Panoramic Perspectives

Evaluating spatial widgets in immersive video through heuristic evaluation

Max Meijers
Stud. No. 6493769

Wolfgang Hürst
Supervisor, First examiner

Egon van den Broek
Second examiner

A thesis presented for the degree of
MSc Game and Media Technology

Utrecht University

Department of Natural Sciences
Utrecht University
The Netherlands

4 July 2023

Contents

PREFACE

This thesis consists of two parts. The first part is a stand-alone scientific paper describing the main contribution, followed by an annotated appendix to reflect more related work, background information and offer insight on aspects that were omitted from the main paper.

Panoramic Perspectives

Evaluating spatial widgets in immersive video through heuristic evaluation

Max Meijers, 6493769
Utrecht University

ABSTRACT

This research explores the value of spatial indicator widgets in 360-degree virtual reality (VR) videos to address the research question of *which spatial widget design, if any, is suited to adoption into a user interface for VR 360-degree video playback*. It proposes and implements three designs for spatial widgets and compares their effectiveness based on a heuristic study involving three experts. The study found that spatial widgets offer welcome orientation aids, especially in visually challenging environments, despite potential issues with motor load and distraction. Among the evaluated designs, the *linear compass* widget emerged as the most promising but requires redesign for optimal functionality. The thesis concludes that the inclusion of spatial navigation widgets in VR 360-degree video players is a worthy consideration and suggests avenues for future research, such as the integration of these widgets into emerging technologies like volumetric video and augmented reality. The results of this research serve as a first step towards designing more user-friendly VR video interfaces.

1 INTRODUCTION

1.1 Motivation

With the surge in consumer VR headsets in recent years, 360-degree video—which is particularly well-suited to the format—has enjoyed increased popularity. Standalone devices and ones based on smartphones, like the Meta Quest, Oculus Go, Samsung Gear VR and Google Cardboard, include apps for watching online video services, such as YouTube, in VR, greatly lowering the barrier of entry into 360-degree video for the average consumer.

Compared to conventional video, which only features time as a dimension to contend with, 360-degree video adds a whole new dimension for the user to navigate; that of space. A limited subset of video, volumetric video, allows for six degrees of freedom in spatial navigation. However, VR video is typically 360-degree video, filmed as a sphere around the viewer, where navigation is limited to rotation of the viewpoint along the three rotational axes in the centre of the captured sphere. This is the type of VR video that is focused on in this research.

A big strength of immersive 360-degree video is that it can really transport the user to a different place. But despite adding a new navigable dimension, player applications for VR video hardly ever provide symbolic visualisations of it, continuing to only feature visualisations for the temporal domain that they share with traditional 2D video. The directorial intent behind a video can greatly influence the way the user interacts with the video and the relevance of having a visualisation of the spatial domain. Four main cases can be distinguished.

In the extreme case, the video director wants the viewer to watch only one point of interest. Examples could be sports videos like downhill skiing, or movie-like experiences with pre-designed choreography and intention. There is no meaningful exploration at all. The video might rotate around the user to keep the forward direction always facing the action, or various visual cues might be provided to draw the user's attention to where the director wants them to look. There is no need for a spatial indicator here, as the system guides you.

Otherwise, there may be a main direction, but freedom is given to the viewer to explore a little bit. This could be a sports match recorded by a member of the audience. You are seated in a stand and can freely look around wherever you want on the field, but also look around and behind you at things happening in the stand. In this case, it may be useful to have a compass depending on whether the viewpoint is dynamic or not. If it is recorded from a fixed spot, the viewer will not have any trouble intuitively keeping their sense of orientation.

A third case is where the user is free to orient themselves in a guided context. This might be a guided city tour. A guide indicates various points of interest all around the user, allowing the user to look around at their own leisure, taking advantage of the video simulating that they are really there, exploring the atmosphere and allowing to build an intuition of the lay of the land in that location. A spatial visualisation would be useful here, to give the user a fixed point to rely on for maintaining their sense of direction and aid in the creation of a mental map of the area.

Finally, at the other extreme, the user is left to do exploration entirely on their own. This would be a video without any direction or guidance, points of interest are to be made up by the viewer themselves. This may be a recording by an individual visiting a natural landmark or wandering through a city, with natural things happening all around. No guidance is given and the user can explore totally on their own. Like the previous case, there would be a real use for a spatial visualisation here, but even stronger as the user is left to their own devices to maintain their orientation.

Within 360-degree video, the user may have a few ways to navigate in the provided spatial domain, to turn the facing direction of their viewpoint. It can be achieved by physically turning their head to look around, or more rarely, a supplemental method is offered to turn the viewpoint without needing to physically move that involves a controller interaction such as a drag gesture, virtual UI button press, or physical controller button/joystick press.

A common problem encountered in VR—in videos as well as in general usage—is the feeling of disorientation, where the user gets lost after looking around and is not able to find back the original forward direction, either in the virtual world they are viewing or in their real-life environment. If a VR application or experience allows the user to manually rotate themselves through methods separate from rotating their head, through button or joystick interactions, the virtual world and the real environment that the user is located in (a rotating chair at a desk, a couch in a room, etc.) may become misaligned, further amplifying the difficulty of finding back the original orientation.

Amenities to solve this problem are often found in video games. In games that portray a three-dimensional world for the player character to explore, there is often some form of compass or map on the screen during gameplay. Using these, the player can keep their bearings within the game world, easily determining the relative position of points of interest in the world and their orientation relative to the cardinal directions. It makes sense that they would feature such visualisations, because without them a player may easily become lost in the potentially vast virtual worlds that have been created. The element of space and interactivity therein is thus shared between 3D video games and 360-degree VR video, and as argued above a case can be made for the relevance of spatial visualisation in 360-degree VR video too. So why do we not see spatial visualisations in 360-degree VR video?

Having spatial indicator widgets present during playback of 360-degree videos in VR is thus valuable to look into, to possibly improve spatial awareness, combat the feeling of disorientation, or even helping users better remember the portrayed environments by providing a sense of direction for them to anchor themselves by. Little focus has been put on spatial indicator designs for 360-degree video as of yet, beyond radar cone displays such as in existing projects and players ImAc and Omnivirt, which will be discussed in Section 2.

Widgets to visualise the spatial element can be translated into virtual reality in a large variety of ways. Because this is a largely untapped field, there is a point in doing a heuristic evaluation to start, because various visualisations may simply be unsuitable in this context and the likely but unproven statement that they are helpful at all needs to be verified. We shall thus be doing an initial study to determine if the visualisations most commonly used for spatial orientation even make sense to use. New approaches to visualising time indicators have not been touched much in VR video either, but one can easily assume that the tried-and-true linear timeline design from traditional video will be the preferred method by virtue of it being the indicator that the whole population has gotten used to over years of experience with it in many other applications. Therefore, the choice is made to focus exclusively on effect of spatial indicators, and take temporal indicators out of the equation.

1.2 Research question

This research proposes development of and subsequent scientific comparison and analysis of a variety of spatial visualisation widgets for VR video. It is not a matter of course that such widgets will bring benefit to the viewer, so taking this into account we created the following research question:

Which spatial widget design, if any, is suited to adoption into a user interface for VR 360-degree video playback?

We will answer this question for the spatial widget designs most commonly used in other interactive contexts, such as 3D gaming and virtual world exploration. See Section 3.2 for a description of these widgets and justification of why they were selected for this study.

Given that it is unclear if such widgets provide a benefit at all, we opted for a qualitative evaluation with a heuristic study to evaluate the usefulness of such interaction components for 360-degree video. See Section 3 for more details on the chosen methodology.

1.3 Following structure

In the rest of the paper, first a summary of related research and existing practical interface implementations is given in Section 2. In Section 3 we introduce our new widget designs, followed by a description of the further design and methodology behind our research. The results of said research are provided in Section 4, and finally conclusions and related work are found in Sections 5 and 2, respectively. This paper features an appendix. This contains a more in-depth evaluation of related work, in Appendix section B. Health issues led to us being unable to put the full design into practice. As such, the design was adapted, while the original design is preserved in Appendix section C.

2 BACKGROUND

2.1 VR interface research

Prior research into interfaces in VR has been conducted, and their design principles from various scientific sources have been evaluated with respect to their relevance for the research presented in this paper. Regular, non-360 video and VR video are quite different, so common design principles for regular video might not apply. Therefore, we look at existing work in the fields of both regular video and VR video to take inspiration from the various problems and designs that they may propose.

The found literature can be divided into several broad categories. These include research into interfaces and control for regular video, VR video and for general interfaces in VR, and research into making 360-degree video more browsable in desktop as well as VR interfaces. These categories will be traversed in the following subsections.

2.1.1 Video interface and control design.

This is the main area of related work. These papers all deal with user interactions relating to videos in VR, and many of them feature novel interface designs, like for timeline design[20][35], the use of picture-in-picture with reduced FOV in certain situations[20], while others feature completely new control methods, like temporal and spatial video control through head rotation[26] or the control over hypervideo[24], which is a network of video that features selectable split paths that all lead to unique video content.

Some notable works are described separately in the following paragraphs.

A large, international project that has common ground with this research is the *ImAc* (Immersive Accessibility) *project*[14]. This is a project registered with the European Commission that seeks to specify the requirements of and develop a cross-platform video player for 360-degree video that incorporates important accessibility features into its user interface.

While the *ImAc* project can be freely contributed to, its scope far outpaces the aim of our research. However, several of their observations and learned lessons pertaining to UI design in VR can be adapted into our work for certain of the tested UI designs. These include an interface design that is narrower than the field of view so as not to hide elements off the edge of the visible area or strain the eyes of the user[17], and a circular radar display that uses an arrow pointing to the middle of the small radar circle to indicate which direction is the 'front' direction of the scene. A screenshot of the player featuring the narrower timeline UI and radar circle can be seen in Figure 1.

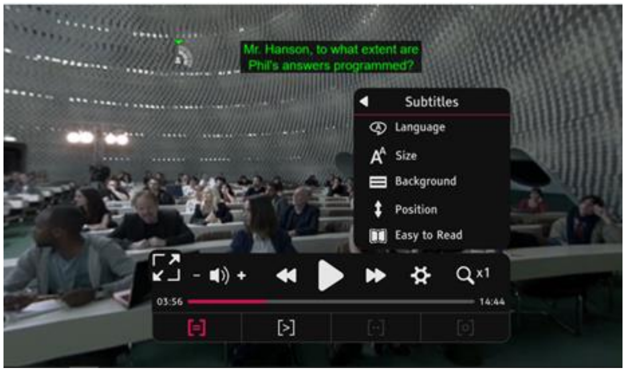

Figure 1: One of the configurations of the ImAc player. A spatial widget can be seen at the top, left of the captions.

