR devices [23].

Summary: Dataspace

Overview

Hybrid environment supporting *touch*, *speech*, *proxemic interaction*, *mid-air gestures*, *gaze*, and *egocentric navigation*.

Classification

Synergistic when used with AR devices, as users can use multiple input modalities in *parallel* and in *combination*.

Lessons Learned

- Immersive technologies are not ready yet to replace physical devices (e.g. touch, keyboard)
- Mid-air gestures and gaze for interacting with UI elements is inaccurate

Applicability

- Support separate areas for workspaces (SEPARATE, WORKSPACES)
- Use proxemics for details on demand (DETAILS)

4.2 Summary

The presented multimodal data analysis projects explore a wide range of modalities and different multimodal interaction approaches (see Table 4.1). Several projects [20, 23, 56, 106, 107] show the benefits of using touch with 2D content or touch gestures for 3D content. Similarly, several projects [20, 56] also use these touch gestures for eyes-free interaction, which can be beneficial for interacting with 3D content. On the other hand, two projects [20, 23] advise against the use of mid-air gestures, as they easily fatigue the user, are hard to learn, and are not recognised well. The use of tangibles [20, 40, 56, 77] is often rated positively; especially as a 'personal' tangible device this can offer new interaction possibilities [56]. Proxemics have to be considered carefully, as their use may not be obvious and requires explicit indicators; similarly, gaze interaction can easily lead to accidental interactions. Yet, all projects use a diverse set of modalities, highlighting the benefits of different uses of modalities as well as different multimodal fusions.

System	Use of Modalities	Fusion	Classification	Modalities
Tangible Data Analysis [40]	Sequential	Independent	Exclusive	Touch; Tangibles
VisTiles [77]	Parallel	Independent	Concurrent	Touch; Tangibles[i]; Proxemics
Proxemic Lens [5]	Parallel	Combined	Synergetic	Gestures; Gaze[ii]; Proxemics
When David Meets Goliath [56]	Sequential	Combined	Alternate	Touch; Gestures[iii]; Tangibles[iv]; Proxemics[v]
Orko [128]	Parallel	Combined	Synergistic[vi]	Touch; Speech
DebugAR [107], **DesignAR** [106]	Parallel	Independent	Concurrent	Pen; Touch; Mid-Air Gestures; Egocentric Navigation
Smartphone-Based Pan and Zoom [20]	—	—	—	Touch; Tangibility[vii]; Mid-Air Gestures; Egocentric Navigation[viii]
Immersive Insights [23]	Parallel	Combined	Synergistic	Touch; Speech; Proxemics; Mid-Air Gestures; Gaze; Egocentric Navigation
ART [21]	Parallel	Independent	Concurrent	Touch; Egocentric Navigation

[i] Mobile devices [ii] Used in initial prototype, but not in the final prototype. [iii] Rough deictic gesture through smartwatch. [iv] Smartwatches [v] Considered for future work. [vi] Synergistic functions were not used, could therefore also be considered *exclusive*. [vii] Smartphone position & rotation [viii] Unused

Table 4.1: Overview over related work, their modalities, and their multimodal interaction classification according to Nigay and Coutaz [94].

5

Design

This chapter presents the design and features of MIDAiR. MIDAiR is a collaborative immersive analytics application that combines a spatially-aware tablet with an AR HMD for multimodal interaction with a 3D visualisation. The application has been developed with the functional requirements in mind, taking the advantages of different input modalities and the lessons learned from related systems into account. This chapter first introduces MIDAiR's visualisation, which is the heart of any visual analytics application. Since MIDAiR was developed for expert users, it is a feature-rich system. Section 5.2 therefore outlines the general interaction concept, while Section 5.3 provides an overview over all of MIDAiR's features. Lastly, Section 5.4 summarises MIDAiR's features in regards to the previously established functional requirements. This chapter has already been covered in the project report [59], but many sections have been updated due to further developments of MIDAiR.

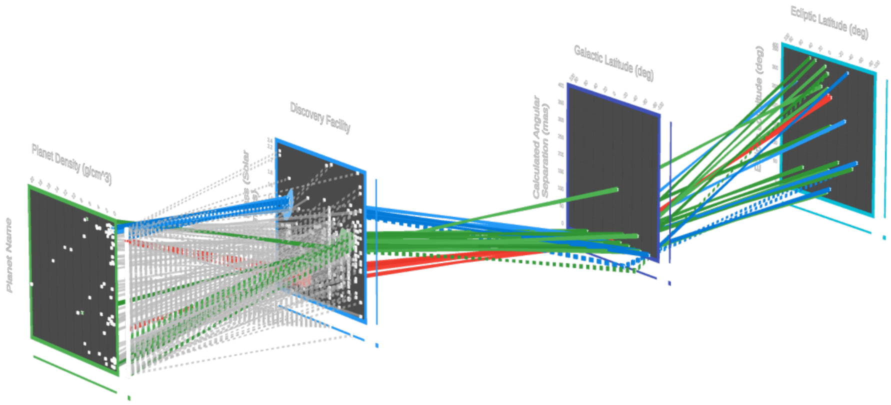


Figure 5.1: MIDAiR employs a 3D parallel coordinates visualisation that is composed of several linked 2D scatter plots. Data is filtered based on the link direction (here: left to right) and can be colourised. Dotted lines indicate null values.

5.1 Visualisation

The visualisation of MIDAiR uses several 2D scatter plots (see Figure 5.2) with linked data points, thus forming a 3D parallel coordinates visualisation (see Figure 5.1). Unlike traditional 2D scatter plots data entries with missing values (i.e. null values) are displayed beneath their respective axes instead of being left out; furthermore, the line connecting a null value is shown as dashed line. This ensures that lines between scatter plots are always visible, allowing the user to track all data entries throughout several scatter plots (akin to a parallel coordinates visualisation). The connections between scatter plots are directional, meaning that data filtered out in an 'upstream' scatter plot will not show up in a 'downstream' scatter plot, but not vice versa.

Thus, users can easily identify clusters, trends, correlations, and outliers in different data dimensions. The familiar 2D scatter plots aid the user in finding clusters and outliers between one or two dimensions: while outliers can be easily identified through their abnormal position on the scatter plot, clusters are recognised by the amount of data points in the same area. Similarly, multidimensional clusters and outliers can be identified through their line behaviour between two scatter plots. Users can track trends within the data set by following the lines across scatter plots, thus observing their differences. Lastly, correlations are visible by the amount of lines with similar behaviour between two scatter plots.

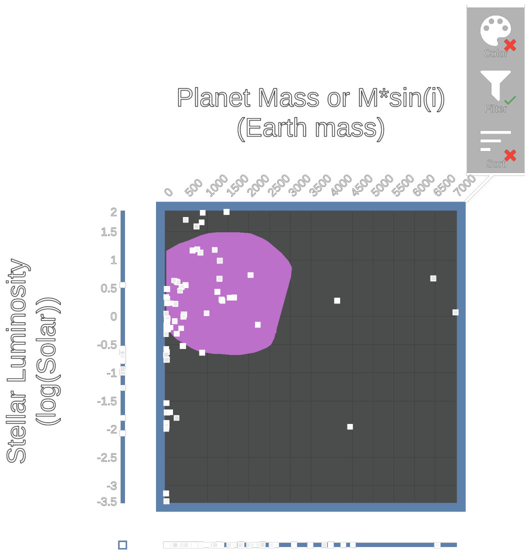


Figure 5.2: Scatter plots in MIDAiR's AR environment. Separate areas below each axis can display null values, while indicators in the top right corner show the scatter plot's attributes. Filters (purple) can be added to remove or colourise data points between linked scatter plots.

5.2 Concept

The requirements have shown that the necessary interactions can be mapped to several input modalities such as touch, mid-air gestures, voice, proxemics, tangibility, egocentric navigation, and gaze. To use all these different input modalities, the MIDAiR system combines an AR HMD with a spatially-aware tablet, providing all the necessary input options (see Figure 5.3): The AR HMD provides egocentric navigation and gaze input, which enable the usage of proxemic interactions. In addition, the HMD captures the user's voice, enabling voice commands, and provides built-in mid-air gesture support. The spatially-aware tablet, on the other hand, further adds touch input as well as tangible input, as certain actions can be mapped to the device's position and rotation within the room. The use of this tablet does, however, impede the users' ability to use mid-air gestures, as users are occupied with holding the devices. Consequently, the MIDAiR system excludes mid-air gesture interaction to focus on the remaining input modalities. A similar spatially-aware tablet has been previously explored in immersive VR environments for the use of 3D solid modelling [133].

The introduction of the tablet also provides an auxiliary output modality. While the AR environment is more suitable for 3D content, the tablet offers a more natural way to view 2D content. Although this is beneficial when interacting with 2D content on the tablet itself, it becomes cumbersome when users want to interact with 3D content in the AR scene through a menu displayed on the tablet: Users then have to simultaneously look at the tablet to find the correct interaction element, and the 3D content to perceive the triggered action. To avoid this conflict, MIDAiR distinguishes between *eyes-free interaction*, where users can operate the tablet without looking at the tablet itself, and *symbolic interaction* for interacting with 2D content on the tablet itself.

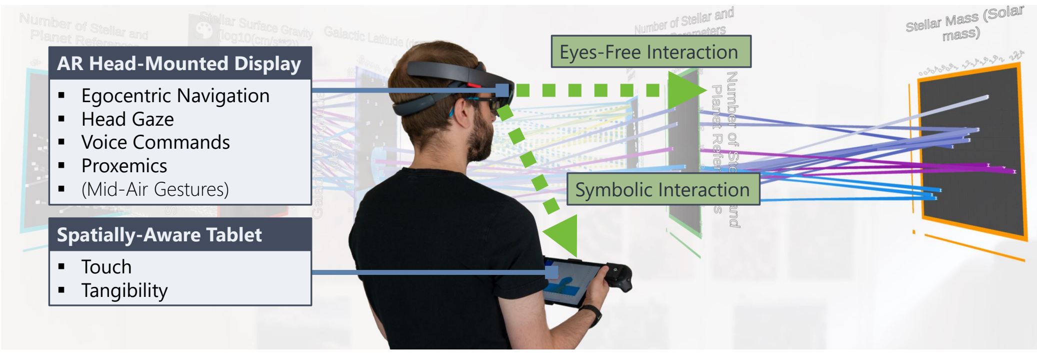


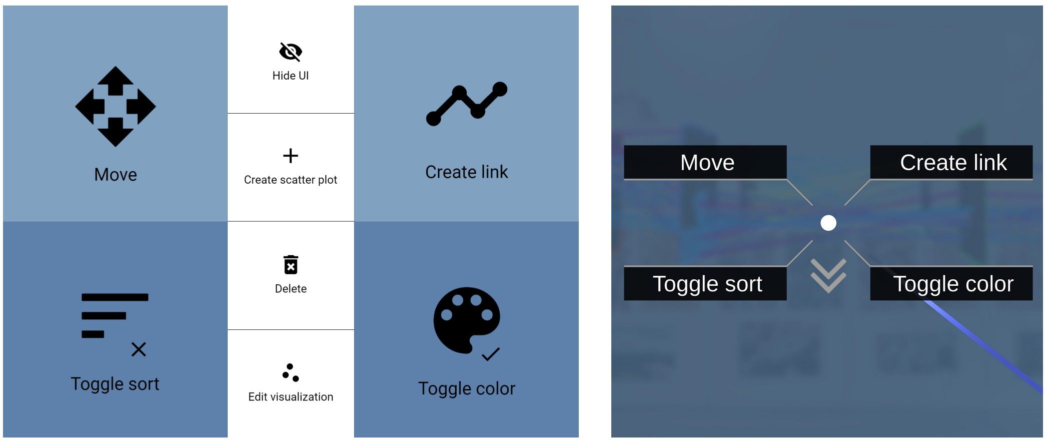
Figure 5.3: The AR HMD and the spatially-aware tablet provide different modalities. This combination also provides both eyes-free interaction for focusing on the AR content and symbolic interaction for looking at the tablet.

5.2.1 Eyes-Free Interaction

The eyes-free interaction offers menu buttons on the tablet that the user should be able to press without looking at the tablet. This is especially useful for actions that change the state of the 3D content, allowing the user to observe the action of the button press on the 3D object itself instead of concentrating on the tablet. Although eyes-free interaction menus have been employed with a variety of mobile devices [10, 11, 56, 88, 95], their use in an immersive AR environment has not yet been researched. To realise this eyes-free menu, the MIDAiR system assumes that the users hold their tablets with each hand on the left or right edge. Therefore, the user should be able to tap either the top or bottom corner of the tablet with their thumbs and without looking at their tablet, resulting in up to four large menu buttons that can be triggered by the user (see Figure 5.4a). These actions are also displayed in an AR HUD (see Figure 5.4b), as the user should not look at the tablet and is not expected to memorise these actions. Once a button has been pressed, the system responds with a confirmation sound as well as visual feedback in the AR HUD, informing the user of a successfully registered tap.

5.2.2 Symbolic interaction

In contrast to the eyes-free interaction where users should focus on the 3D content in AR, the symbolic interaction uses the tablet as both main input and output modality. Thus, users are expected to look at the tablet during interaction (see Figure 5.4a).



(a) Tablet **(b)** AR

Figure 5.4: (a) Tablet menu with four interaction zones in each corner for eyes-free interaction. Further actions are available in the middle, which require the user to look at the tablet. (b) The menu entry of each corner is displayed in an AR head-up display (HUD), so that the user can interact with the tablet without looking at it.

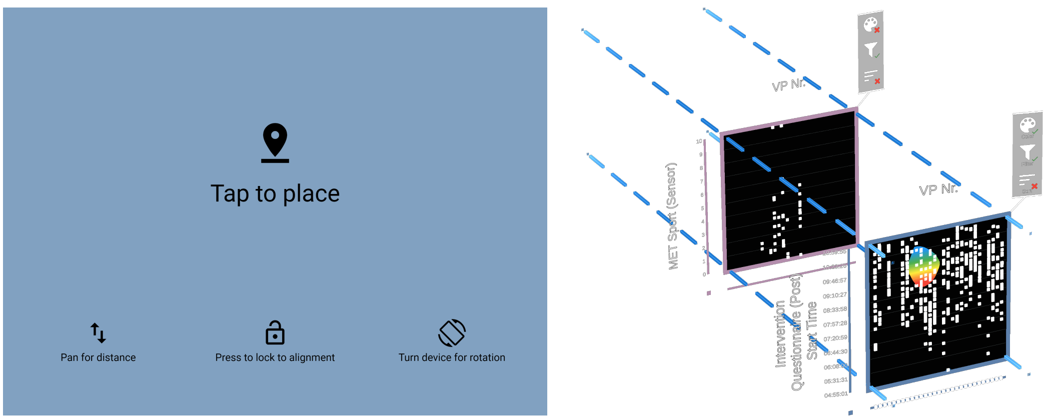
For menu buttons, this increases the effort to trigger a particular action, making it harder to trigger the action unintentionally. This is useful for actions that may significantly change the state of the visualisation (e.g. creating or deleting objects) and are therefore mapped to buttons placed in the middle of the tablet, requiring the user to look at the tablet. Similarly, menu items that lead to further actions on the tablet (e.g. showing a 2D visualisation) are also placed in the middle. Since the user may not interact eyes-free with these menu entries, the AR HUD only displays an indicator that there are more options available on the tablet (see Figure 5.4b).

5.3 Features

The MIDAiR system supports several features that both address the core requirements and aim to offer a fluent interaction with the visualisation itself: *scatter plot creation & placement*, *selection*, several *general actions* for both scatter plots and links, *creating links*, *editing the visualisation*, *tablet lens mode*, *voice commands*, *analysis mode*, and several *implicit proxemic interactions*.

5.3.1 Scatter Plot Creation & Placement

When starting the MIDAiR applications, users can either hide the AR HUD (which will be explained later on), or create a new scatter plot. By clicking on *create*, an empty scatter plot with a randomly coloured frame is created (CREATE), and the placement mode is initiated (see Figure 5.5a). In this mode, the scatter plot's position



(a) Tablet **(b)** AR

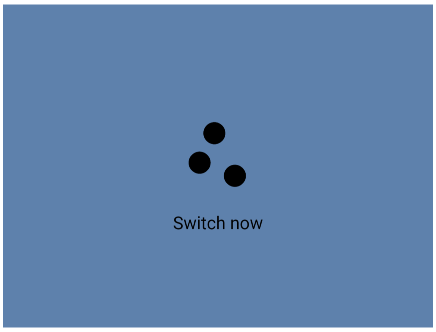
Figure 5.5: Moving a scatter plot in MIDAiR. (a) The tablet contains instructions which interactions are possible. (b) The AR environment shows several dotted lines if the scatter plot is locked to a nearby scatter plot.

is determined by the user's head gaze (Layout). Users can use touch input to pan up and down on the tablet which controls the distance between user and the scatter plot, while the scatter plot's rotation is mapped to the tablet's physical rotation. A short tap on the tablet places the scatter plot at the current position.

If there are any existing scatter plots nearby when placing the scatter plot, the moving scatter plot is automatically aligned to the nearest axis of the existing scatter plot, as indicated by dotted lines (see Figure 5.5b). Users can remove this alignment by moving the plot away from any existing scatter plots. Alternatively, users can also lock the moving scatter plot to this axis by holding one finger on the tablet, thus extending the locked axis. The scatter plot is placed once the user releases their finger from the tablet. This feature makes it easier to organise the visualisation, while still providing the necessary flexibility for any arbitrary layout (Layout). The alignment is also essential when comparing relative differences between two scatter plots, as the user could otherwise introduce a bias based on slight differences in the scatter plot's position when examining the lines (Differences).

5.3.2 Selection

Most of MIDAiR features relate to the selected object (i.e. scatter plot or link, with a link referring to all lines between two scatter plots). Once two or more objects exist, the user can switch the selection between these objects through head gaze: When looking at an unselected object, the user's cursor in AR shows a progress bar (see Figure 5.6b) that slowly fills, as well as a prompt on the tablet to 'switch now' (see Figure 5.6a). If the user keeps looking at the object (thus filling the progress bar) or



(a) Tablet

(b) AR

Figure 5.6: Selecting objects in MIDAiR. (a) The tablet becomes one large button to instantly switch selection. (b) The AR HUD shows a round progress bar around the centre cursor, indicating when the selection changes if the user keeps looking at the object. A green line points towards the currently selected object.

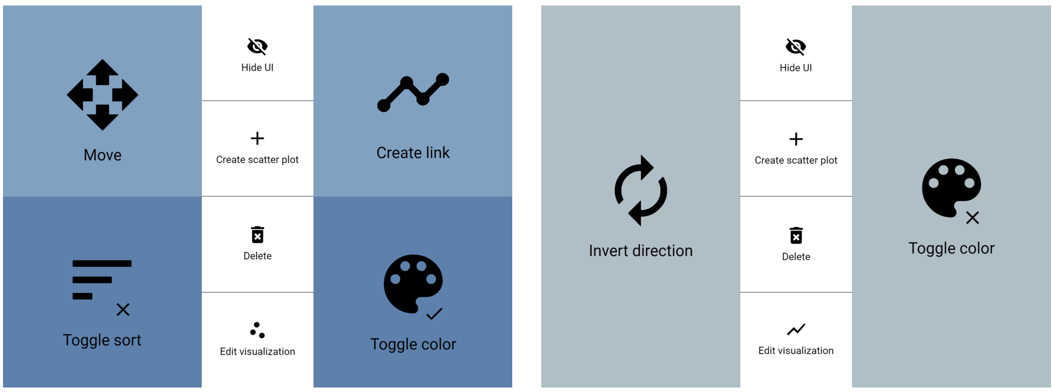
taps on the tablet, the selection is switched. This dwelling time is necessary to avoid the Midas touch problem, while the tablet allows the user to skip this dwell time.

Once an object is selected, a particle effect around the selected object appears in the AR environment, and the menu buttons on the tablet use the colour of the selected object. Furthermore, a selection line points to the currently selected object (see Figure 5.6b). This line becomes more visible the further away the user looks from the selected object and is invisible when the user is looking directly at the selected object.

5.3.3 Menu Actions

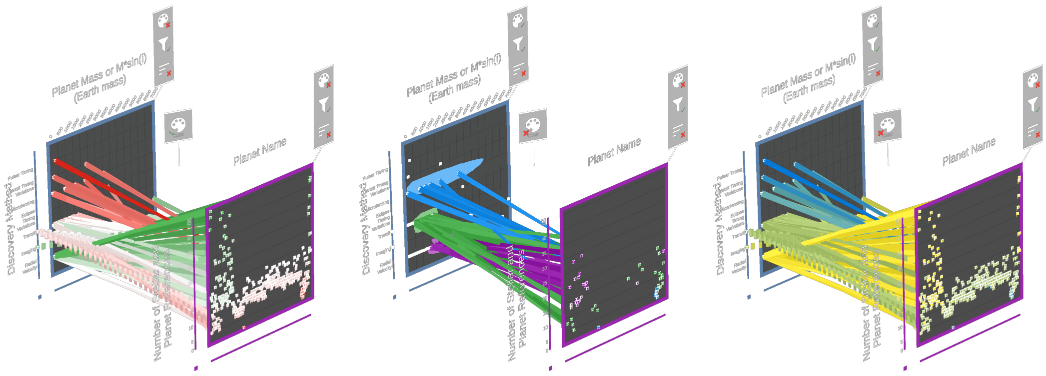
Once an object (i.e. scatter plot or link) has been selected, the application provides a menu for several general actions (see Figure 5.7). While part of this menu uses *eyes-free interaction* (i.e. move, toggle sort, toggle colour, link, invert), the menu has several options in the middle for *symbolic interaction* (i.e. create scatter plot, delete, edit visualisation, hide UI).

Toggle colour. The *colour* attribute of an object determines the data colour for all connected scatter plots and links (Brush). Although the links are directed, the colour is applied regardless of a link's direction. Only one object within a visualisation (i.e. all connected scatter plots) can have the colour attribute active – all other colour attributes are automatically disabled if there is more than one object with an active colour attribute. The colour itself is determined based on the object type and the available filters (see Figure 5.8): If a link has an active colour attribute, the colours indicate the relative differences between the two connected scatter plots (i.e. red for negative, green for positive, white for neutral line inclination,



(a) Scatter plot menu **(b)** Link menu

Figure 5.7: The tablet shows different menus depending on the selected object.



(a) (b) (c)

Figure 5.8: Data colour is determined by one object: (a) If a link determines the colour, every data point is coloured by its relative differences. (b) If a scatter plot determines the colour (here: left scatter plot), the colours of the filters are used. (c) When a scatter plot has no filter, a default gradient is used.

addressing Differences). If a scatter plot has the colour attribute, the colours are determined based on the filters within a scatter plot, or a default gradient (i.e. blue to yellow) if there are no filters available.

Move. The *move* function allows users to change the position and rotation of the selected scatter plot (as already explained for scatter plot creation, see Section 5.3.1).

Toggle sort. When sorting is active, the scatter plot disregards its X-axis dimension and instead orders the data based on its Y-axis values, thus showing a distribution of values on the scatter plot itself. This can be useful to reduce the complexity of the visualisation (see Figure 5.9).

Link. The *link* function creates a new link, originating from the selected scatter plot. This function is explained later on in more detail.

Invert. The *invert* function switches the direction of the selected link. Since links are directional, this allows the user to quickly reconfigure the visualisation without having to recreate a link.

Create scatter plot. With the *create scatter plot* button, a new scatter plot is created (explained previously, see Section 5.3.1).

Delete. The *delete* function removes the selected object from the AR environment. If the selected object is a scatter plot, any connected links are automatically deleted.

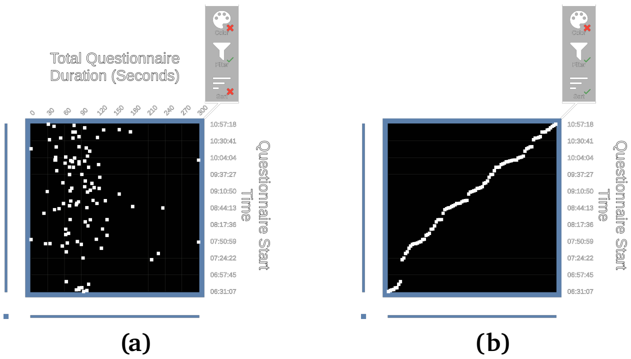


Figure 5.9: A scatter plot without sorting active (a) and with sorting active (b).

Edit visualisation. The *edit* function allows the user to view and edit the visualisation in 2D on the tablet. Given the complexity of this function, it is explained later on in more detail.

Hide UI. Lastly, users can also hide the interaction prompts in AR to focus on the visualisation itself. This function is also explained later on in more detail.

