The Australian National University
2601 ACT | Canberra | Australia

Australian National University

School of Computing

College of Engineering and
Computer Science (CECS)

Multi-View Web Interfaces in Augmented Reality

— 24 pt Honours project (S2 2021–2022)

A thesis submitted for the degree
Master of Computing (Advanced)

By:
Feiyue Tao

Supervisors:
Dr. Matt Adcock
Dr. Charles Martin

June 2022

Declaration:

I declare that this work:

- upholds the principles of academic integrity, as defined in the University Academic Misconduct Rules;

- is original, except where collaboration (for example group work) has been authorised in writing by the course convener in the class summary and/or Wattle site;

- is produced for the purposes of this assessment task and has not been submitted for assessment in any other context, except where authorised in writing by the course convener;

- gives appropriate acknowledgement of the ideas, scholarship and intellectual property of others insofar as these have been used;

- in no part involves copying, cheating, collusion, fabrication, plagiarism or recycling.

June, Feiyue Tao

Abstract

The emergence of augmented reality (AR) is reshaping how people can observe and interact with their physical world and digital content. Virtual instructions provided by see-through AR can greatly enhance the efficiency and accuracy of physical tasks, but the cost of content authoring in previous research calls for more utilization of legacy information in AR. Web information is a great source hosting a wide range of legacy and instructional resources, yet current web browsing experience in AR headsets has not exploited the advantage of 3D immersive space mixing the real and virtual environments.

Instead of creating new AR content or transforming from legacy resources, this research investigates how to better present web interfaces in AR headsets, especially in a physical task instruction context. A new approach multi-view AR web interfaces is proposed, which suggests separating web components into multiple panels that can be freely arranged in the user's surrounding 3D space. The separation and arrangement would allow more flexible combination of web content from multiple sources and with other AR applications in the user's field of view.

This thesis presents a remote and self-guided elicitation user study with 15 participants that derives layout arrangement preferences of the proposed multi-view interfaces. The study uses a VR system developed to simulate three scenarios of performing real-world tasks instructed by multi-view AR web content involving different types of media. The study analyzes how users arrange such web interfaces, and the system also simulates various physical environments and general AR applications to investigate their impact on the virtual interface arrangement.

According to participant survey responses and interface arrangement data, the study identifies patterns in interface layout, grouping relationships between interfaces, physical environment constraints, and relationships between web interfaces and general applications. Then five implementation strategies are suggested based on the design preference findings.

Table of Contents

Table of Contents

List of Figures

List of Figures

List of Tables

x

Listings

Introduction

Augmented reality (AR) technology enhances a human's real-world perception by adding virtual information generated by computers in their normal field of vision (Azuma, 1997). Different from virtual reality (VR) where the user is immersed in entirely digital surroundings, an AR system enables real-time user interaction with virtual objects in a physical environment. However, one challenge is to provide proper AR content with low cost of authoring. This project explores how we can make use of existing 2d content within 3d physical task augmentation.

An AR head-mounted device such as HoloLens superimposes virtual contents on the user's view of the real world. Compared to handheld displays such as mobile AR and monitor-based video AR displays that limit virtual information within the screen (Milgram et al., 1995), AR headsets could provide a more immersive experience, allowing the user to interact with virtual elements in the surrounding physical 3D space. The user would also be able to focus on any physical tasks without their hands occupied. With technologies such as scene understanding and camera tracking, virtual elements can be rendered relative to the physical objects (e.g., by enabling occlusion) to simulate their existence in the real world.

There have been academic and commercial studies utilizing AR headsets to provide virtual task instructions for physical activities in various fields such as manufacturing, education, and medical science (Park et al., 2021). Many of them implemented an application specific to a task to evaluate the efficiency of AR task instruction. Since creating a particular application could be costly, it would be valuable to utilize existing legacy information in AR (Wang et al., 2016). A website is a typical media format to display various types of information, such as text, images, and videos. The Web is also useful for finding instructional information. However, there has so far not been much research done to investigate how to present websites in AR exploiting the advantage of mixing virtual interfaces into the 3D physical space.

This project proposes a new idea of multi-view web interfaces in AR, which is to turn one web page into a set of web components that can be freely arranged in the augmented 3D space. We present a remote and self-guided elicitation user study that analyzes patterns of arranging such web components. To support the remote setting and taking into account different physical environments, the study involves an application deployed on VR headsets to simulate multiple real world scenes and AR interfaces. The application features three scenarios for working on some physical tasks while viewing AR web instructional information. There are different types of information simulated in the application, including texts, images, videos, and general applications other than websites. The study tasks are rearranging those virtual interfaces in each scenario, and data can be collected via virtual object data files generated by the application and an online survey form. The purpose of this study is to analyze how multi-view AR web interfaces could be presented in a suitable layout considering usage of other AR applications and different physical environments or activities.

1.1 Motivation

Previous research has investigated the use of AR to guide physical tasks such as assembly. AR multi-media and interactive instructions has been shown to provide an improvement of task completion efficiency and accuracy compared to traditional instructions presented on paper or video manuals. AR headsets are useful for displaying context-sensitive instructions by overlaying virtual information in the operator's field of view. The operator may understand such instructions more easily and can focus on the ongoing physical task at hand. Some AR applications could detect the operator's action and annotate physical objects with virtual information to provide particular interactive and procedural instructions, yet a common problem for such AR tools is that the authoring process could be complicated and time consuming (Wang et al., 2016) to create tailored instructions. There are studies working on reducing the authoring cost, for example, by automatically segmenting existing videos to generate step-by-step instructional AR animation (Yamaguchi et al., 2020). On the other hand, the authoring cost problem highlights the value of legacy information, including texts, images, and videos that have already been hosting instructional content. Instead of **transforming** legacy information to interactive AR interfaces, this research investigates how to better **present** legacy information, especially in a physical task instruction context.

The design of legacy information presentation for AR needs extra considerations such as fitness to the field of view in the AR display, interaction and overlay between the real world scenes, and the user's cognitive workload to perceive virtual contents. Without appropriate design considerations, real-time virtual information rendered in the user's view may interfere with the physical tasks, make the interface confusing, or even cause adverse health effects such as dizziness (Moro et al., 2017). Design preferences can also vary depending on AR usage scenarios because the different levels of complexity of the real-world environment or tasks can influence how much information the user prefers

to receive from the virtual view (Lindlbauer et al., 2019; Klose et al., 2019). Previous research has investigated AR presentation of a particular data type in details. Take text presentation in AR as an example, existing studies have suggested preferences of the typeface (Bernard et al., 2003), text size (ms, 2019; Renkewitz et al., 2008), text colour (ms, 2019; Debernardis et al., 2014), positioning and anchoring in the field of view (Klose et al., 2019; Fukushima et al., 2020) for better readability and coexistence with real-world tasks. We further explore AR presentation of multiple types of legacy information.

This research focuses on presenting web content with AR in the context of instructing physical tasks. Websites can host different types of textual or visual legacy information at one place, and they provide a wide range of contents easily accessible to general users. There are existing studies investigating web interfaces on mobile AR (Lam et al., 2019) or layouts of presenting multiple general applications on AR headsets (Cheng et al., 2021), but not much research has so far been done particularly for web content presentation on AR headsets. AR headsets such as HoloLens can display a virtual web browser, yet it still follows the traditional flat web browser as on a computer screen (Figure 1.1). The AR web browser still requires the user to view information in a 2D style, switch between tabs or browsers to view different web pages, scroll (with hand or eye) up and down to view full contents, etc. Improvement could be expected to further exploit the features of an AR environment: the user can freely look around their surrounding physical 3D space and interact with both virtual and real objects.
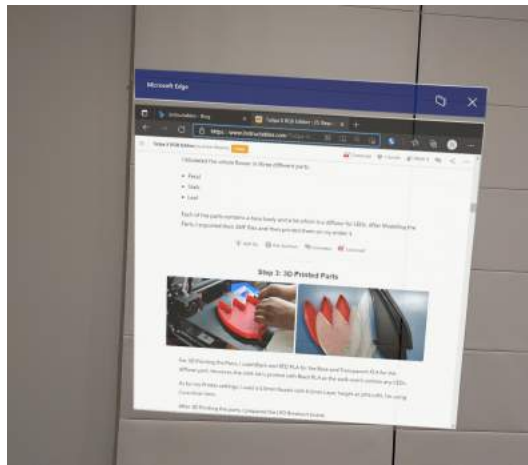

Figure 1.1: Screenshot of a website displayed in HoloLens.

1.2 Research Proposal and Question

This project proposes a new style for presenting web information in AR headsets: multi-view AR web interfaces, which is to separately display web information components and allow free arrangement in the 3D space. The separation could be based on the infor-

mation's media type (e.g., text, image, video) and semantic relationship (e.g., contents related to the same physical object). An example of this idea is illustrated in Figure1.2. In a task instruction scenario, the user may focus on the (either physical or virtual) task-related objects in their view while an AR web page. The crossed rectangle represents objects to operate for the physical task, and the blue frame represents the user's field of view. The diagram on the left side shows the existing web browser style similar to Figure 1.1, where the user would need to interact with the browser as a whole and view all information in the single flat window. The right-side diagram presents a simple example of multi-view web interfaces where the three images and texts on the same web page can be separately arranged. If the user would like to treat those images as more important information, they may place the images at the top of their view along with the task objects so that the images serve as visual step-by-step task instructions. Each image could also be placed relative to some particular physical objects. The long textual content, which may be less significant, can be put out of the current view but still visible if the user turns their head.
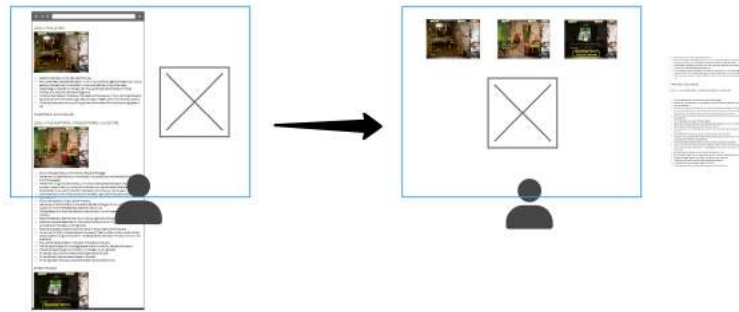
Figure 1.2: Illustration of the idea of multi-view web interfaces. The crossed box represents physical task objects.

Therefore, this proposal of multi-view AR web interfaces emphasizes:

- Flexibility of arranging web components individually. Information is not limited within the original single web page. Virtual contents from the same page could be positioned differently in the user's view; contents from different pages could also be grouped together.

- Selection of key information. There could be an abundance of information to perceive from a web page, but the user's field of view in an AR headset and a comfortable cognitive load is limited. With a multi-view interface, the user may select what seems more important to keep in their view. Less interesting information could be put somewhere less visible or closed directly.

- Placement of particular web information relative to the physical environment. Since the interaction unit is no longer the entire web browser but individual web components, different information could be placed considering different physical

scene constraints. For example, an instruction video could be attached to a wall, while a textual paragraph describing an appliance could be placed near the corresponding physical object.

- Larger user interaction space. Information is no longer confined in the flat 2D browser, but could be moved around the user's surrounding 3D space.

Based on the proposed concept, this research explores the following questions:

1. whether the multi-view AR interfaces could be useful to display web information especially for task instructions;

2. what could be preferred arrangement layouts for different web information components;

3. how the arrangement of AR web interfaces may be affected by different real world scenarios including the physical environment, objects, and activities;

4. how the arrangement of AR web interfaces may be affected by other AR applications that could be used along with websites.

1.3 Contribution

The main contributions of this project are as follows:

1. A proposal of a new approach to present immersive AR web information: multi-view AR web interfaces.

2. A design of a remote and self-guided elicitation user study to elicit multi-view AR interface layouts.

3. An implementation of a VR system that simulates various interfaces of AR websites or apps and physical environments to conduct the remote elicitation study.

4. An analysis of data collected from the VR system and an online survey to evaluate the idea of multi-view AR web interfaces and derive design principles or implementation strategies.

1.4 Thesis Outline

The remainders of this thesis is structured as follows.

Chapter 2 provides background information for the research area and presents a literature review of the relevant work. It investigates existing techniques or principles in relation to AR interface design about layout arrangement, task instructions, and legacy or web information presentation. The review also includes examples of previous studies that applies methodologies relevant to the remote self-guided elicitation user study performed in this project.

Chapter 3 presents the design and implementation of the prototype system developed with VR and used for the user study. A description of the hardware and software platforms is provided, including the justification of utilizing VR techniques. The system simulates multi-view AR web interfaces, general AR application interfaces, and different scenarios involving physical environments and task activities, which are described respectively. Also presented is an explanation of how the system supported the remote and self-guided user study with its UI instruction flow and data saving/loading methods implemented.

Chapter 4 describes the elicitation study conducted. It explains the study set-up particularly considering the remote and user self-guided features. The tasks undertaken by users and the study procedure are presented. The previous chapter explains how the VR system saves virtual interface interaction data, while this chapter describes another data collection method via an online survey.

Chapter 5 discusses the results of the user study, covering quantitative and qualitative data collected from the VR system and the online survey. Based on the interface arrangement layouts generated by users and their opinions about the interaction experience, the analysis derives design and implementation strategy for the proposed multi-view AR web interfaces.

Finally, Chapter 6 provides a conclusion by summarizing the research activities and findings and discussing potential future work.

Background and Related Work

This chapter provides background information and a review of the literature on the research topic explored in this project. Section 2.1 introduces the classification within extended reality which is a superset including AR and VR. Related studies are reviewed in Section 2.2 for AR user interface design in general including layout arrangement, Section 2.3 focusing on web interfaces in AR, and Section 2.4 for using AR to present task instructions.

2.1 Extended Reality Spectrum

There is a spectrum of computer-enhanced reality technology that presents gradually varied levels of virtuality. Milgram et al. (1995) introduced a Reality-Virtuality continuum which defines the boundaries of AR between real and virtual environments (Figure 2.1). Extended Reality (XR) is an umbrella term that encapsulates the continuum from the real to the virtual environment, including AR and VR (Chuah, 2018). The figure also displays examples of related devices: Microsoft HoloLens is an optical see-through AR headset that superimposes holographic content on the user's view of the real world, while Oculus Quest is a typical VR head-mounted display that immerses users in a completely virtual world. The research area of this project locates around AR on the left in the spectrum, but VR techniques on the right are also utilized to aid the study conducted. AR systems combine the real world and virtual information, support real-time interactivity, and use three dimensional registration to align real and virtual objects (Azuma, 1997). The way to augment digital content onto the physical world can be categorized into three targets to augment: the user, the physical object, and the environment (Mackay, 1998). User augmentation perceives physical object information via head-mounted or handheld devices, object augmentation embeds devices such as sensors within the objects, and environment augmentation projects images onto the physical environment and captures user interactions. This project falls into the user augmentation

approach with a head-mounted display (HMD).
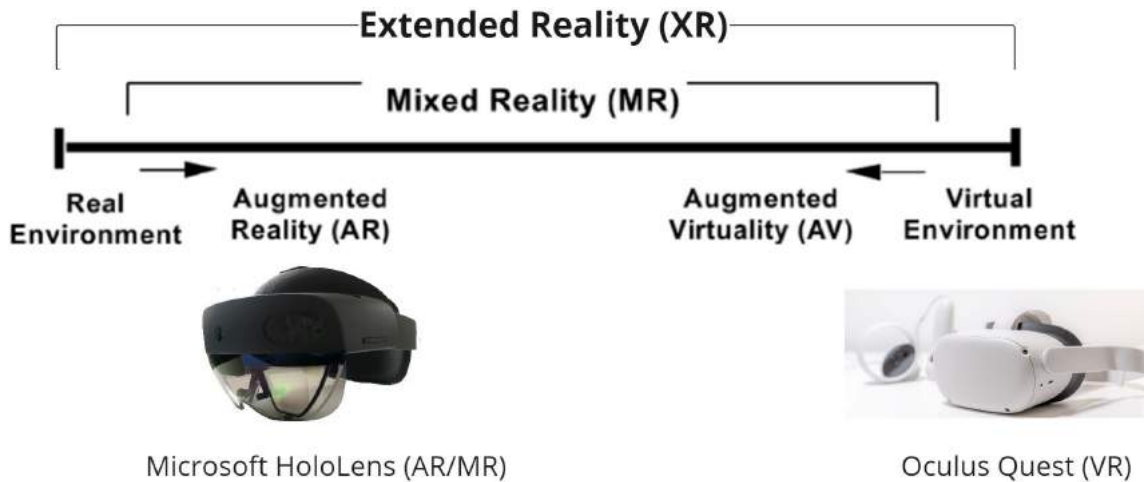
Figure 2.1: The Reality-Virtuality Continuum (Milgram et al., 1995) added with the range of extended reality and typical devices.

2.2 General AR Interface Design and Layout Arrangement

General human-computer interaction (HCI) principles can be developed and applied to AR interface design. Dünser et al. (2007) summarized 8 key design principles or usability heuristics that have been established in HCI research and could be tailored in AR systems:

1. Affordance: an action possibility (Gibson, 1977) formed by the relationship between user and an object. AR objects, or tangible user interfaces in particular, may inherit interaction metaphors from real world examples. Familiar shapes such as "a pen" may indicate the possible user interaction and provide passive tactile feedback.

2. Reducing cognitive overhead: excessive non-automatic effort required by the system could turn into distraction, especially for novice users who are not familiar with novel AR systems yet.

3. Low physical effort: minimize interactions steps to decrease chances of arm fatigue or motion sickness.

4. Learnability: AR presents novel interaction techniques and should support an easy learning process. This could be achieved by adopting intuitive interaction techniques that simulate real-world behaviors or traditional interactions with a computer, self-descriptiveness of UI elements, and the consistency in interface appearance and behaviour.