2.1.2 Making 360-degree video more browsable. A good amount of research has gone into ways for making 360-degree videos more easily watchable in VR[4][29], or even in 2D[10][32], typically by introducing additional UI elements to hint at the location of points of interest. This is work that is relevant to us, because solutions often involve the introduction of new UI elements and the subsequent evaluation of their functionality. Despite possible expectations, results focusing on desktop interfaces for 360-degree video are just as relevant as those focusing on VR interfaces, as solutions and ideas for user interface elements are not necessarily limited to just desktop or VR UIs.

A notable paper in this category is Lin *et al.*'s "Outside-In"[15], that adds so-called spatial picture-in-picture (PIP) previews of regions of interest (RoI) that are outside the current video viewport. These previews are put on the connecting line between the viewport centre and RoI, then rotated and tilted, to intuitively indicate their location relative to the current viewport. Additional rules are in place to prevent overlapping PIPs. They compared their method to arrow-based guidance methods, which it outperformed, and they applied their method to a mobile-based 360-degree video player and a teleconferencing application for use with a regular flat-panel screen. Both applications allow tapping or clicking the spatial PIPs to automatically rotate the camera (virtual camera, in the case of the mobile application) to face their indicated regions.

Notably, however, they did not apply their method to a video player for use with actual VR HMDs, although we believe that their method could be highly important in improving the VR video viewing experience overall if applied there. However, as our research seeks to determine the effects of only the various UI designs, it was decided to not incorporate their findings into the experiment, to avoid situations like, for example, one where the participant is looking for a musician in the video and would have to rotate their view to find them, but a spatial picture-in-picture preview at the edge of their vision alerts them to exactly when and where the musician appears. This removes much of the need to actively browse the video, and thus removing interaction with the tested UIs.

2.2 Practical interface implementations

Due to the area being relatively young, not much specifically pertaining to 360-degree video interface interaction has been scientifically documented yet and most individual persons and companies either do their own private research or base themselves on traditional UI when designing VR interfaces. Hence, direct empirical evaluation of methods in existing products was relevant. An evaluation of various designs currently available in VR and non-VR games and video players which may be used for inspiration follows below. For VR video players, all player applications available on the Meta Quest store at the time of writing, early 2021,were evaluated.

An important thing to note is that all evaluated video players incorporate a traditional, linear timeline along the lower part of the visible area, with clickable buttons near it to control playback, similar to what is seen in Figure 1. However, depending on their remaining characteristics, they are split into the following groups.

2.2.1 Desktop compass players. This group of players features *OmniVirt*[25], the desktop *Facebook* player and *Radiant Media Player*[9]. The main distinction of this group is that they feature a compass in their UI that indicates the current viewport orientation relative to the original 'front' direction and functions as a re-centre button when clicked. Radiant Media Player does not feature an actual compass, but instead has arrow buttons for changing the orientation and a separate re-centre button. A screenshot of the OmniVirt UI can be seen in Figure 2.

Figure 2: The Omnivirt UI. Note the compass in the top right.

Inspiration from this group is to include this compass feature as one of the tested UI elements. A challenge in adapting it to VR would be its placement and sizing.

2.2.2 Simple control in VR. The *Oculus Browser* video player, along with many other VR video player apps on the Oculus platform, including *Oculus TV*, *Neverthink*, *Facebook Watch*, *Red Bull TV* and

Major League Baseball only feature pointer controls, with no spatial visualisation or controls whatsoever, if applicable at all (Neverthink, Facebook Watch and Red Bull TV have no spherical content).

2.2.3 Spatial control in VR. The following group, including the *YouTube VR*, *DeoVR Video Player*[16] and *Within* apps share the characteristic of featuring spatial controls, allowing one to watch everything around them in a 360-degree sphere without being necessitated to move their body. They are the only ones to have this feature. YouTube VR achieves the feature through allowing the player to ‘grab’ the viewport by holding the grip button and moving the controller to move it around themselves. DeoVR offers plenty of adjustment sliders in its interface, including pointer-operated sliders to control all three axes of orientation. A screenshot of the interface is shown in Figure 3 Within features continuous (not step-wise) control over orientation about the vertical axis through the use of the joystick on the handheld controller.

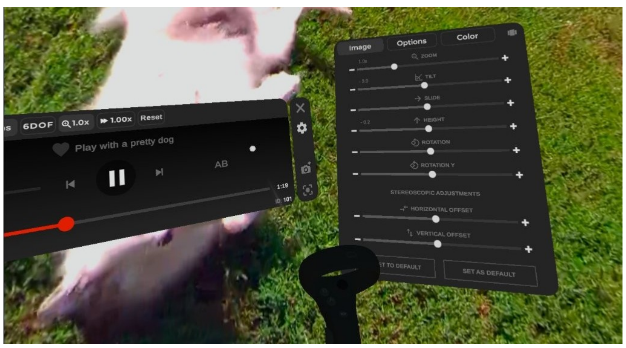

Figure 3: The interface of DeoVR, showcasing its many sliders.

None of the players in this group feature an orientation visualisation.

2.2.4 Quill Theater. The video player in the app *Quill Theater* is wholly unique. Instead of traditional pre-rendered videos, the videos in this app are real-time 3D environments in which the user experiences six degrees of freedom. The controls are unique too. Users find themselves with a PDA-like screen attached to their offhand that houses all the controls. Buttons on the offhand menu are pressed by bringing their main hand close to it and moving their hand forward as if physically pressing a button. There is no timeline on the UI, as videos are split into ‘scenes’ that either repeat or are jumped between by using buttons. A screenshot of this UI can be seen in Figure 4.

Inspiration for the experiment in this paper from this app comes in two forms. First, the UI being attached to the user’s hand is an interesting avenue to explore, as it may solve the common problem when using a video player where the user needs to either wait for the video UI to go away, or it going away when the user does not want it to. With it attached to the user’s hand, they can simply raise their hand when they want the menu and lower it again when they are done.

2.2.5 Video games. As seen in the prior subsections, 360-degree video players, VR or non-VR, do not often feature a visualisation of the spatial domain. A form of media that features spatial visualisations much more commonly is that of video games. This is

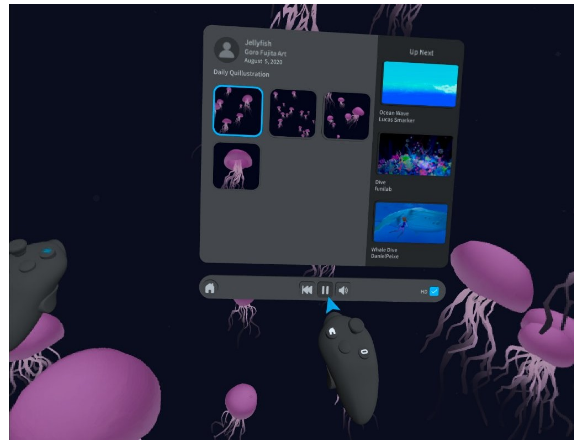

Figure 4: The touch-based Quill interface, attached to the left controller.

most likely to thank to the player being in direct control of the spatial domain in most games, through movement and orientation of their character, making it relevant to display that information to the player. Two principal designs are used for spatial visualisation in video games. They are a circular radar, or a horizontal compass, as described in the following paragraphs.

The most common visualisation of the two is a circular *radar* near a corner of the screen, as found in many games, including the *Grand Theft Auto* and *The Witcher* series. This design can take the form of a plain radar (or sonar) circle, usually with a centre dot to indicate the player character and other-coloured dots to indicate other points of interest. It is more common to see a *minimap* in place of the plain radar. This design functions identically to the radar, but instead of the inner circle merely indicating points of interest over a blank void, as in the case of the radar, it is filled in with a top-down view of the area from the map around the player character, allowing the UI widget to do double duty. Both radar designs tend to rotate their contents in order to stay aligned with the facing direction of the player character. The minimap design also almost always features compass functionality, with an indication of the North that rotates with the map contents. Something to note is that most video game radars are divorced from the usual beam-based scanning that actual radars do, meaning objects of interest around the player character will update in real-time, rather than only when the scanning beam next passes over the object. Of course, there are exceptions to this, especially in simulation-style games featuring radar- or sonar-based equipment. An example of a radar from the game *Grand Theft Auto IV* that is representative of most uses is shown in Figure 5a.

In typical modern 360-degree videos, the viewpoint is not allowed to move laterally, and so there is no point to displaying a map of any kind within the radar circle. One might suggest that a top-down view of the video sphere could work as an approximation of a map around the viewer that could help more intuitively visualise the virtual space around them, but because non-stereoscopic 360-degree videos, which is most of them, do not encode depth, it is typically not possible to separate the ground plane from objects standing on it. While for the downward-facing area of a video this is not a problem because of the perspective, it becomes a problem for the front/back/side-facing faces of a projected video, as a plain wall

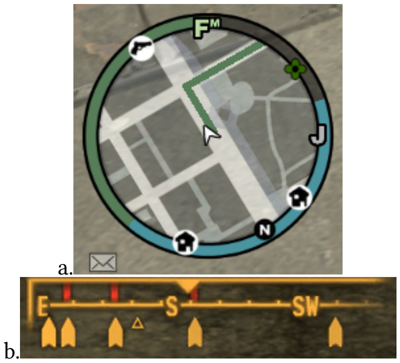

Figure 5: **a.** depicts the radar/minimap combination from GTA IV. Note the 'N' and other icons indicating the relative direction of the North and other points of interest. **b.** shows the compass in Fallout: New Vegas. Letters on the line indicate compass headings, arrows below indicate points of interest, and red lines on top indicate relative enemy directions.

standing straight up is indistinguishable from the ground stretching into the distance. Therefore it makes sense to adopt the traditional radar design with the empty inner circle for use in VR video. In fact, this is the design that is used by video players that contain a spatial visualisation, such as the *ImAc*, *OmniVirt* and *Facebook* 360 players.