5.3.4 Creating Links

New links between scatter plots can be created by activating the *link* function from a selected scatter plot, thus entering link creation mode (CONNECT). In this mode, a preview of new lines is displayed, which is controlled by the user's head gaze and connected to the originating scatter plot. Meanwhile, existing links fade out to make the new link more visible. Once the user looks at a different scatter plot, the lines snap to this scatter plot, showing a preview of the line connection. The user can then confirm the connection on the tablet's menu (provided that the connection does not form an endless loop) or cancel the link creation. Particle effects below a link show the link's direction (see Figure 5.10).

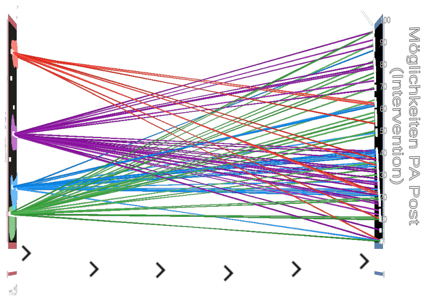
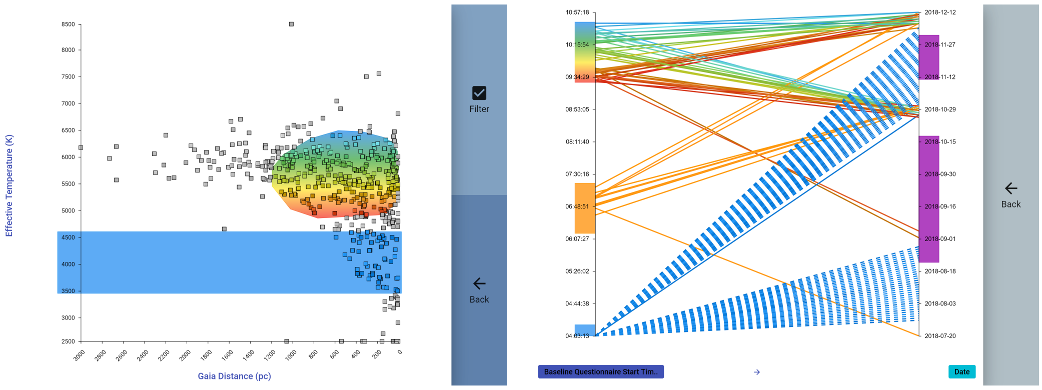


Figure 5.10: Particles beneath each link indicate the link's direction. Arrow size and arrow colour were adjusted for this document.



(a) (b)

Figure 5.11: MIDAiR can show the selected visualisation in 2D. Buttons on the right-hand side require the user to hold the tablet with both hands again (assuming that the user is right-handed). (a) A scatter plot visualisation. (b) A parallel coordinates visualisation showing the user a simplified view of a selected link.

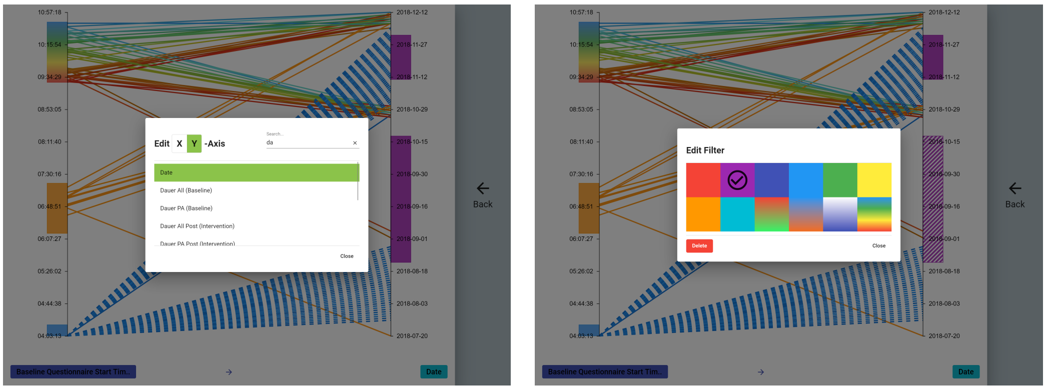
5.3.5 Edit Visualisation

One of the main benefits of using MIDAiR is that the application can display individual components of the 3D visualisation in a *2D visualisation screen*, thus combining the advantages of both worlds. When editing the visualisation the focus lies on symbolic input on the tablet – the AR HUD therefore only displays an arrow pointing towards the tablet. Depending on the selected object, the tablet shows a 2D scatter plot (see Figure 5.11a) or a 2D parallel coordinates visualisation (see Figure 5.11b). Given that the visualisations are also placed in a 3D space, the MIDAiR system adjusts the visualisation according the user's perspective (e.g. the X-axis of a scatter plot may be inverted, depending on the user's position). This is updated in real time, meaning that the visualisation on the tablet may flip around if the user moves around.

Both the scatter plot and the parallel coordinates visualisation contain a menu button to return to the main menu on their right-hand side. Additionally, for scatter plots the user can also turn off any filters, using them only for their colour (BRUSH).

To address the FILTER requirement, users can create filters on both the scatter plot as well as parallel coordinates visualisations. For the scatter plot, users can create filters by drawing them directly into the scatter plot itself. On both visualisations, users can click on an axis label or drag alongside the axis to create a filter containing values from this range on the touched axis. Filters are instantly visible on both the tablet and in the AR visualisation for all users.

Clicking on a dimension label opens up a dimension selection dialog box, in which users are presented with a searchable and scrollable list of available dimensions (see Figure 5.12a), addressing the DIMENSIONS requirement. Similarly, clicking on



(a) (b)

Figure 5.12: (a) A dialog that allows the user to search for and select the X- and Y-axis dimensions. (b) A filter dialog that allows the user to edit the filter's colour or delete the filter entirely.

an existing filter opens up a filter dialog box (see Figure 5.12b), where users can choose from several predefined colours and gradients, or delete the filter entirely. To make the selected filter more distinguishable from other filters, the selected filter is highlighted with a hatched texture on both the tablet and in AR.

5.3.6 Tablet Lens Mode

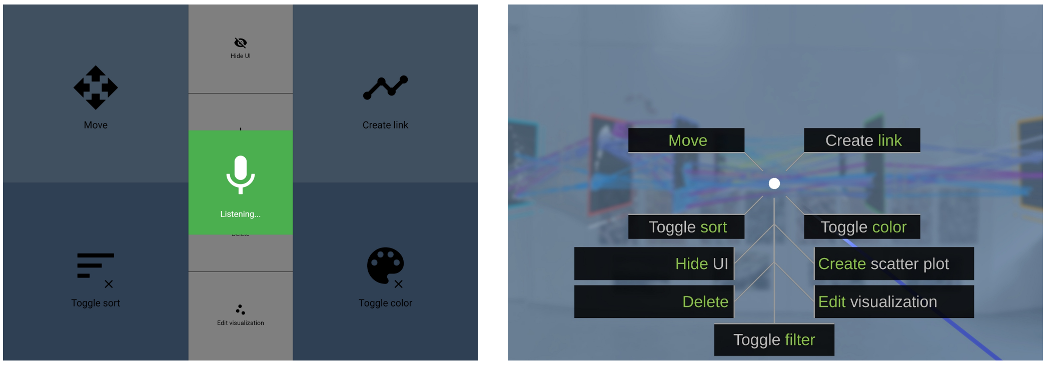
Users may also want to quickly edit a visualisation when analysing the data. However, when editing a visualisation by way of the main scatter plot menu, the user has to first look down at the tablet and can thereby lose track of the visualisation in 3D, hence losing important context of where to place the filter. To alleviate this, the user can bring the tablet up to eye-level and hold the tablet vertically, which automatically shows the selected visualisation on the tablet (see Figure 5.13). The tablet thereby acts as a 2D lens of the 3D content, resulting in a stronger mapping between the data in AR and the data on the tablet without context switch (Context). To further increase this effect, selected scatter plots will rotate towards the user, so that the visualisation in AR overlaps with the scatter plot on the tablet. The tablet thus behaves similar to a transparent prop [118], or a spatial menu [95]. The tablet will revert back to its main menu once the user holds the tablet fully horizontally again.

While this tablet *lens mode* is active, the head-based gaze selection is replaced by a tablet-based gaze selection: Instead of looking at another object via head gaze to switch selections, the user must point the tablet at another object, as if pointing the back-facing camera at the target. To make targeting more obvious, a blue cursor appears where the tablet is pointing at, allowing the user to aim more precisely.

Figure 5.13: When entering *lens mode*, selected scatter plot will rotate towards the user, thus roughly overlapping with the tablet.

5.3.7 Voice Interaction

Voice commands are available for every menu item, allowing users to use their voice instead of touch input – except for actions that require symbolic input (e.g. creating filters). To avoid accidental activation of voice commands (i.e. in collaborative scenarios), MIDAiR only starts listening while the user holds down both thumbs. While the system is actively listening, touch interaction is disabled (see Figure 5.14a), and the AR HUD displays a list of all possible actions with their respective activation keyword highlighted in green (see Figure 5.14b). These keywords help to avoid ambiguities between similar voice commands. Similar to touch, a sound confirmation and visual feedback indicate if a voice command has been activated successfully.



(a) Tablet **(b)** AR

Figure 5.14: MIDAiR adjusts the interface when voice commands are active. (a) On the tablet, an indicator shows that the application is now listening, fading out the other interaction elements. (b) The AR HUD displays all available actions, highlighting the voice activation keywords in green.

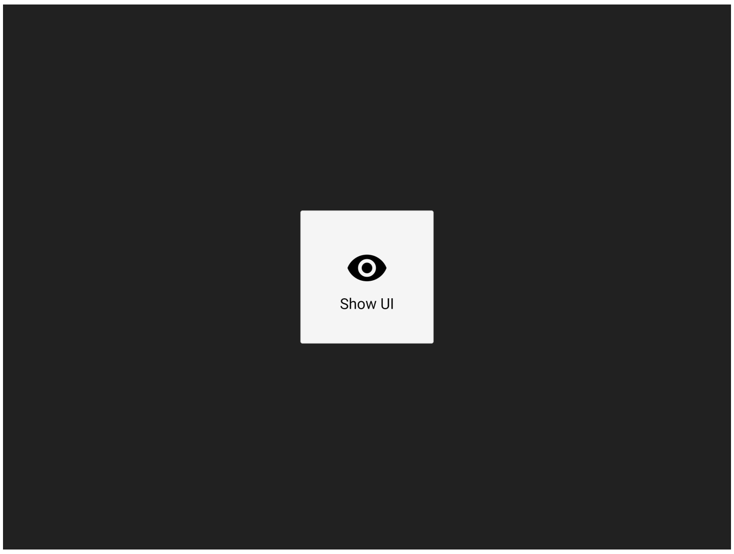


Figure 5.15: In *analysis mode*, the AR HUD is disabled, and the tablet only shows a single button in the middle to exit this mode.

5.3.8 Analysis Mode

Once the user has created the desired visualisation, the interaction prompts of the AR HUD may get in the way of the visual data analysis. To address this, MIDAiR provides an option to hide the interaction prompts, thus disabling both the AR HUD as well as the interaction on the tablet, leaving the user with an almost solely black screen, except for a single button to exit this *analysis mode* (see Figure 5.15). However, while the analysis mode is active, the user can still issue voice commands by pressing both thumbs on the tablet, which briefly shows the AR HUD with all available actions. Similarly, the *tablet lens* is still available when holding the tablet vertically. Both of these features allow the users to quickly make minor changes to the visualisation without disrupting the analysis workflow.

5.3.9 Implicit Proxemic Interaction

MIDAiR also provides several implicit proxemic interactions, which aim to make the analysis more natural. To make the axis labels more readable, the text automatically appears on the same side as the user and rotates towards the user if the user is within vicinity. Similarly, lines automatically vanish if the user is standing inside a link, allowing the user to look at the connected scatter plot. The visualisation on the tablet also automatically changes based on the position of the user in relation to the selected object, thus matching the perspective of a user (i.e. scatter plots and links are flipped if the user is standing behind them).

5.3.10 Collaboration

A substantial advantage of the AR environment is its natural disposition for collaboration. The MIDAiR system has several features to support co-located collaboration between different users, as multiple users can view and interact with the application at the same time. Actions such as positioning a scatter plot are also visible to all collaborators in real time. Furthermore, users see the cursors of other users, allowing them to know where the other users are looking at, thus making it easier to discuss specific features within the data set. The MIDAiR system also supports both *loosely-* as well as *tightly-coupled collaboration*.

Loosely-coupled collaboration is supported as each user can create their own individual visualisation in the same room (Separate, Workspaces). Users can then discuss their created visualisation or connect the visualisations of different users together to share their filters, thus combining their insights (Merge).

In contrast, *tightly-coupled* collaboration is supported as each user can independently work on the same object with their own device (Concurrency). For example, both users can simultaneously create a filter on the same scatter plot, or work on different scatter plots within the same *filter pipeline*.

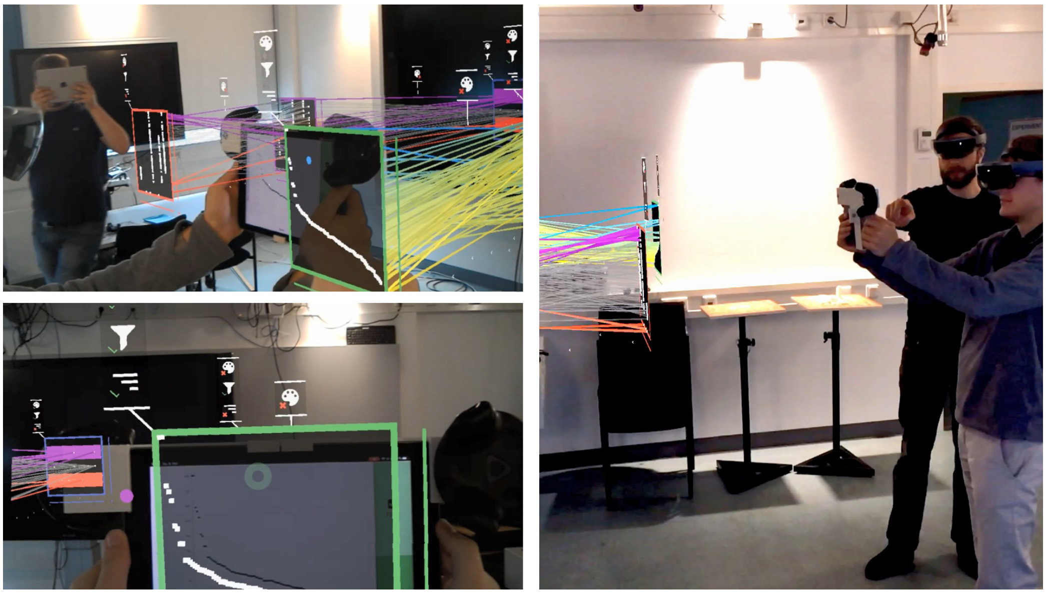


Figure 5.16: MIDAiR supports collaboration with different users: Every user can see the visualisations at the same point in space, thus allowing natural gestures such as pointing at the data. All pictures were taken at the same time from different perspectives.

5.3.11 Viewer Mode

The MIDAiR system also supports handheld AR devices, such as the Apple iPad with ARKit [3]. This mode is mainly intended to observe the active users with HMD, for example to allow study coordinators to observe their participants. Consequently, interaction in this mode is disabled, as users cannot hold a device for viewing and another device for interaction at the same time. To this extent, users may either use their own spatially-aware position to view the available visualisations as if they were using an immersive HMD, or choose the perspective of another user (i.e. independent of the tablet's position).

5.4 Summary

By combining a spatially-aware tablet with an immersive AR environment, MIDAiR offers multimodal interaction consisting of:

- *touch* (e.g. for creating filters, gestures for adjusting distance);
- *voice* commands (as alternative to many touch actions);
- *proxemic interaction* (e.g. for rotating labels towards user);
- *tangibility* (e.g. *lens mode*, rotation during scatter plot placement);
- *gaze* (e.g. for selection);
- and *egocentric navigation* (for viewing the 3D visualisation).

This combination allows for an eyes-free interaction concept using an AR HUD for interacting with the 3D visualisation via the tablet. With this, MIDAiR offers *alternate* (e.g. *combining* gaze, touch or voice, and gaze in *sequence* to create a link), *synergistic* (e.g. *combining* gaze, touch gestures, gaze, egocentric navigation, and the tablet's rotation in *parallel* to place a scatter plot), and *concurrent* (e.g. using egocentric navigation for viewing, touch for interaction in *parallel* but for *independent* actions) multimodal interaction.

Although MIDAiR currently fulfils many of the requirements specified in Section 2.2, there were several requirements that have been left open for future work: These requirements are compatible with the design and concept of MIDAiR, for example by adding another input modality, and can be added in later versions, but have not been implemented as they exceed the scope of this work. Table 5.1 provides an overview over all functional requirements.

Requirement		Description
Visualisation Interaction		
Annotate	✗	Could be realised by adding pen to tablet.
History	✗	Could be added by offering further menu actions.
Details	✗	Could be added via gaze or in the *2D visualisation view*.
Data Manipulation		
Filter	✓	Users can filter data in the *2D visualisation view*.
Brush	✓	A filter toggle may colourise data without filtering.
Derive	✗	Could be added by providing more advanced filter options (e.g. build average instead of filtering).
Scatter Plot Arrangement		
Create	✓	Users can create new scatter plots via menu.
Layout	✓	Scatter plots can be placed with gaze, touch gestures, egocentric navigation, and the tablet's rotation; in addition, the alignment feature allows for more accurate placements next to other scatter plots.
Connect	✓	Users can connect arbitrary scatter plots.
Workspaces	✓	Separate workspaces are implicitly supported as users can work on different unconnected visualisations.
Scatter Plot Configuration		
Dimensions	✓	Users can select dimensions through a touch interface.
Zoom	✗	Could be added by using touch gestures (*pinch-to-zoom*).
Context	✓	Context switches can be avoided by using *lens mode* and *eyes-free interaction*.
Navigation		
Navigate	✓	Egocentric navigation allows users to explore their visualisations, while *analysis mode* prevents distractions from interaction prompts.
Remote	✓	The use of a tablet along with selection through gaze allows for remote interaction.
Move	✗	May be realised by mapping tablet rotation to visualisation.
Inter-Plot Interaction		
Differences	✓	Differences between scatter plots can be visualised by colourising links. In addition, the alignment feature helps to keep scatter plots on the same height. The *2D visualisation view* also allows for an accurate parallel coordinates view, removing any bias due to the 3D perspective, if necessary.
Line-select	✗	Could be realised in *2D visualisation view* of links.
Collaboration		
Concurrency	✓	Separate devices give users opportunity to work on objects concurrently.
Separate	✓	Users can work independently on their own visualisation.
Merge	✓	By dynamically linking and repositioning their visualisations, users can merge and share their results.

Table 5.1: Overview over fulfilled requirements. ✓ denotes a fulfilled requirement, ✗ an opportunity for future work.

6

Implementation

This chapter outlines the prototypical implementation of MIDAiR. Firstly, Section 6.1 describes the hardware that was used to realise this prototype. Secondly, Section 6.2 provides a brief overview of the software implementation. The contents of this chapter were updated and summarised from the project report [59], which discusses the full implementation in more detail.

6.1 Hardware

MIDAiR is a distributed system, with different devices necessary for displaying the immersive environment with an *augmented reality head-mounted display*, a *mobile device* for interacting with the visualisation, and hardware for establishing the *spatial tracking* of the mobile device.

6.1.1 Augmented Reality Head-Mounted Display

Commercially available hardware for immersive AR environments is still limited to a handful of devices. Given that MIDAiR requires the user to move around, it is essential that the users can still see their physical environments unimpaired. The system therefore uses the Microsoft HoloLens (see Figure 6.1), as this HMD does not obstruct the user's view and is not bound to any cables.

The HoloLens is an optical see-through AR HMD with a holographic resolution of 2.3 M total light points (2.5 k light points per radian), and a field of view of about $30° \times 17.5°$, which makes text easily readable [55]. The device uses an internal Intel 32-bit processor (1 GHz, 2 GB RAM) and a 'custom-built Microsoft Holographic Processing Unit'. Environmental tracking is established with a combination of an

Figure 6.1: The Microsoft HoloLens is an optical see-through HMD, allowing users to immerse themselves in an AR environment [55].

inertial measurement unit, a depth camera, and four 'environment understanding' sensors. These sensors use infrared signals to map and track their surroundings.

Since the HoloLens does not include eye-tracking equipment, the gaze input is limited to head-based gaze. Furthermore, as the HoloLens relies on an internal processor, the computational capacity is severely limited, especially compared to other devices that utilise a full desktop computer. As a workaround, the amount of available data was limited, and several performance optimisations were added. Similarly, the field of view for digital content is rather restrictive. To counter this, the visualisations were scaled down to fit well within the field of view of the user.

6.1.2 Mobile Device

The mobile device acts as main form of interaction for MIDAiR. Although a small device (e.g. smartphone) could suffice for *eyes-free interaction* with menu entries, a larger screen provides more space to display and interact with the 2D visualisations. MIDAiR therefore uses Apple iPad Pro devices (see Figure 6.2) that are both lightweight and provide enough display space for the visualisations. The iPad Pro has a resolution of 2224×1668 pixels at 264 pixels per inch, with a size of around 25×17.4 cm and a weight of 477 g [61]. Furthermore, with a camera resolution of 12 megapixel and ARKit [3] support, the iPad Pro is also suitable for MIDAiR's *viewer mode*.

6.1.3 Spatial Tracking

Because the tablet contains no inherent spatial awareness, special tracking hardware was mounted to each mobile device using a custom 3D printed frame (see Figure 6.2). MIDAiR employs HTC Vive Trackers to establish the tablet's spatial position within a room. Specifically, MIDAiR uses the '2018' version, as previous versions lost their position due to infrared interference from the HoloLens [59]. Each tablet holds two trackers – one front-facing, one back-facing – to ensure that at least one tracker can provide the necessary spatial data at all times.

The trackers are approximately 10 cm wide and 5 cm high, with a weight of 89 g [57], making them easily attachable to a 3D printed frame. The Vive Trackers use both an internal accelerometer for fast internal tracking, as well as a more stable but slower external tracking using infrared base stations that are positioned in opposite corners of the room.

Although these trackers add substantial weight to the tablet, this approach allowed for quicker prototyping and provided highly accurate and stable positional data. Furthermore, the weight of the trackers may be reduced by stripping away everything but the necessary hardware, as Quiñones et al. [104] have demonstrated. Alternatively, a camera-based approach similar to Mohr et al. [90] may be used, where the front-facing camera of a device is used to track the device's position relative to a marker on the user's head. However, both approaches would require additional development time and thus exceed the scope of this work.

(a) Front **(b)** Back

Figure 6.2: For interaction, MIDAiR uses an iPad Pro tablet. To make this tablet spatially aware, a custom 3D printed frame was attached to the tablet and HTC Vive trackers were mounted to the frame.

6.2 Software

MIDAiR's software is split into four main modules, with one module for each device type (AR HMD, mobile device, spatial tracking) and a server for establishing communication between all modules (see Figure 6.3). The server also provides a module for a server frontend interface for error management, debugging, and administrative tasks (e.g. importing data) via a JavaScript-based command-line interface.

Both the module for tablet interaction and the server frontend are entirely web-based, and thus written in TypeScript [140] using the Angular [2] framework. This allowed for rapid prototyping thanks to several feature-rich UI libraries such as D3.js [16], which was used for the 2D visualisations. The software for the AR environment and spatial tracking, on the other hand, is written entirely in C# with Unity [136], as this provides the necessary support for each individual hardware and offers powerful tools for developing 3D content. Lastly, the server is written in TypeScript, running in a NodeJS [38] environment on a Windows 10 machine. The asynchronous nature of NodeJS is ideal for creating a prototypical server architecture, and a large community provides native libraries for many third-party dependencies such as SQLite [127]. MIDAiR uses SQLite databases for persistence between application restarts, as well as data store for interaction logs and network packages, allowing for a detailed reconstruction of the system usage later on. Network packages are sent using native TCP sockets between the server and Unity-based systems (i.e. the AR environment and spatial tracking), while the web environments (i.e. tablet interaction, server frontend) use WebSockets for real-time communication.