5. User satisfaction: usability could be evaluated by objective measures regarding efficiency, but also subjective user perception such as the system being engaging or "fun".

6. Flexibility in use: The mixture of reality and virtuality in AR enables integrating different input/output devices or human senses. Multi-modal interactions can complement each other.

7. Responsiveness and feedback: system responsiveness could be challenging for AR if tracking performance is slow. Proper feedback would inform the user about the interaction status and could be multi-sensory as well to increase the system enjoyment.

8. Error tolerance: this could be also challenging for AR due to tracking instability. One AR research area is to develop algorithms or input channels that improve the accuracy of spatial registration of physical and virtual objects.

A study of context-aware MR interfaces (Lindlbauer et al., 2019) investigated the automatic adaptation of UI elements in MR/AR according to the cognitive load related to the usage context. The UI elements are general applications such as email, YouTube, and a task list. The system detects pupil dilation to infer different cognitive load required for the user's current task then adjusts the number, positioning, and level of details of those UI elements. For example, if the user is reading a paper which needs concentration and higher cognitive need, the system only displays a limited number and size of application icons out of the view without additional non-urgent information; on the other hand, when the user is doodling which may require less mental effort, more application icons are displayed in the view with more detailed contents such as video preview image or email excerpts. This study worked on cognitive load computation and automatic placement of general UI elements that may not be relevant to the user's main activity. Our research focuses on placement elicitation of AR web interfaces that are needed for or related to the physical task context, while the simulation of different scenarios also covers daily or more technical use cases to indicate a varied cognitive load.

Multi-view layouts have been applied to 2D maps, and Maps Around Me project (Satriadi et al., 2020) explores the extension of 2D multi-views into 3D immersive spaces with VR. The layout consists of a large overview map and multiple smaller maps that may present different hierarchical levels such as from country to city. The research used VR to create initial multi-view maps and conducted an exploratory study into how users locate and view the immersive maps. Participants were asked to perform map usage tasks such as route planning, then the map arrangement layouts were observed, analyzed, and classified. Based on the results, user preferences were identified regarding map layout geometry, relationships between the overview and detailed maps, and interaction patterns of whether the layout changed during the task. The layout of placing maps in a spherical cap (Figure 2.2) was found to be dominant (over 80%) and convey maximised map visibility . The research further implemented a layout manager system that enables controlling map locality and size to switch between different layout geometries, which

was then evaluated by a few user walk-through sessions. Our idea of multi-view web interfaces also tries to display originally integrated 2D information as multiple panels and explores suitable layouts, but the differences include the following factors:

1. Different web interfaces would not be geographically hierarchical but instead vary semantically. There are also different types of textual and visual information presented. Therefore the relationship between panels could be further explored, which should be expected to be different from the overview-detail relationships for maps. The judgement of placing web panels may also be based on semantic comprehension in addition to visual perception.

2. The arrangement of maps in VR only concerns the visibility or usability of maps themselves, while to investigate AR interfaces, the impact from the physical environment should be considered.

3. For the task instruction context in our research, the panel placement and interrelationships could be affected by the task-related objects that exist either physically or virtually.

Figure 2.2: Maps Around Me (Marner et al., 2013):multi-view maps in 3D space (VR).

2.3   Web Interfaces in AR

In the literature, the term Web-AR (or similarly WebXR including VR (MacIntyre and Smith, 2018)) often refers to processing and displaying AR content based on a web browser or a web-based app in contrast to native applications, which has been investigated for mobile AR (MacIntyre et al., 2011; Qiao et al., 2019; Shepiliev et al., 2021). Such research focused more on real-time image processing on the Web with techniques including computer vision algorithms, JavaScript libraries such as WebGL, and cloud or edge computing (Qiao et al., 2018; Klein et al., 2013). To clarify, our research focuses on user interface design for displaying web content in AR, especially head-mounted hardware-based AR.

Earlier AR browsers were reviewed by Langlotz et al. (2014), who pointed out the need of AR browser interfaces to further support more types of digital media, seamlessly integrate digital content into the real world, and adapt to the user's current context or environment. A prototype browser was implemented to present textual, audio, video, and 3D media. To demonstrate possible context awareness techniques, the study presents video processing to detect important areas and geometric structure, including edge and surfaces, which could be applied to physical scene understanding. The survey also suggested that "upcoming" head-mounted displays would require new characteristics of AR interfaces including web browsers.

There are previous studies transforming existing websites to AR mode. A proof-of-concept system to support AR web content adaption was presented by Srinivasa et al. (2016). Given a demo web page scripted to define the content to be displayed in AR, the proposed AR browser system removes the original background of the web page, sets the video stream of a Quick Response (QR) code captured by a camera as the browser background, and separates and places the specified web content around the QR code, which was mainly achieved by web development techniques (JavaScript/CSS API). They also mentioned depth perception as a possible enhancement compared to the planar layout. The web content arrangement surrounding a QR code is in a 2D style not considering context-based customisation.

Another system was proposed by Lam et al. (2019): Mobile to AR (M2A), a context-aware web content decomposition framework for mobile AR. They introduced Layout CSS as a CSS extension to specify and overlay web content over the physical view. The decomposition style reuses the fluid grid system to separate information into blocks, following the responsive design approach that has been established for mobile web development. Context awareness is demonstrated by detecting identifiable physical objects such as an appointed marker (the restaurant logo in Figure 2.3(a)) or a QR code, also considering the user's GPS location to filter Geo-located content. Their experiment shows that the time taken to find web information with a M2A application appears to be 5 times faster than that with a traditional web browser on the smartphone, and 2 times faster than a simple AR web browser that does not support context awareness or content decomposition.

Compared to the former study, M2A shows more generalized web content transformation framework and a more responsive layout. The two studies demonstrated web development techniques to decompose and adjust the web content to fit AR browsing needs. The evaluation result could support a hypothesis that context-aware multi-view web interfaces could be more efficient than a conventional single-window browser in terms of perceiving web information. The augmented view in the studies mentioned above is displayed on a screen instead of an immersive see-through headset, and the layout is surrounding a specific marker. Our study investigates the design especially the layout of web interfaces to be displayed in a more complex immersive physical environment.

The same research group of M2A published a very recent study (Lam et al., 2021) that

(a) M2A: Web content
decomposition for Mobile AR

(b) Raw web contents.

(c) Context-aware recommendation.

A2W: Web content recommendation for AR headsets

Figure 2.3: Mobile-to-AR framework (Lam et al., 2019) and AR-to-Web recommendation
system (Lam et al., 2021).

moved from mobile AR to AR headsets and with more generalized object detection techniques. They proposed an AR-to-Web (A2W) system, which recommends web contents based on the physical scene in the user's view with context-based and spatial filters. All web contents are stored and classified in a dataset. The recommendation weights would be updated by user interactions (clicking and scrolling) and rating of web contents. The web content dataset involves GPS coordinates to label location-based web content and web genre information for categorization. The study presented a simulation with Google Street View as the physical scene and data from Google Map Place API as the web contents. It transformed original 2D web contents that were directly spread on the screen (Figure 2.3(b)) to a context-aware layout (Figure 2.3 (c)) that filters the content with the proposed recommendation service, modifies the layout based on the buildings' distance and surface angle, and displays icons for the defined categorization.

Web views displayed in A2W individually relates to physical objects in the scene but not to each other. The study focused on the recommendation algorithm to find the right content for each identified physical entity (e.g., a shop on the street). The layout adaption especially considered the user's GPS location and their distance between the physical entity so that the content for a closer building would appear bigger in the view. For our study, web views are more semantically related, which could be different sections or different types of media separated from the single original web page. The context is task instruction where there would be some specific physical objects to concentrate on instead of wandering on the street. Our focus is more on layout arrangement than (automatic) content selection. As stated by the authors and to the best of our knowledge, A2W might be the first (and the most recent) work contributing to context-aware web browsing using AR headsets. The arising interest could also indicate a promising research space of immersive AR web interfaces.

2.4   Task Instructions in AR

Previous research has investigated procedural task instruction in AR. Henderson and Feiner (2011) presented an AR task instruction system that supports detecting the result of user operation and automatically proceeding to the next step. This was achieved by tracking multiple physical objects and the user's head movement, and representing task steps in a finite-state machine. However, the marker-based tracking and prescriptive feedback would require high authoring cost to create instructions specific to a task. Another step-by-step AR guidance tool was presented by Khuong et al. (2014) with a LEGO instruction application. It enables automatic detection of the error or completion state, applying context awareness by recognizing real-time user operations and showing the adaptive instead of predefined next-step instruction. This was achieved by comparing the actual and the target models in a 3D grid space to estimate the next layer to be filled or to detect an incorrectly placed block. The study also compared two AR visualization modes: overlaying virtual guidance information directly on the physical objects, or rendering the instruction as a virtual model adjacent to the real one and following the same orientation changes. The latter method led to better results in the experiment, suggesting that some space would be desired between the virtual content and the physical task to avoid confusion from overlaying.

To reduce content authoring cost, video is a typical legacy resource that can be utilized for task instruction with AR, given that it provides visual and procedural guidance. Goto et al. (2010) presented a task support system by overlaying a segmented instructional video on the user's view of work area in an origami and block placement context. The perception of real world is via a head-mounted camera and blended with the video. Visual effects such as background transparency and contour enhancement are applied to avoid confusion, and the user can follow the instructor's hand in the overlaid video step-by-step. More recently, the similar research group proposed a new system AR Assembly Mirror ((Yamaguchi et al., 2020)), which can automatically translate a given 2D instructional video into a procedural 3D model tutorial labelled with cropped video excerpts in each step. The animated virtual model is displayed on a monitor and aligns with orientation changes of the real object so that the user can see how the parts should be manipulated from the same viewpoint. Another recent study (Lu and Mayol-Cuevas, 2021) also worked on automatic instructional video segmentation, which is based on hand-object interaction detected with Convolutional Neural Network techniques. The first study shows a way to present step-by-step video-based instructional information in an AR headset, and the latter two studies aim to reduce the authoring cost of preparing procedural video instructions ready to be displayed in AR.

2.5   Summary

To provide background information for this research regarding AR web interfaces, this chapter firstly clarified our research area in the extended reality spectrum: AR with user augmentation approach using head-mounted display. We reviewed previous work related

to general AR interface design in terms of HCI principles that could be tailored for AR and studies that worked on AR interface layouts. More specific discussion was made for previous studies around AR web interfaces. Context awareness was found to be one of the most important features for AR interfaces, which could customize UI elements according to the user's physical environment, activity context, or cognitive load. There was study investigated layout arrangement for other types of interfaces, and our research focused on layout design for web interfaces in AR. Previous research has investigated AR web content creation or transformation especially for mobile AR, but the design for immersive web interfaces displayed in AR headset appears to be a research gap.

This motivation led this project to design and conduct an elicitation study to formulate design principles for multi-view AR web interfaces. There has been research using VR systems to conduct remote user studies (Mottelson et al., 2021; Rivu et al., 2021), and our research developed a VR system to simulate task instruction scenarios with multi-view AR interfaces, which supports remote and self-guided elicitation user study. The details will be described in the following two chapters.

System Design and Implementation

A prototype VR system was developed to simulate different physical environments and AR web or app interfaces for performing the elicitation study. There are 3 scenarios of users doing physical tasks while following web instructional information and using general AR applications. Each virtual scene presented in the system consists of three key components: simulated physical environment and objects, simulated AR web or app interfaces, and UI instructions for the study task flow. The user would interact with the system following the UI instructions to freely arrange or modify the simulated AR interfaces in each scenario. The system can remotely save the interface transformation data including position, rotation, and scale, which were then analyzed to explore multi-view AR web interface layouts preferred by users and derive design principles.

This chapter describes the design and implementation of the prototype VR system. Section 3.1 introduces the essential toolkit utilized to develop the system. Section 3.2 describes the implementation details of the AR interfaces simulated in the system, while details of how the system presents different physical environments and scenarios are described in Section 3.3. Since the system was built as the tool for a remote and self-guided user study tool, additional considerations were needed to support smooth self-guidance and remote data collection. Section 3.4 explains how the system guides a remote and self-guided user study with UI instructions, and Section 3.5 describes the remote data saving and loading methods implemented. More details about the user study setup and procedure are presented in Chapter 4.

3.1 Hardware

The hardware platform to deploy the system is Oculus Quest and Oculus Quest 2 (re-branded as Meta Quest 2) headsets. VR headsets can simulate an immersive artificial world where the user can look or move around and interact with virtual items. The

reasons for choosing VR instead of AR devices to build this elicitation study system are:

- To simulate different physical environments. Although a VR system cannot exactly reproduce the interaction between virtual interfaces and real world objects, virtual scenes provide the flexibility to experience multiple environments and task scenarios within one application. Such flexibility enables analysis of how physical environment variation may affect the virtual interface arrangement. By contrast, a study with an AR headset would be hard to customize and diversify the real world space where the participant is physically staying.

- To support remote study. The study was designed to be remote complying with COVID-19 social distancing regulations during the research period. AR headsets such as HoloLens are not owned by individuals as commonly as VR devices due to limited availability.

The system was released in the Oculus App Lab so that test users or subscribed users could easily access and install it on their own Quest headsets. The procedure for users to get access to the system and complete the study is detailed in Section 4.4.

Unity was the primary software platform for building all virtual objects or scenes, associated with Visual Studio to script logical behaviors in C#. Mixed Reality Toolkit (MRTK) was utilized to develop simulated AR web and app interfaces. MRTK provides a set of mixed reality UI building blocks and features that supports spatial interactions and a range of AR/VR platforms including HoloLens and Quest. Therefore, the VR system could also present interfaces similar to what users could see with an AR headset HoloLens (Figure 3.1).

(a) HoloLens interface   (b) VR simulation
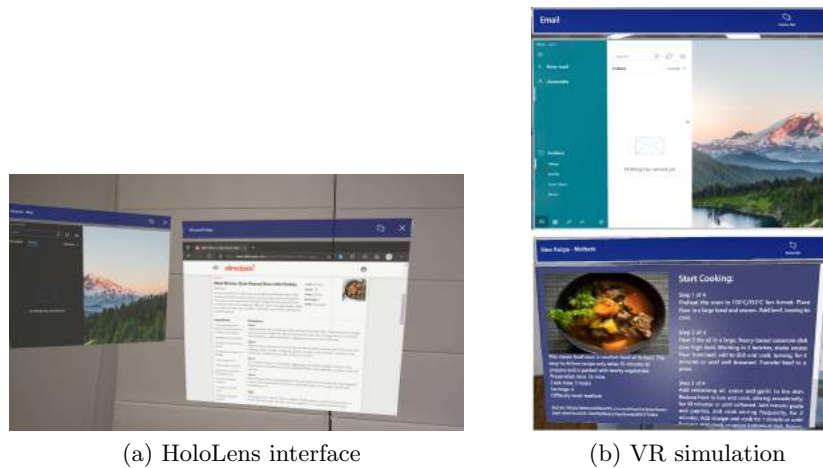
Figure 3.1: Comparison between real AR interfaces of an email application and a website in (a) HoloLens and (b) the implemented VR simulation

3.2 Virtual Interfaces

The system simulated two types of AR interfaces: multi-view web information panels described in Section 3.2.1 and general application UIs described in 3.2.2. Users would be asked to arrange web information panels with or without general applications. All interfaces are initially stacked at back of the user's view in order to increase the incentive to arrange the panels in their preferred ways rather than leaving them as default. The initial order of interfaces' closeness to the user is randomized so that the order of interfaces to be interacted with also varies for different users.

3.2.1 Multi-View AR Web Interface Simulation

The context of "viewing websites" in the system is to work on some physical tasks at hand while browsing websites for instructions or related background knowledge. Figure 3.2 demonstrates the types of legacy and web information presented in the system: text, image, PDF document, and video, which are some common data types used in websites and/or task instruction contexts. Additional types specific to the task context such as a checklist and 3D model were also included.