The second design is a horizontal, linear *compass* along the top or bottom of the screen. This design is more commonly seen in open-world games, like the *Fallout* and the latest games in the *Assassin's Creed* series. This design of compass is visualised as a long, horizontal line that spans a certain distance along the top or bottom of the screen, featuring letters to indicate the compass headings. As the player rotates the camera, the letters on the compass move in and out of view too, in order to stay aligned with the current viewing direction of the player character. Points of interest such as important locations or enemies are commonly indicated on this compass too, so that the player can easily find their way to them through the world, by rotating the camera until the desired point of interest is in the middle of the compass and then moving forward, or get a good idea of what is around them without needing to move the camera. An example from *Fallout: New Vegas* can be seen in Figure 5b.

This compass design is somewhat unusual, as from real life we are used to compasses being round discs, but thanks to the combination of its linear nature, the fact that its horizontal orientation matches up to the primary axis of rotation used in such 3D games, and the fact that its length is often made to match to that of the horizontal field of view in the game, navigation can still become rather intuitive with practice; if something is visible on the compass, it or its relative direction is visible in the 3D view of the game from the current orientation of the camera. Its horizontal, linear design may match up very well with the traditional horizontal design of a video timeline. It is therefore worth implementing into our experiment as a test condition to evaluate in-depth.

3 METHODOLOGY

As discussed above, there are intuitive reasons why adding a spatial widget to the interface of a 360-degree video player could be beneficial. Yet, if this is true and what kind of widget design is most promising needs to be verified. Therefore, we perform a pilot study to answer the research question introduced in Section 1.2.

3.1 Participants and structure

Being a pilot study, the experiment is structured as a *heuristic analysis*, done with experts in the 360-degree VR video field. Heuristic questions are used by the experts during the experiment to help determine the pros and cons of each spatial widget design. Heuristic analysis relies on the experience and expertise of the evaluator to be representative, but it being based on heuristics will always introduce bias and is thus not fully representative of the real world. However, the experts' expertise in the field allows to assume that their assessment holds true for most real-world scenarios with real users.

Heuristics relevant to our research are selected and combined from several established sets of heuristics and guidelines broadly focused on the Human-Computer Interaction field, by the authors Nielsen & Molich[23], Gerhardt-Powals[12], Weinschenk & Barker[36] and Shneiderman[28]. A categorised, literature-wide summary of usability principles by Connell[5] was used to guide and unify the search. The selected heuristics are detailed in Section 3.3.

No matter how effective a certain design may be, if the experience of the end user using the system is negative, it would be ill-advised to recommend said design. Therefore, beyond the established heuristic analysis, emphasis is also placed on each expert's perceived intuitiveness of each system and their overall opinion of working with it, which they are free to remark on during their analysis trials and asked about explicitly near the end of the session.

3.2 Test conditions: Widget designs

The experiment features three distinct test conditions, each representing one unique spatial widget. One distinctive characteristic between widgets for spatial orientation is their placement with respect to the user and environment. For our test, we have therefore chosen three designs that are representative for a typical implementation of widgets that are (a) located in UI space, (b) in world space, and (c) on the user's virtual body.

A schematic representation is provided in Figure 6, illustrating the placement of each widget in relation to the user. The design of each widget is discussed further in the following paragraphs.

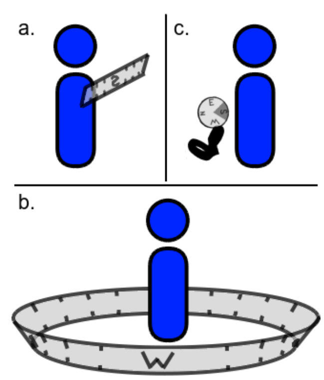

Figure 6: Schematic drawings of the **(a)** UI-, **(b)** world- and **(c)** body-located widgets.

The first widget, the UI-located one, is the "ribbon" or **linear compass**. This is a long strip that can scrolls horizontally as the user turns their head. It features horizontally aligned compass letters and a pair of red arrows indicating the current facing direction. The widget is affixed along the bottom of the user's field of view. A screenshot of the player featuring this widget is depicted in Figure 7.

Figure 7: The linear compass test condition.

The second widget is world-located. This is the "cylindrical" or **ring compass**. It is a compass ribbon that fully surrounds the user, roughly at shin-height. Like the linear compass, this compass lines up the various compass letters along its length, but it does not feature arrows to indicate the faced direction. Instead, it being world-located makes it physically intuitive for the user to read, as their line of sight into the virtual world will naturally line up with the relevant cylindrical compass reading. This widget can be seen in Figure 8.

The third and last widget is located on the user's virtual body, and is given form as a **"wristwatch" compass**. This widget attaches to either of the user's wrists (determined by a questionnaire question to be the wrist they prefer to wear a watch on) and features a circular plate like a physical wristwatch. Instead of numbers, however, it features compass directions on its face, which rotates around its axis according to the user's head orientation. A grey cone affixed to the top of the watch face corresponds to the heading and width of the user's field of view, used to read which way they are looking. This widget can be seen in action in Figure 9.

Each widget has a way for the participant to make it appear and disappear. For the view-mounted widget, the linear compass, it appears when the user rotates their vision a few degrees above a certain speed to make it fade in. The widget fades out automatically after inactivity. For the world-mounted widget, the ring compass, the user pitches their vision to below the horizon to fade it in. It fades in further the further the user looks down. Looking up again makes the widget fade back out. Finally, for the watch widget, the user simply raises or lowers their wrist to check it, like a real watch.

Figure 8: The ring compass test condition.

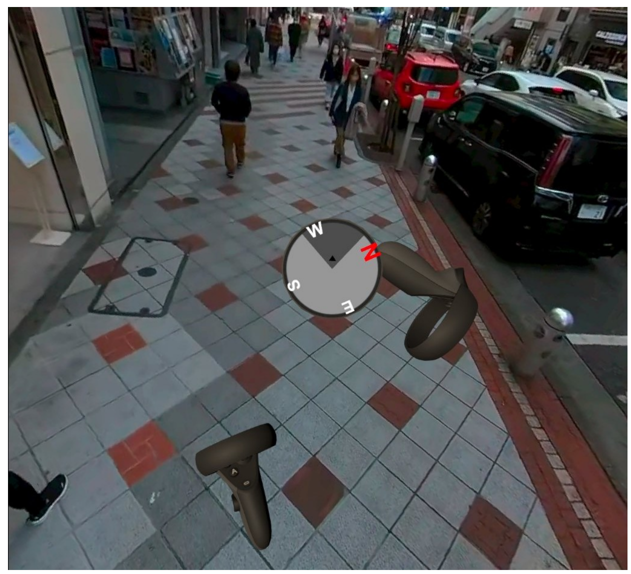

Figure 9: The wristwatch compass test condition.

3.3 Heuristics

The list of heuristics used in this research is compiled from various sources, to include ones relevant to our specific area of research. A list of them follows below. For each listed heuristic, the numbers of the relevant heuristics and principles from the literature are listed within parentheses.

The cited sources indicate it is important to use intuitive visuals and metaphors, and to match real life concepts where possible. The first heuristic we use is therefore:

Heuristic H1: Intuitive readability (Nielsen 2, Powal 2/5, Weinschenk 7, Connell 14/19).

Sub-questions belonging to this heuristic are:

- Are the various widget UIs intuitive to read?
- Do you feel the widget is a suitable metaphor for the directions, even if its depicted directions may not necessarily

match the video content (e.g. ‘north’ in the video actually being east in the real location)?
- Do the different widget UIs give you enough information?

Another important principle shared between all common heuristics sources is that visually, UI elements should not present too much information at once, and avoid obscuring other information in the scene. Therefore, the second heuristic we use is:

Heuristic H2: Perceptual load/clarity/contrast (Nielsen 8, Powal 8, Weinschenk 2/6, Connell 24/27/28).

- Does the design of any of the widgets get in the way during use?
- Do they obscure things from the videos due to their size or placement?

The sources state it is important to ensure the various elements of the UI in the application express a consistent design between one another. This is adapted into our third heuristic.

Heuristic H3: Consistency (Nielsen 4, Shneiderman 1, Weinschenk 16, Connell 21).

- Do the different widgets feel consistently designed to each other?
- If not, what elements stick out as problematic to you in that regard?

It is also important for UI elements to be in locations where users would expect to find them, or allow to move them if not. The fourth heuristic is therefore:

Heuristic H4: Manipulability (Nielsen 7, Weinschenk 12, Connell 3/22).

- Do you feel the widgets are intuitively placed?
- Would you recommend moving them?

The sources agree that it is important to offer multiple ways to navigate the application, to offer accessibility as well as faster workflows for advanced users. This makes our fifth heuristic:

Heuristic H5: Multiple inputs/flexibility/universal usability (Nielsen 7, Shneiderman 2, Weinschenk 4, Connell 18).

- Do you feel there is enough control flexibility and variety to cater to a diverse variety of users, with different experience levels and control wishes?
- How do you feel the implemented joystick rotation factors into this?

Another common principle is that motor load should be minimised. The sixth and final heuristic is therefore:

Heuristic H6: Motor load (Weinschenk 2, Connell 26).

- Is the implemented joystick rotation a good addition to reduce motor load, or do you feel that it rather impacts the proprioceptive intuition of feeling your own body rotate to aid with orientation?
- Are there any instances of problematic motor load that you noticed?

3.4 Materials

Results are gathered using an experiment with a custom-built video player that is used with a VR headset. Materials used in the project are:

- The Unity software development suite to build a custom application;
- The custom application developed for this research. It includes the following in one package:
  - General demographics, 360-degree video-specific and self-evaluation questionnaires;
  - Custom video player and interface with various widgets;
  - Videos to be watched with the video player.
- The VR headset, a Meta Quest 1 and its associated VR handheld controllers.

Additional materials, including implementations made for potential other research directions and follow-up studies can be found in Section C of the Appendix.

3.5 Video player

The experiment takes place entirely within the video player, which is a custom-built Unity application.

3.5.1 *Functionality offered to participants.*

In addition to the spatial visualisation, the video player shows standard widgets, such as play/pause buttons, and ones for temporal control. The reason for this is to make the experience as close to a standard 360-degree video player as possible. The following describes the player and related experience as seen from the participants partaking in the user study.