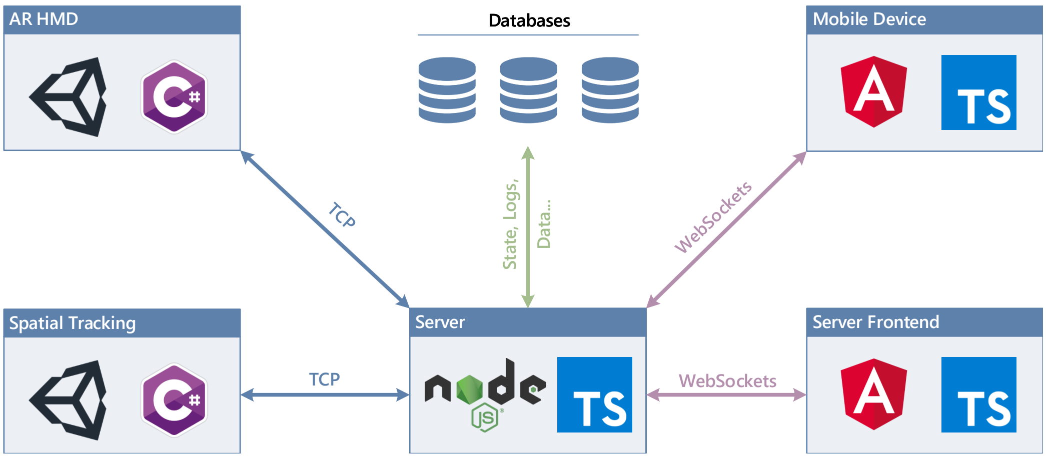


Figure 6.3: Overview over MIDAiR's software architecture with five modules (boxes) and their used frameworks and programming languages. Modules communicate over TCP or WebSockets, and several databases store application data.

7

Evaluation

This chapter reports on the evaluation of MIDAiR through a usability study. Section 7.1 defines research objectives that define concrete goals for the usability study, which is described in Section 7.2. Next, Section 7.3 presents the findings, while Section 7.4 examines the limitations of this study. Section 7.5 then discusses the results with regard to the research objectives. Lastly, Section 7.6 showcases several ideas how MIDAiR can be further improved. All quotes in this chapter were translated from German.

7.1 Research Objectives

This work aims to address three research objectives relating to *multimodal interaction*, the use of *spatially-aware tablet*, and *system usability*. Although the *spatially-aware tablet* objective is a part of *multimodal interaction*, it will be presented in its own objective to allow for a more differentiated discussion.

(RO1) **Research Objective 1: Multimodal Interaction.**

- Can users accurately select objects within the used 3D visualisation?
- Do users utilise all available input modalities?

(RO2) **Research Objective 2: Spatially-Aware Tablet.**

- Do users make use of the tablet's spatial awareness?
- Are users able to interact eyes-free with the tablet?

(RO3) **Research Objective 3: System Usability.**

- Does the system aid the users in their tasks?

7.2 Usability Study

A usability study with 8 participants was performed to address the previously defined research objectives. Participants for this study were not required to have prior knowledge with data analysis or visualisation tools. Thus, the study focuses more on evaluating the interaction with the system through pre-defined tasks, and less on evaluating a genuine data analysis workflow. The resulting user feedback will be integrated for a follow-up study with domain experts in 2020, where a genuine data analysis workflow will be analysed. The following sections describe the *participants*, the general study *procedure*, the used *apparatus*, the *tasks* which the participants had to solve, and the data *logging* used during the study.

7.2.1 Participants

The study was performed with 8 participants (4 female, 4 male) between 21–27 ($M = 24.13$, $SD = 2$), with 7 students (e.g. economics, natural sciences, psychology) and 1 Ph.D. student. None of the participants had physical disabilities hindering their movement, and none had any form of colour blindness. All participants were right-handed.

Participants were asked to rate their own proficiency with several technologies and devices on a scale from 1 (inexperienced) to 5 (very experienced). One participant did not fill out this part of the questionnaire and was thus left out for this paragraph. All participants use a smartphone on a daily basis, with a self-evaluated proficiency between 3–5 ($M = 4.29$, $SD = 0.95$). Most did not use a tablet on a daily basis, only one participant (P3) reported daily usage with a self-evaluated proficiency of 5. Usage of AR and VR devices was mixed: 4 (P2, P4, P6, P7) did not have any experience with AR, whereas the remaining 3 (P1, P3, P5) had moderate (1–3) experience with AR applications. 5 participants (P1–P3, P5, P7) were already familiar with VR devices, with experiences ranging from 1–4 ($M = 2.4$, $SD = 1.14$), mostly from VR games. One participant (P7) also participated in an unassociated data analysis study about heatmaps prior to this study.

Regarding their data analysis experience, 4 participants (P1, P5–P7) were already experienced with tools focusing on data *visualisations*, with a self-reported experience of 3–4 ($M = 3.5$, $SD = 0.58$), and 4 participants (P1, P2, P5, P6) were experienced with tools focusing on data *analysis*, again with a self-reported experience of 3–4 ($M = 3.5$, $SD = 0.58$). Only 2 participants (P1, P7) reported any experience (3–4) with 3D visualisations. In total, 3 participants (P3, P4, P8) had no experience and 5 participants (P1, P2, P5–P7) had a moderate self-evaluated proficiency in either data analysis or data visualisations.

7.2.2 Procedure

Participants were first greeted and provided with several documents that informed them about the general study procedure, a consent form, and a preliminary questionnaire for demographic data (see Appendix A). Once the participants filled out all documents, they were given an introductory presentation about the general concept of MIDAiR, as well as possibly unknown concepts such as augmented reality or scatter plots. They were then instructed to wear the HoloLens and tablet with several reminders that they could adjust the devices at any time if they felt uncomfortable. The participants then had to look at a visual marker to establish their position within the room for the system, which allowed the study coordinator to spectate their progress through the *viewer mode* on a tablet.

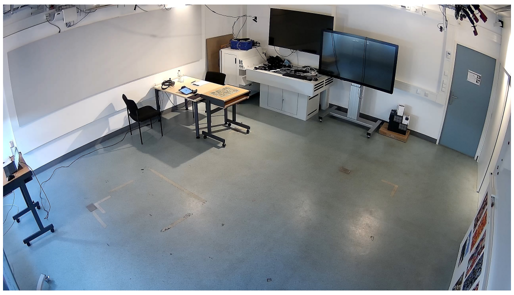
Next, the participants had to solve four artificial data analysis tasks with MIDAiR (see Section 7.2.4), where the next task was only revealed once the previous tasks had been completed. For the initial two tasks, the participants received a detailed tutorial guiding them through the main features of MIDAiR and thereby solving the tasks. The participants had to solve the final two tasks on their own, but could ask for help at any time. They were also encouraged to use the think-aloud method (i.e. to voice their thoughts). The study coordinator was able to follow the participants with a tablet and MIDAiR's *viewer mode*, either by navigating the room through the AR view or by viewing a mirror of their perspective. To solve a task, participants simply had to find a single data point matching several filter criteria and report their solution to the study coordinator.

Once all tasks had been completed, the participants were asked to fill out a questionnaire consisting of a User Experience Questionnaire (UEQ) [119] with key performance indicator (KPI) extension [52], a Simulation Sickness Questionnaire (SSQ) [72], and a Temple Presence Inventory (TPI) questionnaire [84]. A semi-structured interview was then conducted to gather qualitative data (see Appendix A). Lastly, participants received a compensation of 10€ for their efforts.

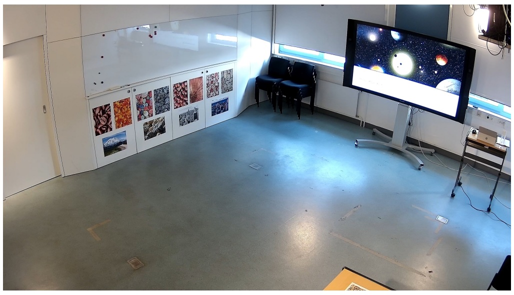
In total, study duration ranged between 40–90 minutes, with task completion time roughly between 24–50 minutes ($M = 31.14$ minutes, $SD = 9.6$ minutes) and one participant aborting the tasks prematurely due to motion sickness. 5 participants answered all tasks correctly; the remaining 2 participants provided an incorrect answer for one task.

7.2.3 Apparatus

The study took place in a spacious room (see Figure 7.1), so as to not restrict the participants in their space for placing the visualisations. The room provided a walkable area of about $4\,\text{m} \times 4\,\text{m}$ and contained a large display for displaying



(a)



(b)

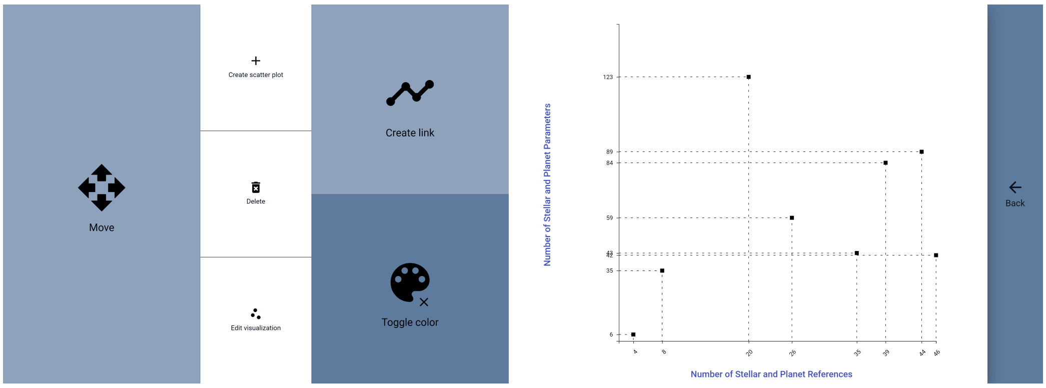
Figure 7.1: Room used for usability study. (a) A table was placed in the corner for filling out questionnaires and conducting the semi-structured interview. (b) An open area provided a large space for creating AR visualisations, while a large display showed the study tasks.

information about MIDAiR and the study tasks, as well as a table where participants could fill out the documents and questionnaires. A visual marker was placed next to the table which served as an anchor for MIDAiR, allowing multiple devices to share the same coordinate system. A small table next to the display was used to place a laptop that allowed the study coordinator to control the study procedure. Participants were given a HoloLens and a spatially-aware iPad (as described in Section 6.1), while the study coordinator used an iPad for the *viewer mode*. All devices were connected to a 5 GHz Wi-Fi network to mitigate network latency.

Because the tasks were focused on evaluating the interaction aspect of MIDAiR, and because participants did not necessarily have any knowledge with data analysis tools, several features were removed from the MIDAiR prototype: data sorting, *analysis mode*, and the filter toggle (used for colourising data without filtering). Thus, the tablet user interface was slightly simplified (see Figure 7.2a). Although the colour toggle was not necessary for the completion of any task, activating this function did not affect the data itself and was therefore kept in the system. To support these specific tasks, the prototype also shows individual data labels on the 2D visualisation once there are less than ten data points available (see Figure 7.2b).

7.2.4 Tasks

As study participants were not required to have any expertise in data analysis, simple tasks were chosen that allow the participants to interact with the system in a consistent manner (see Figure 7.3). The study was also restricted to four tasks with limited scope since the HoloLens is uncomfortable to wear over longer periods of time and could otherwise cause too many participants to abort the tasks prematurely.



(a) (b)

Figure 7.2: Several adjustments were made to MIDAiR: (a) Simplified scatter plot menu screens with missing *toggle sort* and *hide UI* buttons. (b) The 2D scatter plot visualisation displays individual ticks once less than ten data points are available.

Each task revolves around creating filters in different scatter plots until one data point remains, forcing the participants to create several *filter pipelines*.

The data set is the NASA Exoplanet Archive [93] and contains information about known extra-solar planets. This data set is highly dimensional with over 100 dimensions containing over 4000 planets and is thus ideal for MIDAiR. However, due to performance limitations and to not overwhelm participants with too many alien dimensions, the data set was reduced to about 500 planets (removed randomly from the original data set) and 22 hand-picked dimensions.

The tasks were tailored to suit the reduced data set. Furthermore, Task 2 and Task 4 build upon their respective previous tasks, allowing participants to reuse their results from previous tasks. Participants therefore received several hints to not discard their current results. To solve Task 1, participants were instructed to create three scatter plots containing the relevant dimensions (i.e. scatter plots containing *Ecliptic Longitude* & *Ecliptic Latitude*; *Number of Stellar and Planet Parameters*; *Planet Name* & *Effective Temperate*), create the appropriate filters, and link all three scatter plots together. For Task 2, participants were instructed to re-define the filters in the existing scatter plots from Task 1 through the *2D visualisation view* of a link. Participants were then left to solve the remaining two tasks on their own: For Task 3 the participants could create a second *filter pipeline* containing three new scatter plots with the relevant dimensions, similar to Task 1. Lastly, for Task 4 participants were expected to reuse scatter plots and filters from Task 2 and Task 3, since previously established scatter plots may already contain the correct filters; users would mainly need to re-link the existing scatter plots to solve this task.

GUIDED

Task 1: Find the *Planet Name* and *Effective Temperature* of the planet with the following properties:
- *Ecliptic Longitude*: Between **100°** and **200°**
- *Ecliptic Latitude*: Greater than **20°**
- *Number of Stellar and Planet Parameters:* More than **120**

Task 2: Find the *Planet Name* and *Year of Discovery* of the planet with the following properties:
- *Ecliptic Longitude*: Between **100°** and **150°**
- *Number of Stellar and Planet Parameters:* Between **100** and **120**

UNGUIDED

Task 3: Find the *Discovery Facility* and *Discovery Method* of the planet with the following properties:
- *Stellar Surface Gravity*: Greater than **4**
- *Stellar Luminosity*: Greater than **−2**
- *Stellar Mass*: Greater than **1**
- *Planet Letter*: **d**

Task 4: Find the *Planet Name* of the planet with the following properties:
- *Stellar Surface Gravity*: Greater than **4**
- *Stellar Luminosity*: Greater than **−2**
- *Ecliptic Longitude*: Between **100°** and **150°**
- *Year of Discovery*: **2017**

Figure 7.3: Tasks used in the usability study.

7.2.5 Logging

During the study, the participants were recorded in several ways: A microphone recorded audio data during the whole study, and two cameras were set up to record video from opposite corners of the room. In addition, a tablet was set up to record the participants through the *viewer mode*, thus capturing the participants within their AR environment. During task completion, both the tablet used by the study coordinator as well as the tablet used by the participants were recording their screens, the latter of which allows for detailed analysis of the user's interactions with the tablet (e.g. drawing filters). Lastly, MIDAiR logged all user interactions as well as all incoming and outgoing network messages of the server, allowing for a full reconstruction and analysis of a participant's course of action (e.g. user position, touch input coordinates).

7.3 Results

The following sections describe the qualitative and quantitative results from the usability study according to the previously defined research objectives: *multimodal interaction*, *spatially-aware tablet*, and *system usability*. Furthermore, design recommendations and further research directions for immersive analytics tools are provided.

7.3.1 RO1: Multimodal Interaction

Participants generally had no problems using multimodal interaction and liked that they had the option of using several different input modalities:

> *'I liked that [the different input methods], I think they were somewhat intuitive after a while.'* – P5

> *'I also liked that you can speak, that you can do that with the glasses [HoloLens], and with the tablet and all that.'* – P6

Use of Voice Commands

MIDAiR offers voice commands as an alternative way for activating menu items displayed in the AR HUD. Participants were introduced to these voice commands by creating a link via voice command, so each participant had to use voice commands at least once. The study coordinator also provided several reminders that voice commands were available. Generally, participants reported no issues in detection speed or quality for voice commands. However, most participants (P2, P4–P8) strongly preferred touch interaction over voice commands, though several participants (P3, P4, P6) appreciated that they at least had the option of using voice commands:

> *'I'm more of a haptic person, I want to grab things with my hands.'* – P3

> *'I think touch is somehow more intuitive.'* – P5

Only one participant (P1) expressed no preference for either voice or touch. In total, all 8 participants used a total of 36 voice commands ($M = 4.5$, $SD = 4.44$), as opposed to a total of 599 actions triggered by touch ($M = 74.88$, $SD = 37.07$), with touch actions being defined as button clicks on the tablet.

When voice command detection was active, MIDAiR showed all available commands in the AR HUD and highlighted the necessary keywords in green. This was seen as advantageous by one participant:

> *'I liked that the things [hints] were there [. . .] for a voice command that I didn't have present, I liked that I could look at it and there it is.'* – P1

But it also presented too much information and was therefore barely used: Participants often tried to speak out the whole command instead of only the highlighted keyword (i.e. '*create link*' instead of '*link*'), which caused issues due to overlapping commands (i.e. '*create*' would trigger the '*create scatter plot*' command).

Although some participants (P1, P3) did not like the idea of using voice commands in a collaborative environment, others (P6) mentioned that they would like voice commands even more:

> *'You feel weird when you say something into the room and the other one doesn't even see it.'* – P3

> *'Maybe I would even like it better, because then the other person knows exactly what you are doing.'* – P6

Several participants (P1, P3, P5, P7) were also more open to the idea of using voice commands for more complex actions, especially if it saves time. For example, participants suggested voice commands for creating a scatter plot with specific dimensions (e.g. '*Create a scatter plot with planet name and temperature*'), for creating filters with exact values, for searching for dimensions (instead of using keyboard input), or for linking two scatter plots together. While the last command is already possible with MIDAiR, it still requires establishing the relevant context via gaze.

Other Input Modalities

Several participants noted that they were missing other input modalities, such as touch or controller inputs. One participant (P3) mentioned using gesture interaction for creating links between scatter plots:

> *'I always wanted to do that with my hands.'* – P3

However, another participant (P7) voiced concerns that gestures may not be recognised by the system, and would thus take longer. Of note is that this participant did activate the HoloLens bloom gesture on accident during the study, and took two attempts to close the menu via gesture again. Instead of gestures, this participant proposed the use of VR controllers for interaction, which also offer more tangible input modalities (e.g. joystick, triggers) than touch input. This participant did, however, prefer touch interaction for interacting with the 2D scatter plot.

Head Gaze Selection

While participants were generally able to select objects via head gaze input, there were several corner cases that caused problems. Especially in more complex visualisations, selecting objects hidden behind other objects proved difficult, though the physical navigation allowed the participants to circumvent this issue. Other participants (P6, P8) were also confused by the hitboxes for links (e.g. the invisible box around each link to facilitate selection, regardless of how many lines are visible), and thought they had to look at the small, visible lines to select the link. As such, they thought it became harder and harder to select the link once there were less lines visible. Another participant (P4) revealed that they wanted to change selection by swiping on the tablet, which could be offered as alternate input modality instead of gaze selection.

Selection indicators. Most participants (P1, P2, P4–P8) reported that it was generally clear which object was selected, in part thanks to the AR particle effects. Especially the matching colour on the tablet was well-received and helped to confirm which object was selected:

> *'What I liked, that it became instantly clear that you're now on the green scatter plot, because your tablet's background was now green, and then I could instantly switch, I really liked that.'* – P7

Some participants (P5, P8) also wanted additional indicators to make the selection more obvious (e.g. audio feedback).

Midas Touch. Participants reported no unwanted selections (i.e. Midas Touch problem) when selecting objects via gaze. Although participants did unintentionally start the selection progress very often, the participants could quickly look away before

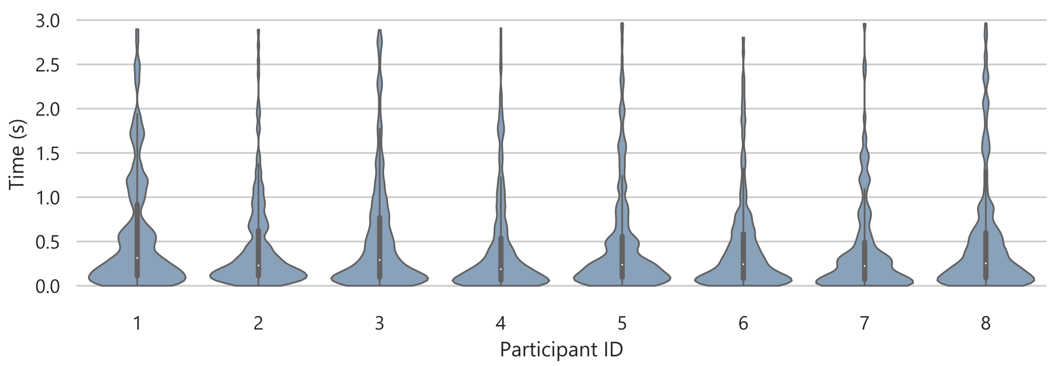


Figure 7.4: Distribution (bandwidth = 0.1) of after how many seconds the selection was aborted by looking away.

the selection was completed. The majority of such selections were therefore aborted very early ($M = 0.48$ s, $SD = 0.59$ s, see Figure 7.4).

While there was no Midas touch problem for unwanted selections, the changing indicators on the tablet (i.e. showing a *switch now* button) caused some distractions:

> *'When I wanted to edit something and looked [at the 3D visualisation] again, then it started to select something different.'* – P2

> *'I was sometimes in the menu, when I wanted to click something it was sometimes briefly red or briefly purple.'* – P6

Furthermore, the long dwell time caused several participants (P5, P6) to look away prematurely, thereby aborting the selection briefly before completion.

Selection skipping. MIDAiR also offers to skip the long dwell time by touching the tablet once the selection has started. All participants tried this selection skipping at least once as part of the tutorial (see Figure 7.5). This skipping was liked by

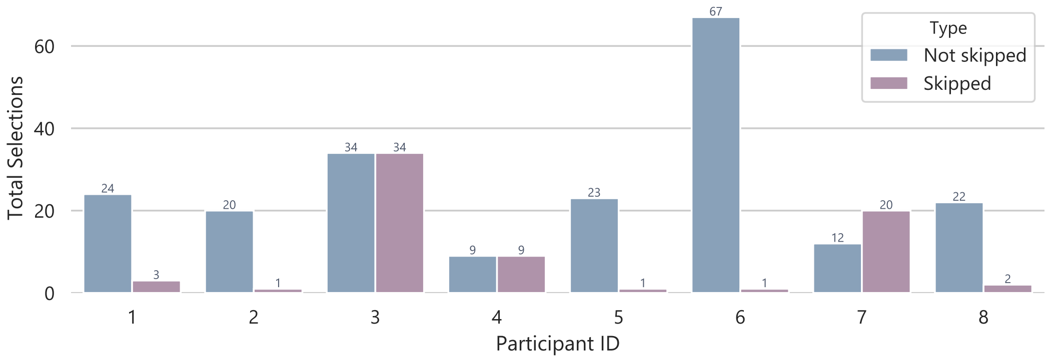


Figure 7.5: Total amount of how often the selection process was (not) skipped.

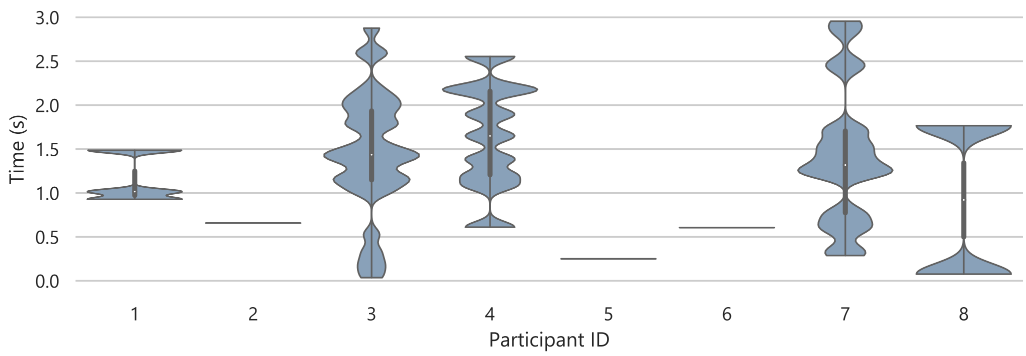


Figure 7.6: Distribution (bandwidth = 0.1) of after how many seconds the selection was completed by tapping on the tablet.

participants who used it, as it allowed for quick error correction even if the participant selected the wrong object:

> *'I liked the dual options of waiting and tapping.'* – P4

> *'That was very cool, that I could tap on that [the tablet] and then it directly finished loading, I liked that.'* – P7

Aside from trying out the manual skip during the tutorial, participants either used the skip function at least as much as waiting for the gaze dwell to complete (P3, P4, P7), or virtually not at all (P1, P2, P5, P6, P8). Most of these manual skips occurred at around 1.5 s ($M = 1.43$ s, $SD = 0.7$ s, see Figure 7.6).

Temple Presence Inventory

The TPI questionnaire [84] was used to measure the perceived realism and immersion of MIDAiR. Since no comparison was made against other systems, the following paragraphs will describe the results of the individual items (see Figure 7.7). One participant was removed from these results due to (presumably) forgetting to fill out this questionnaire, and one participant was removed due to aborting the study.