(a) Text   (b) Image   (c) PDF document

(d) Video   (e) checklist   (f) 3D model

Figure 3.2: Different types of AR web components simulated in VR

The idea of "multi-view interfaces" turns traditional single web pages into separated panels that can be freely arranged. As illustrated in Figure 3.3, a traditional recipe website could contain ingredients, cooking steps, reviews, video demonstrations, etc., on

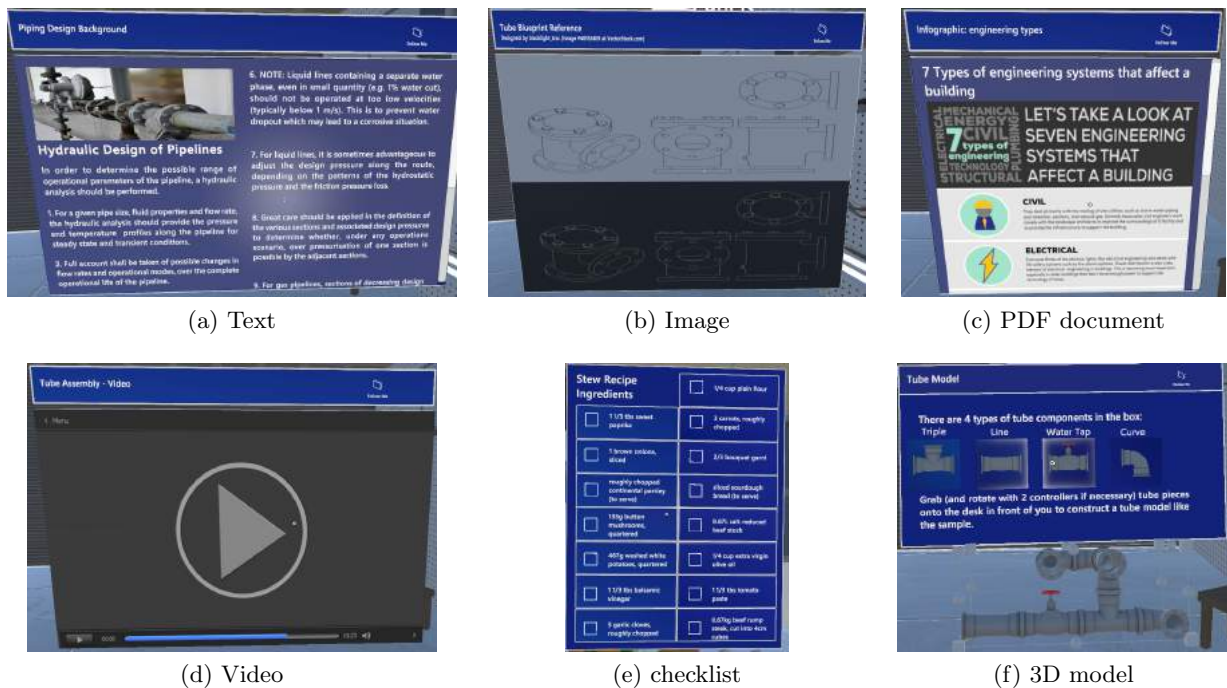the same page; in the kitchen scene developed in the system, the information is divided into multiple views (panels) to be arranged individually. Common interactions with these information panels include:

- Grabbing and moving the panel with one controller;

- Resizing the panel by dragging the panel's corner;

- Rotating the panel backward and forward by dragging the panel's edge;

- Toggling whether the panel follows the user's view by clicking the "Follow Me" button at the top-right corner.

Furthermore, long textual content (Figure 3.2 (a) and (c)) supports vertical scrolling; full image content (Figure 3.2 (b)) allows pan-and-zoom interactions to move or resize the image within the panel.
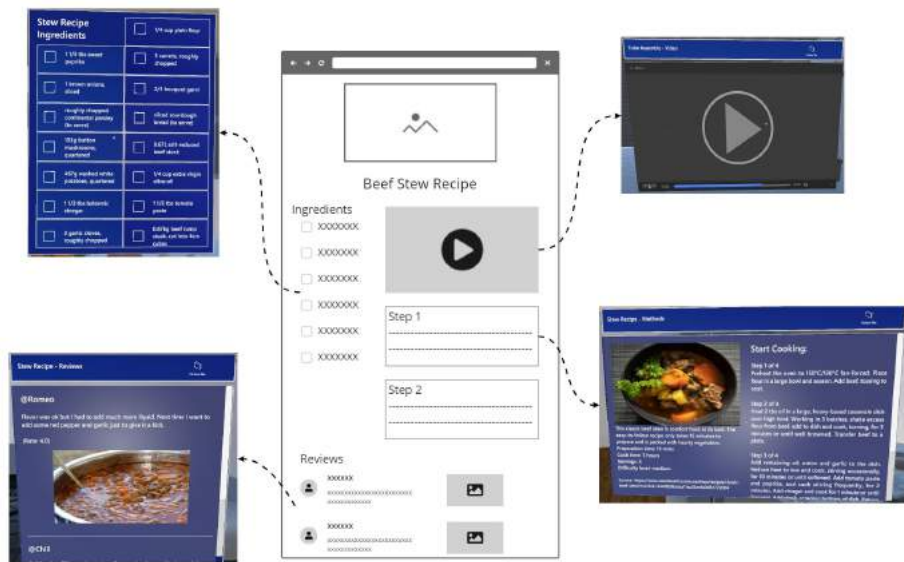
Figure 3.3: Illustration of simulating multi-view AR interfaces for a recipe web page

3.2.2 AR App Interface Simulation

Additional interfaces for general AR applications were also implemented to simulate browsing information on the web with other apps. Figure 3.4 shows all the developed app interfaces, and Table 3.1 lists the interfaces added in each scenario. They were designed to be either common daily-life applications (e.g., clock, email) or some apps specific to the current scenario (e.g., a map app used in the outdoor fieldwork scene). All interfaces can be grabbed and moved, resized, and rotated. Additional scripted specific behaviors were the real-time ticking of the clock, the checklist button interaction, and the pan-and-zoom map interaction; while other interfaces showed images and texts to

illustrate the context. The user would be asked to arrange all web and app interfaces shown in the scene so that the generated layouts could indicate how general applications might affect the arrangement of web contents.

(a) Clock	(b) Email	(c) To-Do List	(d) Phone Call
(e) Music Player	(f) Notebook	(g) Map	(h) Compass
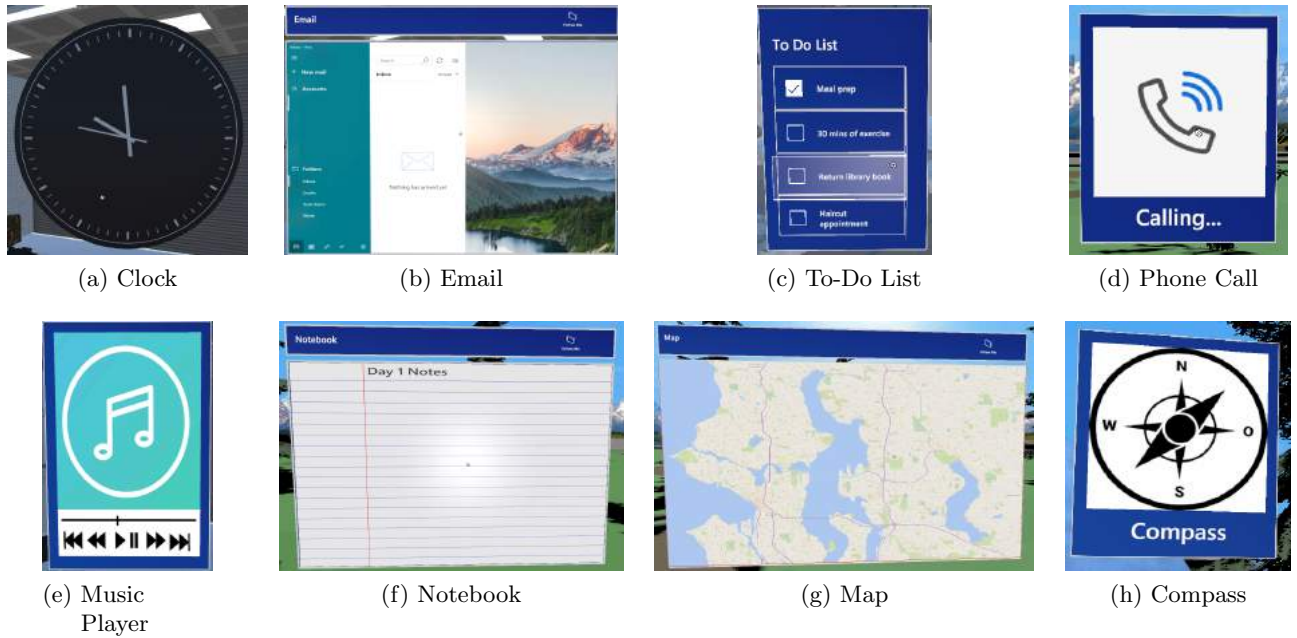
Figure 3.4: Different types of other AR applications simulated in VR

3.3 Physical Environment Simulation

There were three scenarios developed in the system: tube assembly activity in a garage, cooking activity in a kitchen, and fieldwork activity in an outdoor natural environment (Figure 3.5).They were implemented for different purposes and paired with different simulated AR web contents that fit each task context. The idea of having multiple scenarios aims to investigate how physical environments and activities may affect the choice of AR web interface arrangement preference.

In the **tube assembly scenario**, the user's view is surrounded by virtual assets simulating a garage or workshop room. The context is to assemble a tube model while checking web information/instructions about the tube model or background knowledge around piping. The system also shows a work area with a box of interactable tube model pieces. The user could grab or rotate each tube piece with the controller, which is to simulate a real-life assembly activity. The goal of this scenario is to analyze how users would like to organize web information while working on physical handcraft-style tasks.

The **cooking scenario** in a virtual kitchen provides a context of making a stew fol-

Table 3.1: App interfaces in each scenario

Scenario / Apps	Tube Assembly	Cooking	Fieldwork
Clock	✓	✓	✓
Email	✓	✓	✓
To-Do List	✓	✓	✓
Phone Call	✗	✗	✓
Music Player	✓	✓	✗
Notebook	✗	✗	✓
Map	✗	✗	✓
Compass	✗	✗	✓

lowing recipes. The scene consists of assets about the interior of the kitchen or dining room, kitchen appliances, and food ingredients. To let users still focus on interface arrangement instead of the cooking activity, objects except simulated AR interfaces are not interactable but just to render the cooking context. Compared to the previous tube assembly task, this scenario simulates a more casual daily-life activity with a larger focus area (e.g. the kitchen benchtop) to work on the physical task - cooking.

The last scenario changed to an outdoor space with an open view rather than obstructed by room walls/ceilings as in the previous two scenes. The context is fieldwork which includes activities of building a camping tent, taking soil samples, navigating with a compass or map, etc. Virtual assets are used to build a landscape with plants and rocks, a tent for the camping context, and a sky background for the open outdoor view. The previous two scenarios explore AR interface arrangement considering the physical task at hand (assembly or cooking) and the compatibility with real world objects around, while this fieldwork scenario emphasizes freer arrangement in an open space and the compatibility of physical activities such as walking.

3.4 UI Instructions

To guide the user to complete all tasks without real-time supervision from the investigator, the system implemented an introduction scene as an interaction tutorial and a floating user interface panel that displays step-by-step instructions.

The introduction scene (Figure 3.6) firstly asks the user to enter a given username which served as the unique participant ID and requires confirmation of the input. The next step displays the UI instruction panel that would guide the user throughout the system and explains how it works. The instruction panel follows the user's view and has a "Next" button to navigate to the next step. The user can toggle the UI panel to open or hide it by clicking the "Instructions" checkbox below the panel. The last step shows two example AR web and app interfaces, introducing interaction methods available in

(a) Tube assembly scenario

(c) Cooking scenario

(e) Fieldwork scenario
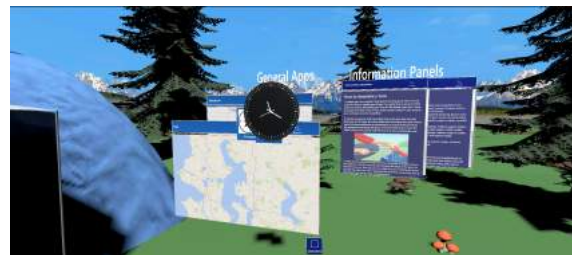
Figure 3.5: Different types of physical environment simulated in VR

this system, then navigating to the first "tube assembly" scenario.

In each of the three task scenarios, the floating UI panel presents three-step instructions that introduce the current scene, ask the user to arrange virtual web interfaces with or without general AR applications, and ask for completion of the corresponding survey section. Section 4.3 presents each step of the user study tasks in more detail. The transition between adjacent scenarios may take time, and the panel displays a message of "loading" as feedback. The last step of the application displays a "Thank you" ending message and a button to exit the system.

3.5 Remote Data Saving and Loading

The system supports both users saving data and the investigator loading data to rebuild the same interfaces arranged by a particular user. All virtual objects implemented in

(a) Username input

(b) Interface introduction
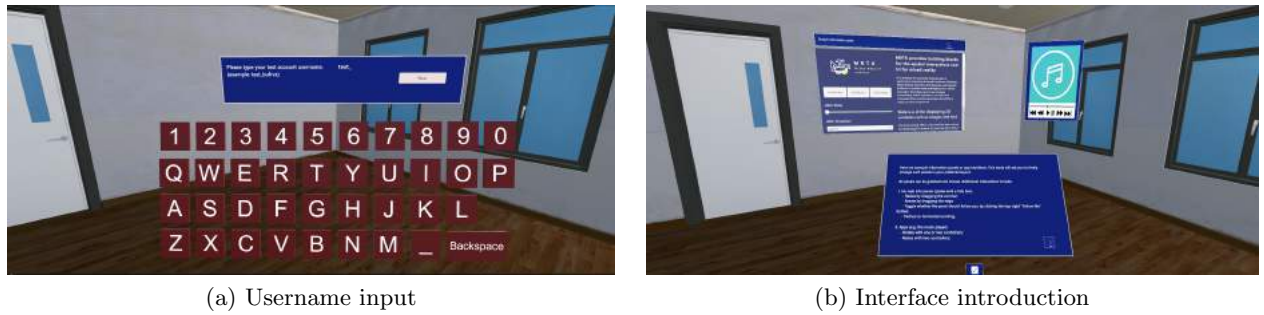
Figure 3.6: Introduction scene

the VR system were Unity GameObjects associated with Transform data. Transform includes the object's position, rotation, and scale at its x, y, z axes, respectively. Transform data for each simulated AR interface in the system were saved into a JSON data file locally then uploaded to a remote FTP server. The data could also be loaded locally and remotely by the investigator using the same system.

All users were given a unique and random username such as "test_bufrvz". The first scene of the system asked for a username input with the default prefix "test_", which means the system would be in the saving mode: all interfaces would be stacked at the back of view to be arranged; their Transform data would be saved at each step and exported into a JSON file with the username ("test_bufrvz") as its filename.

Listing 3.1: Example structure of the saved data

```
1  {"username": "test_bufrvz",
2   "GarageData": {
3     "Step 1": {
4       "PipingDesign": {
5         "position": [1,2,3], "rotation": [4,5,6], "scale": [7,8,9]
6       },
7       "TubeModel":{...},
8     },
9     "Step 2": {
10       "PipingDesign": {
11         "position": [3,2,1], "rotation": [6,5,4], "scale": [9,8,7]
12       },
13       "Email": {...}
14     }
15   },
16   "KitchenData":{...},
17   "FieldworkData":{...}
18  }
```

The system implemented a customized serializable Dictionary and Transform data structure to support saving and loading in the JSON data format, which were handled by a Singleton DataManager class. Listing 3.1 provides a sample structure of the saved
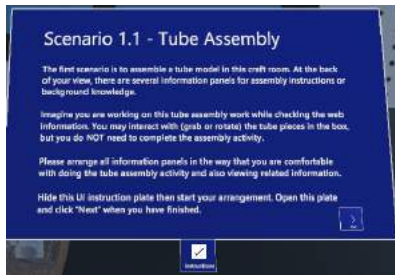
(a) Step 1: scene introduction and initial web interface arrangement
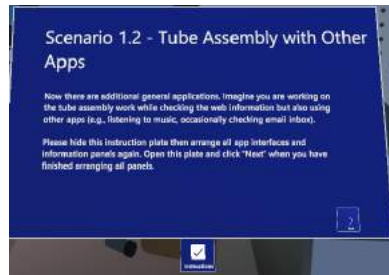
(b) Step 2: arrangement including general applications
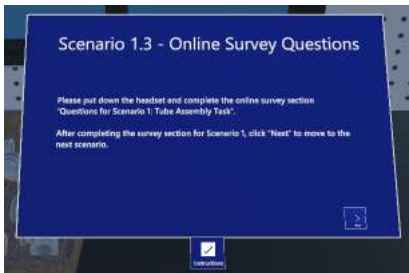
(c) Step 3: online survey section

(d) End of the application and study

Figure 3.7: Step-by-step UI instruction flow

data where interface-Transform pairs are stored for each step in each scenario. The dictionaries were updated when the user finished the corresponding step for arrangement tasks and clicked the "Next" button on the instruction panel. The dictionary for the first step recorded Transform data for all active simulated AR interfaces. In later steps, an object-transform pair would be saved only for new objects or if the object's position, rotation, or scale had changed compared to its status in previous steps. After the last arrangement task at the end of the last scenario, the DataManager class including dictionary data for all scenarios and the username were serialized and written into a local JSON data file. The file was then uploaded to the remote FTP server.

If the username input deleted the default "test_" prefix and instead started with "l_" or "f_", the system would change to the loading mode for loading the data generated by the user "test_bufrvz". The "l_" prefix means loading data from a local JSON file saved in Quest, while the "f_" prefix means firstly downloading the remote JSON file stored in the FTP server then loading its data. These two special prefixes were not shared with users and were used by the investigator only. In loading mode, the system would read from the matching JSON file and deserialize the DataManager class so that the dictionary data for each scenario could be restored. At each step in each scenario, the Transform values stored in the Dictionary were re-applied to active simulated AR interfaces accordingly. The system would not save data again after clicking the "Next" button. Hence the interface changes made by the user at each step could be visualized.

Elicitation User Study

This chapter presents the details of an elicitation user study performed to analyze:

- patterns of users arranging AR web and app interfaces in different physical environments;

- user opinions of the arrangement tasks taken, and the experience of interacting with the VR prototype system developed.

Information about the participants and recruitment is provided in Sections 4.1. Configurations for setting up the study is explained in Section 4.2. As justified in Section 3.1, the study was designed to be self-guided by utilizing the developed VR system so that participants could complete all tasks at their own pace remotely. Chapter 3 described the design and implementation details of the VR prototype system. Users were invited to complete several tasks to arrange virtual interfaces with the VR system and answer online survey questions. These tasks are outlined in Section 4.3. Section 4.4 outlined the overall process of the study.

There were two types of data collected from the user study: an interface data file generated by the system and responses to an online survey. The details about the data file are described in Section 3.5. Section 4.5 in this chapter describes the structure of the online survey form. The two types of data were analyzed to elicit preferred AR web content layouts, evaluate the effectiveness of the proposed multi-view AR web interfaces, and formulate design principles, which will be presented in Chapter 5.