- Participants start the experiment by answering a few necessary questions in a VR environment such as one about the wrist they wear or would wear a watch around in order to determine the placement of certain widgets. Functionality in the questionnaire is as follows.
  - Participants point their handheld VR controller to select buttons/boxes and press the trigger button to tick boxes.
  - They advance to each next page by pressing the “Next” button at the bottom of each page.
- During video playback they can look around and reorient themselves within the video sphere.
- They can also make rotational jumps in 45-degree increments through use of the joystick on either of the handheld controllers.
- They also have access to temporal controls, such as play/pause, fast forward/reverse and clicking the timeline to jump to different points in time in the video. This is to enable them to pause or return to certain problematic spots for demonstration or verification.
- Additionally, the experts have access to a button (on the visual interface as well as on one of the handheld controllers) to switch between spatial widgets at will.
- After testing the interface with a video and still being in the VR environment, participants are asked some questions about its content and relevant heuristics.

The app features various configurations for playing videos, with UI buttons that can be operated using pointer controls with a handheld VR controller. The configurations developed for the experiment are the following:

- A spherical video screen that surrounds the viewer, for watching 360-degree video in the equirectangular projection;

- A setup for watching 360-degree video in Google's Equi-Angular Cubemap projection[3], suitable for playing 360-degree YouTube videos without modification of the video. The video is projected directly onto Unity's skybox, using a shader reworked from prior work done by *hakanai*[13].

These different configurations are used to ensure that the application can be used to experiment with a wide range of videos, as 360-degree video is available in either of the implemented formats on YouTube: new videos nowadays are uploaded in the equi-angular cubemap format, while legacy videos continue to be available in the equirectangular projection format. Both formats are used in the experiment for the various videos.

A screenshot taken during use of the finalised version of the video player can be seen in Figure 10. It showcases a 360-degree video being played in the EAC format, making use of the implemented control UI—a traditional video UI with timeline and buttons, affixed to the bottom of the user's field of view. One of the controllers specific to the used VR hardware—the Meta Quest in this case—is also seen, shown clicking the play button.

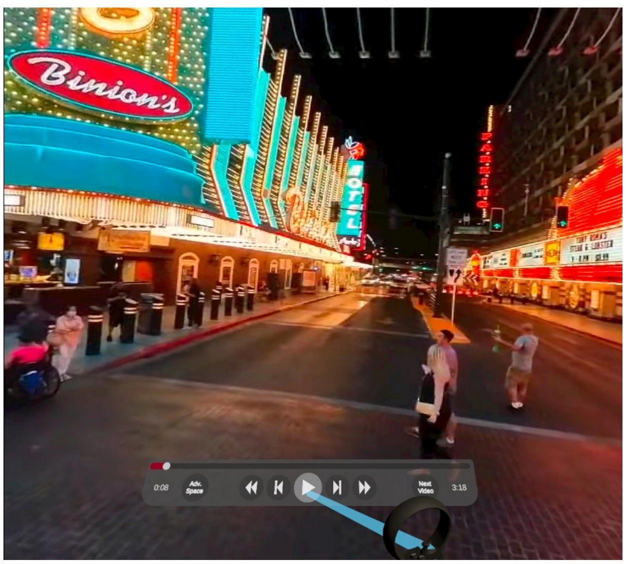

Figure 10: Using the video player on a Meta Quest.

3.6 Dataset (Videos and tasks)

Because the baseline video player has been made to work with the *equirectangular* projection as well as with the *equi-angular cubemap* projection, giving it wide compatibility, it is possible to use videos from a wide variety of sources without requiring any additional conversion.

3.6.1 Videos. Videos selected for use in the experiment are chosen to be conducive to a variety of representative 360-degree video tasks. They contain a variety of things for the user to look at and to draw attention to various different regions of the sphere. In particular, the following types of videos were selected:

- A virtual visit to various colourful locations in Las Vegas. This is the introduction video for the experts to try the different spatial widgets and get their bearings with the system.
- A virtual tour through a 3D-modelled house, including stops in the living room, kitchen and a bedroom.
- A walk down a British shopping street, with many shops to the left and right of the viewer.
- A winding pedestrian walk through Shibuya, Tokyo, past various shop fronts and buildings and down smaller streets.

3.6.2 Video tasks. Typical heuristic analysis calls for the usage of a selection of user personas to steer the evaluation of different aspects of the interface. In the case of this research, it was decided that it is better to instead use a selection of different tasks to outline the user's intent for watching the video. They serve a similar purpose to user personas, but make the intent more explicit.

- House tour: Mental map task–You have seen the layout of the house. Where is the toilet from your current position?
- Parking bike in street: Spatial estimation task–Find a spot in the street so as to minimise the distance between three shops.
- Arcade in city: General orientation and memory task–Give directions to an amusement centre after the fact.

3.7 Measured data

As a heuristic-based expert study, the measured data in this experiment is primarily subjective in nature, encompassing opinions, remarks and advice from the experts, in addition to their answers to questions. Because the experts are wearing a VR headset through the whole heuristic analysis, the choice has been made to orally evaluate the heuristics with them rather than giving them a sheet to write down their comments as they go. In addition, to complement the oral evaluation, a colour-based severity scale is used for any issues encountered. Severity levels are defined as stoplight colours. They are employed to serve as convenient shorthand for the evaluator to use so they do not have to spend too much time describing the issue while using the application, not risking as much distraction and allowing them to list off possible issues more rapidly.

The defined severity colours are as follows:

- Red: show-stopping issue that needs fixing.
- Orange: smaller issue. It gets in the way but is tolerable or only a minor annoyance
- Green: there is no issue. This level is only listed for completeness and would not be remarked on.

3.8 Experiment flow

The participant (the expert) starts by donning the VR headset. They find themselves in the experiment environment and are faced with the experiment introduction and a digital consent form. After this, they fill in a short questionnaire consisting of a set of questions directly relevant to the research (such as experience with VR, 360-degree videos and which wrist they would wear a watch on.).

When the questions are finished, the expert finds themselves in a black scene added specifically for the benefit of the experts while they receive final instructions on the widgets and UI buttons they will be operating, after which the main experiment starts. In this, participants do simple spatial tasks in videos played in the 360-degree VR video player. During these trials the expert can freely toggle the presence and the design of the various UI widgets to visualise the viewer rotation. They are free to make comments while using the experiment application, regarding observations about the widgets, issues, or anything else that may stand out to them. After each video, they are asked a question about the video content they just watched, to guide the evaluation.

After all videos have been watched, the expert is asked questions relating to the heuristics that were selected for this study. During this, they have the freedom to stay in the VR application, allowing them to jump to any of the videos and widget designs to illustrate talking points.

After all questions are answered, some free time is allotted to discuss any potential related topics that came up during testing. Following this, the expert is thanked and leaves the experiment room.

The opening questionnaires are taken in a basic empty grid void. The interface pop-ups following the videos, and the video-related questions are asked while the final frame of the last played video serves as the backdrop, in order to maintain the participant's sense of location and prevent simulator sickness.

Another important thing to note is that because the entire experiment is conducted in VR, participants are offered time to take off the headset and rest at any time if needed. This is to keep the chance of discomfort as low as possible. For the sake of this, they are free to interrupt at any time, not just between videos.

4 EVALUATION

4.1 Participants

A total of three experts were involved as participants in the evaluation. All of them have a similar background to the author of this paper, and all are working on 360-degree video interaction projects, as well, which is why we consider them suitable experts on this topic and this research. Each expert participated individually, over the course of around an hour of evaluation.

4.2 Heuristics and expert opinions

In this subsection each numbered heuristic with its questions is addressed with the summarised experts' opinions and observations.

4.2.1 *Intuitive readability.*

Are the various widget UIs intuitive to read?

The widgets are overall intuitive to read according to experts. The **wristwatch** widget caused slight confusion regarding what its face indicates. Pointing the user's wrist in a direction was expected to change the reading, while in reality, the reading is based on the head rotation. The room for interpretation/confusion that either implementation brings makes it suboptimal in intuitiveness.

Do you feel the compass is a suitable metaphor for directionality, even if its directions may not necessarily match the video content?

The experts felt the compass style of the widgets is suitable. The North-East-South-West (NESW) directions are an instantly recognisable metaphor for directions, while any other design or set of letters would require prior explanation. Furthermore, experts say that in the context of a 360-degree video, it does not really matter if any direction is the true north depicted in that space or not, unless the user is actively intending to use the video as a reference for real-life directions.

Do the different widget UIs give you enough/sufficient information?

The experts found that, generally, they do. However, they provided the caveat that the linear compass felt a bit too narrow and did not display as many direction labels as the expert would have liked in some circumstances. In the current implementation this compass design displays a fixed-angular section of the full 360 degrees that is set to 90 degrees, to approximate the FOV of the Quest headset.

4.2.2 *Perceptual load/clarity/contrast.*

Does the design of any of the widgets get in the way during use? Do they obscure things in the video?

The experts found the **linear compass** to obscure too much of the video sphere area and suggested it be drastically flattened. The **ring compass** was found to be much better, because being located at the user's feet, it will typically not be looked at accidentally. It however does get in the way quite severely when there are points of interest to look at in the lower third of the video sphere. Experts suggested having it fade out when the user keeps their head still looking down for more than a few seconds. The **wristwatch compass** was found to be excellent in staying out of the way, because it resides on the user's wrist and is thus typically held out of view beside the body when not in use.

The experts made several further notes relating to this heuristic. Reading the **wristwatch** or the **ring compass** was noted to lead to disorientation and potential simulator sickness due to needing to shift your gaze back and forth between the video and widget. However, the **ring compass** being fixed in the world was also remarked on as providing a welcome fixed point to focus on to stave off simulator sickness during use that may arise from video content. They compared it to looking at a fixed point in the distance during a boat or car ride to help against sea- or carsickness.

One expert suggested showing just NESW in the widgets was enough, no need to include the in-between compass directions. These were included mostly for the sake of the **linear compass**, because the 90-degree FOV of the Quest would otherwise mean that the user would not be able to see any other direction labels in the visible range of the widget than the one for the cardinal direction they are currently looking in. This same expert actually commented on having this exact issue with that widget in the current version.

4.2.3 *Consistency.*

Do the different widget designs feel consistently designed to each other?

According to the experts, all widgets feel like they are consistently designed, from the same visual language. The **wristwatch compass** was found to be distinct from the other two, being simply a compass in appearance and function. Conversely, for the **ring** and **linear compasses** the experts commented that the **linear compass** feels like a small section cut out of the **ring compass**, saying both widgets give the same feel, since the experience of reading both of them involves reading letters off a horizontally oriented strip.

4.2.4 *Manipulability.*

Do you feel the widgets are intuitively placed, or would you recommend moving them?