Generally, participants felt moderately involved and immersed in the system. Furthermore, participants felt that the experience was more like looking through a window, which is due to the small field of view of the HoloLens. Objects were

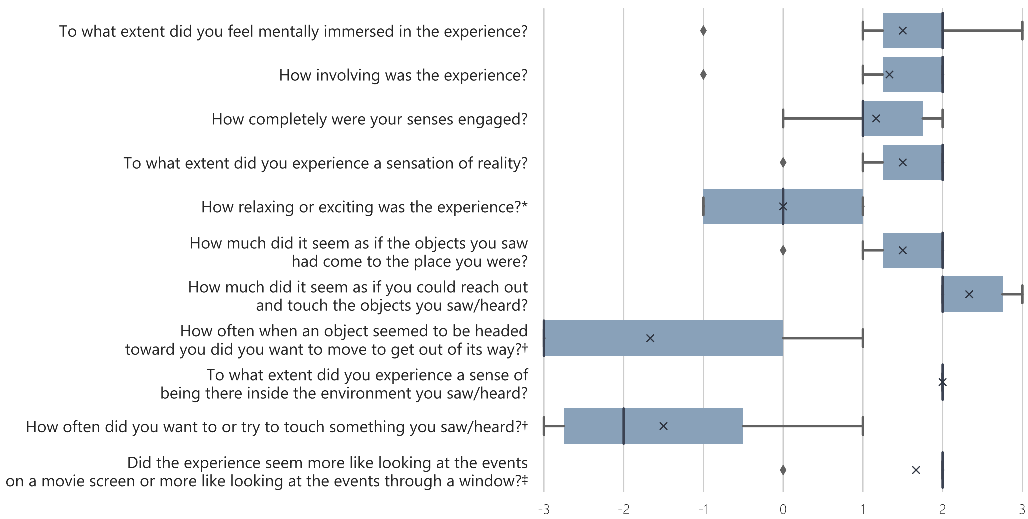


Figure 7.7: Results of the TPI questionnaire. Diamonds indicate outliers, crosses indicate averages. The scale ranges from '*Not at all*' (−3) to '*Very much*' (3), except for: * '*Very relaxing*' (−3) to '*Very exciting*'; † '*Never*' (−3) to '*Always*' (3); ‡ '*Like a movie screen*' (−3) to '*Like a window*' (3).

mostly perceived as *being there*, but participants did not feel that they wanted to move out of the objects' way. Similarly, while participants felt strongly that they could touch the objects, they felt little need to touch something they saw. This may be task related, as the objects were stationary in front of the participants, and they were instructed to interact with the tablet, or because they had to hold the tablet.

Use of Augmented Reality Environment

Participants generally liked the use of AR, and saw no benefit of using a VR environment instead, especially as one does not have to worry about bumping into other people or objects. However, several participants (P5, P6, P7) mentioned that a special room was necessary, as their own desk would be too small or too cluttered. Participants (P3, P4) therefore liked the rather sterile room and felt that a white background was necessary. With AR, there is also an emotional connection to the feeling of a workplace:

> *'You still have the feeling of somehow sitting in a workplace.'* – P7

> *'You get a little bit immersed and concentrate on that [the visualisation], and then I think it's good that you have a relation to the real world.'* – P3

While the participants liked that they still could physically interact with the environment, they saw no benefit of using real-world objects (e.g. tables, walls) as anchors for attaching the visualisations. Only one participant (P7) liked the possibility of realigning a scatter plot's orientation to, for example, match a wall, or use predefined anchors (e.g. visual markers) in a workspace for attaching visualisations.

Egocentric navigation. The movement patterns while solving the tasks also show several differences between participants (see Figure 7.8). There are some indicators where participants with more complex visualisations (P1, P7) moved around more than participants with simpler visualisations (P4, P8), as the latter only moved from side to side. However, with the data currently at hand, no definite statement can be made in this regard. Yet, all participants did move around and did not stay still: This movement sometimes happened explicitly (e.g. participants stepping closer to zoom in), and sometimes implicitly (e.g. slightly moving around to select an object). One participant rated the possibility to move around positively:

> *'I like that you can move around, because otherwise you just sit there. [...] You can actively work with it.'* – P3

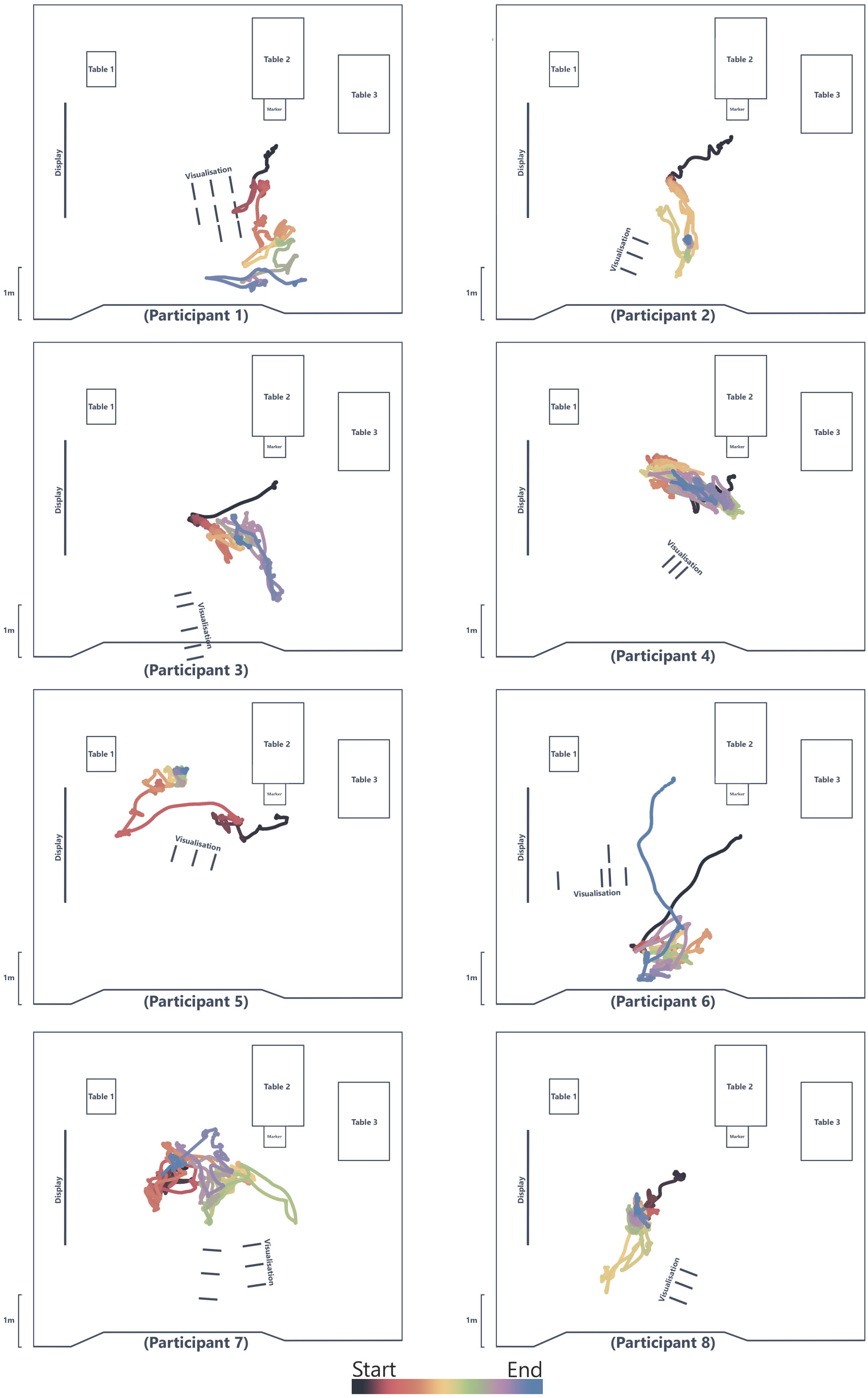


Figure 7.8: Top down view of participants' movement in the room, colourised by time. The participant's final visualisation is marked with black lines.

Design Recommendations

Immediately-visible progress bar for gaze selection
An immediately-visible progress bar of the dwell time helps users to know if an object is being selected, allowing them to abort unwanted selections.

Skipping dwell time
Providing an explicit trigger for skipping dwell time allows expert users to quickly change selection.

Offer large space for complex visualisations
Users need room for egocentric navigation around the visualisation, especially if gaze-based selection is employed.

Further research directions

Selection in complex visualisations
(e.g. selecting objects hidden behind other objects)

Hitbox visibility
(e.g. showing an outline of where users have to look to select an object)

Space requirements for AR visualisations
(e.g. using smaller visualisations in a small office workplace; or offering alternative navigation techniques [20])

Leverage natural language processing for complex voice commands
(e.g. similar to Orko [128], Immersive Insights [23])

7.3.2 RO2: Spatially-Aware Tablet

Participants generally liked the tangibility of the tablet, and that it allowed them to manipulate objects from a distance:

> *'That you can link this with your tablet, that's like a kind of remote control, that's really good.'* – P8

Although one participant liked the tablet interaction, the participant preferred to put the tablet away when not in use.

Simulator Sickness Questionnaire

One concern was that users constantly switched between the AR visualisation and the tablet. Participants thus had to move their head up and down as well as change their focus, which could lead to motion sickness. The SSQ [72] describes three different dimensions of simulator sickness (nausea, oculomotor issues, and disorientation) by measuring the severeness of 16 symptoms on a scale from 0 (*not at all*) to 4 (*strong*). The measured dimension scores range from 0–200.34 for nausea, 0–159.18 for oculomotor issues, 0–292.32 for disorientation, and 0–235.62 for the total score.

In general, most participants only suffered from *general discomfort*, *eyestrain*, *difficulty focusing*, *sweating*, *fullness of head*, and *dizziness (eyes closed)* (see Figure 7.9). The results show a nausea score of 23.85 ($SD = 38.16$), an oculomotor score of 23.69 ($SD = 27$), and a disorientation score of 57.42 ($SD = 72.69$), with a total score of 36.47 ($SD = 46.62$). One participant experienced strong motion sickness and had to abort the study prematurely. However, it is unclear whether these results are caused by the AR hardware itself, or if they are exacerbated by the use of the tablet. For example, the high scores for *general discomfort* and *fullness of head* can be attributed to the uncomfortable AR hardware itself, especially as the used translation of *fullness of head* ('*Kopfdruck*') may have been misunderstood as *head pressure*.

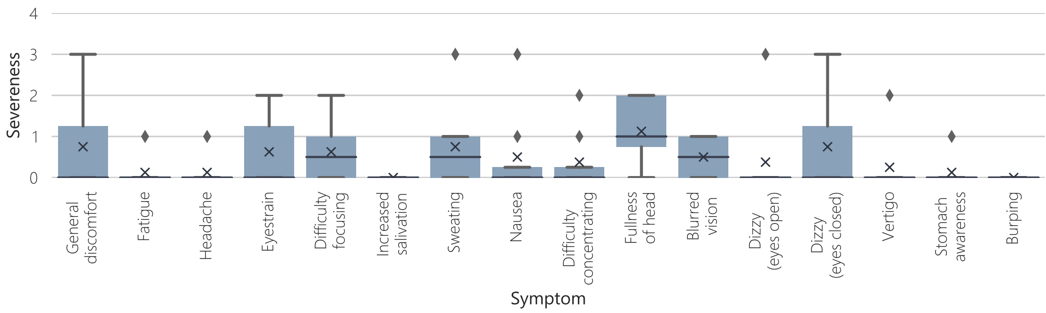


Figure 7.9: Distribution of results for the SSQ from 0 (*not at all*) to 4 (*strong*). Crosses mark average values, diamonds mark outliers.

Fear of Dropping Tablet

Another concern was the relatively heavy weight of the tablet due to the added Vive Tracker hardware, as this could tire out the users. However, most (P1–P7) participants were not particularly bothered by the extra weight and did not feel tired after completing the tasks. Several participants (P1, P6, P7) did find the tablet somewhat heavy when holding the tablet in one hand, for example when drawing filters or using the on-screen keyboard, while one participant (P7) was irritated by the tracker hardware.

Several participants (P1, P3, P4, P7) also voiced concerns about accidentally dropping the tablet, especially when holding the tablet in one hand:

> *'You have to be careful not to drop it [the tablet] when you interact with the 2D visualisation.'* – P4

The participants indicated that this problem was not related to the weight of the tablet, though that a lighter tablet would curb this problem. One participant (P7) therefore preferred VR controllers as input modality, while others suggested a shoulder strap, which could also free up their hands when the tablet is not in use.

Tablet Rotation

MIDAiR allows users to activate lens mode for quickly accessing 2D visualisations of selected objects by holding the tablet vertically, and to rotate the scatter plots during placement by rotating the tablet. For both features, it is essential to analyse how participants used the tablet's rotation to their advantage.

Participants generally did not use the rotation during placement after placing the first scatter plot. Rather, they used the existing scatter plots as anchors for the alignment feature, ignoring the tablet's rotation entirely. One participant (P7) did, however, use the rotation feature extensively for placing two scatter plots, but found touch interactions awkward to use when trying to rotate the scatter plot (and thereby the tablet) at a 90° angle.

The spatial awareness also did not take any contextual information into account, which caused problems for one participant (P7): Instead of holding the tablet horizontally during placement, the participant tried to match the tablet's physical rotation with the scatter plot (i.e. holding the tablet vertically). This activated MIDAiR's *lens mode* and aborted the scatter plot placement, which confused the participant.

While interacting with MIDAiR, most participants (P2–P8) generally held the tablet in a horizontal or moderately angled (roughly <45°) position, but did not vary too much when holding the tablet (see Figure 7.10). Users therefore did not accidentally activate *lens mode*, except when moving a scatter plot.

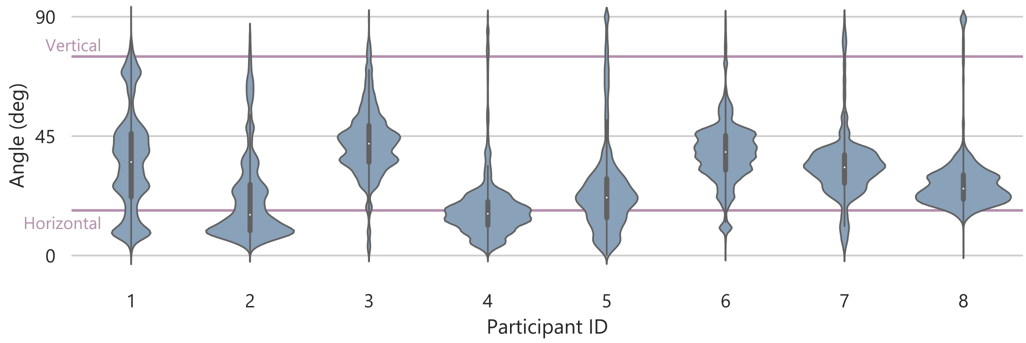


Figure 7.10: Distribution (bandwidth = 0.1) of tablet rotation: 0° means tablet's display is facing the ceiling (tablet is horizontal), 90° means the display is facing the user (tablet is vertical). Purple lines indicate thresholds for activating *lens mode* (vertical), or resetting the *lens mode* and going back to the menu (horizontal).

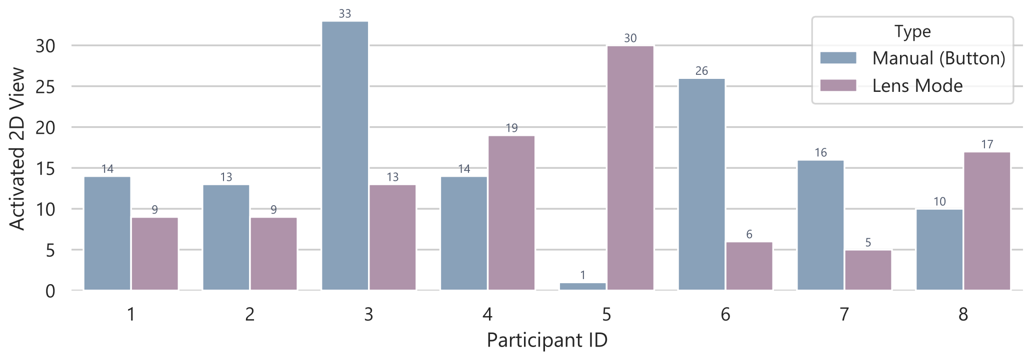


Figure 7.11: Total amounts of how often participants switched to the *2D visualisation view* manually via touch, or physically through *lens mode*.

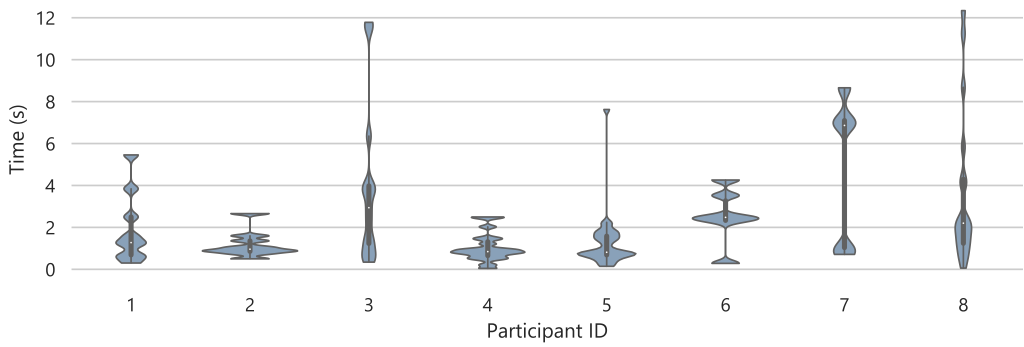


Figure 7.12: Distribution (bandwidth = 0.1) showing how much time each participant spent holding the tablet vertically (i.e. in *lens mode*).

Lens mode

MIDAiR allows users to enable the *2D visualisation view* by holding the tablet vertically, thereby activating *lens mode*. Most users frequently made use of this mode (see Figure 7.11), with one participant (P5) only using this mode to switch to the *2D visualisation view*.

> *'I think the tilting was intuitive, because the motion was virtually reproduced in the three-dimensional [space].'* – P4

> *'Once you know that when you tilt it, it activates the 2D view, then you eventually do that automatically [...] in the end, it was actually totally normal.'* – P3

However, most participants did not like holding the tablet vertically for too long, and reverted back to holding the tablet at a more comfortable position. Rather, participants used the *lens mode* only briefly, similar to a trigger for activating the *2D visualisation view* on the tablet (see Figure 7.12).

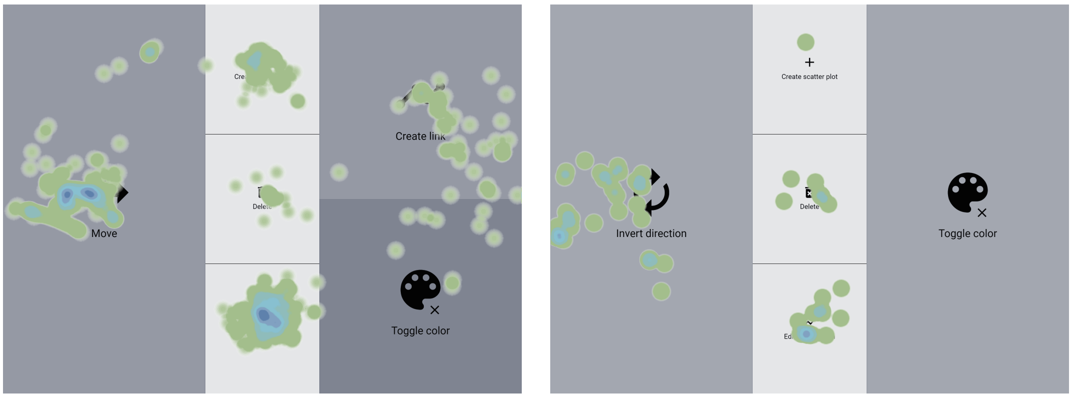
Eyes-Free Interaction

MIDAiR allows users to interact eyes-free with the tablet by offering various prompts in an AR HUD. It was uncertain whether participants would be able to use the eyes-free interaction, or if they still preferred to look at the tablet during interaction. Although the concept was generally clear to all participants, several participants (P4, P6) reported looking at the tablet during interaction out of habit. Yet, most participants liked the eyes-free interaction, especially due to the oversized buttons:

> *'Conciously I never had to look at the tablet [...] I thought the two big buttons on the left and right side were really intuitive and practical.'* – P1

> *'To fully operate this blindly you would need more time to get used to it, because I still had the feeling that I had to look at it [the tablet] to click on the right point, but actually it's made so that you don't need it.'* – P5

When using the tablet eyes-free, participants had no problems pressing on the right buttons, thanks to their large size (see Figure 7.13). Although clicks naturally gravitated towards the button's icon, most of the eyes-free touches were located in the outer half of these buttons. There are, however, several outliers: While most participants initially held the tablet with both hands during the menu screens (see Figure 7.14a), several participants increasingly held the tablet with one hand as the study progressed (see Figure 7.14b). These participants therefore interacted with the menu using their index finger as opposed to their thumbs.



(a) Scatter Plot Menu **(b)** Link Menu

Figure 7.13: Heatmap showing where participants tapped on the tablet.

(a) **(b)**

Figure 7.14: Participants holding tablets during menu screens. (a) In the expected pose, participants hold the tablet firmly in both hands, interacting mainly with their thumbs. (b) Over time, some participants shifted to holding the tablet with one hand and only using their index finger for interaction.

When looking at the tablet, some participants (P1, P6–P8) peeked at their tablet beneath the HoloLens, or held the tablet in front of them (see Figure 7.15), while others (P2–P5) often measurably looked down at their tablets.

Lastly, MIDAiR also displayed the available interaction possibilities in an AR HUD, which should facilitate eyes-free interaction. Although participants initially still had to look at the tablet to view which commands were available, the HUD served as useful reminder later on:

> *'I did actually use it [the eyes-free interaction] once I knew [. . .] where each button is located.'* – P7

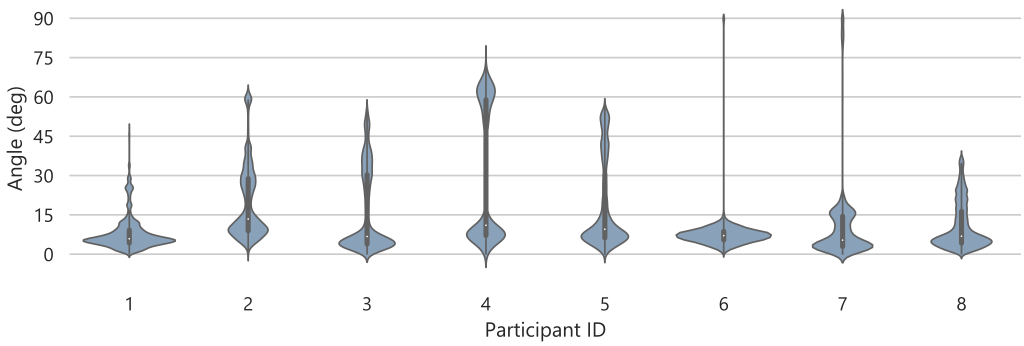


Figure 7.15: Distribution (bandwidth = 0.1) of each participant's HoloLens angle, as measured between the upward vectors of the HoloLens and the scene. An angle of 0° therefore means looking forward, while 90° means looking at the floor.

Design Recommendations

Use spatial triggers for actions
Flipping the tablet was used as an explicit action, replacing the touch interaction (high degree of compatibility [20]).

Use eyes-free interaction for 3D content
Especially for scatter plot placement and linking, all participants used eyes-free interaction for interacting with 3D content.

Further research directions

Investigate motion sickness due to switch between tablet and AR
(i.e. was motion sickness caused due to AR HMD, or due to context switch?)

Support of more spatial gestures with the tablet
(e.g. flipping tablet as menu action, see Surale et al. [133])

Ad hoc tablet creation
(i.e. creating interactive surfaces when necessary from arbitrary objects, using the HMD's gesture sensors to detect touch; see e.g. Corsten et al. [30])

Longitudinal study for eyes-free interaction
(i.e. break habit of users looking at tablet during interaction to measure practicality of eyes-free interaction)

Comparison between eyes-free interaction techniques
(e.g. using gestures [20, 56], radial menus [11], or spatial menus [81, 95])

7.3.3 RO3: System Usability

In general, participants were positive about using the system, and felt that it had 'much potential' (P3, P6) and that it was 'interesting' (P3, P4, P6), 'cool' (P7), but also 'complex' (P8):

> *'It was so much fun to go through these things [tasks] with it.'* – P1

The system's quick reaction to all commands was also appreciated:

> *'What I liked, it reacted very quickly and was always there wherever I brought it with me.'* – P6

Habituation Period

Despite the positive consensus, all participants mentioned that the system initially took some time to get used to:

> *'Like learning how to ride a bike, in the beginning you'll have to look, but then it's fine.'* – P3

> *'As soon as you performed each action twice it actually was very intuitive and you could get a feel for it, I thought that was great.'* – P7

This habituation period can be partially explained by the amount of new terms that the participants had to learn before solving the tasks: Several participants mentioned that there were 'many new terms' (P4), and that it was 'initially overwhelming' (P6). Another factor is the combination of new technologies (i.e. an AR HMD) which most participants have not used before, as well as a novel interaction concept (i.e. eyes-free using a spatially-aware tablet) that culminated in an overload of information for all participants. Several participants (P5, P8) therefore valued the initial introduction (i.e. presentation and guided tasks), but wished for an even more interactive tutorial.