4.1 Participants

Target participants were adults who can communicate in English and preferably have access to an Oculus Quest or Quest 2 headset. The main recruitment channel was email lists. Since social distancing regulations were loosen during the development of the

prototype system, and thus in-person research activities became possible, participants who do not have access to a Quest headset were also involved when the study was running. Such participants were invited to join a scheduled in-person session at the Australian National University (ANU) campus to use a provided device.

To aid the recruitment process and provide tailored information for remote and in person users accordingly, an online registration form (Appendix A) was linked in the recruitment texts. The form checks:

- if the interested participant has access to a Quest device:

  - if yes, whether the participant would like to use their personal Oculus account on Quest to install the app, then collect their account (or not) accordingly;

  - if no, whether the participant could visit the ANU campus for an in-person session and their availability;

- the participant's email address for sending further details;

- if the participant has any questions.

We recruited 15 participants in total, 10 of which joined remotely and 5 of which attended in-person sessions for a provided device. The online survey form contains a pre-study section asking the user to rate their familiarity with AR and VR respectively from 1 to 5. According to the responses, the number of experts and novices was balanced: 8 participants indicated that they were familiar with AR, and all except one rated high familiarity with VR as well; the other 7 participants indicated that they were relatively new to both AR and VR. The average familiarity ratings for VR and AR were both 3.2. There were another 5 remote potential users who expressed the interest, which means they submitted the registration form and received further details, but did not complete the study by the end of data collection.

4.2   Study Setup

There were several additional problems to consider to enable a remote and self-guided VR user study:

1. the user should be given sufficient yet easy-to-understand instructions to complete all tasks without the need of investigator's real-time supervision;

2. there should be some communication channels to ensure users' questions about the study or VR system can be answered in a timely manner;

3. the investigator should be able to collect all data needed remotely. The data gathering methods are described in Section 4.5.

A full online written instruction (Appendix B) was prepared and shared with the participants. It attaches the participant information sheet, explains how to join the study

according to whether the participant has a Quest, and lists step-by-step instructions for the user study process. These details are described in Section 4.4.

The VR system and online survey also included detailed instruction to ensure user study tasks were completed in the expected order. As explained in Section 3.4, the VR system implemented a floating UI instruction panel that can be toggled, specifying how to interact with the system and the goal of each step in each scenario. The online survey form contains several sections of questions: pre-study questions, questions for each scenario in the VR system, and post-study questions. Before each section, there is a separate area instructing when to move to the next section: e.g., answer questions for scenario 1 after completing scenario 1 tasks in the VR system.

In addition to emails, online meeting option was provided in case remote participants have problems with the VR system and need real-time help. A booking link was added to the written instruction. In-person users could communicate promptly during the session, but they were still able to complete study tasks following the UI of the system or given written instructions. No additional interviews were conducted so that data collection methods were consistent between remote and in-person users.

The VR system was developed and released on Oculus App Lab. To allow remote participants to install the app on their own Quest device, they were given two options to use a test account or their personal Oculus account to log into the Quest:

- Test account: all participants would receive a fake, unique, and random test account including a username, email address, and password. Once the participant log into the Quest with such test account, they could find the app in the app library ready to be locally installed.

- Personal account: the participant was still given a test account, but in addition, their personal account email address would be added as a subscribed user for the app. The participant would receive an invitation email. Once accepted, the app would be added in the user's app library for installation.

The random test account username also served as a unique ID to represent each participant without any personal or identifiable information.

4.3 Study Tasks

The VR system implemented three scenarios as described in Section 3.3. Guided by UI instructions (Figure 3.7), each scenario consists of three steps for user study tasks:

1. Introduced with the scenario (including the environment and context of some physical tasks such as cooking) and simulated AR web information panels, arrange those panels in the way that the user could conduct the physical task while checking AR web information comfortably. This step explores how the user would arrange multi-view AR web contents considering the physical environment and activities,

deriving insights into scene understanding and spatial awareness constraints for AR web views.

2. Arrange AR web information panels along with general AR applications. The web contents in the previous step would be reset and stacked at the back of the user's view. Several new panels would be added to simulate general AR applications next to web content. All interfaces' initial closeness to the user were randomized again to mitigate potential bias from the order of seeing the interfaces. This step investigates how the user would arrange multi-view AR web information while using other applications, indicating AR web design and development consideration of external virtual interface constraints.

3. Answer the corresponding survey section for this scenario. The user was asked to pause at the end of each scenario in the VR system, put down the headset and answer survey questions for that scenario, then continue with the VR system again to move to the next scenario. This design, compared to completing all tasks in the VR system and then answering all survey questions, could let the user have a break from the VR headset, which might cause motion sickness after long-time interaction. It also ensures that the user could remember the details for each scenario. The online survey also includes a few pre-study and post-study questions. The structure of the survey is detailed in Section 4.5.

4.4 Process

The study involved one participant at a time either remotely or in-person. After preparation, including obtaining human ethics approval and participant recruitment, the study process consisted of the following steps:

1. The participant submitted the registration form whose link was contained in the recruitment text.

2. An email response was sent to the participant to inform Oculus test account details, share the written instruction including the participant information sheet, schedule an in-person session if applicable, and address any questions that the participant had about the study.

3. The remote participant installed the VR system on their headset following the written instructions. The in-person participant would be using a provided device that had been set up.

4. The participant completed the first section of the online survey to confirm research participation consent.

5. The participant completed the second section of the online survey to enter the given unique test account username and collect information on their familiarity with AR and VR.

6. The participant launched the VR application installed on the Quest headset. They were presented with the introduction scene to enter the username and practise interactions with the virtual interfaces.

7. The participant completed scenario one arrangement tasks including initial arrangement of simulated multi-view AR web contents and the second round with general application interfaces added.

8. The participant completed the next survey section - questions for scenario one to gather their opinions of the arrangement tasks they had taken.

9. Repeat steps 7-8 for scenarios two and three implemented in the VR system. These steps were guided by the in-app UI instruction panel and emphasized again in the online survey form between each section. At the end of scenario three, the VR system would save the JSON data file (described in Section 3.5) remotely and locally then display an ending message.

10. The participant completed the last survey section for post-study questions about their overall interaction experience with the VR system.

11. Remote users could unsubscribe from the application following the provided written instructions.

Based on the observation from in-person sessions, it took around 45 minutes for a participant to complete the study tasks including VR system interaction and survey answering. Remote participants could complete all tasks at their own pace; in-person participants were able to indicate if they needed a break during the study session.

4.5   Survey Structure

The user study collected a range of qualitative and quantitative data from the data files generated by the VR system and the responses to the online survey. Section 3.5 detailed the system implementation of saving all simulated AR interfaces' position, rotation, and scale into a JSON data file remotely and locally, which gathered quantitative data and could be used to rebuild interface arrangement generated by each participant. This section details the structure of the online survey form hosted on Microsoft Forms (Appendix C). The survey mainly consisted of three parts: a pre-study section to understand the participant's expertise in AR and VR, scenario-related sections to gather the participant's description and justification about their interface arrangement, and a post-study section to collect the participant's opinions about the overall interaction experience with the VR system. The content for each section is described in detail below.

The survey began with asking for research participation consent, followed by a pre-study section containing:

- entry of the participant's test account username which was regarded as their unique ID, and

- two 5-point Likert-scale questions for the participant's familiarity with AR and VR, respectively.

Then the participant would be asked to undertake tasks in the VR system and complete the same set of survey questions for each scenario. The set of scenario-related questions were designed to gather data for the following aspects:

1. How the participant arranged simulated AR web views in the first step (that is, without additional general app interfaces) and their thoughts behind such an arrangement. This question was to understand why the user may prefer a certain layout to view AR web contents in a physical task instruction context.

2. The interface arrangement at step 2 with general apps added: whether the layout for web views were changed compared to step 1 and why the participant made such changes. This question was to analyze how external virtual interfaces may affect AR web browsing preference.

3. If the participant wanted to move any simulated physical objects provided as part of the environment background - what the objects were and why the participant wanted to move them. This question was to further explore the two-way relationship between physical environment constraints and AR web view layouts. Users would often try to fit AR objects into the physical scene. On the other hand, the virtual content might also affect the placement of real-world objects if possible.

4. Any additional virtual information or objects that might be helpful in the scenario and where to place them. This question looked for potential amendment to the system and task instructional information.

5. A 5-point Linkert-scale rating of the difficulty of completing the arrangement tasks in the current scenario.

6. Following the difficulty rating, the reason why the participant thought the tasks were easy or hard. These two questions contributed to evaluate the complexity of the study design and usability of the VR system implementation .

The last section of the survey contained post-study questions, which are 5-point Linkert-scale ratings to evaluate:

1. How immersive the VR system was to present virtual information and physical environments

2. Efficiency of AR app or web interface displaying legacy information including texts and images

3. Efficiency of AR interfaces conveying task instruction information.

4. Usefulness of separating information into multiple interactable panels compared to a single window or screen.

There was a major update during the study and VR simulation system design. The survey form originally included reading comprehension questions based on textual content in the simulated AR web interfaces. The questions were straightforward without complex reasoning, but the participant would need to look for the answer from the information panels. The participant could change the interface arrangement in order to search the answer, and the system would also save the interface Transform changes made at this third step for survey answering. Such questions were removed from the survey to reduce the complexity of user study tasks. The system was also updated to save only data for the first two steps for interface arrangement. We considered that if a simple question was designed, it may only relate to information on one panel, which would be easier for the user to just glance or turn their hand to find the information rather than make new arrangement changes. However a more complex question involving information from multiple panels is not suitable either for this study, since it had already contained enough tasks and information for the participant to handle. We decided to make sure the current system focuses on visual and spatial layout arrangement elicitation instead of mixing with information efficiency concerns.

Results

This chapter presents the data gathered from the user study described in Chapter 4 and discusses design implications derived from the result analysis. Section 5.1 discusses the analysis of interface data collected, which include virtual interface arrangement data saved by the VR application and online survey responses that described or justified the arrangement. The section explains the data processing method undertaken and presents findings in four aspects: overall interface layouts, grouping between inner interfaces, the relationship between virtual interfaces and the physical environment, and patterns to combine the layout of web interfaces with general applications. Section 5.2 discusses the results from other survey questions for evaluation. The analysis is summarised in Section 5.3 to answer the research questions defined initially. This section also contains a reflection on the study tasks and the VR system implemented. Section 5.4 further suggests potential design or implementation strategy for multi-view AR web interfaces.

5.1 Interface Arrangement Analysis

Participants were asked to complete virtual interface arrangement tasks in three scenarios with the VR system, and the online survey form contained a set of questions for all scenarios. The arrangement tasks in each scenario involved two steps: 1) to arrange only simulated AR web interfaces; 2) to arrange both web interfaces and general AR apps. The survey questions asked for the participant's description and justification for their arrangement made in the two steps, followed by a question asking what physical objects they may want to move. These are detailed in Section 4.3 for study tasks and Section 4.5 for survey questions.

The layout generated in each step was saved into the JSON data file named with the participant's username. The username was also required on the online survey form to distinguish matching responses. In this dissertation, individual participants will be

referred to with a number from 1 to 15 instead of randomly created usernames. The total number of layouts generated is 90 (15 participants×3 scenarios×2 steps). The number of simulated AR interfaces involved in each layout is between 4 and 12:

- Tube assembly scenario: 8 web interfaces in step 1, plus 4 general apps in step 2;

- Cooking scenario: 4 web interfaces in step 1, plus 4 general apps in step 2;

- Fieldwork scenario: 4 web interfaces in step 1, plus 6 general apps in step 2.

Therefore, the total number of simulated AR interfaces presented to each user and saved in each data file is 30. The sum of all interfaces collected is 30 (interfaces/user) ×15 (participants) = 450.

This section presents results for each scenario, combining user arrangement layouts saved by the VR system and subjective survey responses regarding the arrangement. This is followed by a summary of the common patterns discovered.

5.1.1 Data Processing

Section 3.5 describes that the released system enabled both saving and loading data for individual users, which means the same arrangement (i.e., Transform values consisting of position, rotation, and scale) of all interfaces by one user could be reconstructed on any headset that had installed the application. After reconstruction, screenshots of each participant's generated layouts in each step were also saved and analyzed along with the corresponding textual descriptions collected from the survey. These data were firstly classified according to the average distance between virtual interfaces, then manually adjusted, categorized, and labeled to derive common properties regarding virtual interface arrangement. The properties cover three aspects: overall interface layouts, interface grouping relationships, and virtual-physical object relationships. The extracted properties are discussed in the following three subsections in order.

In addition to rebuilding each user's individual layouts, layout aggregation was implemented in an unreleased version of the system for analysis purposes. In each scenario, the system could display the arrangement for a list of selected interfaces completed by all users. Therefore, it could indicate commonly performed transformations with each interface (Figure 5.1). This was achieved by merging the individual JSON data files, parsing it as a list of data handling classes, cloning multiple instances for each selected interface, then applying each set of saved Transform values to a cloned interface object. The panels could be abstracted as colored blocks for clearer comparison.

Six invalid individual layouts were identified and discarded. The user perhaps skipped a step or misunderstood the task, leaving interfaces at the default place or with slight change only. A small amount of invalid data is within the expectation considering the study's remote and self-guided nature. For each discarded case, the corresponding username, scenario number, and step number were stored in the system. When loading aggregated patterns for selected interfaces, the system would ignore those invalid cases
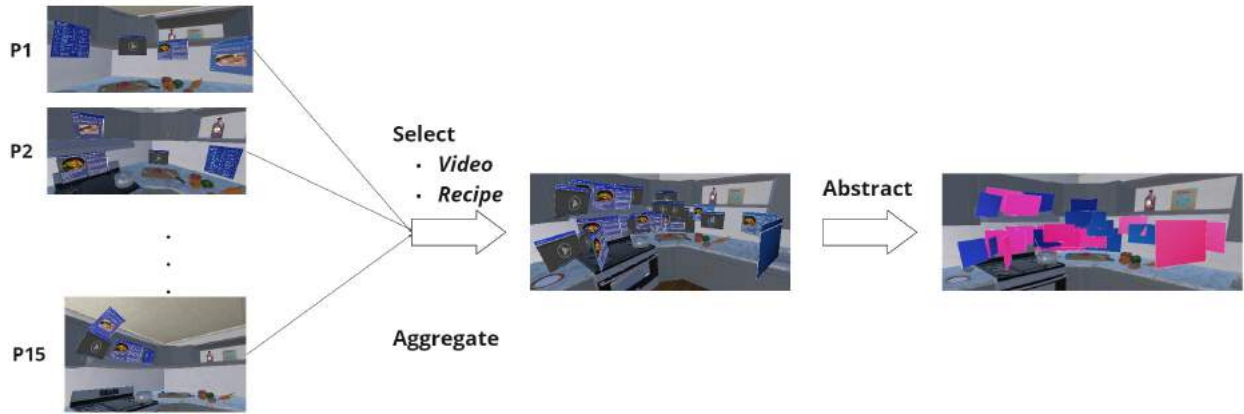
Figure 5.1: Diagram of the implementation of layout aggregation

to mitigate impact from noisy data. The number of valid layouts remained for analysis is 90 - 6 = 84.

5.1.2 Overall Layouts

Inspired by Maps Around Me (Satriadi et al., 2020) which analysed map layout geometries, we categorised common patterns of how participants arranged all simulated AR interfaces in the space. Based on virtual interface data and survey descriptions, four types of overall multi-view interface layouts are defined (Figure 5.2):

1. **Cylinder**: interfaces are arrayed into a curved grid.

2. **Sphere**: interfaces are spread around the user's surroundings in 360 degrees.

3. **Hemisphere**: interfaces are spread around a spherical cap.

4. **Cube**: interfaces are separated into different areas, and arrayed into a relatively flat grid on each side.

In general, the most popular layout was Hemisphere (41.7%), followed by Cylinder (23.8%), Cube (17.9%), and Sphere (16.7%) (Figure 5.3). The Hemisphere layout was the most common in both step 1 with web interfaces only and step 2 including general apps. Participants who chose this layout stated in the survey responses that they centered the most important panels (e.g., video instructions) and placed less significant content on the left or right side. Interfaces were spread spherically in front of the user so that they were easily accessible but also avoid crowding the field of view. The user would focus on the physical task while browsing key interfaces in the view, occasionally turning their head to check other less frequently needed content. General app interfaces were added in step 2 to be arranged along with web interfaces. In each scenario, around 5 participants out of 15 changed the overall interface layouts between the two steps. Figure 5.4 displays a transition diagram for the layout changes. Due to the increased

(a) Cylinder

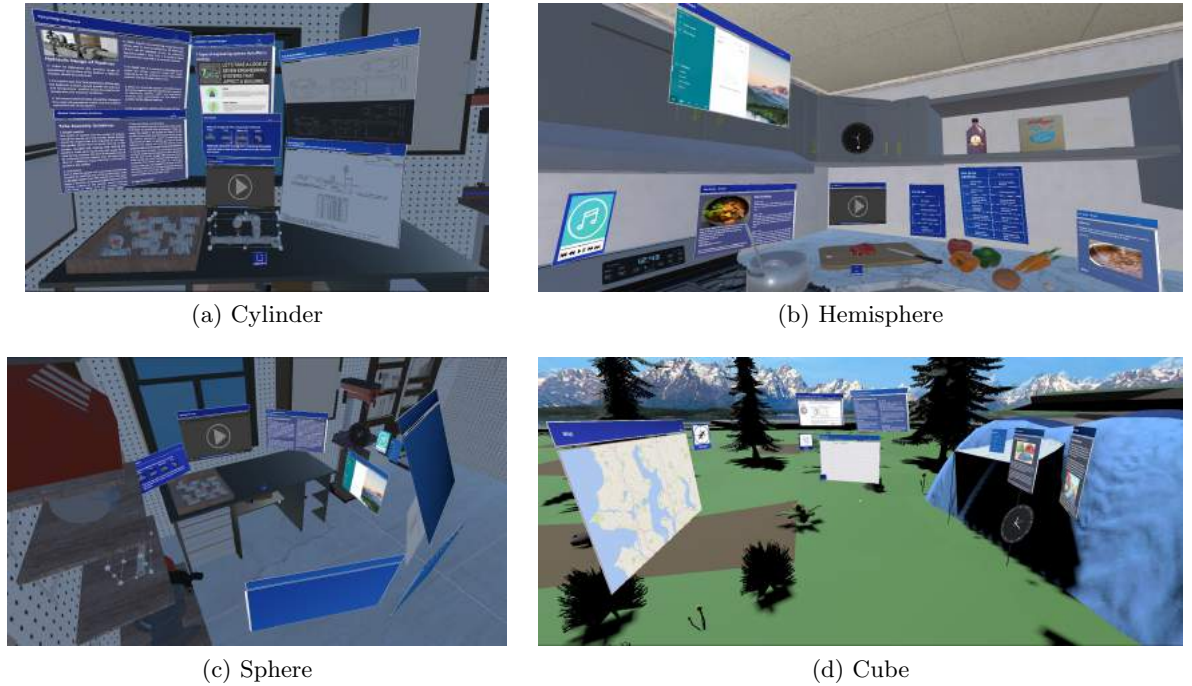(b) Hemisphere

(c) Sphere

(d) Cube

Figure 5.2: Four types of overall multi-view interface layout geometries.

number of panels, some layouts may change from others to Sphere or Hemisphere for more space.