While seeing the **linear compass** in the pre-experiment black scene, before it was seen in combination with the player UI during use, experts' first impressions unanimously were that this widget should be moved up to the top of the field of vision to match its placement to that of similar compasses in open-world video games (The Witcher, Fallout, Skyrim, Halo, etc.). After the player UI was introduced in the actual video part of the experiment, however, the experts suggested instead fusing the **linear compass** into the playback controls UI to form one self-contained interface unit. Further details about this suggested redesign are discussed in Section 4.4.

The placement of the **ring compass** felt intuitive to the experts, too, because it being placed around the user's feet physically relates it to the user's orientation within in the real world and video. The on-wrist placement of the **wristwatch compass** was found to be very intuitive and impressive, just like checking a real watch.

4.2.5 Multiple inputs/flexibility/universal usability.

Do you feel there was enough control flexibility and variety to cater to a diverse variety of users, with distinct levels of experience and control wishes?

Experts noted that the inclusion of the comfort turning jump using either of the joysticks is a very welcome feature, specifically for people that do not want to or are unable to physically turn their body. They found, however, that the default comfort turning jump angle is too large at 45 degrees, leading to visual disconnect between viewing angles, which might be too much for inexperienced VR users. Following from this, they offered a suggestion to increase accessibility: include a control element that will allow the user to adjust the turning jump angle to suit their preference. Alternatively, the turning jump angle could be reduced to a fixed angle of 30 degrees, to suit as many users as possible, mirroring the implementation in popular, established VR experiences such as VRChat.

When asked to compare against the YouTube VR method of the user grabbing the video sphere and moving it around themselves to adjust their viewing angle as an alternative implementation, an expert noted that it would present bigger disadvantages than advantages, pointing out the increased motor load of the much bigger motion and the risk of damage or injury that it carries, pointing out that conversely snap turn lowers motor load as much as possible.

4.2.6 Minimise motor load.

The **wristwatch compass** presented motor load issues to the experts. There are two possible ways to read the watch: the user either needs to look down to read the watch, straining their neck with time, or raise their arm to read it, instead straining their arm with time. The **ring compass** had the neck strain issue as well, as pointed out by one of the experts, arising from having to physically crane their neck down and up again repeatedly during the play time of a video, because of its feature of fading in and out depending on headset pitch angle. In contrast to the other two widgets, the **linear compass** was found to be mostly free of motor load issues, as it comes up automatically while the user is already moving their head to look around. One motor load issue remains, though, in the case when a user wants to check the compass when their head is still. Then they have to effectively shake their head back and forth to make the compass appear. This problem is addressed in a suggested redesign where the compass would be made to always be visible, which is discussed in Section 4.4.

The video player supports rotation of the viewpoint through manipulation of any of the thumbsticks on the controller, in addition to turning by physical rotation of the user's body. Is thumbstick rotation a good addition, or do you feel that it rather impacts the proprioceptive intuition?

According to the experts, the snap turning feature is a valuable addition which does not affect the feeling of orientation. One expert suggested this would hold true as long as *two* compasses are provided, one representing the orientation within the video (which would spin along with the user virtually rotating themselves with the thumbsticks) and the other being anchored to the real world (which does not) are provided in the UI. Having these two compasses together might eliminate any possible negative effects of the snap turning feature according to them.

4.3 Widget assessment

Following from assessment by the experts, three distinct cases regarding the use of spatial widgets emerge. Firstly, when merely recreationally watching 360-degree VR videos, experts argue spatial widgets are not very useful for navigation, because using visual points of reference in the video to keep oneself oriented is typically sufficient. Likewise, secondly, widgets may provide only minimal benefit in contexts where users want to actively explore the environment as long as points of reference are available. However, thirdly, in the absence of points of reference, like in flat open areas, forests or very dense cities, a spatial widget can help the user navigate and also help them determine whether they have made a full revolution when exploring the depicted space or searching for things. In a case like this, they argue a spatial widget could be a valuable addition.

Irrespective of the above, all experts recommended for a spatial widget to be implemented into current video players as even when it isn't necessary, in their opinion having a widget present does not negatively impact the usability or immersion of the app and can help people that would prefer to use one. Two out of the three experts independently suggested the implementation of a particular redesign for one of the widgets.

According to one expert, a spatial widget would also be useful for retaining one's grip on their orientation relative to the real world, as this is a problem that they in particular run into while using VR, which could be improved further through a suggested redesign of a widget. Both suggested redesigns are detailed in Section 4.4.

Each expert preferred a different widget design but agreed on many points regarding each. In the following, we summarise the main advantages and disadvantages of each design.

Pros of wristwatch:

- + Out of the way when not in use.
- + Intuitive to summon.

Cons of wristwatch:

- - Motor load issues, both of the neck (looking down) and of the arm (raising wrist).
- - Distracts from video to focus on wrist close to face.

- – Can lead to simulator sickness.

Pros of linear compass:

- + Placement makes sense.
- + Easy to read.
- + Familiar to users from similar designs seen in video games, and fits in with linear design of timeline.
- + Always in view
- + No additional motor load

Cons of linear compass:

- - Obscures more video space than necessary
- - Automatic fading in and out gets annoying and confusing
- - The compass displaying only a small part of the full 360 degrees leads to insufficient visual information on directionality at times.

Pros of ring compass:

- + Very intuitive to read (fast and little mental piecing together) because it is world-located and the user is put right inside it
- + Out of the way during normal use

Cons of ring compass:

- - Motor load issues (neck when looking down)
- - Forces looking away from video content, leading to distraction or missing of critical object
- - Fades in over too long of a distance
- - Gets in the way of looking at things below the user

4.4 Suggested redesigns

A common complaint the experts had about the **linear compass** widget was that it takes up more space in the user's field of view than necessary, while it fading in and out could make reading it difficult. To address these issues, the experts suggested to integrate the linear compass into the main player UI, attaching it to its top and spanning it to fully fill the whole horizontal space of the player UI. That way it can make whatever contribution to navigation, intuition, or real-world localisation that the user may desire, while staying out of the way enough for people who do not make use of it to not find it obstructive. Along with these changes, they suggested the displayed angle be made slightly wider so that more direction labels will visibly fit on it at any time to aid in orientation, perhaps with darker areas towards the sides to signify which directions are outside the user's FOV. A conceptualised mock-up of this redesign has been made and can be seen in Figure 11.

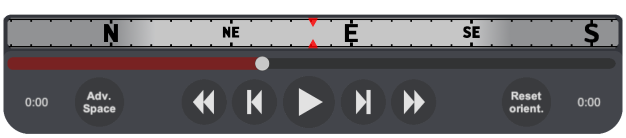

Figure 11: The linear compass redesign, integrated into the main player UI.

Another redesign was put forward by one of the experts to exclusively address the struggle of keeping oneself oriented properly within their play-space and not risk bumping into things or tumbling down a flight of stairs. Their concept could be applied to any of the widgets and consisted of having just a singular arrow present that indicates the initial (or "North") direction of the video to serve as a point of reference within the world. However, after consideration and discussion, they deemed it a potential problem to still not know at a glance exactly which direction a user is facing in some situations, so concluded that it may be best to include the lettered directions, too, after all, bringing their suggestion back in line with the already-implemented widgets.

Given the integration of the described changes into the **linear compass**, the experts suggest that it would be a good fit and the most promising to be implemented into current video players.

5 CONCLUSION

Intuitively, we can argue that spatial widgets are not very useful for navigation when merely recreationally watching 360-degree VR videos and may provide only minimal benefit in contexts where users want to actively explore the environment but sufficient points of reference are present. However, in visually challenging environments with no clear points of reference, we speculated that a spatial widget can aid navigation and help maintain a sense of orientation while exploring. Widgets could also potentially help a user retain their grip on their positioning in the real world and combat simulator sickness.

Our results suggest that these intuitive assumptions are correct, since all three experts confirmed them and provided evidence for the usefulness of such widgets. While each expert personally preferred a different widget, they agreed on various common points regarding them.

The **wristwatch** and **ring compass** widgets, while out of the way and intuitive to use, presented significant issues with motor load if they were to be used over longer periods of time, while also distracting from the video while the user reads them.

Out of the three widget designs, experts found the **linear compass** widget to be the most intuitive and least intrusive in its usage. They felt its drawbacks could be addressed with a redesign and in this hypothesised redesign found it the best suited to adaptation into current video players.

In this work we showed that it is worthwhile to consider adding spatial navigation widgets to VR 360-degree video players. This follows from how even in situations where the widgets may not contribute much, they do not take away from the user experience. Meanwhile, in situations where they would provide benefits, their presence is welcome. The experts saw benefits in the various widgets but no clear solution out of the ones implemented since they all have significant pros and cons. Discussions with the experts revealed a way to refine one of the widgets to be the definitive solution for implementation in video players. Future research into elements of this redesign, and other potential avenues is therefore relevant to engage in. This is discussed in the following section.

6 FUTURE WORK

The study revealed several ways to go and areas to explore for future research in various related fields.

6.1 Linear compass redesign

The experts' suggested redesign of the **linear compass** widget is a compelling option to delve into for future research. Various elements of it could be subjected to experimental research, such as the value of featuring distinct internal and external compasses. Research could be put into whether having an external compass widget would help the user keep their bearings in the real world, or if it would lead to information overload in the UI for the user. Further, an aspect worth experimenting with is whether showing a faded-out area on the widget to represent the directions that may be outside of the field-of-view of the user in the headset is relevant.

After experimentation, the refined widget could then be implemented into a video player for wider deployment, following the experts' suggestion, to verify its added value and usefulness at a larger scale, utilising remote data collection as enabled by the implementation detailed in Appendix section C.

6.2 Volumetric video

A possible reason why spatial widgets are in games but not typically found in VR video is that in traditional videos, spherical or not, the user is merely a spectator, unable to decide where to go. This is in contrast to video games, where the user has full agency in moving around the world, only then making a spatial widget like a compass truly necessary.

A notable middle ground between the two mediums, though, is volumetric video. This type of video came up during discussions with one of the experts as an area where the implementation of a spatial widget would potentially be even more relevant than in spherical video. Volumetric videos capture a true, three-dimensional space over the course of a certain stretch of time and in videos like this, the user would be able to move through and explore the recorded space interactively. This user agency would warrant the presence of navigational aids such as spatial widgets, which would be a relevant avenue of research.

The field of volumetric video is still in its early days, so determining ways to intuitively represent the navigable space to the user and help them keep their bearings is relevant research to pursue.