User Experience Questionnaire

The UEQ with KPI extension was used to gain more insight into the system's usability. One participant was left out of this evaluation due to a printing error, while another participant was removed from the data set due to aborting the study prematurely. Analysis of the UEQ shows excellent results in stimulation and novelty, good to average results in attractiveness and efficiency, and below average results for dependability and perspicuity (see Figure 7.16), and a KPI of 1.54 ($SD = 0.4$). Although attractiveness, stimulation, and novelty received good scores, this could also be due

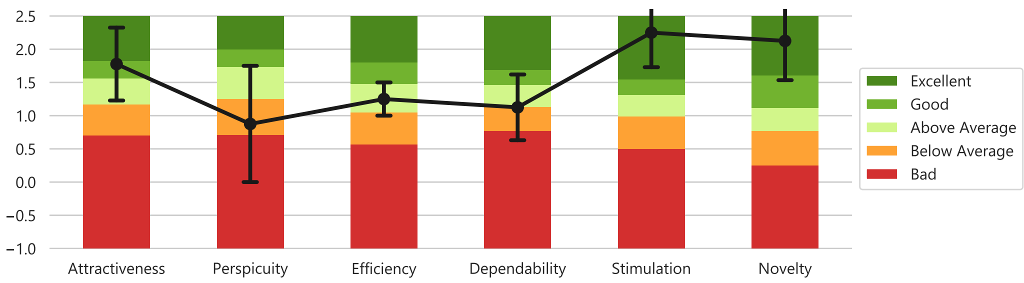


Figure 7.16: Results of the UEQ questionnaires. Points indicate mean value; error bars show standard deviation. Adapted from the UEQ data analysis tool [142].

to the novelty of AR that may wear off over time: Participants with prior experience in AR or VR ($n = 4$) consistently scored the system lower in these three categories than participants without such experience ($n = 2$), though no definite statement can be made due to the small sample size. The relatively low efficiency score may be due to the fact that more traditional 2D visualisations are more suited towards this study's tasks, whereas MIDAiR is tailored for exploratory data analysis, which could not be covered with this study. Similarly, the low perspicuity score could also be related to the previously mentioned habituation period, but also shows a high variance ($SD = 0.96$): Notably, participants with experience in either data analysis or data visualisations ($n = 4$) rate the perspicuity much lower (0.44) than participants ($n = 2$) without such experience (1.75). This could again be related to the mismatch between MIDAiR's intended use case and the study's tasks, for which MIDAiR is too convoluted for more experienced data analysts. Lastly, several network connectivity issues and the missing history function may contribute towards the low dependability score. The history function in particular was necessary as several participants accidentally deleted a scatter plot during the study.

Missing features

Both observation and user feedback revealed two essential features that could further improve the system usability: First, a history feature (i.e. undo, redo) was requested by several participants (P3, P7) due to accidental deletion and creation of scatter plots. Second, an annotation feature could have helped the participants to keep their tasks in front of them; one participant in particular (P3) was bothered by constantly switching between the visualisation and the screen, since this participant placed the visualisation far away from the screen.

Viewer mode

Although the *viewer mode* used by the study coordinator was not subject of this usability study, it proved very helpful in assisting the participants. The study coordinator thus did not have to wear another AR HMD but could join the analysis ad hoc. One participant commented positively on this aspect:

> *'I found it really cool that you could look at what I was doing with my HoloLens by simply coming to me with a tablet, and that you did not have to also put on a HoloLens or boot up something else; but that you could simply look into it "quick and dirty", I thought that was cool. Because assuming you just discovered something really great and you want to show it to someone else, then they can just look via the tablet.'* – P7

Visualisation

All but one participant (P8) quickly understood the visualisation after being guided through the first two tasks:

> *'First it looked very complex, but as soon as you I understood the filter feature [...] I thought it was actually very understandable.'* – P4

Many participants (P2, P4, P5) were, however, reluctant to use additional scatter plots for Task 3. Rather, they reused existing scatter plots and links by reconfiguring their dimensions and filters, despite several hints that they should not discard their current results. Since their visualisations only occupied a small corner of the room, they were also not limited with regards to space. It is possible that these participants tried to avoid a complex 3D visualisation:

> *'It [the visualisation] is only complex if you make it complex.'* – P1

Scatter Plot Placement. During the tasks, participants were instructed to place the first scatter plot at a position of their liking anywhere within the room, and to place subsequent scatter plots in a straight line using the alignment feature. With these instructions, all participants positioned the scatter plots slightly below eye level: The average scatter plot position (measured from the middle of a scatter plot) was 10 cm ($SD = 0.14$ cm) below the participant's eye level – since the scatter plot's height is 19 cm in total, this offset is roughly reduced to 0 cm.

Generally, most participants had no problems using several different input modalities (gaze, tangibility, touch, egocentric navigation) for placing the scatter plots. One participant (P3) was not able to place a scatter plot to close to themselves due to software limitations, but just took a step to resolve this issue. Still, the multimodality was well-received:

> *'For stuff like moving the thing [scatter plot] [...] that you can move it here and there, that was eventually intuitive.'* – P3

Similarly, while the alignment was not necessary for these tasks (as there was no comparison of relative differences between scatter plots), the alignment received positive feedback:

> *'Especially this [the alignment] was really good, because [...], once I opened a second layer, I could align the scatter plots with the ones next to them, and then it's clearly arranged.'* – P1

> *'I thought that was a good feature, too, that it was aligned, that you have a clear line on which it's [the scatter plot] is located, because that feels more pleasant than if it's randomly in the room and not parallel.'* – P5

> *'I was very very happy about the alignment function.'* – P7

However, participants could only move one scatter plot at a time, which was restricting for at least one participant (P7), since this participant wanted to place their second visualisation above the first visualisation. For such scenarios, a multi-selection could be useful for moving multiple objects at the same time.

Linking. While linking two scatter plots together generally did not pose a problem, the links themselves were often not well understood by several participants: The link direction was not always obvious, and it was hard to tell which two scatter plots were actually connected. Especially for the latter point, one participant (P3) tried to place a scatter plot on an existing link (i.e. in-between two connected scatter plots) and expected that the link now connects to the newly-placed scatter plot.

Despite these problems, the links were generally appreciated as filtering tool, especially since they served as visual indicator of how much data is still left for each scatter plot. However, given that the lines only served as visual indicator, several participants requested an option to reduce the line's prominence (e.g. by greying out the lines) which could make selection easier.

3D visualisation. The 3D visualisation proved very helpful as an overview, with several participants (P1, P4, P5) mentioning that the 3D visualisation helped them keep track of each individual filter step and see how the data is reduced. Furthermore, the potential for different workflows (i.e. visualisations) was also appreciated by several participants (P1, P7):

> *'[...] but the more elaborate and the more dimension you want to look at, I think it's much more [...] comprehensible.'* – P1

One participant (P2) liked the coloured borders for differentiating between scatter plots, but remarked that it was important to keep the colours distinct, especially for scatter plots that are close together. Another participant (P3) saw much potential for 3D visualisations in AR, either in the form of scientific visualisations, for displaying related, non-abstract data (e.g. displaying planets), or for building a scatter plot matrix (i.e. similar to ImAxes [29]).

2D visualisation. The *2D visualisation view* is an essential tool for assigning dimensions, filtering data, and viewing the end results. Therefore, participants had to use this view several times to solve the tasks. However, several participants (P5, P6, P8) reported that it was often unclear how the access this *2D visualisation view* on the tablet. Furthermore, the majority of participants (P1, P2, P4, P5, P7) expected common touch gestures (e.g. *pinch-to-zoom*) to work:

> *'Intuitively I thought you could zoom in somewhere on the tablet.'* – P4

Filtering. Similar to the 2D visualisations, filtering data was essential for solving the tasks. There were, however, many issues when using these filters:

- **Unclear which data is contained in filter:** Several participants (P1, P5) mentioned that it is unclear which data points are actually contained within a filter, which can be unclear for points directly on a filter's edge.
- **Filter overlap:** As a result from the previous point, several participants (P1, P6, P7) initially thought that only data not contained in *all* filters would be removed, whereas MIDAiR removes all data point not contained in *any* filter.
- **Insufficient accuracy:** Although filtering via touch input can be sufficiently accurate for exploratory tasks [21], the given tasks gave the impression that filters had to be drawn accurately. Participants therefore noted that touch input was too inaccurate for these tasks, although they agreed that zooming could alleviate this issue.
- **Missing tick lines:** Several participants (P1, P5) noted that tick lines could have helped them both in terms of accuracy, and also to draw straighter lines.
- **Filter adjustments:** Participants (P1, P7) also wanted to readjust the filters once they were drawn and suggested text input for adjusting filters created by drawing on the axis, allowing the user to pinpoint the filter's range.

Design Recommendations

Combination of 3D and 2D visualisations
3D visualisations serve as useful overview for individual 2D visualisations.

Multimodal object placement
Combining gaze, touch, and egocentric navigation allows users to easily place objects within a room.

3D object alignment
Allow users to align objects to existing objects is useful for avoiding disorder.

Further research directions

Mixed device collaboration
(e.g. co-located collaboration between tablet and AR HMD)

Proxemics for linking scatter plots
(e.g. automatically create or split existing links based on proxemic distance; should display feedback in AR [5, 77])

7.4 Limitations

There are several limitations to the previously presented usability study. A major limitations of the usability study is that many participants were not data experts, and thus were easily overwhelmed by the amount of information. The participants therefore had a different view on the presented data and may not care about features that could be important to actual data analysts, such as taking screenshots for research papers, collaboration, or interoperability with other applications. In addition, the limited amount of participants (especially as several participants did not fill out parts of the questionnaires or had to abort the study prematurely) reduces the significance of these results.

The tasks used in this study are very artificial and do not necessarily represent the real-world use case of MIDAiR. Thus, only a subset of features was tested, with key features such as colours, sorting, and relative differences missing. This may influence the usability of this system, as 2D visualisations (e.g. parallel coordinates) are more convenient for these tasks. The results of this study should thus be regarded as an examination of the general usability of the spatially-aware tablet and multimodality.

Furthermore, MIDAiR is still a prototype and thus has several hardware and software limitations. Firstly, many participants (P1, P2, P6, P8) found the HoloLens rather uncomfortable to wear, which might have influenced the SSQ scores. Secondly, the additional trackers made the tablet significantly heavier, which may have impacted its handiness. Lastly, the software experienced several minor outages due to connectivity issues, HoloLens system gestures, or remaining bugs in the software.

7.5 Discussion

The following sections discuss the results with regard to the previously defined research objectives of *multimodal interaction*, usage of *spatially-aware tablet*, and *system usability*. Lastly, an overview over the evaluation concerning each modality and each requirement is provided (see Table 7.1).

7.5.1 RO1: Multimodal Interaction

In general, users liked the different features that were presented and had no issues in using them. Although the system was unusual at first, users quickly got used to it and were able to use the input options on their own.

Can users accurately select objects within the used 3D visualisation?

For the most part, all users were able to select the objects they intended. Thanks to the long dwell time with instantly visible progress bar and instant colour feedback on the tablet, users also did not accidentally select the wrong object. The long dwell time also did not cause any problems, as several users were able to skip the selection with an explicit trigger (i.e. touching the tablet).

The selection was, however, often limited to simple visualisations. Several indicators show that the selection became more problematic as the visualisation's complexity increased. While the problem was alleviated by the egocentric navigation, this approach may not be sufficient for more complex visualisations.

Users also requested more feedback when selecting an object. This is especially apparent for selecting the links in-between scatter plots, which often caused confusion due to the invisible hitboxes. More research is required on how to make the hitbox more apparent, yet non-distracting.

Do users utilise all available input modalities?

Users generally used all modalities without issues and made frequent use of touch, gaze, tangibility, and egocentric navigation to achieve their tasks. Only voice commands were seldomly used; this is in part due to the user's reluctance to use voice commands, but also because these voice commands were seen as inferior to touch:

> *'When I have an idea and say that I want to create a new scatter plot... To execute that, I either need one button click to do that, or I briefly hold that, say the words, wait until the system understood that; so I feel like I'm slower with that.'*
> – P7

Users were, however, more open to the idea of more complex voice commands (e.g. as used in Orko [128] or Immersive Insights [23]), especially if these commands are more efficient or more accurate than their alternatives.

Several users also indicated that several tasks, such as linking scatter plots, were suitable for gesture interaction. On the other hand, results from the TPI questionnaire indicate that users did not feel the need to touch the virtual objects – hinting that gesture interaction may not be fitting for directly interacting with visualisations. However, this may have been distorted as the user's hands were already preoccupied and they thus did not want to use gestures.

7.5.2 RO2: Spatially-aware Tablet

Although a significant number of users feared that they would drop the tablet, the users appreciated the relatively large screen for interacting with the 2D visualisation. It is unclear whether the context switch causes simulation sickness issues, especially as other studies [23] observe similar discomfort with current AR devices, and similar use of a spatially-aware tablet [133] in VR found no indications of motion sickness.

Do users make use of the tablet's spatial awareness?

Users made little use of the tablet's spatial awareness while placing scatter plots, but frequently used this awareness as trigger for activating the *2D visualisation view*.

Placing scatter plots. While placing scatter plots, the tablet's spatial awareness was seldomly used to rotate the scatter plot. Users only changed the rotation for the first scatter plot, and aligned the subsequent scatter plots to the existing scatter plots using the alignment feature. Furthermore, the rotation is only useful for smaller rotations: rotating a scatter plot by 90° makes the tablet uncomfortable to hold, making it hard to tap on the tablet.

Trigger 2D visualisation view. Although users made frequent use of the tablet's tangibility by activating *lens mode*, they often only used it briefly to trigger the *2D visualisation view*. *Lens mode* allows users to match the tablet's screen with the AR scatter plot, which was not necessary for the given study tasks. Similar to gestures, users may get tired quickly when trying to hold the tablet in *lens mode* for a long time. Therefore, the use as a trigger to activate the *2D visualisation view* was an unexpected, yet well-used application for the spatial awareness. Further elaboration using different angles for different triggers (e.g. see Oakley and Park [95]) may prove beneficial.

Are users able to interact eyes-free with the tablet?

Usage of *eyes-free interaction* received mixed results. Although all users needed some time to get used to the system, several users were able to interact eyes-free with the system, while others fell back into their usual habit of looking at the tablet while interacting. The large buttons on the menu screen were helpful during eyes-free interaction and made the user feel more secure for interacting eyes-free, but could also be made about half as small without impacting the user's accuracy. Further research is required to what extent the buttons can be made smaller – maybe users would also be able to hit much narrower buttons that are located at the edge. Several users also fell back into the habit of interacting with their index fingers instead of holding the tablet with both hands; it is unclear whether they could still interact eyes-free, or how often they misclicked while interacting eyes-free. The AR HUD was not a sufficient replacement for looking down at the tablet. It was, however, a useful reminder once the participants were used to the system.

7.5.3 RO3: System Usability

Although some users were enthusiastic about the system, the results also indicate several problems concerning MIDAiR's usability. Due to the system's novelty, all users mentioned that they needed some time to get used to the system, which may contribute to a relatively low perspicuity. However, it is also unclear to what extent these problems are caused by a mismatch of the participants and the used tasks: MIDAiR was designed for exploratory data analysis, while the study's tasks define explicit characteristics. Moreover, the participants were not required to have any prior knowledge in data analysis. As a result, little can be said about the system's actual usability for its intended use case; the follow-up study with domain experts may yield clearer results. Instead, the current study focused on testing the interaction concepts, which all participants were quickly able to learn.

Does the system aid the users in their tasks?

Most users were able to quickly and accurately solve the study's tasks. The 3D visualisation served as an overview, while the links themselves were useful indicators of how much data had been filtered. The different input options allowed users to easily navigate, place, and connect the different scatter plots.

There were also several problems, most of which are centred around the usage of the *2D visualisation view*. Since the tasks were formulated with very specific requirements, users often wanted to emulate these strict characteristics in their filters. However, this was not possible as the touch input was not accurate enough. In addition, several features were missing that would have helped the users when drawing these filters (e.g. zooming, tick lines). The mode switch between the *2D visualisation view* and the main menu was also confusing for some users.

Users with prior knowledge in data analysis found the system too convoluted (i.e. imperspicuous): This may be because existing data analysis tools (e.g. using a 2D parallel coordinates visualisation) are much simpler to use for the given tasks, hence further studies are required. Still, the system performed well in the attractiveness and hedonic qualities. Although some of this can be contributed to the general novelty of immersive AR environments and may wear off over time, participants generally enjoyed interacting with the system. Further studies will show if the pragmatic qualities can be further improved by applying MIDAiR to its intended tasks.

Identifier	Modality	Evaluation
Data Manipulation		
Filter	Touch	Requires more accuracy through other features (e.g. zoom); users also wanted more complex voice commands for creating precise filters.
Scatter Plot Arrangement		
Create	Touch / Voice	Users only used touch to create scatter plots, but also requested more complex voice commands.
Layout	*Multiple*	Users liked using multiple modalities to position scatter plots.
Connect	Gaze & Touch	All users were able to link scatter plots with gaze and touch; one user wanted to create links with mid-air gestures.
Scatter Plot Configuration		
Dimensions	Touch	Users also wanted speech interaction in addition to touch.
Zoom	Touch	Users expected *pinch-to-zoom* gesture.
Context	Tangibles	Users were able to use *lens mode* and eyes-free interaction.
Navigation		
Navigate	Ego. navigation	Users naturally moved around.
Remote	Gaze	Object selection through gaze worked well.

Table 7.1: Evaluation of mapping between modalities and requirements. Requirements that were not implemented or not subject of this study were omitted.

7.6 Design Improvements

The following design improvements are based on the feedback and observations presented in Section 7.3. Each section will address shortcomings regarding the research objectives *multimodal interaction*, *spatially-aware tablet*, and *system usability*.

7.6.1 RO1: Multimodal Interaction

All participants were generally able to use the multimodal interaction techniques. Yet, the selection via head gaze could see further improvements:

- **Skipping dwell time**
  Several users liked and made good use of the option to skip the dwell time when selecting an object. However, the current implementation's problems are twofold: (a) The sudden screen change of the tablet can be distracting; and (b) triggering the explicit action relies on too much on the user's reflexes. Instead of relying on a short action, the system should offer a continuous action similar to activating voice commands (i.e. holding both thumbs on the tablet) which causes the selection progress bar to fill up quicker. Therefore, when users quickly want to change their selection, they can premeditate their intention to skip (i.e. significantly shorten) the dwell time. This can also be an opportunity for further multimodal interaction, for example by employing a pointing gesture for quicker selections.

- **Selection indicators**
  Some users also looked away shortly before the dwell time was complete, thereby aborting the selection prematurely. The system should offer more feedback in the form of auditory and visual indicators that the selection has been completed. Furthermore, the system may also internally complete the selection a little bit earlier (e.g. at 90% completion), and still give users feedback that the selection is still in progress: thus, even if the user looks away at the last moment, the selection is not aborted. However, further user tests are required to evaluate if this works as intended.

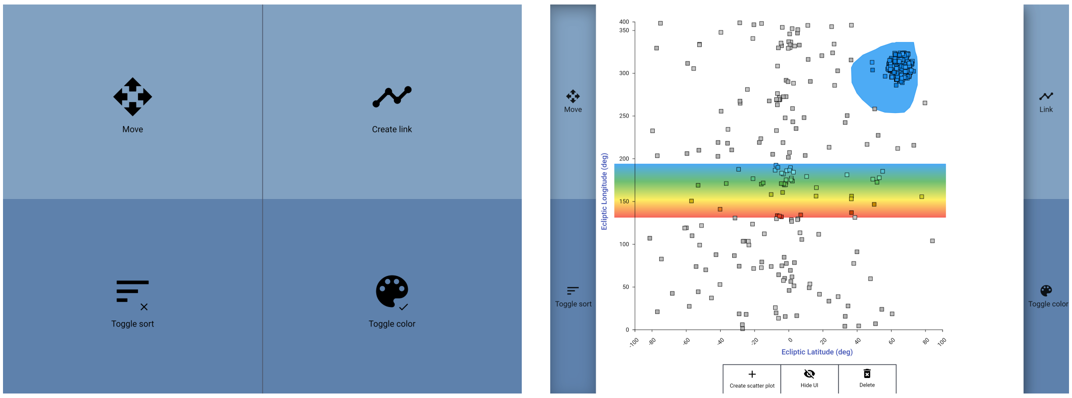
7.6.2 RO2: Spatially-Aware Tablet

The evaluation shows that while the spatially-aware tablet was useful in many ways, it also had several problems, especially relating to the *eyes-free interaction* and *tablet rotation*.

Eyes-Free Interaction

- **Gaze-based UI adjustments**
  While the analysis shows that users were generally able to use the concept of *eyes-free interaction*, some users still looked at the tablet out of habit and thus did not always interact with their thumbs as intended. The UI could be therefore dynamically adjusted depending on whether the user is currently looking at the tablet or at the AR visualisation (see Figure 7.17): When looking at the tablet, the buttons can become smaller, allowing for more options and supporting the user with a more traditional tablet interface; when looking at the AR visualisation, the buttons could become much bigger, thus facilitating *eyes-free interaction*. These changes should be subtle, so as to not attract the user's attention. Although this requires proper eye tracking (as many users kept the HoloLens facing forward while peeking at the device), it exemplifies the *orientation* dimension of proxemic interaction and thus further complements the multimodal interaction of MIDAiR.

- **Scrolling menu**
  While the large buttons for *eyes-free interaction* were appreciated, they also severely limit the menu screens to a maximum of four eyes-free actions. Al-



(a) (b)

Figure 7.17: Readjusting tablet UI based on gaze. (a) Large buttons when focusing on the AR environment. (b) Smaller buttons when looking at the tablet.

(a) (b)

Figure 7.18: A scroll menu could provide more than four menu entries. Users can be shown hints upon scrolling for both (a) the tablet UI and (b) the AR HUD.

though other eyes-free interaction techniques [56, 81, 95] offer more possible actions, they are also limited in the total amount of actions. To overcome this limitation, a scrolling menu can be employed (see Figure 7.18). Users may scroll through separate entries with their thumbs, while the AR menu provides a preview of the next entries within the scroll list. However, it is unclear whether splitting up the menu items between the left and right side, or scrolling through a shared pool of menu items is more perspicuous and requires further usability studies.

Tablet Rotation

- **Contextual awareness during rotation**
  Participants tried to emulate the scatter plot's rotation with their tablet by holding the tablet vertically. Instead of activating *lens mode*, MIDAiR should be more aware of the context and fulfil the user's expectation by rotating the scatter plot correctly to match the tablet's rotation, especially when the tablet is held vertically.

- **Offer alternative for rotation during placement**
  MIDAiR should also provide an alternative way of rotating the scatter plot, for example by offering the established two-finger rotation gesture. This can be helpful for larger rotations, thus preventing the users from having to hold the tablet in an awkward position.

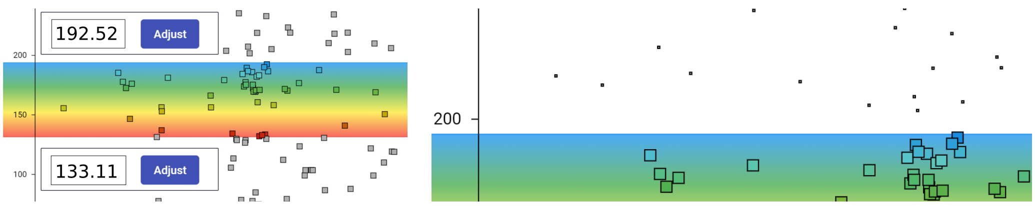
7.6.3 System Usability

Although users were able to complete the study's tasks, there are several points that could improve the general usability of the system.

- **Support established touch gestures**
  Several users mentioned that they expected established touch gestures, such as *pinch-to-zoom*, to work in the *2D visualisation view*. To avoid an overlap with the voice activation (i.e. holding both thumbs on the display), voice commands could only be activated by touching the edges of the display, leaving the middle of the screen free for other touch gestures.

- **Avoid separate 2D visualisation screens**
  The separation of the *menu screen* and the *2D visualisation screen* caused confusion among some users. By reducing the size of the large menu buttons (e.g. with the previously mentioned *gaze-based UI adjustments*), such a separation is no longer necessary (see Figure 7.17b).