The layout distribution in the kitchen cooking scenario was similar to the overall pattern, which means that the Hemisphere layout was dominant. In this scene, key physical items such as stove, cutting board, and food ingredients were placed along the kitchen counter so that the focus areas were spread horizontally. The interfaces were placed in a larger sphere aligned with the relevant physical objects, while other panels that were less important to the cooking activity were placed on the edge of the sphere cap.

The preferences towards the three layouts other than Hemisphere were balanced in step 2, but the Cylinder layout would be more preferred in step 1, where the number of views was smaller and interface contents were all relevant to the physical task. Especially in the tube assembly scenario, the most popular layout was Cylinder (40.7%), and the distributions of Hemisphere (29.6%) and Sphere (25.9%) layouts were similar. Interfaces were more compact in a Cylinder layout so that the user could view all panels layered in a curved grid on top of the working table in front of them. 2 participants mentioned that this layout was similar to multiple computer monitors that they had been using in real life. The Cylinder layout could be useful when there is only one focus area for the physical task and a small number of virtual interfaces. Interfaces are all arrayed near the focus area so that the user could concentrate on one direction viewing both the physical

Figure 5.3: Pie charts for overall layout distribution in different steps or scenarios
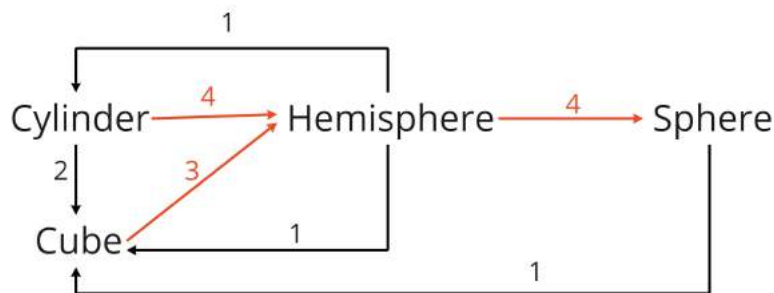
Figure 5.4: Diagram for the layout transition between the two-step arrangement tasks, labelled with the number of occurrence, i.e., the number of participants who made the corresponding change.

task and virtual information.

The preference towards the Cube layout increased in the fieldwork scenario, and it became the second most common layout (32.1%) following Hemisphere (46.4%). The example in Figure 5.2 (d) shows that the participant placed tent-related instruction panels near the physical tent, navigation panels (compass and map) on the opposite side in the open area, and reading/writing materials on another side. This scenario simulated an open space involving multiple physical activities (building a tent, taking soil samples, etc.). There were fewer physical surfaces constraining the position of virtual interfaces and more than one work areas to focus on, which could be a reason why an increased number of participants separated interfaces in different areas forming the Cube layout.

According to the arranged panel positions and participant descriptions mentioning which panels were more important, the layouts also indicated that different areas in the user's

field of view were associated with different levels of significance (Figure 5.5). When the user was concentrating on a physical task in front of them, the related physical items were often centered in the view. Therefore, the space around the center was where the user would pay more attention. The most important interfaces were commonly placed in this center space, but with some distance between the operated physical objects to avoid blocking the view to the task itself. The next significant area was from the top of the view to the right side of the front view. 4 participants mentioned that they put secondly important interfaces slightly right in their view for easy reference while focusing on the task. One of them explicitly mentioned that it was due to being right-handed. It is possible that the preference meant to be the dominant hand side rather than always right, but because of the lack of data regarding participants' dominant hands and side preference, this remains as an open question. Less important panels were placed further away from the center of the front view. There could be different centers for more than one key physical item in different places, and the relevant interfaces were placed around the items accordingly.
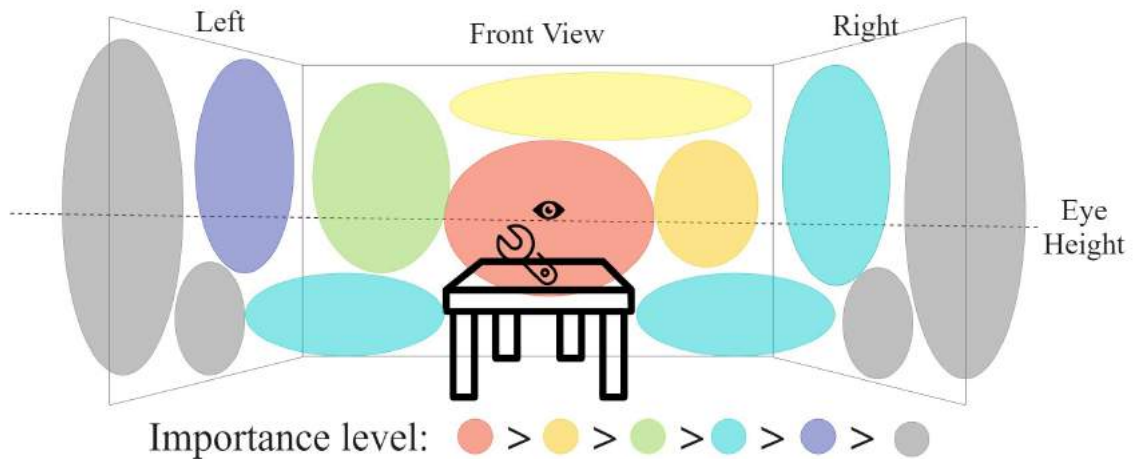
Figure 5.5: Importance levels in the field of view.

The importance of one interface and thus its position in the user's view could be affected by the following factors:

1. Relevance to the task or physical objects to concentrate on.

2. Value of the content, for example, whether it contains step-by-step instructions for the task that the user could not complete by themselves.

3. Cognitive load of the content, for example, whether it contains a large amount of text. Interfaces with visual instructions (video, image, 3D model, etc.) were commonly placed at the center of the work area. Several participants placed long textual content on the side or back of view for occasional reference and less distraction from the physical task.

4. Frequency of direct interactions. For example, some participants preferred positioning the clickable to-do checklist closer to them and moving textual panels further away. Although the textual panels were more relevant to the physical task, the reading activity could be done at a distance; while ticking off items on the checklist would require more UI interactions with a reachable panel.

5.1.3   Interface Grouping Relationships

Within the overall layouts described above, it was also found that certain interfaces were commonly put adjacent to each other. This argument is based on survey descriptions/justifications and visual investigation of individual and aggregated layouts. We named three types of the relationship between interfaces that could be often grouped together:

1. **Semantic**: the interface contents are semantically related. As shown in Figure 5.6 (a), the video and recipe panels were placed together on top of the cooking area, both containing cooking instruction information.

2. **Functional**: the functionalities or presentations style (text, image, video, etc.) of the interfaces are similar. As shown in Figure 5.6 (b), there were groups of textual panels, pdf/image panels, and model visualisation.

3. **Visually Balanced**: the number or size of the interfaces is balanced within the overall layout. As shown in Figure 5.6 (c), the interfaces were separated into two sides that took a similar amount of space. Both sides contained textual and visual information and tended to place textual panels higher.

The three grouping relationships are not exclusive. It was common for a layout to feature multiple relationships at the same time. In Step 1 with web interfaces only, 53% layouts contained Semantic grouping, and 44% contained Functional grouping; in Step 2 with added general applications, the percentages increased to 70% and 92% respectively. With more views with various content or functionalities, it was easier to identify grouping patterns. The Visually Balanced or unidentified grouping was 11% and 17% respectively considering both steps.

The layout of a group of interfaces may be different from the overall layout. The names of overall layouts defined in Section 5.1.2 can be reused for interface grouping, and an additional one is defined as **Plane**: interfaces are arrayed into a flat grid. For Figure 5.6 (c), the overall layout is defined as a Cube containing two sides, and each side consists of a group of interfaces arranged in a Plane layout. One scene may contain multiple interface groups in different layouts. Among the scenes involving grouping relationships, the most popular grouping layout was Cylinder (40%) arranging panels closely in a curved grid, followed by Plane (26%) and Hemisphere (12%). All Visually Balanced interfaces were arranged in either Cylinder or Plane layout with a balanced number of grid rows and columns aligned for each group.

(a) Semantic

(b) Functional

(c) Visually Balanced
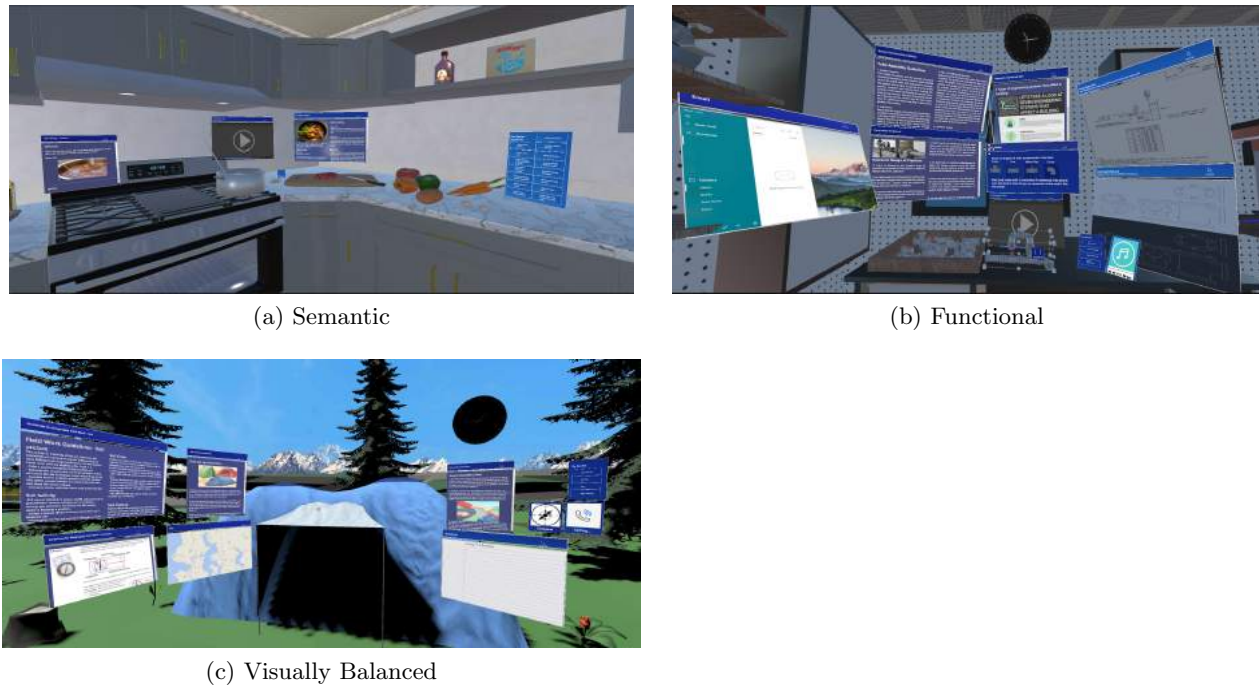
Figure 5.6: Three types of interface grouping relationships, also demonstrating spatial and contextual physical environment constraints.

5.1.4 Virtual-Physical Object Relationships

The three scenarios in the VR system simulated different physical environments containing real-world items and provided contexts to perform physical tasks or activities. The cooking and tube assembly scenarios added spatial limitations in purpose: positioning the main work area near walls, adding cabinets on one side of the work area, etc., to explore how physical obstruction could affect the virtual interface arrangement. The fieldwork scenario simulated an open space as a comparison. The position and orientation of virtual interfaces were commonly constrained by the physical environment. Two characteristics of such constraints are named:

1. **Spatial**: physical surfaces confine the position, size, and orientation of virtual interfaces.

2. **Contextual**: interfaces are placed near the relevant physical objects.

Figure 5.6 can also demonstrate the relationships between virtual interfaces and the physical scene. In the cooking scenario (a), participants restricted the interfaces within the space between the cupboards and the counter top, which was an example of Spatial relationship. Another common pattern was to attach all interfaces to the cupboard surfaces. The positions also related to Contextual constraints: the recipe and video

instruction panels were placed near the main work area around the stove and cutting board, while the ingredient checklist was near the food.

Key task-related physical items (food or kitchen appliances on the counter, tube pieces on the table, tent on the ground) were treated as reference points or the center of task instructional interfaces, which could affect the overall interface layout as explained in Section 5.1.2. For a physical task that especially involves the user's frequent movement in a large open space, the follow-me option was useful to keep a panel following the user's view.

The virtual items including interfaces could indicate similar affordance to physical ones. In the cooking scenario, 4 participants tended to attach checklists for to-do items or ingredients to the fridge at the back of their view, which was similar to real-life usage of paper checklists. The video panels were commonly placed behind the focus area and near the wall; the clock app interface was instead often attached to the top of wall, simulating an actual clock. The tube assembly scene contained a shelf on the left side of the working table, and one of its layers was empty. 3 participants put the 3D model or video panel on the shelf to reproduce their physical existence, and another 2 participants mentioned that they would clear the occupied shelf layers for enough space to fit more interfaces there. In the fieldwork scene, it was popular to place tent building instruction panels around but not overlapping with the physical tent entrance, although a virtual panel would not affect the entrance in reality. The preference of interface orientation may also change according to the interaction status: a participant stated that they placed the notebook interface matching the eye height for reading and would put it down to be ground-parallel for hand-writing if possible, simulating writing on paper in real-life.

5.1.5 General App Addition Patterns

Moreover, the patterns of adding general applications were also categorized into three types as follows, which means how the participant placed the general apps into the overall layout compared to the layout in step 1 with web interfaces only.

1. **Insertion**: the arrangement of web interfaces remains nearly unchanged, while general apps are inserted into the layout of the web interfaces.

2. **Separation**: the arrangement of web interfaces remains nearly unchanged, while general apps are placed into another area separated from the web interfaces.

3. **Rearrangement**: the layout and/or grouping of web interfaces show obvious changes when general apps are added, which means new interface groups are formed involving both web interfaces and general apps.