6.3 Augmented reality

A final avenue for future research that came up during testing, is that while the implemented widget designs were intended just for VR videos, they could be applied to wearable augmented reality devices, like for instance Google Glass, too. In these, rather than helping navigate in videos or virtual spaces, the widgets would aid with real-life navigation, as one of the experts gave a personal anecdote to illustrate. They mentioned they tend to struggle with knowing which way is, for example, north when following directions and would often have to get out their phone to check a map app, which they note is a time-consuming and bothersome process. Having a spatial widget present in their head-mounted AR device would solve this problem without taking any extra actions and within half a second, as it could be accomplished by just glancing over at the spatial widget that is already in your field of view.

For this the widgets would have to be provided with the true north from the real-life location the user is in, but could otherwise function identically to their current VR counterparts. Further implementation details could be determined through further research and subjected to experimentation with various parameters on a selection of different AR devices.

REFERENCES

[1] 2019. *Modular Testbed for KPI Monitoring in Omnidirectional Video Streaming Scenarios.* Zenodo. https://doi.org/10.5281/zenodo.3755819

[2] Dmitry Alexandrovsky, Susanne Putze, Michael Bonfert, Sebastian Höffner, Pitt Michelmann, Dirk Wenig, Rainer Malaka, and Jan David Smeddinck. 2020. Examining Design Choices of Questionnaires in VR User Studies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* 1–21.

[3] Chip Brown. 2017. Bringing pixels front and center in VR video. *Retrieved March* 19 (2017), 2019.

[4] Seunghoon Cha, Jungjin Lee, Seunghwa Jeong, Younghui Kim, and Junyong Noh. 2020. Enhanced Interactive 360° Viewing via Automatic Guidance. *ACM Transactions on Graphics (TOG)* 39, 5 (2020), 1–15.

[5] Iain W Connell. 2000. Full Principles Set. Set of 30 usability evaluation principles compiled from the HCI literature. (2000). http://www0.cs.ucl.ac.uk/staff/i.connell/DocsPDF/PrinciplesSet.pdf

[6] Philip Day. 2016. Whirligig Media Player. (Apr 2016). http://www.whirligig.xyz/

[7] Google Developers. 2017. Designing Screen Interfaces for VR. (May 2017). https://www.youtube.com/watch?v=ES9jArHRFHQ

[8] Tilman Dingler, Kai Kunze, and Benjamin Outram. 2018. Vr reading uis: Assessing text parameters for reading in vr. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems.* 1–6.

[9] Radiant Media Player EURL. 2016. (2016). https://www.radiantmediaplayer.com/360-video.html

[10] Hannes Fassold and Barnabas Takacs. 2019. Towards Automatic Cinematography and Annotation for 360° Video. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video.* 157–166.

[11] Roberto Iraja Tavares Da Costa Filho, Marcelo Caggiani Luizelli, Stefano Petrangeli, Maria Torres Vega, Jeroen Van der Hooft, Tim Wauters, Filip De Turck, and Luciano Paschoal Gaspary. 2019. Dissecting the Performance of VR Video Streaming through the VR-EXP Experimentation Platform. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 15, 4 (2019), 1–23.

[12] Jill Gerhardt-Powals. 1996. Cognitive engineering principles for enhancing human-computer performance. *International Journal of Human–Computer Interaction* 8, 2 (1996), 189–211. https://doi.org/10.1080/10447319609526147 arXiv:https://doi.org/10.1080/10447319609526147

[13] hakanai. 2018. EACSkyboxShader: Panoramic skybox shader for Unity. (May 2018). https://github.com/hakanai/EACSkyboxShader

[14] Chris Hughes and Mario Montagud. 2020. Accessibility in 360-degree video players. *arXiv preprint arXiv:2005.03373* (2020).

[15] Yung-Ta Lin, Yi-Chi Liao, Shan-Yuan Teng, Yi-Ju Chung, Liwei Chan, and Bing-Yu Chen. 2017. Outside-in: Visualizing out-of-sight regions-of-interest in a 360 video using spatial picture-in-picture previews. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology.* 255–265.

[16] InfoMediji LTD. 2018. DeoVR. (2018). https://deovr.com/

[17] M Montagud, I Fraile, E Meyerson, M Genís, and S Fernández. 2019. ImAc Player: Enabling a Personalized Consumption of Accessible Immersive Content. *ACM TVX* (2019).

[18] Afshin Taghavi Nasrabadi, Anahita Mahzari, Joseph D Beshay, and Ravi Prakash. 2017. Adaptive 360-degree video streaming using scalable video coding. In *Proceedings of the 25th ACM international conference on Multimedia.* 1689–1697.

[19] Afshin Taghavi Nasrabadi, Aliehsan Samiei, Anahita Mahzari, Ryan P McMahan, Ravi Prakash, Mylene CQ Farias, and Marcelo M Carvalho. 2019. A Taxonomy and Dataset for 360○ Videos. ACM, 273–278. https://github.com/acmmmsys/2019-360dataset

[20] Cuong Nguyen, Stephen DiVerdi, Aaron Hertzmann, and Feng Liu. 2017. CollaVR: Collaborative in-headset review for VR video. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology.* 267–277.

[21] Cuong Nguyen, Stephen DiVerdi, Aaron Hertzmann, and Feng Liu. 2017. Vremiere: in-headset virtual reality video editing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.* 5428–5438.

[22] Cuong Nguyen, Stephen DiVerdi, Aaron Hertzmann, and Feng Liu. 2018. Depth conflict reduction for stereo vr video interfaces. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* 1–9.

[23] Jakob Nielsen. 2005. Ten usability heuristics. (2005).

[24] Antoni Oliver, Javier del Molino, and Antoni Bibiloni Coll. 2020. Palma360: A 360-degree Web-based Multi-device Hypervideo. In *ACM International Conference on Interactive Media Experiences.* 165–169.

[25] OmniVirt. 2018. OmniVirt VR Player. Meet The Best HTML5 VR Player. (2018). https://www.omnivirt.com/vr-player/

[26] Yun Suen Pai, Benjamin I Outram, Benjamin Tag, Megumi Isogai, Daisuke Ochi, and Kai Kunze. 2017. GazeSphere: navigating 360-degree-video environments in VR using head rotation and eye gaze. In *ACM SIGGRAPH 2017 Posters.* 1–2.

[27] Susanne Putze, Dmitry Alexandrovsky, Felix Putze, Sebastian Höffner, Jan David Smeddinck, and Rainer Malaka. 2020. Breaking The Experience: Effects of Questionnaires in VR User Studies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* 1–15.

[28] Ben Shneiderman. 1986. Eight golden rules of interface design. *Disponible en* 172 (1986).

[29] Marco Speicher, Christoph Rosenberg, Donald Degraen, Florian Daiber, and Antonio Krúger. 2019. Exploring visual guidance in 360-degree videos. In *Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video*. 1–12.

[30] Jasper Stephenson. 2018. (Mar 2018). https://medium.com/@jasper.stephenson/a-ux-analysis-of-first-person-shooter-damage-indicators-59ac9d41caf8

[31] Thomas Stockhammer. 2011. Dynamic adaptive streaming over HTTP– standards and design principles. In *Proceedings of the second annual ACM conference on Multimedia systems*. 133–144.

[32] Yu-Chuan Su and Kristen Grauman. 2017. Making 360 video watchable in 2d: Learning videography for click free viewing. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 1368–1376.

[33] Vildan Tanriverdi and Robert JK Jacob. 2001. VRID: a design model and methodology for developing virtual reality interfaces. In *Proceedings of the ACM symposium on Virtual reality software and technology*. 175–182.

[34] Source Technology. 2017. SKYBOX VR. (Oct 2017). https://skybox.xyz/en/

[35] Balasaravanan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Björn Hartmann. 2019. TutoriVR: A Video-Based Tutorial System for Design Applications in Virtual Reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[36] Susan Weinschenk and Dean T Barker. 2000. *Designing effective speech interfaces*. John Wiley & Sons, Inc.

APPENDIX

A VIDEO PLAYER IMPLEMENTATION

For this research, a 360-degree video player was implemented that should also enable follow-up projects as well as easy extension and adjustment for future research in this area. The basic VR video player as described in Section 3.5 was developed in a first phase to verify the feasibility of the practical aspects needed for this research and as a base to build the evaluation framework upon during the second phase of the research, which was originally supposed to use remote evaluations. To accommodate the remote design of the study, the basic version was specified to need to run on multiple platforms and VR hardware brands without additional implementation work and to be able to play 360-degree videos in various projections, such as the equirectangular projection (ERP) and (equi-angular) cubemap (EAC) projection.

These two main features were realised using the Unity platform, and the structure of the UI was made open-ended and easily extendable with various UIs and control widgets. An HTML server and necessary networking code in the experiment application was also set up to allow for automatic data and result collection in a flexible way, helping simplify the experiment experience for participants as much as possible. Although the remote evaluation-focused elements of the implementation were ultimately not used, the implemented framework and features will allow for easy usage or adjustment for follow-up projects.

B LITERATURE SUMMARY

To educate our research question and find the state of the art, literature research was conducted. The research that we conducted for this subject is twofold and not entirely limited to just scientific research. This is because little formal scientific research has yet gone into the exact research field of UI design for VR video players. Therefore, beyond scientific papers, a significant amount of insight has to be gained from existing 360-degree video players in VR and on the desktop.

Because of this, the literature summary is split into two major parts: an evaluation of scientific papers, and an evaluation of the various existing 360-degree video players that were found.

B.1 Summary of found papers

Found literature has been divided into six broad categories. These include research into interfaces and control for regular video, for VR video, for general interfaces in VR, research into making 360-degree video more browsable in desktop as well as VR interfaces, solutions for online video streaming, and research pertaining to relevant evaluation techniques. Paper categories and separate papers that were not treated in the main article are gathered here.

B.1.1 Video interface and control design. This is the main area of related work. A work that was not mentioned in the main article, Oliver *et al.*'s work, *Palma360*[24] stands apart from the rest as it tackles hypervideo: a network of videos that is traversed by the viewer, with various branching points throughout that lead into other videos. Interaction methods that select video branches are dependent on the used device: desktop computer with mouse, mobile devices by tapping, mobile VR HMDs by lingering their gaze on video hyperlink spots during a *dwell time* and desktop VR HMDs by pointing a laser line from the handheld controller at the hyperlink and pulling the trigger. There is no temporal control while playing video, the only available control is over the hyperlink splits. Their analysis method makes use of the *Experience API* standard evaluation, traditionally used for e-learning applications, which automatically collects user movement data during VR use.