- **Filter adjustments**
  Another problem was the insufficient accuracy of touch, especially for precise tasks as used in this study. Several users therefore wanted to readjust their filters, which could be supported by offering controls for fine-tuning the created filters (see Figure 7.19a). Users could also move the filters by dragging the existing filters via touch. In addition, there should be more visual feedback whether a given data point is contained within a filter (see Figure 7.19b).



(a) (b)

Figure 7.19: (a) A pop-up menu allows for exact adjustments of filters. (b) Data values not contained within the filter are clearly marked by becoming small dots.

8

Conclusion

This work presents a multimodal interaction approach for visual data analysis in AR that is based on prior work in the form of ART. MIDAiR is an immersive analytics tool for the exploratory analysis of multidimensional abstract data in an AR environment. MIDAiR combines a spatially-aware tablet with an immersive AR HMD, thus allowing for multimodal interaction consisting of touch, head gaze, voice, tangibility, proxemics, and egocentric navigation. This combination also produced a novel eyes-free interaction technique where users can interact with the tablet while focusing on the AR environment. MIDAiR employs a 3D parallel coordinates visualisation consisting of several linked 2D visualisations that allows users to identify clusters, outliers, and trends within the data. The system supports full collaboration between several users with AR HMD, as well as limited collaboration between users with mixed devices (e.g. AR-capable tablet and HMD).

The interaction concepts of MIDAiR are based on functional requirements for interacting with visualisations, a literature research on current input modalities, as well as lessons learned from related work. An initial usability study with 8 participants revealed that, although users needed some time to get used to MIDAiR, they could quickly utilise the full range of input modalities to interact with the 3D visualisation and complete their tasks. While both the use of multimodality and the use of the spatially-aware tablet can be further improved, the current prototype already offers a feature-rich framework for interacting with 3D visualisation through a spatially-aware tablet.

8.1 Contributions

This work thus contributes three aspects to the area of *immersive analytics* tools:

1. **Design recommendations & further research directions**
   This work provides several design recommendations and further research directions for the use of spatially-aware tablets, as well as for the use of 3D visualisations within immersive AR environments.

2. **Novel interaction concept using spatially-aware tablet**
   Using a spatially-aware tablet with an AR HMD allows for novel interaction concepts that employ an AR HUD to allow users to interact eyes-free with the tablet while looking at the AR scene, or allow the user to hold the tablet vertically to avoid context switches.

3. **MIDAiR**
   MIDAiR demonstrates how a multimodal interaction approach can be employed for interacting with 3D visualisations in immersive AR environments.

8.2 Future Work

Aside from the research directions offered in Section 7.3, there are several opportunities for future work.

Collaboration. While MIDAiR does support local collaboration, the study was intentionally restricted to a single user scenario due to time constraints. Further studies with multiple collaborators might reveal further insights, especially when collaborators can use different devices (e.g. AR HMD and AR-capable tablet) for interacting with the AR scene. The AR environment also allows for remote collaboration that could provide additional research opportunities.

Foldable devices. Foldable devices like the Samsung Galaxy Fold [117] are now commercially available and can be an interesting alternative to the tablet used in MIDAiR. Opening or folding such a device provides chances for a better interaction, for example by showing MIDAiR's *2D visualisation screen* when opening the device. Furthermore, such devices combine the handiness of smaller devices with the advantages of a big screen, which can prove especially useful for MIDAiR.

Integration of gestures. Although gesture interaction was considered for MIDAiR, current technological restrictions as well as time constraints prevented MIDAiR from taking gestures into account. Yet, devices like the Microsoft HoloLens 2 [89] or

Google Soli [83] promise several improvements in gesture interaction, which could overcome the current technological limitations and make gesture interaction a viable choice. While the use of a spatially-aware tablet might restrict the usage of gestures, the previously mentioned *foldable devices* could work well with additional gesture interaction.

Pen Input. Pens as additional input device for tablets are becoming increasingly common and can offer a natural way to input text or sketches. This is especially useful for creating annotations, which can further improve the workflow of a data analysis tool and facilitate collaboration.

Integration of statistical measures. MIDAiR currently enables users to filter the data on each scatter plot, allowing for the creation of a *filter pipeline*. This could be extended further by integrating more statistical measures: Instead of only filtering data, the marked data could be aggregated or otherwise transformed, allowing for more complex scenarios.

Visualisations. MIDAiR uses a 3D visualisation that is suited for 2D manipulation on tablets. However, other visualisations may provide interesting use cases by offering interaction with a familiar touch display, while still offering the complexity of a 3D visualisation. Especially in the realm of scientific visualisations where 3D visualisations are commonplace, this could provide further research opportunities.

Bibliography

[1] John R. Aiello. *Human Spatial Behavior*. 1987. 389–448.

[2] *Angular*. URL: https://angular.io/ (visited on 2019-07-16).

[3] *ARKit*. URL: https://developer.apple.com/augmented-reality/ (visited on 2019-07-01).

[4] Myroslav Bachynskyi, Gregorio Palmas, Antti Oulasvirta, Jürgen Steimle and Tino Weinkauf. 'Performance and Ergonomics of Touch Surfaces: A Comparative Study Using Biomechanical Simulation'. In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*. The 33rd Annual ACM Conference. Seoul, Republic of Korea: ACM Press, 2015, pp. 1817–1826. ISBN: 978-1-4503-3145-6. DOI: 10.1145/2702123.2702607.

[5] Sriram Karthik Badam, Fereshteh Amini, Niklas Elmqvist and Pourang Irani. 'Supporting Visual Exploration for Multiple Users in Large Display Environments'. In: *2016 IEEE Conference on Visual Analytics Science and Technology, VAST 2016 - Proceedings* (2016), pp. 1–10. DOI: 10.1109/VAST.2016.7883506.

[6] Aruna D. Balakrishnan, Susan R. Fussell and Sara Kiesler. 'Do Visualizations Improve Synchronous Remote Collaboration?' In: *Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems - CHI '08*. Proceeding of the Twenty-Sixth Annual CHI Conference. Florence, Italy: ACM Press, 2008, p. 1227. ISBN: 978-1-60558-011-1. DOI: 10.1145/1357054.1357246.

[7] Aruna D. Balakrishnan, Susan R. Fussell, Sara Kiesler and Aniket Kittur. 'Pitfalls of Information Access with Visualizations in Remote Collaborative Analysis'. In: *Proceedings of the 2010 ACM conference on Computer supported cooperative work - CSCW '10* (2010), p. 411. DOI: 10.1145/1718918.1718988.

[8] Till Ballendat, Nicolai Marquardt and Saul Greenberg. 'Proxemic Interaction: Designing for a Proximity and Orientation-Aware Environment'. In: *ACM International Conference on Interactive Tabletops and Surfaces - ITS '10* (2010), p. 121. ISSN: 10725520. DOI: 10.1145/1936652.1936676.

[9] Andrea Batch, Andrew Cunningham, Maxime Cordeil, Niklas Elmqvist, Tim Dwyer, Bruce H. Thomas and Kim Marriott. 'There Is No Spoon: Evaluating Performance, Space Use, and Presence with Expert Domain Users in Immersive Analytics'. In: *IEEE Transactions on Visualization and Computer Graphics* (2019), pp. 1–1. ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2019.2934803.

[10] Jens Bauer and Achim Ebert. 'Virtual Buttons for Eyes-Free Interaction: A Study'. In: *Human-Computer Interaction – INTERACT 2015*. Ed. by Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque and Marco Winckler. Vol. 9296. Cham: Springer International Publishing, 2015, pp. 47–54. ISBN: 978-3-319-22700-9. DOI: 10.1007/978-3-319-22701-6_4.

[11] Jens Bauer, Achim Ebert, Oliver Kreylos and Bernd Hamann. 'Marking Menus for Eyes-Free Interaction Using Smart Phones and Tablets'. In: *Availability, Reliability, and Security in Information Systems and HCI*. Ed. by Alfredo Cuzzocrea, Christian Kittl, Dimitris E. Simos, Edgar Weippl and Lida Xu. Red. by David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Madhu Sudan, Demetri Terzopoulos, Doug Tygar, Moshe Y. Vardi and Gerhard Weikum. Vol. 8127. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 481–494. ISBN: 978-3-642-40510-5. DOI: 10.1007/978-3-642-40511-2_35.

[12] Dominikus Baur, Bongshin Lee and Sheelagh Carpendale. 'TouchWave: Kinetic Multi-Touch Manipulation for Hierarchical Stacked Graphs'. In: *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces* (2012), pp. 255–264. DOI: 10.1145/2396636.2396675.

[13] Hans-Joachim Bieg, Lewis L. Chuang, Roland W. Fleming, Harald Reiterer and Heinrich H. Bülthoff. 'Eye and Pointer Coordination in Search and Selection Tasks'. In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10* (2010), p. 89. DOI: 10.1145/1743666.1743688.

[14] Dan Bohus and Eric Horvitz. 'Facilitating Multiparty Dialog with Gaze, Gesture, and Speech'. In: *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction on - ICMI-MLMI '10* (2010), p. 1. DOI: 10.1145/1891903.1891910.

[15] Richard A. Bolt. '"Put-That-There": Voice and Gesture at the Graphics Interface'. In: *Proceedings of the 7th annual conference on Computer graphics and interactive techniques - SIGGRAPH '80* (1980), pp. 262–270. ISSN: 00978930. DOI: 10.1145/800250.807503.

[16] Mike Bostock. *D3.Js - Data-Driven Documents*. URL: https://d3js.org/ (visited on 2019-07-16).

[17] Peter Brandl, Clifton Forlines, Daniel Wigdor, Michael Haller and Chia Shen. 'Combining and Measuring the Benefits of Bimanual Pen and Direct-Touch Interaction on Horizontal Interfaces'. In: *Proceedings of the working conference on Advanced visual interfaces* (2008), pp. 154–161. DOI: 10.1145/1385569.1385595.

[18] Scott Brave, Hiroshi Ishii and Andrew Dahley. 'Tangible Interfaces for Remote Collaboration and Communication'. In: *Proceedings of the 1998 ACM conference on Computer supported cooperative work - CSCW '98* (1998), pp. 169–178. DOI: 10.1145/289444.289491.

[19] Gerd Bruder, Frank Steinicke and Wolfgang Stuerzlinger. 'Touching the Void Revisited: Analyses of Touch Behavior on and above Tabletop Surfaces'. In: *Proceedings of the IFIP Conference on Human-Computer Interaction (INTERACT '13)*. Vol. 8117 LNCS. Berlin, DE: Springer, 2013, p. 868. ISBN: 978-3-642-40482-5. DOI: 10.1007/978-3-642-40483-2_19.

[20] Wolfgang Büschel, Annett Mitschick, Thomas Meyer and Raimund Dachselt. 'Investigating Smartphone-Based Pan and Zoom in 3D Data Spaces in Augmented Reality'. In: *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '19*. The 21st International Conference. Taipei, Taiwan: ACM Press, 2019, pp. 1–13. ISBN: 978-1-4503-6825-4. DOI: 10.1145/3338286.3340113.

[21] Simon Butscher, Sebastian Hubenschmid, Jens Müller, Johannes Fuchs and Harald Reiterer. 'Clusters, Trends, and Outliers: How Immersive Technologies Can Facilitate the Collaborative Analysis of Multidimensional Data'. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. New York, New York, USA: ACM Press, 2018, pp. 1–12. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3173664.

[22] Marco Cavallo, Mishal Dholakia, Matous Havlena, Kenneth Ocheltree and Mark Podlaseck. 'Dataspace: A Reconfigurable Hybrid Reality Environment for Collaborative Information Analysis'. In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). Osaka, Japan: IEEE, 2019-03, pp. 145–153. ISBN: 978-1-72811-377-7. DOI: 10.1109/VR.2019.8797733.

[23] Marco Cavallo, Mishal Dolakia, Matous Havlena, Kenneth Ocheltree and Mark Podlaseck. 'Immersive Insights: A Hybrid Analytics System forCollaborative Exploratory Data Analysis'. In: *25th ACM Symposium on Virtual Reality*

Software and Technology on - VRST '19. 25th ACM Symposium on Virtual Reality Software and Technology. Parramatta, NSW, Australia: ACM Press, 2019, pp. 1–12. ISBN: 978-1-4503-7001-1. DOI: 10.1145/3359996.3364242.

[24] Li-Wei Chan, Hui-Shan Kao, Mike Y. Chen, Ming-Sui Lee, Jane Hsu and Yi-Ping Hung. 'Touching the Void: Direct-Touch Interaction for Intangible Displays'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)* (2010), pp. 2625–2634. DOI: 10.1145/1753326.1753725.

[25] Andy Cockburn and Bruce McKenzie. 'An Evaluation of Cone Trees'. In: *People and Computers XIV — Usability or Else!* Ed. by Sharon McDonald, Yvonne Waern and Gilbert Cockton. London: Springer London, 2000, pp. 425–436. ISBN: 978-1-85233-318-8. DOI: 10.1007/978-1-4471-0515-2_28.

[26] Christopher Collins and Sheelagh Carpendale. 'VisLink: Revealing Relationships Amongst Visualizations'. In: *IEEE Transactions on Visualization and Computer Graphics* 13.6 (2007), pp. 1192–1199. ISSN: 10772626. DOI: 10.1109/TVCG.2007.70521. pmid: 17968064.

[27] Christopher G. Coogan and Bin He. 'Brain-Computer Interface Control in a Virtual Reality Environment and Applications for the Internet of Things'. In: *IEEE Access* 6 (2018), pp. 10840–10849. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2018.2809453.

[28] Maxime Cordeil, Andrew Cunningham, Benjamin Bach, Christophe Hurter, Bruce H. Thomas, Kim Marriott and Tim Dwyer. 'IATK: An Immersive Analytics Toolkit'. In: *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). Osaka, Japan: IEEE, 2019-03, pp. 200–209. ISBN: 978-1-72811-377-7. DOI: 10/gf6296.

[29] Maxime Cordeil, Andrew Cunningham, Tim Dwyer, Bruce H. Thomas and Kim Marriott. 'ImAxes: Immersive Axes as Embodied Affordances for Interactive Multivariate Data Visualisation'. In: *Proceedings of the ACM Symposium on User Interface and Software Technology (UIST '17)*. New York, NY, USA: ACM Press, 2017, pp. 71–83. ISBN: 978-1-4503-4981-9. DOI: 10.1145/3126594.3126613.

[30] Christian Corsten, Ignacio Avellino, Max Möllers and Jan Borchers. 'Instant User Interfaces: Repurposing Everyday Objects as Input Devices'. In: *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces - ITS '13*. The 2013 ACM International Conference. St. An-

drews, Scotland, United Kingdom: ACM Press, 2013, pp. 71–80. isbn: 978-1-4503-2271-3. doi: 10.1145/2512349.2512799.

[31] Ciro Donalek, S. G. Djorgovski, Alex Cioc, Anwell Wang, Jerry Zhang, Elizabeth Lawler, Stacy Yeh, Ashish Mahabal, Matthew Graham, Andrew Drake, Scott Davidoff, Jeffrey S. Norris and Giuseppe Longo. 'Immersive and Collaborative Data Visualization Using Virtual Reality Platforms'. In: *Proceedings of the IEEE International Conference on Big Data*. Washington, DC, USA: IEEE, 2014-10, pp. 609–614. isbn: 978-1-4799-5666-1. doi: 10.1109/BigData.2014.7004282.

[32] Steven M. Drucker, Danyel Fisher, Ramik Sadana, Jessica Herron and M.c. Schraefel. 'TouchViz: A Case Study Comparing Two Interfaces for Data Analytics on Tablets'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13* (2013), p. 2301. doi: 10.1145/2470654.2481318.

[33] Andrew T. Duchowski. *Eye Tracking Methodology*. London: Springer-Verlag GmbH, 2007. isbn: 978-1-84628-608-7. doi: 10.1007/978-1-84628-609-4.

[34] Bruno Dumas, Denis Lalanne and Sharon Oviatt. 'Multimodal Interfaces: A Survey of Principles, Models and Frameworks'. In: *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, 2009, pp. 3–26. doi: 10.1007/978-3-642-00437-7_1.

[35] Patrick Ehlen and Michael Johnston. 'Multimodal Local Search in Speak4it'. In: *Proceedings of the 15th International Conference on Intelligent User Interfaces - IUI '11*. The 15th International Conference. Palo Alto, CA, USA: ACM Press, 2011, p. 435. isbn: 978-1-4503-0419-1. doi: 10.1145/1943403.1943486.

[36] Jacob Eisenstein and Randall Davis. 'Visual and Linguistic Information in Gesture Classification'. In: *Proceedings of the 6th International Conference on Multimodal Interfaces - ICMI '04*. Proceedings of the 6th International Conference on Multimodal Interfaces - ICMI '04. ACM Press, 2004, p. 113. doi: 10.1145/1027933.1027954.

[37] George W. Fitzmaurice, Hiroshi Ishii and William Buxton. 'Bricks: Laying the Foundations for Graspable User Interfaces'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '95*. SIGCHI Conference on Human Factors in Computing Systems. ACM Press, 1995, pp. 442–449. doi: 10.1145/223904.223964.

[38] Node js Foundation. *Node.Js*. url: https://nodejs.org/en/ (visited on 2019-07-16).

[39] Mathias Frisch, Jens Heydekorn and Raimund Dachselt. 'Investigating Multi-Touch and Pen Gestures for Diagram Editing on Interactive Surfaces'. In: *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces - ITS '09*. Proc. ITS 2009. ACM Press, 2009, pp. 149–156. DOI: 10.1145/1731903.1731933.

[40] Johannes Fuchs, Roman Rädle, Dominik Sacha, Fabian Fischer and Andreas Stoffel. 'Collaborative Data Analysis with Smart Tangible Devices'. In: *Visualization and Data Analysis 2014*. Ed. by Pak Chung Wong, David L. Kao, Ming C. Hao and Chaomei Chen. San Francisco, California, USA: Society for Imaging Science and Technology, 2013-12-23. DOI: 10.1117/12.2040011.

[41] Tong Gao, Mira Dontcheva, Eytan Adar, Zhicheng Liu and Karrie G. Karahalios. 'DataTone: Managing Ambiguity in Natural Language Interfaces for Data Visualization'. In: *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology - UIST '15*. Daegu, Kyungpook, Republic of Korea: ACM Press, 2015, pp. 489–500. ISBN: 978-1-4503-3779-3. DOI: 10.1145/2807442.2807478.

[42] James J. Gibson. *The Ecological Approach To Visual Perception*. Routledge, 1986-09. ISBN: 978-0-203-76776-4. DOI: 10.4324/9780203767764.

[43] Michael A. Grasso, David S. Ebert and Timothy W. Finin. 'The Integrality of Speech in Multimodal Interfaces'. In: *ACM Transactions on Computer-Human Interaction* 5.4 (1998-12-01), pp. 303–325. ISSN: 10730516. DOI: 10.1145/300520.300521.

[44] Saul Greenberg, Nicolai Marquardt, Till Ballendat, Rob Diaz-Marino and Miaosen Wang. 'Proxemic Interactions: The New Ubicomp?' In: *Interactions* 18.1 (2011), p. 42. ISSN: 10725520. DOI: 10.1145/1897239.1897250.

[45] Edward T. Hall. *Handbook for Proxemic Research*. Vol. 36. 2. Wiley, 1974-02. 40-40. DOI: 10.1111/an.1995.36.2.40.

[46] Edward T. Hall. *The Hidden Dimension*. Doubleday, 1966. 201 pp.

[47] Rex Hartson and Pardha Pyla. *The UX Book*. Elsevier, 2018. ISBN: 978-0-12-805342-3. DOI: 10.1016/C2013-0-19285-9.

[48] Jeffrey Heer, Michael Bostock and Vadim Ogievetsky. 'A Tour through the Visualization Zoo'. In: *Communications of the ACM* 53.6 (2010-06), p. 59. ISSN: 00010782. DOI: 10.1145/1743546.1743567. pmid: 15224215.

[49] Jeffrey Heer and Ben Shneiderman. 'Interactive Dynamics for Visual Analysis'. In: *Queue* 10.2 (2012-02-01), p. 30. ISSN: 15427730. DOI: 10.1145/2133416.2146416.

[50] Ken Hinckley, Randy Pausch, John C. Goble and Neal F. Kassell. 'A Survey of Design Issues in Spatial Input'. In: *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology - UIST '94*. The 7th Annual ACM Symposium. Marina del Rey, California, United States: ACM Press, 1994, pp. 213–222. ISBN: 978-0-89791-657-8. DOI: 10.1145/192426.192501.

[51] Ken Hinckley, Koji Yatani, Michel Pahud, Nicole Coddington, Jenny Rodenhouse, Andy Wilson, Hrvoje Benko and Bill Buxton. 'Pen + Touch = New Tools'. In: *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology - UIST '10*. Uist '10. ACM Press, 2010, p. 27. DOI: 10.1145/1866029.1866036.

[52] Andreas Hinderks, Martin Schrepp, Francisco José Domínguez Mayo, María José Escalona and Jörg Thomaschewski. 'Developing a UX KPI Based on the User Experience Questionnaire'. In: *Computer Standards & Interfaces* 65 (2019-07), pp. 38–44. ISSN: 09205489. DOI: 10.1016/j.csi.2019.01.007.

[53] Lars Erik Holmquist, Johan Redström and Peter Ljungstrand. 'Token-Based Access to Digital Information'. In: *Proceedings Handheld and Ubiquitous Computing*. Springer Berlin Heidelberg, 1999, pp. 234–245. DOI: 10.1007/3-540-48157-5_22.

[54] Kenneth Holmqvist, Marcus Nyström, Richard Andersson, Richard Dewhurst, Halszka Jarodzka and Joost van de Weijer. *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford University Press, 2011-10-11. ISBN: 978-0-19-969708-3.

[55] *HoloLens Hardware Details*. URL: https://docs.microsoft.com/en-us/windows/mixed-reality/hololens-hardware-details (visited on 2019-07-01).

[56] Tom Horak, Sriram Karthik Badam, Niklas Elmqvist and Raimund Dachselt. 'When David Meets Goliath: Combining Smartwatches with a Large Vertical Display for Visual Data Exploration'. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. New York, New York, USA: ACM Press, 2018, pp. 1–13. ISBN: 978-1-4503-5620-6. DOI: 10.1145/3173574.3173593.

[57] *HTC Vive Tracker (2018) Developer Guidelines*.

[58] Sebastian Hubenschmid. 'ART – Augmented Reality above the Tabletop: An Immersive Analytics Tool for the Visual Analysis of Mobile Health Data'. Bachelor's thesis. University of Konstanz, 2017.

[59] Sebastian Hubenschmid. *MIDAiR: Multimodal Interaction for Visual Data Analysis in Augmented Reality*. Master Project Report. University of Konstanz, 2019-09-16.

[60] Sebastian Hubenschmid. *Multimodal Interaction for Visual Data Analysis in Augmented Reality*. Seminar to the Master Project. University of Konstanz, 2018-12-03.

[61] *iPad Pro (10.5-Inch) - Technical Specifications*. URL: https://support.apple.com/kb/SP762 (visited on 2019-07-01).

[62] Petra Isenberg, Danyel Fisher, Meredith Ringel Morris, Kori Inkpen and Mary Czerwinski. 'An Exploratory Study of Co-Located Collaborative Visual Analytics around a Tabletop Display'. In: *2010 IEEE Symposium on Visual Analytics Science and Technology*. 2010 IEEE Symposium on Visual Analytics Science and Technology (VAST). Salt Lake City, UT, USA: IEEE, 2010-10, pp. 179–186. ISBN: 978-1-4244-9488-0. DOI: 10.1109/VAST.2010.5652880.

[63] Petra Isenberg, Tobias Isenberg, Tobias Hesselmann, Bongshin Lee, Ulrich Von Zadow and Anthony Tang. 'Data Visualization on Interactive Surfaces: A Research Agenda'. In: *IEEE Computer Graphics and Applications* 33 (April 2013), pp. 16–24. ISSN: 0272-1716. DOI: 10.1109/MCG.2013.24.

[64] Hiroshi Ishii. 'Tangible Bits: Beyond Pixels'. In: *Proceedings of the 2nd International Conference on Tangible and Embedded Interaction - TEI '08*. The 2nd International Conference. Bonn, Germany: ACM Press, 2008, p. xv. ISBN: 978-1-60558-004-3. DOI: 10.1145/1347390.1347392.

[65] Rob Jacob and Sophie Stellmach. 'What You Look at Is What You Get'. In: *Interactions* 23.5 (2016), pp. 62–65. ISSN: 10725520. DOI: 10.1145/2978577.

[66] Robert J. K. Jacob. 'What You Look at Is What You Get: Eye Movement-Based Interaction Techniques'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Empowering People - CHI '90*. Vol. 23. New York, New York, USA: ACM Press, 1990, pp. 11–18. ISBN: 0-201-50932-6. DOI: 10.1145/97243.97246.