In Figure 5.7 (a) to (b) Insertion pattern, the web interface layout remained as Cylinder and the grouping relationship is also Functional similarly based on the information type. According to the corresponding survey description, the email app was placed at the left hand side for occasional check, the virtual clock was attached to the high wall similar to real-life usage, and more interactive and smaller apps are put near the right hand when

(a) Tube Assembly - Original

(b) Insertion

(c) Cooking - Original

(d) Separation

(e) Fieldwork - Original

(f) Rearrangement
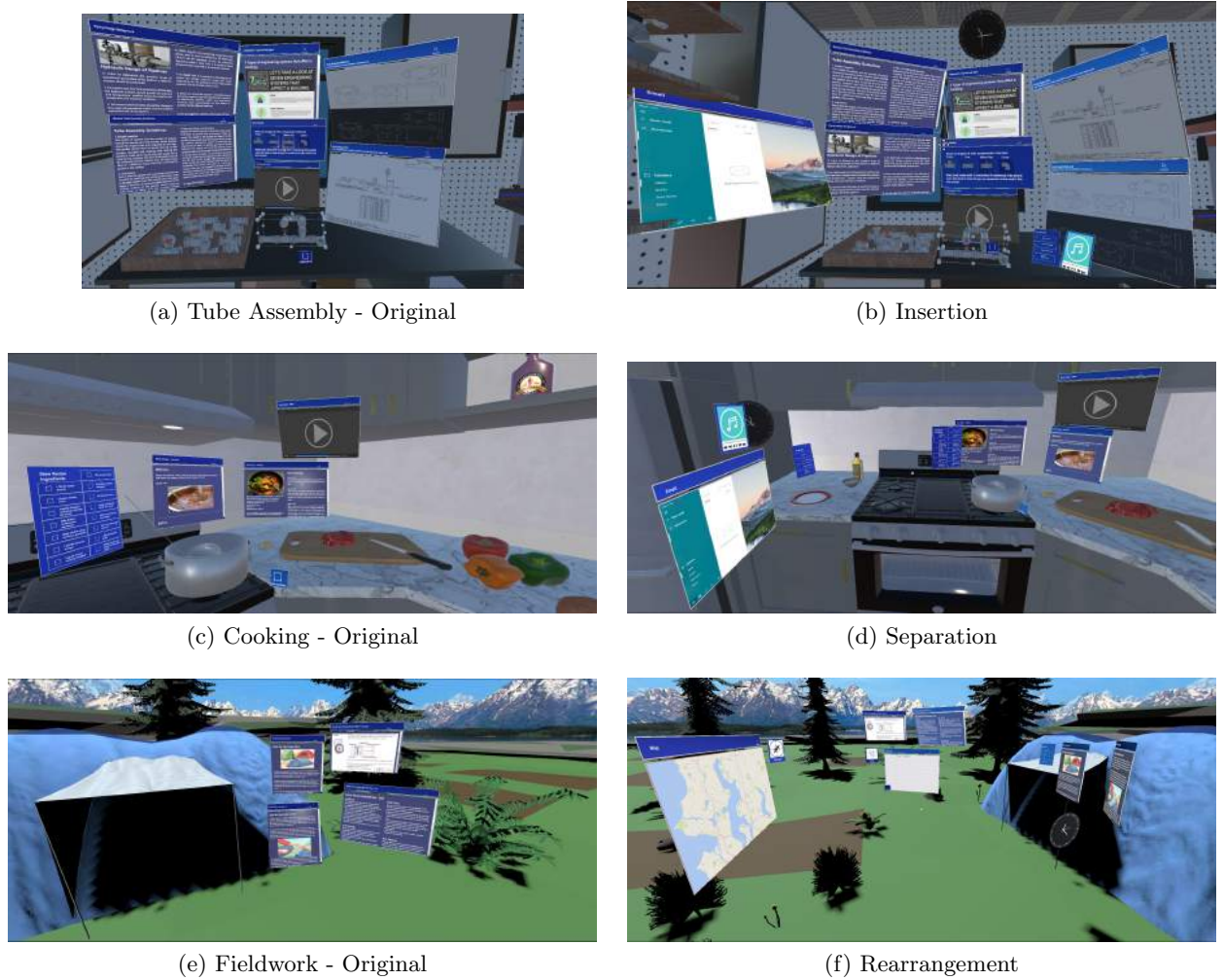
Figure 5.7: Three ways to add general applications into the original web interface layout.

working on the tubes on the table. The Separation pattern from (c) to (d) ensured that the user would still concentrate on the cooking task, while apps were left at another side without interfering the focus area. For the Rearrangement pattern from (e) to (f), panels for tent building instructions remained near to the physical tent, but other panels were grouped with the notebook and phone call apps instead. This is related to the change of focus areas and task activities, where note taking became another task to be worked on in addition to tent building and navigation.

As shown in Table 5.1, in most cases, general apps did not affect the arrangement of web interfaces, but the larger number of Insertion than Separation indicated that a mixture of web and general app interfaces would be more preferred. The preference might also be affected by the physical environment constraints. There was no Separation pattern

Table 5.1: The occurrences of each app addition pattern in each scenario.

Patterns / Scenarios	Tube Assembly	Cooking	Fieldwork	Total
Insertion	8	6	9	23
Separation	0	6	1	7
Rearrangement	5	2	4	11
Total	13	14	14	41

identified in the tube assembly scenario which has only one focus area and with limited space; grouping all panels closer in front would reduce the need to distract away from the focus area, even though not all panels are necessary for the task. The cooking scenario instead has an equal number of Insertion and Separation. It has a larger space with a broader focus area stretched along the counter, which means the apps can still be naturally accessible in another area close to the counter. The fieldwork scenario features the largest space but also multiple activities. The apps provide more important functionalities such as navigation and note taking, which serve as additional tasks requiring attention and can be combined with the web interfaces. Insertion is easy especially when the distance between web interfaces is large enough; if not, rearrangement could be needed to form new groups.

5.2 Survey Analysis

In addition to questions asking how the participant interacted with the system, each scenario-specific section of the online survey asked for potential amendment (e.g., additional virtual content) desired for the corresponding scenario and user opinions on the difficulty of the arrangement tasks. The post-study section asked for ratings on their overall experience with the study. These survey responses were analyzed to evaluate the implementation of the VR system and the design of the user study.

Average difficulty ratings from 1 to 5 for the three scenarios were 3.6 for tube assembly (std 0.51), 4.27 for kitchen cooking (std 0.88), and 3.87 for fieldwork (std 1.06). The tube assembly scenario was more difficult to complete. Since it was the first scenario presented to the user, and half of the participants were relatively new to VR or AR, they had not yet been accustomed to the Quest headset, controller interactions, and the goal of arranging the panels with the assumption of working on a physical assembly task. There were panels containing technical background knowledge about the tube assembly activity. Although it was not necessary to understand all of the panel content to complete interface arrangement tasks, complex information could increase the difficulty of deciding which panel was more important. The ratings for the fieldwork scene show a larger standard deviation. The follow-up answers indicate that some participants preferred the open space to have more freedom to arrange the panels, while some other participants felt the opposite way. Given an open space and multiple activities as the

context, the latter participants would hesitate about an optimal arrangement layout. On the contrary, the cooking scenario simulated a daily activity so that participants could understand the context and judge the significance of each virtual interface more easily. Moreover, 6 participants acknowledged that "intuitive controls" were helpful for free interface arrangement, which involves movement, rotation, and scaling.

5 participants suggested improvement of the follow-me option implementation. Textual or visual web interfaces were implemented with a follow-me button to position the panel to the bottom center of the view and follow the user's sight movement. It was suggested that the follow-me option would also be useful for other interactive interfaces, such as the ingredient checklist and the map, which may be used when the user was walking around. It would be better to enable customizing the position of following panels, considering that the user may want to move multiple panels at the same time, and not all panels are necessary to stay within the view. Another suggestion was enabling optional audio reading of textual content to allow perceiving web information while looking at physical task objects.

The post-study evaluation was a set of 5-point rating questions, and the results were overall positive. The VR system's quality of being immersive was rated as 4 in average (std 0.76). 80% of the participants agreed with the statement "AR apps/websites are an efficient way to display legacy information (e.g., texts, images).", and the remaining 20% selected "Neutral"; 13 out of 15 participants agreed that AR apps or websites contained sufficient task information (Figure 5.8). The last question asked the usefulness of separating information into multiple interactable panels compared to a single window or screen. The average rating was 4.47 out of 5 (std 0.64), indicating a positive impression of the multi-view interfaces.
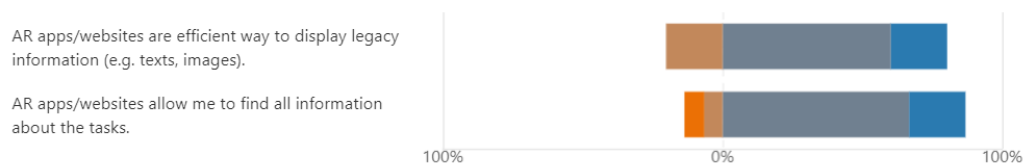

Figure 5.8: Evaluation of AR apps or websites for task instruction

5.3   Discussion

In summary, virtual interface arrangement patterns generated by participants were reconstructed, classified or aggregated, and combined with descriptions gathered from the online survey for analysis. Common properties were identified regarding the overall lay-

out of the interfaces in the 3D space, the layout and relationship of forming interface subgroups, the way to add general applications into the interface layout, and the relationship between virtual interfaces and the physical environment. To reiterate, the research questions have been defined in Section 1.2 to explore:

- RQ1: The usefulness of multi-view AR web interfaces.

- RQ2: The preferred layout for different web components.

- RQ3: The impact from the real-world environment (including physical objects and the task activity context) to the multi-view web interface arrangement.

- RQ4: The impact from general AR applications to the multi-view web interface arrangement.

The first question connects to survey results as part of evaluation. Participants expressed an overall positive opinion about the simplicity of the VR system or study tasks, the VR system's quality of being immersive, the efficiency of AR apps or websites to display legacy or task instruction information, and the usefulness of separating information into the proposed multi-view style. Although the mean goal of this study was eliciting preferred layouts instead of evaluation, the results could be concluded as an initial positive impression of the multi-view style. Based on the implementation strategy suggested later in Section 5.4, further evaluation could involve more objective measures such as task completion time and information search accuracy to compare with a traditional 2D style browser.

To answer the second question regarding the preferred layout, 5 types of interface layouts are defined: Cylinder, Hemisphere, Sphere, Cube, and Plane. The Plane layout only applied to interface grouping, while the other four types were discovered in both grouping and overall layouts. Hemisphere was the most popular overall layout, especially when there were multiple focus areas in the physical environment and the importance of interfaces was clearly varied. On the other hand, the Cylinder layout would be more common for interface subgroups or when there was only one focus area for the physical task. This result as well as the number of total layouts analyzed is similar to the finding from Maps Around Me (Satriadi et al., 2020) which suggested that the spherical cap layout would maximise multi-view interface visibility, but the dominance of Hemisphere (spherical cap) layout preference is much smaller in our result (41.7%) compared to Maps (81.5%). The difference may indicate that visibility could be a less determining factor for web AR interfaces, which should also consider the information significance and physical environment constraints. Therefore, the implementation of multi-view web interfaces could further consider automatic semantic analysis, scene understanding, support for user customisation.

Interfaces containing important information or interacted frequently would often be placed closer to the center or right (possibly dominant hand) side of the view. Interface grouping relationships are categorized as Semantic, Functional, and Visually Balanced. This also relates to Research Question (RQ) 2 regarding preferred layouts.

The characteristics of the relationship between virtual and physical objects are named as Spatial and Contextual, which relates to RQ 3 for physical environment constraints. These relationships (except Visually Balanced) were commonly presented in more than half of the layouts gathered. Interfaces could be grouped according to semantic relation or similarity in functionalities. The panel position and orientation could be adapted to physical surfaces and the position of relevant real-world objects to concentrate on. More important panels were put closer to the top center and secondly top right areas relative to the key physical task objects.

The last question concerns the impact from general applications, and we identified 3 patterns of changes made when adding general applications: Insertion, Separation, and Rearrangement. Insertion was found to be most popular, which means the web interface arrangement itself may not be affected much by the apps, but the mixture of web and general app interfaces would be preferred. This supports the proposal of multi-view web interfaces that enables combination of individual web components and external interfaces. However, this result might be biased because the system only implemented "web interfaces only" as step 1 and "adding general apps" as step 2. It is possible that participants tended to keep the same web interface layouts that they had already made once. A refinement would be randomizing the order of arrangement tasks for these two types of interfaces.

An additional problem to consider could be the occlusion issue. The VR system simulated different physical environments and AR web or app interfaces, but the way to handle occlusion might be different between the VR simulation and the actual AR presentation. This could be taken into account when considering spatial constraints from the physical environment. During the VR tasks, many participants placed some interfaces behind the relevant physical objects simulated. For example, the ingredient list and recipe panels in the cooking scenario may be put behind the food on kitchen counter. Another common pattern was aligning panels with flat physical surfaces such as cupboards or walls, avoiding the interface from "inserting" into the physical surface. Similarly to the real-world existence, the obstruction between virtual interfaces and physical objects could be controlled by modifying their distance and orientations. However, with an optical see-through AR headset such as HoloLens, the user views the real world through the transparent glasses while holographic virtual images are projected onto the retina. The occlusion between physical and virtual objects needs to be handled explicitly. Figure 5.9 shows photos taken with HoloLens that demonstrate two examples of occlusion issues. In subfigure (a), the virtual web browser interface has been "attached to the wall", but the physical object held closely in front of the user's eyes is still obstructed by the virtual interface. This is undesired when the physical object should be the focus (e.g., food and kitchen appliances in the cooking scenario). On the other hand, in sub-figure (b), the virtual interface is blocked by physical surfaces, which could happen especially when the follow-me option is enabled. This should also be avoided and can relate to the tube assembly scene: the cabinet/shelf on the left hand side (Figure 5.6 (b)) limits the space where the participant can put instruction panels. The virtual interfaces would be more

important and can be prioritized to render in this case.

To mitigate the occlusion problem, existing development techniques, such as spatial mapping meshes provided in MRTK, can detect physical surfaces (e.g, the cabinet door). Further implementation options include "magnetically" attaching the virtual panel to the physical surface with adaptive angles to avoid occlusion, hiding the virtual panel behind real-world objects, or rendering the virtual holographic interface even when it is occluded by physical surfaces. A challenge would be to have a combined solution for both occlusion problems argued above, which means the priority between virtual and physical objects can be conditional. Physical surfaces that can be ignored might be large and static, and hence treated as part of the background behind virtual interfaces; while important physical objects could be small and manipulated frequently (e.g., ingredients), which could be more difficult to detect and have more latency concerns. Figure 5.10 displays some of the layouts collected that may mitigate the complex occlusion problem. Virtual content may surround the physical work area, attached to real-world surfaces such as wall and cupboard, but not filling the sight to key physical objects (e.g., tubes, kitchen bench). The occlusion problem should be easier to solve on a video-see-through AR device, since the real-world view would be captured by cameras and combined with virtual content to form an integrated video presentation on the device screen (Rolland et al., 1995).

(a)   (b)
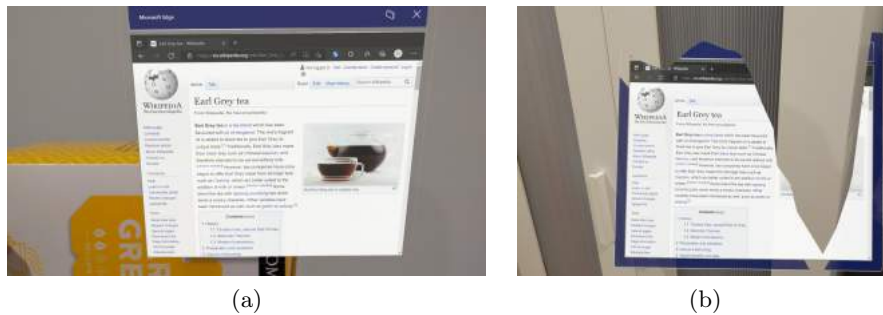
Figure 5.9: Two AR occlusion issues shown on HoloLens: (a) physical object blocked by virtual interface; (b)virtual interface blocked by physical surface

5.3.1 Reflection on System and Study Design

The original implementation of introducing information panels was a list of checkboxes toggling the display of one panel, and the user was asked to active all views then start arranging. The following problems were considered: 1) the user might tend to leave the panels at the default places, and thus the results may be biased; 2) the user might quickly go to the next step with too few panels activated, which would affect deriving insights into the proposed multi-view style. The implementation of checkboxes was updated to initially stacking all panels in a randomized order behind the user to motivate manual arrangement of all interfaces. As a result, common multi-view interface layouts and

(a)   (b)

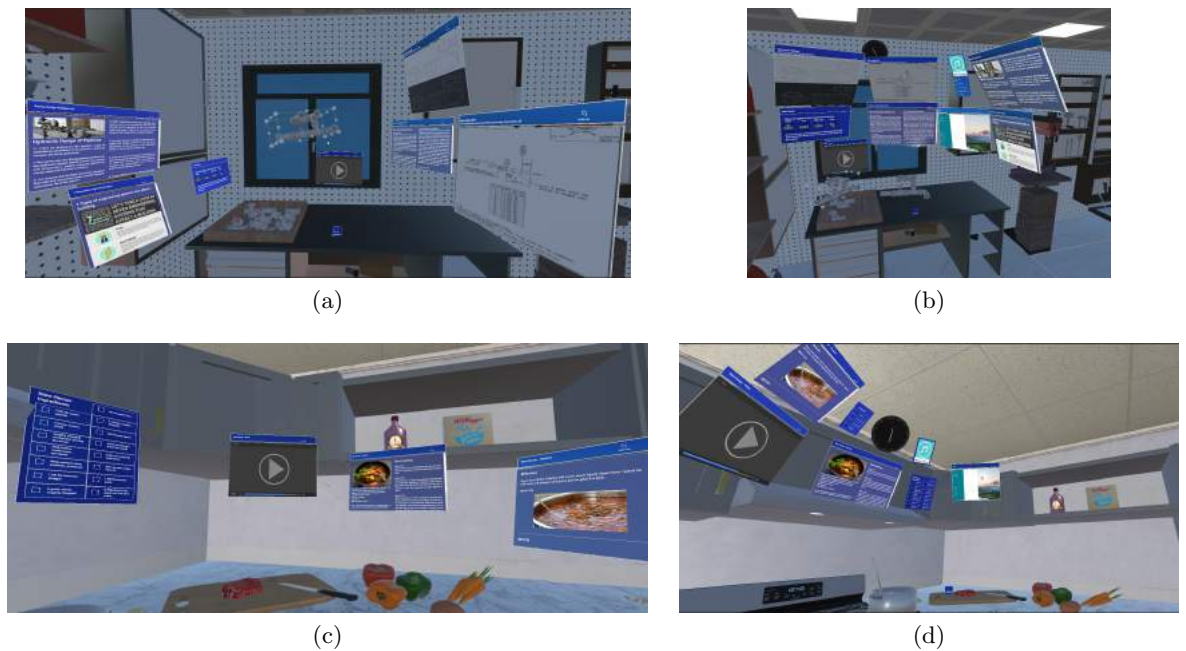(c)   (d)

Figure 5.10: Layouts from participants that may mitigate complex occlusion issue

grouping relationships were identified, but there was also feedback suggesting that the participant could be confused at first with all panels presented at once. The first scenario was the tube assembly context involving technical content, which may further require a higher cognitive load. Apart from modification within the scene, randomization of the order of scenarios may be also helpful to mitigate bias from the unfamiliarity with VR at the first scenario.