Finally, Nguyen *et al.*'s Vremiere[21] is an entire 360-degree video editing suite that is used while in VR. While its UI makes use of a traditional video timeline for editing, it does feature various control widgets with helpful features unique to VR video, and automatic conveniences such as picture-in-picture when previewing a little planet (stereographic) projection or vignetting to reduce simulation sickness when previewing a fast-moving scene. Something to note is that all control is done while seated at a desk. The head is rotated to look around the VR scene, but a standard mouse and keyboard are used to do all editing.

B.1.2 General VR interfaces. These papers all deal with various considerations or determine guidelines when developing general interfaces in VR, a considerable part of the development that went into the practical implementation of this project. The papers include guidelines for making comfortably readable text in VR[8], a solution to reduce depth conflicts between player UI and video content in stereoscopic videos[22] and two particularly relevant papers by Putze *et al.*[27] and Alexandrovsky *et al.*[2] that demonstrate that to avoid biases in questionnaire results for VR experiments the questionnaires need to be taken in the same VR environment as the experiment, and evaluate the practicality of various interface designs for in-VR questionnaires, respectively. The findings from these two papers proved indispensable when designing the questionnaires to use with the experiment.

An early work in the field of VR interaction is Tanriverdi *et al.*'s VRID. In this work they seek to specify a framework for designing interactions in VR[33]. This framework, however, is principally concerned with interactions with physically simulated virtual objects, rather than simple UI interactions.

B.1.3 Solutions for online video streaming. The seminal video streaming work, DASH[31]—Dynamic Adaptive Streaming over HTTP—has been gratefully adopted as the main driving force behind making streaming video with such high bandwidth requirements as spherical VR video and even stereoscopic spherical VR video with good picture quality a possibility. Traditional solutions for VR video using DASH involve predicting the viewer's viewport movements and buffering high-quality video only for that region, to drastically reduce the amount of needed bandwidth for smooth playback.

Much of the work done in 360-degree video players in VR also hopes to further development of effective online video streaming solutions for 360-degree video, such as Nasrabadi *et al.*, who proposed a new streaming method that is tile-based like the rest, but involves scalable video coding to reduce dependence on viewport prediction and effectively reduce buffering[18], or by proposing standardised evaluation systems for the quality of experience when watching streamed 360-degree video[11][1].

B.2 Summary practical implementations

B.2.1 Desktop compass players. *OmniVirt*[25] and *Facebook* both have 360-degree video players that have traditional timeline design, but feature a small radar near the upper corner of vision to indicate current orientation relative to the centre, which can be clicked to reset orientation. OmniVirt offers access to their player source code to anyone for development purposes, including a Unity project, however, in order to upload your own videos to play an account is required, registrations for which have been limited to just enterprise users. They are both desktop-based players. The Facebook player is not accessible in VR to verify its VR functionality, but the OmniVirt player can be opened in VR from a desktop computer. However, doing so does not move the existing UI, which is placed at the very edge of the screen, meaning both the timeline and the radar are outside the field of view in VR and thus unusable. *Radiant Media player*[9] is a desktop player as well and has simple controls for changing orientation, zooming and re-centring in its interface, but no radar.

B.2.2 Spatial control in VR. The *YouTube VR* app has a traditional timeline and temporal controls are only available through pointing and clicking. There is no visual representation of spatial domain in UI. However, the unique point of this app is that manipulation of orientation is available by holding the grip button and moving the controller. While doing this with spherical content the FOV is reduced to a small PIP over a black void, presumably to reduce simulation sickness. What this looks like can be seen in Figure 12. The same can be done with non-VR content, which then moves the virtual screen to where the user drags it. For non-spherical content there is a UI button to re-centre the view, which is not available for spherical content.

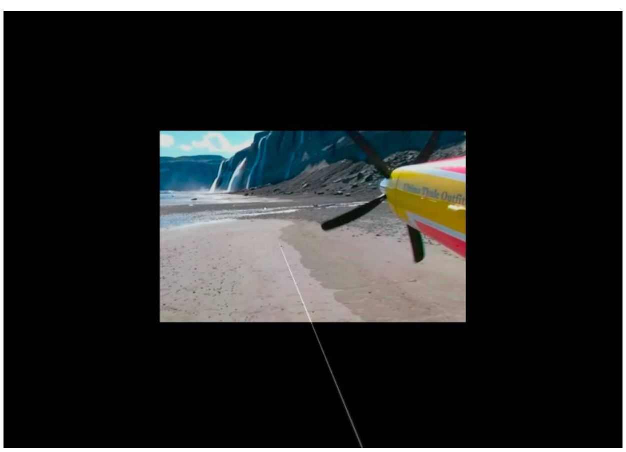

Figure 12: Dragging the viewport in YouTube VR.

The free player *DeoVR Video Player*[16] is remarkable in two aspects. The first is that it features many controls for tweaking the picture, as visible in Figure 3 in the main article. The controls, like the buttons near the traditional timeline, can be interacted with through point-and-click actions, but this player features direct control over various elements for manipulating them, which is the second remarkable aspect of it. All controls are done with only the currently active controller. The trigger button shows the UI and can be held to drag it to any desired location. Pressing the primary face button plays and pauses the video, without bringing up the UI. The left and right directions on the joystick jump forward and back through the video in steps of 10 seconds, while the up and down directions tilt the viewpoint up and down. The grip button acts as a modifier button. While holding it, the primary button and joystick directions get different functions. While held, the primary button now mutes and unmutes the device audio, the left and right joystick directions now zoom the viewpoint in and out, and the up and down directions now navigate to the previous and next videos in the folder or playlist. Controls over the orientation are available in the adjustments panel, but a somewhat disappointing aspect is that despite all the button controls and shortcuts, there is no direct button control for it. All made adjustments, including the orientation, can be returned to their original values with a button in the UI.

The video player in the app *Within*, like DeoVR, allows pausing and resuming playback with a simple press of the primary face button. Unlike DeoVR, however, Within *does* allow direct control over the orientation by using the joystick of the currently active controller, with which rotation is continuous rather than being done in stepwise increments, which might induce simulation sickness. There are no other adjustments to make. Neither DeoVR nor Within feature an orientation visualisation. The joystick is the only way to change the orientation in this app and there is no orientation reset button.

B.2.3 Other players. A desktop VR video player application that is flexible and customisable is *Whirligig*[6]. This video player was unfortunately outside the scope of evaluation as we only had access to an Meta Quest standalone Android-based headset during this research. Despite that, though, its various adjustment settings (similar to DeoVR) and its support for a very wide range of controllers, including mouse, keyboard, traditional gamepads and handheld VR controllers, may make it valuable to look into if an opportunity presents itself. The SKYBOX VR[34] player, too, could not be evaluated in depth, but for want of a free trial version instead. This player has a traditional timeline design and no spatial indicators in UI, as seen in Figure 13. The presence of button or joystick controls is unknown. User reviews suggest there are only pointer controls.

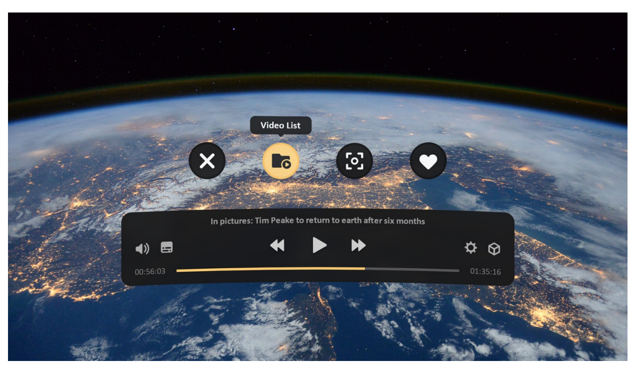

Figure 13: The interface of SKYBOX VR.

B.2.4 Spatial indicators from video games. The two main spatial indicators seen in video games, the radar/minimap and the linear compass have been discussed in the main article. A third, equally prominent but less noticeable spatial indicator often occurs in the genre of first-person shooter (FPS) games.

This indicator is a circle around the middle of the player's viewpoint that appears for a split second whenever the player character takes damage from an in-game source, indicating the direction that the damage came from. A symbol, often a line or arrow, appears at a certain location around the perimeter of the circle for each source of damage. A symbol at the top of the circle corresponds to the damage coming from the front, a symbol at the bottom of the circle corresponds to damage from behind, and so forth. This visualisation in action in *Overwatch* is shown in Figure 14. As the player responds to the damage and turns their character around to face the source, the lines rotate around the circle accordingly, to match up with the player's current viewpoint. The combination of the brief flash of visibility and the directional nature of the symbols around the circle leads to a subliminal effect, where after some practice with the system, it allows for very quick evaluation of and response to damage.

Figure 14: Taking damage from various directions in Overwatch.

A non-scientific evaluation of the various ways to indicate damage, its direction and its strength in video games and the effect they have on the user experience is done by Jasper Stephenson on his Medium blog[30]. The blog post discusses the above indicator, among others, including a 3D version of it to indicate damage coming from multiple elevations, and even a simple flash upon damage.

Functionally this indicator is very similar to the radar/minimap, though it is distinguished by the fact that it always appears around the middle of the screen and is only visible for a very brief period at a time. It is worth mentioning for its unique usage scenario, but in the context of video playback, a situation divorced from urgency in responding to momentary impulses, it is hard to find a use for it, for if it were displayed over a longer period it would start obscuring significant portions of the middle of the user's vision.

Another visualisation similar to the radar/minimap can be found in *Assassin's Creed: Syndicate*. This spatial visualisation manifests itself as a circle in the horizontal plane around the player character, that depicts the spatial relationship between the player and enemy characters. The visualisation can show two types of symbols. If the enemies have not noticed the player, then sound waves with amplitude indicating their awareness of the player character are shown in their direction. If they have noticed the player character, the circle shows arrows instead, in white, yellow or red, depending on their alertness level. The symbol indication system is three-dimensional: it shows general direction through symbol position around the circle as well as relative elevation through the symbol being above, on or below the middle circle. An example can be seen in Figure 15. The indicator being physically displayed in world space around the player character may be an interesting avenue to explore in 360-degree video interfaces, though this is different enough on a fundamental level to not be comparable to other spatial and temporal designs.

Figure 15: A scene in AC:Syndicate with four unaware enemies around the player at a lower elevation. One of the indicated enemies can be seen on the right, highlighted in yellow. Also visible is the in-game minimap display in the lower left.