[67] Robert J. K. Jacob and Linda E. Sibert. 'The Perceptual Structure of Multidimensional Input Device Selection'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '92*. The SIGCHI Conference. Monterey, California, United States: ACM Press, 1992, pp. 211–218. ISBN: 978-0-89791-513-7. DOI: 10.1145/142750.142792.

[68] Hans-Christian Jetter, Jens Gerken, Michael Zöllner, Harald Reiterer and Natasa Milic-Frayling. 'Materializing the Query with Facet-Streams: A Hybrid Surface for Collaborative Search on Tabletops'. In: *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems - CHI '11*. ACM Press, 2011, p. 3013. ISBN: 978-1-4503-0228-9. DOI: 10.1145/1978942.1979390.

[69] Ed Kaiser, Alex Olwal, David McGee, Hrvoje Benko, Andrea Corradini, Xiaoguang Li, Phil Cohen and Steven Feiner. 'Mutual Disambiguation of 3D Multimodal Interaction in Augmented and Virtual Reality'. In: *Proceedings of the 5th International Conference on Multimodal Interfaces - ICMI '03*. Proceedings of the 5th International Conference on Multimodal Interfaces - ICMI '03. ACM Press, 2003, p. 12. DOI: 10.1145/958436.958438.

[70] Maria Karam and m. c. Schraefel. 'A Taxonomy of Gestures in Human Computer Interactions'. In: *Technical Report, Eletronics and Computer Science.* (2005), pp. 1–45. ISSN: 0854328335.

[71] Daniel A. Keim. 'Information Visualization and Visual Data Mining'. In: *IEEE Transactions on Visualization and Computer Graphics* 8.1 (2002), pp. 1–8. ISSN: 10772626. DOI: 10.1109/2945.981847.

[72] Robert S. Kennedy, Norman E. Lane, Kevin S. Berbaum and Michael G. Lilienthal. 'Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness'. In: *The International Journal of Aviation Psychology* 3.3 (1993-07), pp. 203–220. ISSN: 1050-8414, 1532-7108. DOI: 10.1207/s15327108ijap0303_3.

[73] Sanshzar Kettebekov. 'Exploiting Prosodic Structuring of Coverbal Gesticulation'. In: *Proceedings of the 6th International Conference on Multimodal Interfaces - ICMI '04*. Proceedings of the 6th International Conference on Multimodal Interfaces - ICMI '04. ACM Press, 2004, pp. 105–112. DOI: 10.1145/1027933.1027953.

[74] David B. Koons and Carlton J. Sparrell. 'ICONIC: Speech and Depictive Gestures at the Human-Machine Interface'. In: *Conference Companion on Human Factors in Computing Systems - CHI '94*. Conference Companion on Human Factors in \ldots. ACM Press, 1994, pp. 453–454. DOI: 10.1145/259963.260487.

[75] M. Kraus, N. Weiler, D. Oelke, J. Kehrer, D. A. Keim and J. Fuchs. 'The Impact of Immersion on Cluster Identification Tasks'. In: *IEEE Transactions on Visualization and Computer Graphics* (2019), pp. 1–1. ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2019.2934395.

[76] Denis Lalanne, Laurence Nigay, Philippe Palanque, Peter Robinson, Jean Vanderdonckt and Jean-François Ladry. 'Fusion Engines for Multimodal Input'. In: *Proceedings of the 2009 International Conference on Multimodal Interfaces - ICMI-MLMI '09*. Proceedings of the 2009 International Conference on Multimodal Interfaces - ICMI-MLMI '09. ACM Press, 2009, p. 153. DOI: 10.1145/1647314.1647343.

[77] Ricardo Langner, Tom Horak and Raimund Dachselt. 'VisTiles: Coordinating and Combining Co-Located Mobile Devices for Visual Data Exploration'. In: *IEEE Transactions on Visualization and Computer Graphics* 24.1 (2018), pp. 626–636. ISSN: 10772626. DOI: 10.1109/TVCG.2017.2744019.

[78] Joseph J. LaViola, Ernst Kruijff, Ryan P. McMahan, Doug A. Bowman and Ivan Poupyrev. *3D User Interfaces: Theory and Practice*. Second edition. Addison-Wesley Usability and HCI Series. Boston: Addison-Wesley, 2017. 591 pp. ISBN: 978-0-13-403432-4.

[79] Bongshin Lee, Arjun Srinivasan, John Stasko, Melanie Tory and Vidya Setlur. 'Multimodal Interaction for Data Visualization'. In: *Proceedings of the 2018 International Conference on Advanced Visual Interfaces - AVI '18*. Proceedings of the 2018 International Conference on Advanced Visual Interfaces - AVI '18. ACM Press, 2018, pp. 1–3. DOI: 10.1145/3206505.3206602.

[80] Byungjoo Lee, Qiao Deng, Eve Hoggan and Antti Oulasvirta. 'Boxer: A Multimodal Collision Technique for Virtual Objects'. In: *Proceedings of the 19th ACM International Conference on Multimodal Interaction - ICMI 2017*. The 19th ACM International Conference. Glasgow, UK: ACM Press, 2017, pp. 252–260. ISBN: 978-1-4503-5543-8. DOI: 10.1145/3136755.3136761.

[81] Frank Chun Yat Li, David Dearman and Khai N. Truong. 'Virtual Shelves: Interactions with Orientation Aware Devices'. In: *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology - UIST '09*. The 22nd Annual ACM Symposium. Victoria, BC, Canada: ACM Press, 2009, p. 125. ISBN: 978-1-60558-745-5. DOI: 10.1145/1622176.1622200.

[82] Sukun Li, Avery Leider, Meikang Qiu, Keke Gai and Meiqin Liu. 'Brain-Based Computer Interfaces in Virtual Reality'. In: *2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud)*. 2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud). New York, NY, USA: IEEE, 2017-06, pp. 300–305. ISBN: 978-1-5090-6644-5. DOI: 10.1109/CSCloud.2017.51.

[83] Jaime Lien, Nicholas Gillian, M Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja and Ivan Poupyrev. 'Soli: Ubiquitous Gesture Sensing with Millimeter Wave Radar'. In: *ACM Transactions on Graphics* 35.4 (2016-07), pp. 1–19. ISSN: 07300301. DOI: 10.1145/2897824.2925953.

[84] Matthew Lombard, Theresa B Ditton and Lisa Weinstein. 'Measuring Presence: The Temple Presence Inventory'. In: *Proceedings of the 12th Annual International Workshop on Presence*. 2009, pp. 1–15.

[85] Pedro Lopes, Alexandra Ion, Sijing You and Patrick Baudisch. 'Adding Force Feedback to Mixed Reality Experiences and Games Using Electrical Muscle Stimulation'. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. ACM Press, 2018, pp. 1–13. DOI: 10.1145/3173574.3174020.

[86] Nicolai Marquardt and Saul Greenberg. 'Informing the Design of Proxemic Interactions'. In: *IEEE Pervasive Computing* 11.2 (2012), pp. 14–23. ISSN: 15361268. DOI: 10.1109/MPRV.2012.15.

[87] Kim Marriott, Falk Schreiber, Tim Dwyer, Karsten Klein, Nathalie Henry Riche, Takayuki Itoh, Wolfgang Stuerzlinger and Bruce H. Thomas, eds. *Immersive Analytics*. Lecture Notes in Computer Science 11190. Cham: Springer, 2018. 357 pp. ISBN: 978-3-030-01387-5.

[88] Fabrice Matulic, Maria Husmann, Seraiah Walter and Moira C. Norrie. 'Eyes-Free Touch Command Support for Pen-Based Digital Whiteboards via Handheld Devices'. In: *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces - ITS '15*. The 2015 International Conference. Madeira, Portugal: ACM Press, 2015, pp. 141–150. ISBN: 978-1-4503-3899-8. DOI: 10.1145/2817721.2817728.

[89] *Microsoft HoloLens | Mixed Reality Technology for Business*. URL: https://www.microsoft.com/en-us/hololens (visited on 2019-08-31).

[90] Peter Mohr, Markus Tatzgern, Tobias Langlotz, Andreas Lang, Dieter Schmalstieg and Denis Kalkofen. 'TrackCap: Enabling Smartphones for 3D Interaction on Mobile Head-Mounted Displays'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. The 2019 CHI Conference. Glasgow, Scotland Uk: ACM Press, 2019, pp. 1–11. ISBN: 978-1-4503-5970-2. DOI: 10.1145/3290605.3300815.

[91] *MultimodalVis 2018*. URL: https://sites.google.com/view/multimodalvis (visited on 2019-11-08).

[92] Tamara Munzner. *Visualization Analysis and Design*. A K Peters/CRC Press, 2014-12. 117–130. ISBN: 978-1-4665-0893-4. DOI: 10.1201/b17511.

[93] *NASA Exoplanet Archive*. URL: https://exoplanetarchive.ipac.caltech.edu/ (visited on 2019-07-20).

[94] Laurence Nigay and Joëlle Coutaz. 'A Design Space for Multimodal Systems: Concurrent Processing and Data Fusion'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '93*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '93. ACM Press, 1993, pp. 172–178. DOI: 10.1145/169059.169143. pmid: 11798387.

[95] Ian Oakley and Junseok Park. 'Motion Marking Menus: An Eyes-Free Approach to Motion Input for Handheld Devices'. In: *International Journal of Human-Computer Studies* 67.6 (2009-06), pp. 515–532. ISSN: 10715819. DOI: 10.1016/j.ijhcs.2009.02.002.

[96] Sharon Oviatt. 'Multimodal Interfaces'. In: *The human-computer interaction handbook: fundamentals, evolving technologies and emerging applications* 47.03 (2012), pp. 405–430. ISSN: 00094978. DOI: 10.5860/CHOICE.47-1469.

[97] Sharon Oviatt. 'Ten Myths of Multimodal Interaction'. In: *Communications of the ACM* 42.11 (1999), pp. 74–81. ISSN: 00010782. DOI: 10.1145/319382.319398.

[98] Sharon Oviatt and Philip Cohen. 'Perceptual User Interfaces: Multimodal Interfaces That Process What Comes Naturally'. In: *Communications of the ACM* 43.3 (2000-03-01), pp. 45–53. ISSN: 00010782. DOI: 10.1145/330534.330538.

[99] Sharon Oviatt, Antonella DeAngeli and Karen Kuhn. 'Integration and Synchronization of Input Modes during Multimodal Human-Computer Interaction'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '97*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '97. ACM Press, 1997, pp. 415–422. DOI: 10.1145/258549.258821.

[100] Sharon Oviatt, Rebecca Lunsford and Rachel Coulston. 'Individual Differences in Multimodal Integration Patterns: What Are They and Why Do They Exist?' In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '05*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '05. ACM Press, 2005, pp. 241–249. DOI: 10.1145/1054972.1055006.

[101] Biswaksen Patnaik, Andrea Batch and Niklas Elmqvist. 'Olfactory Analytics: Exploring the Design Space of Smell for Data Visualization'. In: *MultimodalVis '18 Workshop*. AVI 2018. 2018, p. 4.

[102] William A. Pike, John Stasko, Remco Chang and Theresa A. O'Connell. 'The Science of Interaction'. In: *Information Visualization* 8.4 (2009), pp. 263–274. ISSN: 14738716. DOI: 10.1057/ivs.2009.22.

[103] Francis Quek, David McNeill, Robert Bryll, Susan Duncan, Xin-Feng Ma, Cemil Kirbas, Karl E. McCullough and Rashid Ansari. 'Multimodal Human Discourse: Gesture and Speech'. In: *ACM Transactions on Computer-Human Interaction* 9.3 (2002), pp. 171–193. ISSN: 10730516. DOI: 10.1145/568513.568514. pmid: 25151251.

[104] Darío R. Quiñones, Gonçalo Lopes, Danbee Kim, Cédric Honnet, David Moratal and Adam Kampff. 'HIVE Tracker: A Tiny, Low-Cost, and Scalable Device for Sub-Millimetric 3D Positioning'. In: *Proceedings of the 9th Augmented Human International Conference on - AH '18*. The 9th Augmented Human International Conference. Seoul, Republic of Korea: ACM Press, 2018, pp. 1–8. ISBN: 978-1-4503-5415-8. DOI: 10.1145/3174910.3174935.

[105] Leah M. Reeves, Jean-Claude Martin, Michael McTear, Tv Raman, Kay M. Stanney, Hui Su, Qian Ying Wang, Jennifer Lai, James A. Larson, Sharon Oviatt, T. S. Balaji, Stéphanie Buisine, Penny Collings, Phil Cohen and Ben Kraal. 'Guidelines for Multimodal User Interface Design'. In: *Communications of the ACM* 47.1 (2004-01-01), p. 57. ISSN: 00010782. DOI: 10.1145/962081.962106.

[106] 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*. The 2019 ACM International Conference. Daejeon, Republic of Korea: ACM Press, 2019, pp. 29–41. ISBN: 978-1-4503-6891-9. DOI: 10.1145/3343055.3359718.

[107] Patrick Reipschläger, Burcu Kulahcioglu Ozkan, Aman Shankar Mathur, Stefan Gumhold, Rupak Majumdar and Raimund Dachselt. 'DebugAR: Mixed Dimensional Displays for Immersive Debugging of Distributed Systems'. In: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*. Extended Abstracts of the 2018 CHI Conference. Montreal QC, Canada: ACM Press, 2018, pp. 1–6. ISBN: 978-1-4503-5621-3. DOI: 10.1145/3170427.3188679.

[108] Jan Richter, Bruce H. Thomas, Maki Sugimoto and Masahiko Inami. 'Remote Active Tangible Interactions'. In: *Proceedings of the 1st International Conference on Tangible and Embedded Interaction - TEI '07*. Proceedings of the 1st International Conference on Tangible and Embedded Interaction - TEI '07. ACM Press, 2007, p. 39. DOI: 10.1145/1226969.1226977.

[109] Alexander Rind, Michael Iber and Wolfgang Aigner. 'Bridging the Gap Between Sonification and Visualization'. In: *MultimodalVis '18 Workshop*. AVI 2018. 2018, p. 4.

[110] Gabriel Robles-De-La-Torre. 'The Importance of the Sense of Touch in Virtual and Real Environments'. In: *IEEE Multimedia* 13.3 (2006-07), pp. 24–30. ISSN: 1070-986X. DOI: 10.1109/MMUL.2006.69.

[111] Ronald Rosenfeld, Dan Olsen and Alex Rudnicky. 'Universal Speech Interfaces'. In: *interactions* 8.6 (2001-10-01), pp. 34–44. ISSN: 10725520. DOI: 10.1145/384076.384085.

[112] Jeffrey M. Rzeszotarski and Aniket Kittur. 'Kinetica: Naturalistic Multi-Touch Data Visualization'. In: *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI '14*. New York, New York, USA: ACM Press, 2014, pp. 897–906. ISBN: 978-1-4503-2473-1. DOI: 10.1145/2556288.2557231.

[113] Dominik Sacha, Andreas Stoffel, Florian Stoffel, Bum Chul Kwon, Geoffrey Ellis and Daniel A. Keim. 'Knowledge Generation Model for Visual Analytics'. In: *IEEE Transactions on Visualization and Computer Graphics* 20.12 (2014-12), pp. 1604–1613. ISSN: 1077-2626. DOI: 10.1109/TVCG.2014.2346481.

[114] Ramik Sadana, Meeshu Agnihotri and John Stasko. 'Touching Data: A Discoverability-Based Evaluation of a Visualization Interface for Tablet Computers'. In: 2018. DOI: arXiv:1806.06084v1.

[115] Ramik Sadana and John Stasko. 'Expanding Selection for Information Visualization Systems on Tablet Devices'. In: *Proceedings of the 2016 ACM on Interactive Surfaces and Spaces - ISS '16*. Proceedings of the 2016 ACM on Interactive Surfaces and Spaces - ISS '16. ACM Press, 2016, pp. 149–158. DOI: 10.1145/2992154.2992157.

[116] Ramik Sadana and John Stasko. 'Interacting with Data Visualizations on Tablets and Phones : Developing Effective Touch-Based Gestures and Operations (Poster)'. In: *IEEE Vis* (2013), pp. 7–8.

[117] *Samsung Galaxy Fold*. URL: https://www.samsung.com/global/galaxy/galaxy-fold/ (visited on 2019-11-09).

[118] Dieter Schmalstieg, L. Miguel Encarnação and Zsolt Szalavári. 'Using Transparent Props for Interaction with the Virtual Table'. In: *Proceedings of the 1999 Symposium on Interactive 3D Graphics - SI3D '99*. The 1999 Symposium. Atlanta, Georgia, United States: ACM Press, 1999, pp. 147–153. ISBN: 978-1-58113-082-9. DOI: 10.1145/300523.300542.

[119] Martin Schrepp, Andreas Hinderks and Jörg Thomaschewski. 'Construction of a Benchmark for the User Experience Questionnaire (UEQ)'. In: *International Journal of Interactive Multimedia and Artificial Intelligence* 4.4 (2017), p. 40. ISSN: 1989-1660. DOI: 10.9781/ijimai.2017.445.

[120] Vidya Setlur, Sarah E. Battersby, Melanie Tory, Rich Gossweiler and Angel X. Chang. 'Eviza: A Natural Language Interface for Visual Analysis'. In: *Proceedings of the 29th Annual Symposium on User Interface Software and Technology - UIST '16*. Vol. 10. Proceedings of the 29th Annual Symposium on User Interface Software and Technology 16. ACM Press, 2016, pp. 365–377. DOI: 10.1145/2984511.2984588.

[121] Orit Shaer. 'Tangible User Interfaces: Past, Present, and Future Directions'. In: *Foundations and Trends® in Human–Computer Interaction* 3.1-2 (2009), pp. 1–137. ISSN: 1551-3955. DOI: 10.1561/1100000026.

[122] B. Shneiderman. 'The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations'. In: *Proceedings 1996 IEEE Symposium on Visual Languages*. 1996 IEEE Symposium on Visual Languages. Boulder, CO, USA: IEEE Comput. Soc. Press, 1996, pp. 336–343. ISBN: 978-0-8186-7508-9. DOI: 10.1109/VL.1996.545307.

[123] Ronell Sicat, Jiabao Li, Junyoung Choi, Maxime Cordeil, Won-Ki Jeong, Benjamin Bach and Hanspeter Pfister. 'DXR: A Toolkit for Building Immersive Data Visualizations'. In: *IEEE Transactions on Visualization and Computer Graphics* 25.1 (2019-01), pp. 715–725. ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2018.2865152.

[124] Katie A. Siek, Yvonne Rogers and Kay H. Connelly. 'Fat Finger Worries: How Older and Younger Users Physically Interact with PDAs'. In: *Human-Computer Interaction - INTERACT 2005*. Springer Berlin Heidelberg, 2005, pp. 267–280. DOI: 10.1007/11555261_24.

[125] Maurício Sousa, Daniel Mendes, Soraia Paulo, Nuno Matela, Joaquim Jorge and Daniel Simões Lopes. 'VRRRRoom: Virtual Reality for Radiologists in the Reading Room'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '17)*. New York, NY, USA: ACM Press, 2017, pp. 4057–4062. ISBN: 978-1-4503-4655-9. DOI: 10.1145/3025453.3025566.

[126] 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*. ACM International Conference on Interactive Tabletops and Surfaces - ITS '10. ACM Press, 2010, p. 157. doi: 10.1145/1936652.1936684.

[127] *SQLite Home Page*. url: https://www.sqlite.org/index.html (visited on 2019-07-16).

[128] Arjun Srinivasan and John Stasko. 'Orko: Facilitating Multimodal Interaction for Visual Exploration and Analysis of Networks'. In: *IEEE Transactions on Visualization and Computer Graphics* 24.1 (2018), pp. 1–1. issn: 1077-2626. doi: 10.1109/TVCG.2017.2745219.

[129] Arjun Srinivasan and John T. Stasko. 'Natural Language Interfaces for Data Analysis with Visualization: Considering What Has and Could Be Asked'. In: *EuroVis 2017 - Short Papers*. The Eurographics Association, 2017. isbn: 978-3-03868-043-7. doi: 10.2312/eurovisshort.20171133.

[130] Sophie Stellmach and Raimund Dachselt. 'Look & Touch : Gaze-Supported Target Acquisition'. In: *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems - CHI '12*. ACM Press, 2012. doi: 10.1145/2207676.2208709.

[131] Sophie Stellmach and Raimund Dachselt. 'Still Looking: Investigating Seamless Gaze-Supported Selection, Positioning, and Manipulation of Distant Targets'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13. ACM Press, 2013, p. 285. doi: 10.1145/2470654.2470695.

[132] Sophie Stellmach, Sebastian Stober, Andreas Nürnberger and Raimund Dachselt. 'Designing Gaze-Supported Multimodal Interactions for the Exploration of Large Image Collections'. In: *Proceedings of the 1st Conference on Novel Gaze-Controlled Applications - NGCA '11*. Proceedings of the 1st Conference on Novel Gaze-Controlled Applications - NGCA '11. ACM Press, 2011, pp. 1–8. doi: 10.1145/1983302.1983303.

[133] Hemant Bhaskar Surale, Aakar Gupta, Mark Hancock and Daniel Vogel. 'TabletInVR: Exploring the Design Space for Using a Multi-Touch Tablet in Virtual Reality'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. The 2019 CHI Conference. Glasgow, Scotland Uk: ACM Press, 2019, pp. 1–13. isbn: 978-1-4503-5970-2. doi: 10.1145/3290605.3300243.

[134] *Tableau: Business Intelligence and Analytics Software*. URL: https://www.tableau.com/ (visited on 2019-09-10).

[135] Slawomir Konrad Tadeja, Timoleon Kipouros and Per Ola Kristensson. 'Exploring Parallel Coordinates Plots in Virtual Reality'. In: *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems - CHI EA '19*. Extended Abstracts of the 2019 CHI Conference. Glasgow, Scotland Uk: ACM Press, 2019, pp. 1–6. ISBN: 978-1-4503-5971-9. DOI: 10.1145/3290607.3313068.

[136] Unity Technologies. *Unity*. URL: https://unity.com/frontpage (visited on 2019-07-16).

[137] Theophanis Tsandilas. 'Fallacies of Agreement: A Critical Review of Consensus Assessment Methods for Gesture Elicitation'. In: *ACM Transactions on Computer-Human Interaction* 25.3 (2018-06-28), pp. 1–49. ISSN: 10730516. DOI: 10.1145/3182168.

[138] Matthew Turk. 'Multimodal Interaction: A Review'. In: *Pattern Recognition Letters* 36.1 (2014), pp. 189–195. ISSN: 01678655. DOI: 10.1016/j.patrec.2013.07.003. pmid: 1000198543.

[139] Jayson Turner, Andreas Bulling and Hans Gellersen. 'Combining Gaze with Manual Interaction to Extend Physical Reach'. In: *Proceedings of the 1st International Workshop on Pervasive Eye Tracking & Mobile Eye-Based Interaction - PETMEI '11*. Proceedings of the 1st International Workshop on Pervasive Eye Tracking & Mobile Eye-Based Interaction - PETMEI '11. ACM Press, 2011, p. 33. DOI: 10.1145/2029956.2029966.

[140] *TypeScript - JavaScript That Scales*. URL: https://www.typescriptlang.org/ (visited on 2019-07-16).

[141] B. Ullmer and H. Ishii. 'Emerging Frameworks for Tangible User Interfaces'. In: *IBM Systems Journal* 39.3.4 (2000), pp. 915–931. ISSN: 0018-8670. DOI: 10.1147/sj.393.0915.

[142] *User Experience Questionnaire (UEQ)*. URL: https://www.ueq-online.org/ (visited on 2019-10-05).

[143] Daniel Vogel and Ravin Balakrishnan. 'Distant Freehand Pointing and Clicking on Very Large, High Resolution Displays'. In: *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology - UIST '05*. Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology - UIST '05. ACM Press, 2005, p. 33. DOI: 10.1145/1095034.1095041.

[144] Daniel Vogel and Ravin Balakrishnan. 'Interactive Public Ambient Displays: Transitioning from Implicit to Explicit, Public to Personal, Interaction with Multiple Users'. In: *Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology - UIST '04*. Vol. 6. Proceedings of the 17th Annual ACM Symposium on User Interface Software and Technology (UIST '04) 2. ACM Press, 2004, pp. 137–146. DOI: 10.1145/1029632.1029656.

[145] Jagoda Walny, Bongshin Lee, Paul Johns, Nathalie Henry Riche and Sheelagh Carpendale. 'Understanding Pen and Touch Interaction for Data Exploration on Interactive Whiteboards'. In: *IEEE Transactions on Visualization and Computer Graphics* 18.12 (2012), pp. 2779–2788. ISSN: 10772626. DOI: 10.1109/TVCG.2012.275.

[146] Matthew O. Ward, Georges Grinstein and Daniel Keim. *Interactive Data Visualization*. A K Peters/CRC Press, 2015-06. ISBN: 978-1-4822-5738-0. DOI: 10.1201/b18379.