Putting interfaces in the back may also distract the user from the context of working on a physical task. In a few collected layouts, all panels were arranged into certain patterns but remained at the back. They still contributed to the identification of common interface layout properties, but lacked meaningful relationships with the physical environment. If not described in the survey response, it was unclear whether the participant kept panels at the back on purpose or just neglected the task-related physical items at the front. Another possible issue of initially stacking all panels is that the interactable areas for each panel may also overlap or be too close, making it harder to select the intended panel which is not necessary the top one. To revise, the panels could be initially positioned at the center of the front view and displayed one by one in a randomized order along with a brief description. Arrangement would still be required to avoid stacking all panels at the center, and the user may understand the purpose of each panel more clearly.

The physical task context in each scenario could be divided into smaller step-by-step

goals and present interfaces related to the activities in each step. As mentioned in Section 5.2, the kitchen cooking scenario was rated the highest with respect to the straightforwardness of completing interface arrangement tasks. Given a daily-life scenario, users can understand better what they would exactly do and what information would be necessary to complete the physical task. A task context can involve multiple steps, e.g., from ingredient preparation to serving the meal, and the most relevant interfaces for each step could be different. A step-by-step guidance simulating real-world activities would be more important for a more complex and less intuitive context like the tube assembly or fieldwork scenarios.

There are other possible refinements on the VR elicitation system implemented. As mentioned previously, randomization of arrangement steps could help remove potential bias for the results regarding general app impact. Both the cooking and tube assembly scenarios simulated a work area near the corner of a room to add more physical spatial constraints; yet one of them could be changed to another location in the room for more variety. The introductory and tutorial scene displayed sample interfaces and textual explanation of available interactions, which could be further enhanced with virtual hand coach animations to help VR/AR novice participants be familiar with the controller interactions. The categorization of interface layouts considered the classification based on the average interface distance, visual investigation of the rebuilt arrangement, and user descriptions from the survey responses. If the number of participants was larger, more advanced automated techniques may be utilized to cluster and identify similar groups of layouts, involving data of both virtual interfaces and the simulated physical objects.

5.4 Design and Implementation Strategy

Based on the results of the study, the following design implications or implementation suggestions are derived to develop multi-view AR web information interfaces.

Adaptive Hemisphere or Cylinder overall layouts. The Cylinder layout presents all panels as a compact curved grid, which would be preferred for a single focus area to view all interfaces at once; when the number of interfaces and/or focus areas increases in the view, the Hemisphere layout would be preferred to spread interfaces in a spherical cap relative to focus areas. The two layouts could be inter-convertible: a Cylinder layout can be stretched to Hemisphere by enlarging the distance between panels along a curve, or vice versa. The implementation could consider defining an initial Cylinder multi-view layout centered in the field of view and indicating the minimal distance between panels, then enabling stretching or compressing the curved inter-panel distance (e.g., manually with a slider control or automatically relative to the number of view presented). The stretching may eventually produce a Sphere layout if panels are spread in 360 degree. In addition, an option of increasing the distance change between certain parts and adding angles could also achieve the Cube layout by separating panels into multiple distinctive areas.

Semantic and Functional information separation and grouping. In the study, interfaces containing semantically related content, sharing similar functionalities, or presented as similar information types (e.g., text, image, video) tended to be placed closer within the overall multi-view layout. Visual instructional information was considered to be the most important content and was often centered among all panels. For an application holding specific information, since the content would be designed in particular, the way to separate information and group relevant parts could be predefined. For a general multi-view interface to present existing web information without particular authoring, a basic automatic information separation method may base on the HTML tags that would have already defined sectioning and multimedia types. A threshold value could be defined to set the depth of sections to be split into different panels. Media such as images and videos may be treated as separate interactable panels but initially placed within the belonging section, which would allow the user to browse all sectioned information together and drag out particular visual content if needed. To support further customization, Layout CSS proposed by Lam et al. (2019) could be an inspiration to specify more properties based on HTML/CSS code for each web content element (e.g., whether the element should be split into a separate panel, the index of the Cylinder grid cell that holds the panel).

Manual adjustment of individual and grouped interfaces. The three strategies described above would present a default multi-view layout, and user modification should also be supported to customize the view. The interactions that had been implemented in the VR elicitation system were panel movement, resizing, rotation, and view following, which were applied to each individual panel. Some participant feedback reported that it would be helpful to enable modifying a group of interfaces together. In the VR system, each web panel had a title bar that contained the follow-me button, which might make the view more cluttered. To enable easier group operations, the title bar for each panel may be replaced by a top-right checkbox to be displayed when hovered. A menu can be added at the bottom center of the view, containing buttons for additional or modified interactions:

- Hiding the selected panel. The menu can contain a drop-down list of hidden panel titles to display the panel again once clicked.

- Toggling the "Follow-me" option for the selected panel. According to participant feedback, there should be an offset range for the portion of the panel staying in the user's view while following. The implementation in the VR system kept the panel at bottom center of the view, which may block other essential objects and cause distraction. It could instead follow at the right or left side of the body.

- Aligning the panel with physical surfaces around. Existing development techniques such as "tap to place" or surface magnetism provided in MRTK allow moving and placing a virtual object along a physical surface detected by spatial meshes.

- Grouping or ungrouping selected panels. Grouped panels are treated as a single larger panel that still enables panel interactions mentioned above. They can return

to the individual status by ungrouping.

- Reset all panels. Bring all panels back to the initial layout (e.g., the compact Cylinder layout) at the center of the current view. For a task that involves multiple focus areas, the user may need to walk around the physical space and move the interfaces to different places. Resetting allows the user to view all of the information together again.

Scene understanding to detect salient physical object and spatial constraints. Especially to augment physical task instruction, it would be important to ensure virtual interfaces do not distract the user from the main task but still provide reachable relevant information. The importance levels of areas in the field of view were proposed in Figure 5.5. Participants tended to put the most important virtual content on top of the key physical object, and place secondly needed content slightly on the right. The preference of aligning panels with flat physical surfaces or restricting them within an empty space was also demonstrated in the study. To apply those constraints, more advanced techniques involving object detection would be required to decide where the panels should be placed in relation to the real world. More than spatial surface alignment, the contextual constraints also expect understanding of the semantic relationships between the web content and physical objects in the scene to place more relevant information near the key task.

Support for multi-sensory interactions. Further interactions free from hand interactions may be enabled. Some participants reported that audio reading of textual content (text-to-speech) could be helpful to let the user focus on the physical task at hand. In addition, voice commands can replace hand gestures with verbal input; eye navigation can scroll a panel following the eye movement. These features have all been provided in MRTK to be configured with customized implementation. The utilization of multiple human senses would also follow the general HCI principle "flexibility in use" mentioned in Section 2.2.

Conclusion

In this thesis we have explored the presentation of legacy information, especially web content in AR to augment physical task instruction. This research contributed by proposing a new approach as immersive multi-view AR web interfaces and formulate a set of design principles regarding layout arrangement of such interfaces.

During this research, a prototype system was developed with VR to simulate three scenarios of performing real-world tasks, each with a set of multi-view AR web interfaces and general AR applications to be arranged. A remote and self-guided elicitation study was designed and conducted using the developed VR system. The study analyzed user interface arrangement data and survey responses, where participants expressed an overall positive impression of the multi-view interface style regarding its usefulness and the efficiency of conveying legacy task instructions. The elicitation study observed and named common patterns for multi-view web interface arrangement as follows:

- The interface layouts: Sphere, Cube, Hemisphere, Cylinder, Plane;

- The closer grouping relationships between interfaces: Semantic, Functional, Visually Balanced;

- The physical environment constraints: Spatial, Contextual;

- The combination between web interfaces and other applications: Insertion, Separation, and Rearrangement.

The most popular overall interface layout was Hemisphere, especially for multiple focus areas in the environment; while the Cylinder layout would be more common for interface grouping or a single focus area. Participants tended to cluster interfaces that show similarities in information types, functionalities, or semantic content. More important panels were put closer to the top center and secondly top right areas relative to the key physical task objects. Virtual interfaces were often aligned with physical surfaces or

confined within an empty space. The participants also preferred mixing web interfaces and general application for an integrated experience.

Based on the arrangement properties elicited, we have also suggested the following design or implementation strategies to develop multi-view AR web interfaces:

- Adaptive Hemisphere or Cylinder overall layouts;

- Semantic and Functional information separation and grouping;

- Interactions for manually adjusting individual and grouped interfaces;

- Scene understanding to detect salient physical object and spatial constraints;

- Support for multi-sensory interactions.

The next section discusses possible future work based on our research findings.

6.1   Future Work

The most obvious extension of this research would be to evaluate the suggested implementation strategies that apply our conclusions to the design preferences. More objective measures could be adopted to evaluate the efficiency (e.g., accuracy and time taken) of finding web information or instructing physical tasks with the designed multi-view interface layout. In addition, support for personalization could be explored for adapting the multi-view layout to the user's habits. The potential preference of placing secondly important content towards side of the dominant hand has been mentioned in Section 5.1.2, but it lacked enough participant information to prove. The adaption could also consider whether there is any particular layout that has been often made by the user in the past.

The elicitation study explored arrangement preferences given a set of predefined multi-view interfaces. It focused more on interface layouts than content creation or selection. Future research can be undertaken to investigate web information processing to prepare the content ready to be displayed in a multi-view style. This could involve two parts: web content separation and combination, which are discussed below.

Web content separation. In our VR simulation system, an assumption was made to separate web information based on content sections or media types (Section 3.2.1). Previous work discussed in Section 2.3 that decomposed content for mobile or web based AR also hard-coded the rule for decomposition (Lam et al., 2019; Srinivasa et al., 2016) . To formulate the principles of separating web content into multiple parts, future study can explore user customized web component separation. For example, given several conventional single-view web pages and a task requiring reading comprehension of the web information, users could be asked to split the web pages in their preferred way.

Web content combination. The multi-view style would enable grouping web components from different sources, which has been demonstrated in our VR simulation system.

In addition to the investigation of web component separation from a single page, more research could be conducted for web component combination that could involve a large number of sources. Techniques such as information retrieval and/or natural language processing could be required to analyze the semantic relationships across different websites, and their content can be separated and regrouped. The criteria for combination may vary depending on the purpose of viewing information: distinctive content could also be desired to group closer for easy comparison.

Particularly for augmentation of physical task instruction with multi-view interfaces, future work may consider procedural layout adaption to provide step-by-step guidance. As mentioned in Section 5.3.1, the scenario of performing a task can be divided into finer sub-steps, each with a particular action goal and possibly different web content needed. Existing web information often describes instructions as a numbered list, which could be one of the criteria to implement information separation and combination. Then interfaces can be wrapped into different steps for the user to switch between.

In summary, the potential future directions may investigate content extraction and adaption for multi-view web interfaces. This involves separating and grouping web components that may originate from different sources, where an AR or VR environment may not be necessary to be involved. Then our research findings may help with design decisions of the arrangement for the separated and combined web content in AR.

56

Participant Registration Form

Registration Form for Multi-View AR Web Interface User Study [1]

This form is to inform the research team about your interest in joining the Multi-View AR Web Interface User Study. After registration, we will email you with further details about this study. It takes around 45 minutes to complete all tasks and an online survey.

Please note that this quick form is only used to organise the study sessions. All information entered here is only for communication purposes rather than research data to be collected.

1. Do you have access to an Oculus Quest (or Quest 2) device? *

- Yes

- No

2. Are you living in Canberra and able to visit The Australian National University campus? *

We may schedule an in-person study session where you can use a provided device.

- Yes

- No - not eligible for this study

3. Do you want to use your personal Facebook/Oculus account for app installation on Quest?*

For all users, we provide a fake and random Facebook/Oculus test account subscribing to our VR application. You can add this test account on your Quest to access the application

---

[1] https://forms.office.com/r/hERdpuKXM2

(we will send you instructions). The test account username also serves as your unique participant ID.

If you would like to stay with your personal account on Quest to install the app, we will need to know your account email address to add our application to your app library.

- No

- Yes/Maybe

4. Please enter your Facebook/Oculus account email address: *

The email address you use to log into your Quest headset.

This information is only used for app installation, not research data to be collected.

5. When will you be available by for an in-person session at ANU? *

We will email you with further details for completing this study.

6. Please enter your email address for contact: *

7. Please let us know if you have any questions:

Written Study Instructions

The instructions are shared online [1]. The printed version is attached at the end of this thesis document.

---

[1] https://feiyue-tao.notion.site/feiyue-tao/User-Study-Participation-Instructions-ad59 71a4834b4de9a14eec5b8c6f5e21

Survey Questions

User Study Survey - Multi-view AR Web Interface [1]

Research Participation Consent

Please provide your consent to participate in this study. 1. I have read and understood the provided Information Sheet about the research project, and I have had any questions and concerns about the project addressed to my satisfaction.

The questions I have had (if any) are noted here:

2. I agree to participate in the project. *

- Yes

- No

3. I agree to be identified in the following way within research outputs: *

- No attribution

- Pseudonym and/or no attribution

Pre-study Questions

Please answer the following questions before interacting with the application.

4. Please enter your Oculus test account username. *

The test account information is included in our "User Study Details" email.

Example: test_bufrvz

5. How familiar are you with augmented reality (AR)?

Very unfamiliar 1() 2() 3() 4() 5() Very familiar

6. How familiar are you with virtual reality (VR)?

Very unfamiliar 1() 2() 3() 4() 5() Very familiar

In the VR Application: Introduction and Scenario 1 - Tube Assembly Task

Please put on the headset and launch our VR application.

The first scene will ask for your username and show a brief introduction. Please enter the same username as you entered in this form (in Question 4).

Move to the next scene to start Scenario 1 - Tube Assembly Task. Complete steps 1.1 and 1.2 following the in-app UI instructions.

At step 1.3 which will ask you to complete Survey Scenario 1, put down the headset and answer questions in the next survey section.

Questions for Scenario 1: Tube Assembly Task

7. Please describe how you initially arranged the panels in this scenario (step 1.1) and why you arranged the panels in such format.

8. Please describe if (and why) you changed panel arrangement at step 1.2 which added other general apps.

9. To better view the information, did you want to move any physical objects (other than the information panels or apps) in the room? Please describe what you would have liked to move and why.

10. Would you have liked to add any additional virtual information/objects to help you complete this task? Please describe what you would have liked to add, where to place them, and the reasons.

11. How hard do you think it was to complete this arrangement task?

Very hard 1() 2() 3() 4() 5() Very easy

12. What factors make you think it was hard/easy to complete the arrangement task?

— (Repeat similar instructions and questions for Scenario 2 - Cooking Task and Scenario 3 - Fieldwork Task) —

Post-study Questions

You have completed all in-app tasks and can exit the VR application.

Please answer the next survey section for the last few questions.

25. How immersive did you feel when browsing information in our VR application?

Least immersive 1() 2() 3() 4() 5() Very immersive

26. To what extent do you agree with the following statements?

- AR apps/website s are efficient way to display legacy information (e.g. texts, images).

- AR apps/web sites allow me to find all information about the tasks.

- Strongly Disagree - Disagree - Neutral - Agree - Strongly Agree - I Don't Know

27. Please rate the usefulness of separating information into multiple interactable panels compared to a single app window/screen:

Not useful 1() 2() 3() 4() 5() Very useful

Submission

By clicking 'Submit' on this survey, you consent to participate in the research

Bibliography

2019. Typography - mixed reality. https://docs.microsoft.com/en-us/windows/mixed-reality/typography. [Cited on page 3.]

AZUMA, R. T., 1997. A survey of augmented reality. *Presence: teleoperators & virtual environments*, 6, 4 (1997), 355–385. [Cited on pages 1 and 7.]

BERNARD, M. L.; CHAPARRO, B. S.; MILLS, M. M.; AND HALCOMB, C. G., 2003. Comparing the effects of text size and format on the readability of computer-displayed times new roman and arial text. *International Journal of Human Computer Studies*, 59 (2003). doi:10.1016/S1071-5819(03)00121-6. [Cited on page 3.]

CHENG, Y.; YAN, Y.; YI, X.; SHI, Y.; AND LINDLBAUER, D., 2021. Semanticadapt: Optimization-based adaptation of mixed reality layouts leveraging virtual-physical semantic connections. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, 282–297. [Cited on page 3.]

CHUAH, S. H.-W., 2018. Why and who will adopt extended reality technology? literature review, synthesis, and future research agenda. *Literature Review, Synthesis, and Future Research Agenda (December 13, 2018)*, (2018). [Cited on page 7.]

DEBERNARDIS, S.; FIORENTINO, M.; GATTULLO, M.; MONNO, G.; AND UVA, A. E., 2014. Text readability in head-worn displays: Color and style optimization in video versus optical see-through devices. *IEEE Transactions on Visualization and Computer Graphics*, 20 (2014). doi:10.1109/TVCG.2013.86. [Cited on page 3.]

DÜNSER, A.; GRASSET, R.; SEICHTER, H.; AND BILLINGHURST, M., 2007. Applying hci principles to ar systems design. (2007). [Cited on page 8.]

FUKUSHIMA, S.; HAMADA, T.; AND HAUTASAARI, A., 2020. Comparing world and screen coordinate systems in optical see-through head-mounted displays for text readability while walking. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 649–658. IEEE. [Cited on page 3.]

GIBSON, J. J., 1977. The theory of affordances. *Hilldale, USA*, 1, 2 (1977), 67–82. [Cited on page 8.]

Bibliography

Goto, M.; Uematsu, Y.; Saito, H.; Senda, S.; and Iketani, A., 2010. Task support system by displaying instructional video onto ar workspace. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, 83–90. IEEE. [Cited on page 13.]

Henderson, S. J. and Feiner, S. K., 2011. Augmented reality in the psychomotor phase of a procedural task. In *2011 10th IEEE International Symposium on Mixed and Augmented Reality*, 191–200. IEEE. [Cited on page 13.]