C ORIGINAL STUDY, EXPERIMENT DESIGN AND EVALUATION PLAN

The original plan for this thesis was to do a more performance-related study focused on quantitative measures. For reasons beyond our control, this had to be replaced with the more qualitative study presented in the scientific paper. In the following, the original plans are described. In addition to the different focus, the original study was also intended to be done remotely and self-supervised by subjects in order to reach more participants and cope with the at that time still ongoing COVID restrictions. Although ultimately unused, the study design as well as the related implementations can be of high value for future follow-up work, such as what is described in Section 6 of the scientific paper.

C.1 Experiment Design

Participants undergo the following steps during the experiment:

- Motivation is explained: going on a trip to a city with strongly differing areas (Tokyo).
- Participants are introduced to the experiment with the narrative that they are travelling to Tokyo soon and are exploring possible locations to visit when there.
- To decide on which area to look for a place to stay, the participant is watching a guide video that takes them to all the areas. They get x number of seconds to observe each scene carefully and internalise what is there.
- This hopefully stimulates them to carefully look around in the video and retain as much information as possible.
- The video automatically moves on to each next location when time expires.

- After all locations have been seen, participants are asked to allot a certain amount of time out of their y-hour schedule to each location based on how interesting the things they observed they found.
- At the end, participants are presented with images of certain objects and are asked to remember if these images were taken from the actual videos or not, testing how well they remembered their environments.
- As an additional statistic, while the experiment goes on the amount of looking around they do (total degrees of HMD rotation) is measured and logged.
- Finally, participants are asked several comfort and self-assessment questions to determine differences in comfort levels and feelings like being lost between widget conditions.

C.2 Randomisation

The test condition is randomly selected. For ease of use and deployment: one version of the application that generates a unique 32-bit identifier per participant. For reproducibility, a participant identifier is stored and used as random seed, to generate the test condition and order of videos for them. The design of this procedure is of high relevance to allow a thorough evaluation of the data while maintaining proper anonymity and privacy for the test subjects. While it was not needed for the ultimately executed heuristic evaluation, the implemented approach will be of great value for future remote studies.

C.3 Materials

Results will be gathered using an experiment with a custom-built video player that is used with a VR headset. Given the current pandemic, the experiment will be conducted remotely. Because the number of people with VR sets is still relatively limited, limiting the participants to volunteers from university and their circles is expected to lead to insufficient participants. Therefore, in the interest of finding as many participants and generating as much data as possible, participants will be gathered through postings on various places on the internet, such as Reddit. To ensure the highest possible participant retention during the experiment, care must be taken to ensure that the experiment does not go on for too long and that videos that might induce simulator sickness are kept to a minimum.

C.4 Materials used in the project

- Unity software development suite to build the custom application;
- Custom application. Includes the following in one package:
  - General demographics, 360-specific and self-evaluation questionnaires;
  - Custom video player and interface;
  - Video to be watched with player.
- The VR headset of the remote participant;
- The computer or standalone VR device of the remote participant;
- A web server with FTP access for the remote collection of experiment results;
  - Uses a temporary account with access rights limited to just one folder, to ensure the security of both the collected experiment data and the server files stored outside of the experiment folder.
  - Files are uploaded directly from memory, to require no local storage permissions on the target devices.
  - Once a file is created on the server, it can be directly appended to from the experiment application without requiring it to be downloaded temporarily.
  - File structure: dateTime and a unique 32-bit identifier on the first line, then each following line has simple numbers to indicate the index of the cell that was ticked per questionnaire section.
- Spreadsheet software for analysis of results.

For the purpose of making the remote experiment work as well as possible, the used video player is a custom Unity application that fulfils two main criteria. The first is that it has support for all common VR headsets and handheld controllers to ensure that it can be deployed on as many different hardware configurations as possible, maximising the potential number of participants. As such, it has been designed to and has been verified that the application can be used on multiple platforms, using multiple different VR headsets and controllers with no additional implementation work.

The second is that the application has code that automatically records and sends participant performance data after every experiment trial, for statistical evaluation.

As a further foundation for the experiment part of the research, the UI buttons can be either positioned in the world within a floating canvas at a fixed position, or be attached to the user's viewpoint to ensure they are always visible. The distance and scaling of UI elements is based on Google's Distance-independent millimeters research[7].

C.5 Functionality offered to participants

In the original experiment design, only spatial visualisation was intended to be tested, as temporal visualisation is left out and even temporal control is removed from the user in order to not distract from the spatial challenge. However, upon turning the research into a heuristic evaluation with experts, increased control was deemed necessary, leading to re-enabling the full player controls to aid in evaluation in full depth. More aspects of the original design follow below.

- Participants start the experiment by answering questions in a VR environment.
  - They point their handheld VR controller to select buttons/boxes on the virtual form and press the trigger button to tick boxes.
  - They advance to each next page by pressing the "Next" button at the bottom of each page.
- During video playback they can only look around and reorient themselves within the video sphere.
  - A visualisation of orientation is visible in the scene, the type of which is determined by the participant's random ID.
  - No temporal controls.
  - Controller/on-screen button to reset orientation.

  – Rotational jumps in 45-degree increments allowed through use of joystick.
- After video: the participant answers questions about the video content in a questionnaire format.

C.6 Further detailing of test conditions

This subsection contains further notes and considerations on the various spatial widgets, including a fourth widget that was a predecessor to the final **ring compass**.

- **Linear compass** attached to viewpoint.
  – Requires no active participation to read, allowing for raised general directional awareness through osmosis.
  – Possible drawback: it being on-screen at all times might be too obtrusive and require fade-in/out logic.
- **Nadir compass** at feet of viewer.
  – Is static. Allows for natural, hands-free viewing, while staying out of the user's view when not expressly looking down.
  – Possible drawback: requires looking down quite sharply, possibly inhibiting awareness of surroundings while using it.
- Semi-transparent, cylindrical, **ring compass** surrounding viewer. This is a revised iteration on the nadir compass.
  – Is static. Allows for natural, hands-free viewing, while staying out of the user's view when they are not expressly looking down.
  – Does not require user to look down as far as nadir compass, inhibiting awareness of video less.
  – Also allows for easier, lower-threshold (possibly subconscious) checking of current direction.
  – Possible drawback: May obstruct view a bit more than nadir compass due to being situated higher in field of view.
- **Watch-like compass** attached to non-dominant wrist.
  – Possibly a good way to make reading the radar/compass mentioned above intuitive and non-disruptive
  – No need to press buttons, simply raise arm.
  – Natural, intuitive motion.
  – Stays out of view when not desired.
  – Wrist to put it around is determined in questionnaire question before experiment.
  – Possible drawback: requires conscious effort to check direction, likely leading to less general directional awareness. This is because the user needs to decide to check the orientation, and has no chance to just happen to see it at other times.

C.7 Measured data

- Experiment structure: between participants.
- Structure motivated by fact that memory-based evaluation can only be done once, because after one trial, the participant will know that it is coming and will pay attention to things more consciously.

Because the experiment will involve strangers on the internet, care must be taken that the second point, the automatic collection of data, is emphasised and an exact specification of all the collected data is provided to them in the consent form. The data that will be collected during the experiment is, at the current time, expected to be limited to:

- The randomly generated user ID that is used to store the experiment data with on the FTP server;
- The names of the VR hardware that was used (headset and controllers might be from separate products/brands and are thus recorded separately);
- The video IDs associated with the taken experiment trials and the order that they were done in;
- The time in milliseconds that it took the participant to answer each trial;
- The index of the frame that the participant indicated as their answer in each trial;
- The number of times that playback went past the correct answer in each trial;
- Non-identifying demographic data such as age and degree of experience with various VR uses like video viewing or game playing.

Because participants are essentially limited to people who own VR sets, the average age might be biased towards the younger end. This must be considered in the analysis of the data.

No names, IP addresses or geographical data will be recorded.

Measured data is sent to and stored on a remote FTP server with a secure user system, which is where the experiment application itself is also hosted. Data measured after each questionnaire section and experiment trial is stored right away so that even if a participant stops partway through, the results that they did provide can contribute to the statistics. Experiment trials are in a randomised order so participants stopping early will not bias the results towards being only from a certain subset of videos.

C.8 Dataset

Videos used in the experiment should display a clear object of interest or occurrence in the video that can be the target of a task, such as "Navigate to the time in the video where you can see a man holding up a pink sign". A balanced set of tasks is desired, to ensure that every test condition has at least one task of each kind, so that there is no need to repeat videos between conditions and learning effects can be avoided. Additionally, videos should contain various things to look at and draw attention in various places of the sphere, in order to keep participants looking around and observing much of the sphere surface throughout the video.

A candidate dataset is Nasrabadi *et al.*'s dataset[19]. This set contains a large variety of 360-degree videos, divided into categories based on objective factors such as type of camera movement and number of objects of interest visible in the video, which is a good fit four our purposes. These videos are all available in the equirectangular projection and EAC, straight from YouTube.

C.9 Experiment flow

Users will start by donning their VR headset. The entire experiment will be in VR. They will find themselves in the experiment environment and are faced with the experiment introduction and consent form. After this they fill in a general demographics questionnaire, all on separate pages, followed by a set of questions directly relevant to the research (experience with VR, 360 videos, etc.).

When the questionnaires have been finished the main experiment starts. In this, users will do simple navigation tasks in videos played in the 360-degree VR video player. During these experiments the independent variable will be the presence and the design of various UI amenities to visualise and/or modify the viewer rotation. The dependent variable is the user's performance under these conditions, expressed in time taken and absolute distance from correct answer (in video frames), or the ability to correctly answer questions about the video content they just watched.

After the main experiment a small questionnaire about the participant's experience and opinions is filled in. Finally, the participant is thanked, takes off their VR headset and leaves the experiment webpage.

Something to emphasise is that the questionnaire and post-experiment recap/opinion all take place in the same VR environment as the main experiment section, in order to prevent the rift in user experience that can bias results, as found by Putze *et al.*[27]. The research done by Alexandrovsky *et al.*[2] further sheds light onto which design principles and interaction methods to follow for optimal 'usefulness', as they put it, when taking questionnaires in VR.

Another important thing to note, is that because the entire experiment is conducted in VR, participants need to be offered time to take off the headset and rest at any time. This is to keep the chance of discomfort as low as possible. Therefore, after each experiment trial, a press on a button labelled "I am ready for the next video" will be required before the next experiment trial is started.