[147] Colin Ware and Glenn Franck. 'Viewing a Graph in a Virtual Reality Display Is Three Times as Good as a 2D Diagram'. In: *Proceedings of the IEEE Symposium on Visual Languages (VL '94)*. Washington, DC, USA: IEEE Computer Society, 1994, pp. 182–183. ISBN: 0-8186-6660-9. DOI: 10.1109/VL.1994.363621.

[148] Colin Ware and Harutune H. Mikaelian. 'An Evaluation of an Eye Tracker as a Device for Computer Input'. In: *ACM SIGCHI Bulletin* 17 (SI 1987), pp. 183–188. ISSN: 07366906. DOI: 10.1145/30851.275627.

[149] Mark Weiser. 'The Computer for the 21st Century'. In: *Scientific American* 265.3 (1991-09), pp. 94–104. ISSN: 0036-8733. DOI: 10.1038/scientificamerican0991-94.

[150] Malte Weiss, Florian Schwarz, Simon Jakubowski and Jan Borchers. 'Madgets: Actuating Widgets on Interactive Tabletops'. In: *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology - UIST '10*. The 23nd Annual ACM Symposium. New York, New York, USA: ACM Press, 2010, p. 293. ISBN: 978-1-4503-0271-5. DOI: 10.1145/1866029.1866075.

[151] Alan Wexelblat. 'Natural Gesture in Virtual Environments'. In: *Virtual Reality Software and Technology*. Proceedings of the VRST '94 Conference. Singapore: WORLD SCIENTIFIC, 1994-08, pp. 5–16. ISBN: 978-981-02-1867-6. DOI: 10.1142/9789814350938_0002.

[152] Alan Wexelblat. 'Research Challenges in Gesture: Open Issues and Unsolved Problems'. In: *Lecture Notes in Computer Science*. Vol. 1371. Springer Berlin Heidelberg, 1998, pp. 1–11. DOI: 10.1007/BFb0052984.

[153] Daniel Wigdor and Dennis Wixon. *Brave NUI World*. Elsevier, 2011. 264 pp. ISBN: 978-0-12-382231-4. DOI: 10.1016/C2009-0-64091-5.

[154] Benfang Xiao, Cynthia Girand and Sharon Oviatt. 'Multimodal Integration Patterns in Children.' In: *Proceedings of ICSLP 2002*. 2002, pp. 629–632.

[155] Benfang Xiao, Rebecca Lunsford, Rachel Coulston, Matt Wesson and Sharon Oviatt. 'Modeling Multimodal Integration Patterns and Performance in Seniors: Toward Adaptive Processing of Individual Differences'. In: *Proceedings of the 5th International Conference on Multimodal Interfaces - ICMI '03*. Proc. 5th International Conference on Multimodal Interfaces. ACM Press, 2003, pp. 265–272. DOI: 10.1145/958468.958480.

[156] Michiya Yamamoto, Munehiro Komeda, Takashi Nagamatsu and Tomio Watanabe. 'Development of Eye-Tracking Tabletop Interface for Media Art Works'. In: *ACM International Conference on Interactive Tabletops and Surfaces - ITS '10*. ACM International Conference on Interactive Tabletops and Surfaces - ITS '10. ACM Press, 2010, p. 295. DOI: 10.1145/1936652.1936723.

[157] Yalong Yang, Tim Dwyer, Bernhard Jenny, Kim Marriott, Maxime Cordeil and Haohui Chen. 'Origin-Destination Flow Maps in Immersive Environments'. In: *IEEE Transactions on Visualization and Computer Graphics* (2018), pp. 1–1. ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2018.2865192.

[158] Ji Soo Yi, Youn ah Kang, John T. Stasko and Julie A. Jacko. 'Toward a Deeper Understanding of the Role of Interaction in Information Visualization'. In: *IEEE Trans Vis Comput Graph* 13.6 (2007), pp. 1224–1231. ISSN: 1077-2626. DOI: 10.1109/TVCG.2007.70515. pmid: 17968068.

[159] ByungIn Yoo, Jae-Joon Han, Changkyu Choi, Kwonju Yi, Sungjoo Suh, Dusik Park and Changyeong Kim. '3D User Interface Combining Gaze and Hand Gestures for Large-Scale Display'. In: *Proceedings of the 28th of the International Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '10*. Proceedings of the Conference on Human Factors in Computing Systems - CHI '10. ACM Press, 2010, pp. 3709–3714. DOI: 10.1145/1753846.1754043.

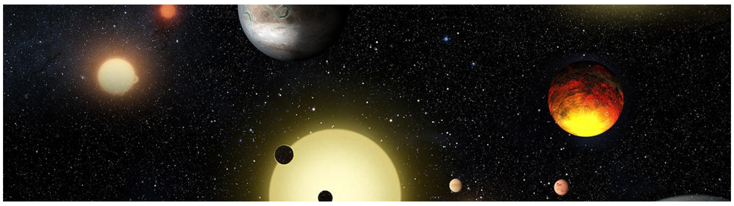
[160] Shumin Zhai, Carlos Morimoto and Steven Ihde. 'Manual and Gaze Input Cascaded (MAGIC) Pointing'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems the CHI Is the Limit - CHI '99*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems the CHI Is the Limit - CHI '99. ACM Press, 1999, pp. 246–253. DOI: 10.1145/302979.303053.

A

Usability Study Material

The following pages contain the documents used within the usability study, in particular:

- Welcome letter
- Consent form
- Preliminary demographic questionnaire
- Post-study questionnaire
- Questions for semi-structured interview



Herzlich willkommen!

Vielen Dank, dass Sie sich dazu bereit erklärt haben, an unserer Studie teilzunehmen. Sie unterstützen damit unsere Forschung maßgeblich! Bevor es losgeht, möchten wir Ihnen kurz vermitteln, um was es bei der Untersuchung geht und welche Rolle Sie dabei spielen.

Ziele und Ablauf der Studie

Im Rahmen dieser Studie untersuchen wir ein neues Interaktionskonzept, um mit einer Datenvisualisierung innerhalb einer Augmented Reality Umgebung zu interagieren. Dazu werden Ihnen mehrere Datenanalyseaufgaben gestellt, die Sie anhand der Anwendung beantworten sollen. Hierbei kommt es nicht auf die Schnelligkeit, sondern mehr auf die Genauigkeit und Vollständigkeit der Antwort an. Falls Ihnen eine Funktion der Anwendung unklar ist können Sie jederzeit nachfragen. Zum Schluss bitten wir Sie, uns von Ihren Erfahrungen und Eindrücken zu berichten.

Um möglichst umfassende Erkenntnisse zu erhalten, zeichnen wir die Studie zusätzlich in Bild und Ton auf. Für diese Aufzeichnungen ist Ihr Einverständnis erforderlich. Im Gegenzug verpflichten wir uns dazu, das Material lediglich zu pseudonymisierten Auswertungszwecken zu verwenden. In diesem Zusammenhang haben wir eine Einverständniserklärung vorbereitet, die diesem Schreiben beiliegt. An dieser Stelle möchten wir darauf hinweisen, dass wir nicht Sie oder Ihre Leistung bewerten, sondern ausschließlich an der Tauglichkeit der Anwendung interessiert sind.

Zeitrahmen und Entlohnung

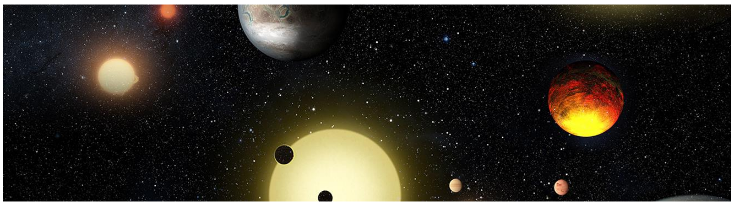
Die Dauer der Studie beträgt insgesamt ca. 1 Stunde. Falls Sie sich zu irgendeinem Zeitpunkt unwohl fühlen und Ihre Teilnahme beenden möchten, ist das selbstverständlich auch ohne Angabe von Gründen möglich. Bitte wenden Sie sich dann an den Versuchsleiter.

Nach der Durchführung der Studie werden Sie für Ihre Hilfe mit 10 Euro entlohnt. Wir wünschen Ihnen jetzt gutes Gelingen und bedanken uns noch einmal recht herzlich für Ihre Unterstützung!

Sebastian Hubenschmid

Arbeitsgruppe Mensch-Computer-Interaktion,
Fachbereich Informatik und Informationswissenschaft,
Universität Konstanz

Figure A.1: Welcome letter.



Einverständniserklärung

ID: ____

Informationen zur Studienleitung

Studienleiter: Sebastian Hubenschmid
Institution: Arbeitsgruppe Mensch-Computer Interaktion, Fachbereich Informatik und Informationswissenschaft, Universität Konstanz

Erklärung

Über das Ziel, den Inhalt und die Dauer der Studie wurde ich informiert. Im Rahmen dieser Studie werden in Fragebögen personenbezogene Daten erhoben. Zusätzlich wird die Studie auf Video aufgezeichnet, es werde Audioaufnahmen gemacht und Bewegungsdaten erfasst.

Hiermit bin ich darüber aufgeklärt, dass die personenbezogenen Daten vertraulich behandelt werden. Die Ergebnisse der Analyse der Video-, Audio- und Bewegungsdaten werden eventuell in späteren Publikationen pseudonymisiert veröffentlicht. Wir garantieren dabei absolute Diskretion. Es wird zu keinem Zeitpunkt Rückschluss auf Sie als Person möglich sein.

Optionale Punkte *(Bei Zustimmung bitte ankreuzen)*

☐ Ich bin damit einverstanden, dass meine Video- und Bewegungsdaten zusätzlich zu internen Präsentationszwecken genutzt werden können.

Hiermit erkläre ich mich mit den unter „Erklärung" genannten Punkten und den angekreuzten optionalen Punkten einverstanden:

	Konstanz,	
(Name)	(Ort, Datum)	(Unterschrift)

Hiermit verpflichtet sich die Studienleitung, die Video- und Audioaufzeichnung sowie sämtliche sonstigen gewonnenen Daten lediglich zu Auswertungszwecken im Rahmen dieser Untersuchung zu verwenden:

Sebastian Hubenschmid	Konstanz,	
(Name)	(Ort, Datum)	(Unterschrift)

Figure A.2: Consent form.

Participant ID: ____

Vorab Fragebogen

Herzlichen Dank, dass Sie sich bereit erklärt haben an dieser Untersuchung teilzunehmen.
Bevor wir anfangen, benötigen wir von Ihnen noch einige Angaben zu Ihrer Person.
Wir möchten Ihnen hiermit noch einmal mitteilen, dass alle Daten vertraulich behandelt werden.

Alter: ________

Geschlecht: ☐ weiblich ☐ männlich ☐ sonstiges ☐ keine Angabe

Ihre momentane Tätigkeit: ____

1) Sind Sie Links- oder Rechtshänder?
   - ☐ Linkshänder
   - ☐ Rechtshänder

2) Leiden Sie unter einer Farbsehschwäche (z.B. Rot-Grün-Sehschwäche)?
   - ☐ nein
   - ☐ ja, unter _____________________________________

3) Besitzen Sie physische Einschränkungen, die Sie bei der Bewegung behindern?
   - ☐ nein
   - ☐ ja

Bitte wenden

1

Figure A.3: Preliminary demographic questionnaire (page 1).

Participant ID: ____

4) Nutzen Sie im Alltag ein Smartphone?

☐ nein

☐ ja, ich schätze mich in der Benutzung damit ein als:

unerfahren ☐ ☐ ☐ ☐ ☐ sehr erfahren

5) Nutzen Sie im Alltag ein Tablet?

☐ nein

☐ ja, ich schätze mich in der Benutzung damit ein als:

unerfahren ☐ ☐ ☐ ☐ ☐ sehr erfahren

6) Haben Sie bereits Erfahrung mit „Augmented Reality" (AR) Anwendungen?

☐ nein

☐ ja, ich schätze mich in der Benutzung damit ein als:

unerfahren ☐ ☐ ☐ ☐ ☐ sehr erfahren

Ich habe Erfahrung mit folgenden AR Anwendungen:

____

____

7) Haben Sie bereits Erfahrung mit „Virtual Reality" (VR) Technologien/Apps?

☐ nein

☐ ja, ich schätze mich in der Benutzung damit ein als:

unerfahren ☐ ☐ ☐ ☐ ☐ sehr erfahren

Ich habe Erfahrung mit folgenden VR Anwendungen:

____

____

2

Figure A.4: Preliminary demographic questionnaire (page 2).

Participant ID: ____

8) Haben Sie bereits Erfahrung mit Anwendungen zur **Visualisierung** von Daten?

☐ nein

☐ ja, ich schätze mich in der Benutzung damit ein als:

unerfahren ☐ ☐ ☐ ☐ ☐ sehr erfahren

Ich habe Erfahrung mit folgenden Anwendungen:

____

____

9) Haben Sie bereits Erfahrung mit 3-dimensionalen Datenvisualisierungen?

☐ nein

☐ ja, ich schätze mich in der Interpretation damit ein als:

unerfahren ☐ ☐ ☐ ☐ ☐ sehr erfahren

10) Haben Sie bereits Erfahrung mit Anwendungen zur **Analyse** von Daten?

☐ nein

☐ ja, ich schätze mich in der Benutzung damit ein als:

unerfahren ☐ ☐ ☐ ☐ ☐ sehr erfahren

Ich habe Erfahrung mit folgenden Anwendungen:

____

____

3

Figure A.5: Preliminary demographic questionnaire (page 3).

Participant ID:___

Bitte kreuzen Sie an wie sehr die folgenden Symptome Sie **jetzt gerade** betreffen.

Allgemeines Unwohlsein	○ gar nicht	○ leicht	○ mäßig	○ stark
Müdigkeit	○ gar nicht	○ leicht	○ mäßig	○ stark
Kopfschmerzen	○ gar nicht	○ leicht	○ mäßig	○ stark
Überanstrengung der Augen	○ gar nicht	○ leicht	○ mäßig	○ stark
Probleme scharf zu sehen	○ gar nicht	○ leicht	○ mäßig	○ stark
Erhöhter Speichelfluss	○ gar nicht	○ leicht	○ mäßig	○ stark
Schwitzen	○ gar nicht	○ leicht	○ mäßig	○ stark
Übelkeit	○ gar nicht	○ leicht	○ mäßig	○ stark
Konzentrationsschwierigkeiten	○ gar nicht	○ leicht	○ mäßig	○ stark
Kopfdruck	○ gar nicht	○ leicht	○ mäßig	○ stark
Verschwommenes Sehen	○ gar nicht	○ leicht	○ mäßig	○ stark
Schwindel (Augen auf)	○ gar nicht	○ leicht	○ mäßig	○ stark
Schwindel (Augen zu)	○ gar nicht	○ leicht	○ mäßig	○ stark
Gleichgewichtstörung	○ gar nicht	○ leicht	○ mäßig	○ stark
Magen macht sich bemerkbar	○ gar nicht	○ leicht	○ mäßig	○ stark
Aufstoßen	○ gar nicht	○ leicht	○ mäßig	○ stark

Figure A.6: Post-study questionnaire (page 1).

Participant ID:___

Setzen Sie ein "X" auf das Item, das Ihre Erfahrungen am besten widerspiegelt. Entscheiden Sie möglichst spontan. Es ist wichtig, dass Sie nicht lange über die Begriffe nachdenken, damit Ihre unmittelbare Einschätzung zum Tragen kommt.
Beispiel:

nie O O O X O O O immer

---

unerfreulich	O	O	O	O	O	O	O	erfreulich
unverständlich	O	O	O	O	O	O	O	verständlich
kreativ	O	O	O	O	O	O	O	phantasielos
leicht zu lernen	O	O	O	O	O	O	O	schwer zu lernen
wertvoll	O	O	O	O	O	O	O	minderwertig
langweilig	O	O	O	O	O	O	O	spannend
uninteressant	O	O	O	O	O	O	O	interessant
unberechenbar	O	O	O	O	O	O	O	voraussagbar
schnell	O	O	O	O	O	O	O	langsam
originell	O	O	O	O	O	O	O	konventionell
behindernd	O	O	O	O	O	O	O	unterstützend
gut	O	O	O	O	O	O	O	schlecht
kompliziert	O	O	O	O	O	O	O	einfach
abstoßend	O	O	O	O	O	O	O	anziehend
herkömmlich	O	O	O	O	O	O	O	neuartig
unangenehm	O	O	O	O	O	O	O	angenehm
sicher	O	O	O	O	O	O	O	unsicher
aktivierend	O	O	O	O	O	O	O	einschläfernd
erwartungskonform	O	O	O	O	O	O	O	nicht erwartungskonform
ineffizient	O	O	O	O	O	O	O	effizient
übersichtlich	O	O	O	O	O	O	O	verwirrend
unpragmatisch	O	O	O	O	O	O	O	pragmatisch
aufgeräumt	O	O	O	O	O	O	O	überladen
attraktiv	O	O	O	O	O	O	O	unattraktiv
sympathisch	O	O	O	O	O	O	O	unsympathisch
konservativ	O	O	O	O	O	O	O	innovativ

Figure A.7: Post-study questionnaire (page 2).

Participant ID:___

Bitte lesen Sie sich die folgenden Fragen aufmerksam durch. Setzen Sie ein "X" auf das Item, das Ihre Erfahrungen am besten widerspiegelt. Beispiel:

Nie ○ ○ ○ X ○ ○ ○ Immer

Die Anwendung soll attraktiv, angenehm und sympathisch wirken.

Überhaupt nicht wichtig ○ ○ ○ ○ ○ ○ ○ Sehr wichtig

Die Anwendung soll mir helfen, meine Aufgaben schnell, effizient und pragmatisch zu erledigen.

überhaupt nicht wichtig ○ ○ ○ ○ ○ ○ ○ sehr wichtig

Die Anwendung soll übersichtlich, verständlich und leicht zu lernen sein.

überhaupt nicht wichtig ○ ○ ○ ○ ○ ○ ○ sehr wichtig

Die Bedienung der Anwendung soll vorhersehbar und gut kontrollierbar sein.

überhaupt nicht wichtig ○ ○ ○ ○ ○ ○ ○ sehr wichtig

Das Arbeiten mit der Anwendung soll interessant, spannend und aktivierend sein.

überhaupt nicht wichtig ○ ○ ○ ○ ○ ○ ○ sehr wichtig

Die Anwendung soll originell, innovativ und kreativ gestaltet sein.

überhaupt nicht wichtig ○ ○ ○ ○ ○ ○ ○ sehr wichtig

Figure A.8: Post-study questionnaire (page 3).

Participant ID:___

Inwieweit fühlten Sie sich geistig in das Erlebnis versunken?

Überhaupt nicht ○ ○ ○ ○ ○ ○ ○ Sehr

Inwieweit fühlten Sie sich in das Erlebnis hineingezogen?

Überhaupt nicht ○ ○ ○ ○ ○ ○ ○ Sehr

Wie umfassend waren Ihre Sinne einbezogen?

Überhaupt nicht ○ ○ ○ ○ ○ ○ ○ Sehr

Wie sehr erfuhren Sie ein Gefühl der Realität?

Überhaupt nicht ○ ○ ○ ○ ○ ○ ○ Sehr

Wie entspannt oder aufregend war Ihr Erlebnis?

Sehr entspannt ○ ○ ○ ○ ○ ○ ○ Sehr aufregend

Inwieweit hatte es den Anschein, dass die Objekte, die Sie sahen, zu Ihnen kamen?

Überhaupt nicht ○ ○ ○ ○ ○ ○ ○ Sehr

Inwieweit hatte es den Anschein, dass Sie die Objekte, die Sie sahen, erreichen und berühren konnten?

Überhaupt nicht ○ ○ ○ ○ ○ ○ ○ Sehr

Figure A.9: Post-study questionnaire (page 4).

Participant ID:___

Wie häufig wollten Sie einem Objekt, das auf Sie ausgerichtet zu sein schien, aus dem Weg gehen?

Nie ○ ○ ○ ○ ○ ○ ○ Immer

Inwieweit empfanden Sie das Gefühl, in der Umgebung zu sein, die Sie sahen/hörten?

Überhaupt nicht ○ ○ ○ ○ ○ ○ ○ Sehr

Wie oft wollten Sie etwas, das Sie gesehen haben, versuchen anzufassen?

Nie ○ ○ ○ ○ ○ ○ ○ Immer

Hatte Sie den Eindruck, dass Sie die Ereignisse, die Sie sahen,
eher auf einem Videobildschirm oder durch ein Fenster wahrnahmen?

Wie ein Video-Bildschirm ○ ○ ○ ○ ○ ○ ○ Wie ein Fenster

Figure A.10: Post-study questionnaire (page 5).

Session ID: ___

Interview

Studienbeginn: ____________ Studienende: ____________

1) Wie ist Ihr erster Eindruck des Systems?
2) Worin sehen Sie die Stärken des Systems?
3) Worin sehen Sie die Schwächen des Systems?
4) Sind Sie gut mit den zwei Geräten zurechtgekommen?
5) Sind Sie gut mit den verschiedenen Inputmöglichkeiten zurechtgekommen?
6) Konnten Sie die Blindbedienung benutzen, oder haben Sie viel auf das Tablet schauen müssen?
   a) Haben Sie eine Eingewöhnungszeit gebraucht, oder war von Anfang an alles klar?
   b) *Falls nicht klar*: Fanden Sie das Konzept allgemein unverständlich? Was genau hat Ihnen Probleme gemacht?
7) War die Aufteilung klar, bei welchen Aufgaben man auf das Tablet schauen soll, und bei welchen Aufgaben man das Tablet blind bedienen kann?
8) Konnten Sie die Datenfilter so setzen, wie sie wollten? Hatten Sie Probleme bei der Genauigkeit beim einzeichnen des Filters?
9) War ihnen immer klar welches Objekt gerade ausgewählt war?
   a) Sie haben auch manchmal eine Linie gesehen, die zum ausgewählten Objekt gezeigt hat. Haben Sie diese wahrgenommen, hat diese Linie Sie gestört?
   b) Hat die Selektion von verschiedenen Objekten funktioniert, oder haben Sie immer aus Versehen etwas ausgewählt was Sie nicht wollten?
10) Gab es allgemein Sachen, die Sie bei der Interaktion gestört hat?
11) Gab es allgemein Sachen, die Ihnen bei der Interaktion unklar waren?

Figure A.11: Questions for semi-structured interview (page 1).

Session ID: ___

12) Bezüglich Sprachsteuerung:
   a) Falls oft benutzt: Haben die Voice commands für Sie gut funktioniert, oder hatten Sie Probleme, z.B. mit der Erkennung?
   b) Falls oft benutzt: Würden Sie die Voice commands auch noch in einem kollaborativen Szenario benutzen, also wenn Sie mit anderen Leuten zusammenarbeiten?
   c) Falls nicht benutzt: Gab es einen Grund warum ihr Voice commands vermieden habt? (Oder hattet ihr einfach nicht präsent, dass ihr auch Sprachbefehle benutzen konntet?)
   d) Welche Interaktionsmöglichkeit bevorzugen Sie? (Touch oder Voice?)
13) War die Visualisierung verständlich?
14) Von den Funktionen, die Sie ausprobiert haben, hat irgendwas gefehlt, was Ihnen die Analyse noch erleichtert hätte? (Etwas wo Sie gesagt hätten, es wäre gut, wenn ich **das** jetzt machen könnte)?
15) Hatten Sie Probleme beim Positionieren der Scatter Plots?
   a) Konnten Sie die Scatter Plots an die Position setzen, die Sie wollten?
   b) Hätte es Ihnen geholfen, die Scatter Plots an echte Orte zu binden (z.B. den Scatter Plot an die Wand zu kleben), oder finden Sie es besser den Scatter Plot in der Luft zu positionieren?
16) Wie erschöpft fühlen Sie sich vom Tragen des iPads, könnten Sie das noch ein paar Stunden mehr tragen (vor allem wenn es leichter wäre)?
17) Hat es ihnen geholfen, die echte Welt noch um sich zu sehen, oder fanden Sie das eher ablenkend?
18) Haben Sie sonstige Anmerkungen?

Figure A.12: Questions for semi-structured interview (page 2).

B

Enclosure

The attached SD-card includes all relevant files for this work, in particular:

- Digital version of this document
- Source code for the MIDAiR system, consisting of:
    - Server
    - Microsoft HoloLens AR application
    - Vive Tracker capture software
    - Tablet web application
- Source code for the *viewer mode* of the MIDAiR system
- Slightly altered source code of MIDAiR used in the usability study
- Binary file for deployment of MIDAiR on the Microsoft HoloLens
- 3D model for custom tablet frame
- Image files for visual Vuforia markers