Khuong, B. M.; Kiyokawa, K.; Miller, A.; La Viola, J. J.; Mashita, T.; and Takemura, H., 2014. The effectiveness of an ar-based context-aware assembly support system in object assembly. In *2014 IEEE Virtual Reality (VR)*, 57–62. IEEE. [Cited on page 13.]

Klein, F.; Rubinstein, D.; Sons, K.; Einabadi, F.; Herhut, S.; and Slusallek, P., 2013. Declarative ar and image processing on the web with xflow. In *Proceedings of the 18th International Conference on 3D Web Technology*, 157–165. [Cited on page 10.]

Klose, E. M.; Mack, N. A.; Hegenberg, J.; and Schmidt, L., 2019. Text presentation for augmented reality applications in dual-task situations. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 636–644. IEEE. [Cited on page 3.]

Lam, K. Y.; Lee, L. H.; Braud, T.; and Hui, P., 2019. M2a: A framework for visualizing information from mobile web to mobile augmented reality. In *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom*, 1–10. IEEE. [Cited on pages vii, 3, 11, 12, 50, and 54.]

Lam, K. Y.; Lee, L. H.; and Hui, P., 2021. A2w: Context-aware recommendation system for mobile augmented reality web browser. In *Proceedings of the 29th ACM International Conference on Multimedia*, 2447–2455. [Cited on pages vii, 11, and 12.]

Langlotz, T.; Nguyen, T.; Schmalstieg, D.; and Grasset, R., 2014. Next-generation augmented reality browsers: rich, seamless, and adaptive. *Proceedings of the IEEE*, 102, 2 (2014), 155–169. [Cited on page 11.]

Lindlbauer, D.; Feit, A. M.; and Hilliges, O., 2019. Context-aware online adaptation of mixed reality interfaces. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, 147–160. [Cited on pages 3 and 9.]

Lu, Y. and Mayol-Cuevas, W. W., 2021. The object at hand: Automated editing for mixed reality video guidance from hand-object interactions. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 90–98. IEEE. [Cited on page 13.]

MacIntyre, B.; Hill, A.; Rouzati, H.; Gandy, M.; and Davidson, B., 2011. The argon ar web browser and standards-based ar application environment. In *2011*

10th IEEE International Symposium on Mixed and Augmented Reality, 65–74. IEEE. [Cited on page 10.]

MACKAY, W. E., 1998. Augmented reality: linking real and virtual worlds: a new paradigm for interacting with computers. In *Proceedings of the working conference on Advanced visual interfaces*, 13–21. [Cited on page 7.]

MACLNTYRE, B. AND SMITH, T. F., 2018. Thoughts on the future of webxr and the immersive web. In *2018 IEEE international symposium on mixed and augmented reality adjunct (ISMAR-Adjunct)*, 338–342. IEEE. [Cited on page 10.]

MARNER, M. R.; IRLITTI, A.; AND THOMAS, B. H., 2013. Improving procedural task performance with augmented reality annotations. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 39–48. IEEE. [Cited on pages vii and 10.]

MILGRAM, P.; TAKEMURA, H.; UTSUMI, A.; AND KISHINO, F., 1995. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, vol. 2351, 282–292. International Society for Optics and Photonics. [Cited on pages vii, 1, 7, and 8.]

MORO, C.; ŠTROMBERGA, Z.; RAIKOS, A.; AND STIRLING, A., 2017. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anatomical sciences education*, 10, 6 (2017), 549–559. [Cited on page 2.]

MOTTELSON, A.; PETERSEN, G. B.; LILIJA, K.; AND MAKRANSKY, G., 2021. Conducting unsupervised virtual reality user studies online. *Frontiers in Virtual Reality*, 2 (2021), 66. [Cited on page 14.]

PARK, S.; BOKIJONOV, S.; AND CHOI, Y., 2021. Review of microsoft hololens applications over the past five years. *Applied Sciences*, 11, 16 (2021), 7259. [Cited on page 1.]

QIAO, X.; REN, P.; DUSTDAR, S.; AND CHEN, J., 2018. A new era for web ar with mobile edge computing. *IEEE Internet Computing*, 22, 4 (2018), 46–55. [Cited on page 10.]

QIAO, X.; REN, P.; DUSTDAR, S.; LIU, L.; MA, H.; AND CHEN, J., 2019. Web ar: A promising future for mobile augmented reality—state of the art, challenges, and insights. *Proceedings of the IEEE*, 107, 4 (2019), 651–666. [Cited on page 10.]

RENKEWITZ, H.; KINDER, V.; BRANDT, M.; AND ALEXANDER, T., 2008. Optimal font size for head-mounted-displays in outdoor applications. doi:10.1109/IV.2008.73. [Cited on page 3.]

RIVU, R.; MÄKELÄ, V.; PRANGE, S.; RODRIGUEZ, S. D.; PIENING, R.; ZHOU, Y.; KÖHLE, K.; PFEUFFER, K.; ABDELRAHMAN, Y.; HOPPE, M.; ET AL., 2021. Remote

vr studies–a framework for running virtual reality studies remotely via participant-owned hmds. *arXiv preprint arXiv:2102.11207*, (2021). [Cited on page 14.]

ROLLAND, J. P.; HOLLOWAY, R. L.; AND FUCHS, H., 1995. Comparison of optical and video see-through, head-mounted displays. In *Telemanipulator and Telepresence Technologies*, vol. 2351, 293–307. SPIE. [Cited on page 47.]

SATRIADI, K. A.; ENS, B.; CORDEIL, M.; CZAUDERNA, T.; AND JENNY, B., 2020. Maps around me: 3d multiview layouts in immersive spaces. *Proceedings of the ACM on Human-Computer Interaction*, 4, ISS (2020), 1–20. [Cited on pages 9, 35, and 45.]

SHEPILIEV, D. S.; SEMERIKOV, S. O.; YECHKALO, Y. V.; TKACHUK, V.; MARKOVA, O.; MODLO, Y. O.; MINTII, I. S.; MINTII, M.; SELIVANOVA, T. V.; MAKSYSHKO, N.; ET AL., 2021. Development of career guidance quests using webar. In *Journal of Physics: Conference Series*, vol. 1840, 012028. IOP Publishing. [Cited on page 10.]

SRINIVASA, R. R.; VELUCHAMY, U. P.; AND BOSE, J., 2016. Augmented reality adaptive web content. In *2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, 107–110. IEEE. [Cited on pages 11 and 54.]

WANG, X.; ONG, S. K.; AND NEE, A. Y., 2016. A comprehensive survey of augmented reality assembly research. *Advances in Manufacturing*, 4, 1 (2016), 1–22. [Cited on pages 1 and 2.]

YAMAGUCHI, M.; MORI, S.; MOHR, P.; TATZGERN, M.; STANESCU, A.; SAITO, H.; AND KALKOFEN, D., 2020. Video-annotated augmented reality assembly tutorials. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, 1010–1022. [Cited on pages 2 and 13.]

User Study Participation Instructions

Project title: Multi-View Web Interfaces in Augmented Reality

How to join this study

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/0959089c-46da-4252-98d8-d53d1fa7f838/Participant_Information_Sheet.pdf

1. There are different ways to participate in this study according to your preference:

   - If you have access to an Oculus Quest or Quest 2 device:

     - You will be joining this study remotely with your own device. The study is designed to be self-guided (online meeting is also available if you need real-time assistance).

     - The "*User Study Details*" email we sent to you includes a random and unique fake Oculus/Facebook test account username, email address, and password. In this study, we will use this username to represent you without any identifiable/personal information.

     - For app installation, you can use either the test account or your personal account on Quest. More details are in the "*Install the app*" instructions below.

   - If you do not have a Quest:

- You may visit ANU campus in Canberra to use our provided device. We will email you to confirm availability.

- When you attend the in-person session, start from step 2 in the following instructions.

> ✉ If you have any questions at any stage of this study, please contact feiyue.tao@anu.edu.au.

How to complete this study with Oculus Quest (or Quest 2)

Please follow these steps after you receive our email reply including test account details. It may take around 40 minutes to complete all steps.

1. Install the app:

    You can use either your personal Facebook/Oculus account or our provided test account on your Quest to download the app.

    ▼ (Click to open/hide) Option 1: Add the test account to your Quest device

    i. Make sure you have logged in with the device owner account.

        (Please note that you can only add up to 3 additional accounts on Quest. If you have reached the limit, you may need to remove an existing account on your device first)

    ii. Press the Oculus button on your right Touch controller to pull up your universal menu.

    iii. Hover over the clock on the left side of the universal menu. When "*Quick Settings*" appears, select it to open the "*Quick Settings*" panel.

    iv. Select *"Settings"* in the top right corner, then select **"Accounts"**.

    v. Select **"Add Account**" in the top right corner.

vi. Log in with the provided email address and password. This test account works as both Facebook account and Oculus account. Complete the headset setup.

https://s3-us-west-2.amazonaws.com/secure.notion-static.com/5780d6f8-3215-4ca1-8e43-683e6a19ed60/Add_Account.mp4

Video for adding a new account

▼ (Click to open/hide) Option 2: Stay with your personal account on Quest

If you have not provided your personal account email address in the submitted Registration Form, please re-submit the form or email us (reply to our "User Study Details" email or contact feiyue.tao@anu.edu.au) to let us know your account email address.

1. We will add your account as a subscribed user for our app. You will then receive an invitation email. Click the link to accept:
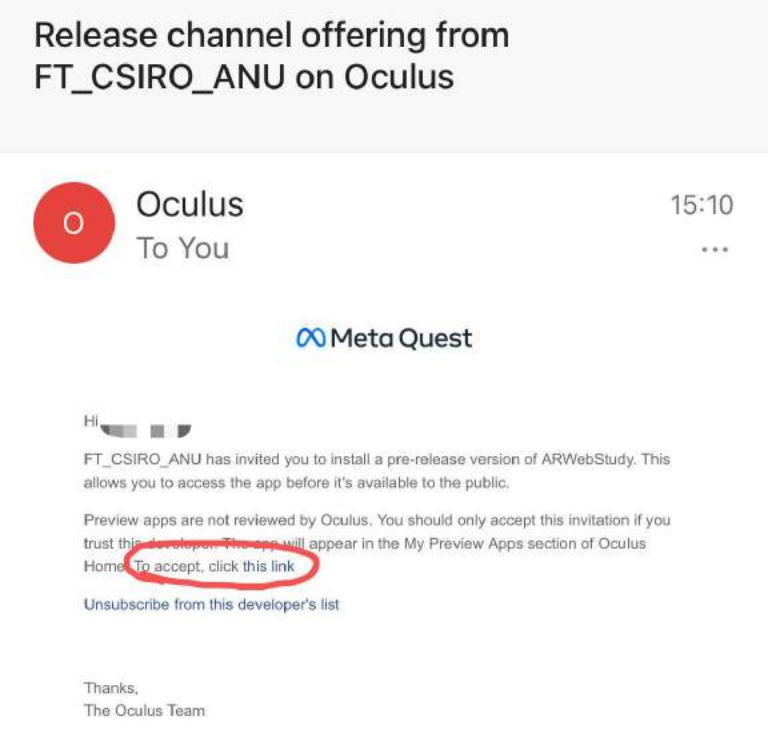
Invitation email

2. You might need to restart Quest to see the app showing up in your app library.

After adding the test account or accepting your personal account as a subscribed user, you should see the app "**ARWebStudy**" showing in *Apps* (or *Library*). Click to install.

Find ARWebStudy in Apps and install

📌 Option for online meeting:
If you need real-time assistance to set up and complete this study, you can book an Online User Study session <u>here</u>.

▼ (Click to open/hide) Cast and share your Quest view in an online meeting**:**

1. On your computer, go to <u>oculus.com/casting</u> with Google Chrome or Microsoft Edge and log in to your Oculus account.

2. Put your headset on and press the Oculus icon on your controller to open the universal menu.

3. Select the Arrow icon to open ”*Sharing***”**, then select ”*Cast***”**.

4. Select ”*Computer*” then select ”*Next*”.

    - If you don't see your computer listed, make sure that your headset and computer are on the same Wi-Fi network.

5. Select ”*Done”*.

6. On your computer, in our Teams meeting room, select "*Share"* in your meeting controls, then select your browser window that is casting the Quest view.
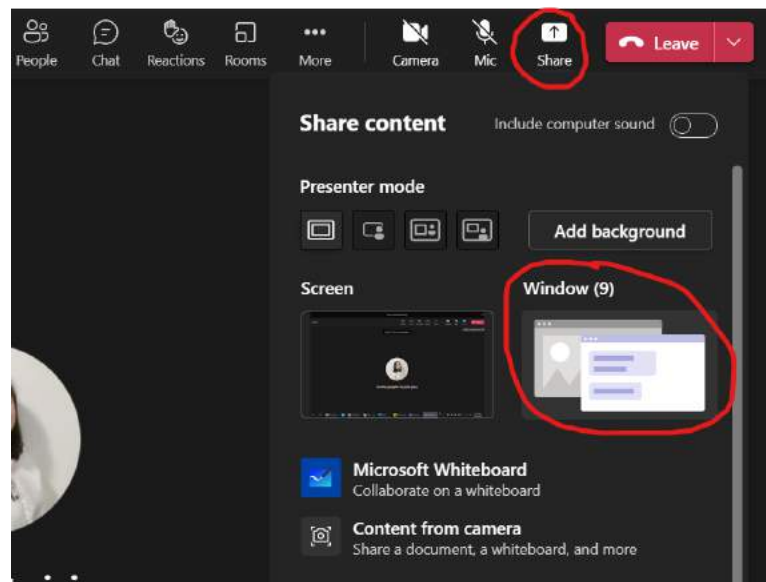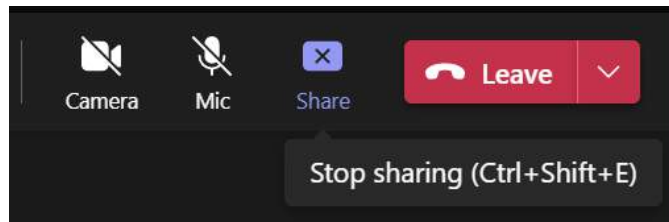

To stop casting:

1. Press the Oculus icon on your controller to open the universal menu.

2. Select the Arrow icon to open "*Sharing***"**, then select "*Cast***"**.

3. Select "*Stop Casting***"**.

4. On your computer, you may select "*Stop sharing*" in Teams meeting controls to stop your screen sharing in our meeting room.


---

2. Open this **online survey form**.

   Fill in the first 2 sections to provide your participation consents and answer pre-study questions. The form will ask for your **test account username** that we provided in our email.

3. Put on the Quest headset. Make sure it is connected to the Internet.

   Click and launch the app "ARWebStudy" that you have installed. The app will also ask for your **test account username**.

   - You can interact with the app while seating or standing.

   - Press the controller's Trigger button for interactions (click buttons, select and move objects, etc.). Avoid using thumbstick for best experience.
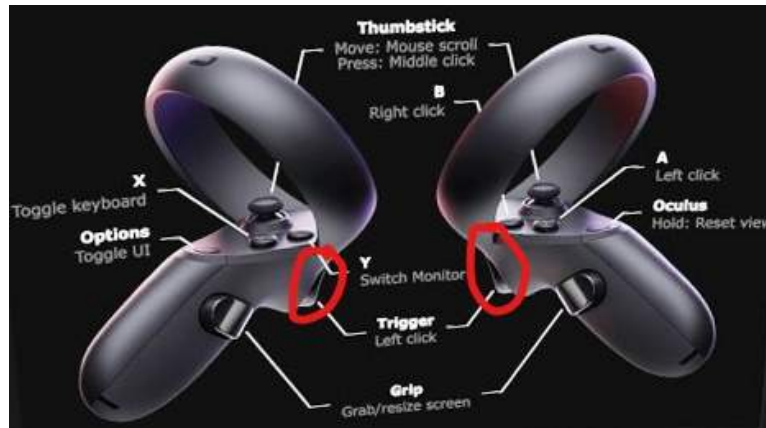
4. Follow the in-app user interface instructions to complete the tasks, including answering the rest of the survey form.

   You will see 3 scenarios in the VR application: "tube assembly", "cooking", and "fieldwork", where you will be asked to arrange a set of virtual interfaces simulating AR contents. The interfaces are initially put at the back of your view.

   After completing tasks in one scenario, put down the headset to answer survey questions for that scenario, then return to the app for the next scenario. Repeat the process until you see a "Thank you" message in the end. The interface arrangement data will be saved remotely.

5. Exit the app. Submit the online survey form.

How to remove the test account and/or uninstall the app

Remove test account:

After completing the study, you can remove the test account including the app from your device:

1. Log in to your device with your device owner account.

2. Press the Oculus button on your right Touch controller to pull up your universal menu.

3. Hover over the clock on the left side of the universal menu. When "*Quick Settings*" appears, select it to open the "*Quick Settings*" panel.

4. Select *"Settings"* in the top right corner, then select **"Accounts"**.

5. Select "*Remove***"** next to the test account and confirm.

Uninstall/Unsubscribe the app:

If you used your personal account, you can uninstall the app.

Click the three dots. Select "Uninstall" from drop-down menu.

We will remove all subscribed users after data collection. You can also manually unsubscribe the app to completely remove it from your library:

1. Log in to https://secure.oculus.com/

2. Preview Apps > Unsubscribe

Redeem Code

Quest

Rift

Gear VR/Oculus Go

Payment Methods

Devices

Cloud backups

Security

Friends

Friends

Notifications

Email Preferences

Privacy Centre

Facebook Settings

Preview Apps

Subscribed

ARWebStudy - RC (Quest)

You do not have any pending invitations to preview apps.

Select the Preview Apps tab to unsubscribe the app