Eye Movements in VR

Walter F. Bischof, Nicola C. Anderson and Alan Kingstone

W. F. Bischof (✉) · N. C. Anderson · A. Kingstone

Department of Psychology, University of British Columbia, Vancouver, Canada

e-mail: wfb@ualberta.ca

N. C. Anderson

Cognitive3D, Vancouver, Canada

Contents

1 Introduction and learning objectives

2 Historical annotations

3 Fundamentals

- 3.1 Hardware
  - 3.1.1 Presentation hardware
  - 3.1.2 Eye tracking hardware
  - 3.1.3 Head and body tracking hardware
- 3.2 Eye Tracking
  - 3.2.1 Fixations and saccades
  - 3.2.2 Calibration
  - 3.2.3 Reference Frame
  - 3.2.4 Data quality
  - 3.2.5 Visualization
- 3.3 Head Tracking
- 3.4 Data Processing
  - 3.4.1 Software
  - 3.4.2 Synchronization of eye movements with other sensors
- 3.5 Summary

4 Panoramic virtual environments

- 4.1 Panoramic analysis
- 4.2 Basic terms
- 4.3 Eyes-in-head analysis
- 4.4 Head analysis
- 4.5 Eyes-in-space analysis
- 4.6 Summary

5 3D virtual environments

- 5.1 Eyes-in-head analysis
- 5.2 Head and eyes-in-space analysis
- 5.3 Visualization of gaze in 3D VR environments
- 5.4 Analysis of eyes-in-space in 3D VR environments
- 5.5 Summary

6 Applications

6.1 Attention research
6.2 Clinical applications
6.3 Consumer research
6.4 Architecture
6.5 Summary
7 General Summary
8 Suggested Readings
9 Questions
References

Abstract

This chapter introduces the use of eye tracking in virtual reality (VR) environments. We begin by reviewing VR hardware and software, as well as technologies for tracking the eyes, head, and body, followed by a discussion of key concepts underlying eye tracking in VR. We then examine a restricted category of VR -- panoramic virtual environments -- and present analytical methods and representative results. This introduction is extended to general 3D environments where observers can move freely. We compare analytical techniques and results across these environments, highlight the limitations of methods developed for panoramic settings when applied to general 3D worlds, and introduce novel visualization techniques for these contexts. We conclude by highlighting several applications of VR eye tracking: attention research, clinical practice, consumer research, and architecture.

Abbreviations

AOI	Area of interest
AR	Augmented reality
HMD	Head-mounted display
VR	Virtual reality

1 Introduction and Learning Objectives

Virtual Reality (VR) creates an immersive visual environment for users, which is usually presented with a head-mounted display (HMD) and sensors to track the orientation of the HMD. Popular VR applications include games, health, security, behavioral studies and more. Related to VR is Augmented Reality (AR), which integrates digital overlays with the real world through a display in order to seamlessly integrate projected computer graphics with the perceived physical world. Here, applications include education, computer-assisted surgery and more (see Arnaldi et al., 2018).

Tracking the eyes in VR and AR permits one to determine where users are looking, thus providing several key advantages over systems that do not include eye tracking. For example, it can improve performance through foveated rendering, improve user interaction (Pai et al., 2019), and enhance realism by aligning the presented virtual world to the user's gaze. It does, however, require substantial computing power so that head movements can be compensated with low latency to improve comfort and reduce motion sickness. In research applications, tracking the eyes in VR provides several advantages over traditional computer-based or mobile eye tracking research that we touch on throughout this chapter. Notably, VR eye tracking allows simultaneous tracking of the eyes, head and body movements with respect to a common reference frame, thus allowing one to determine the relative contributions of the different movement systems in attention.

Much of what we know about visual attention and eye movement control is derived from studies that require people to look at images presented on a computer monitor while their head is restrained. There is, however, growing recognition that eye movements measured in the lab when an observer's head movements are discouraged or restrained are not representative of how people move their eyes in everyday life when their head is free to move (e.g., Backhaus et al., 2020; Foulsham & Kingstone, 2017; Hooge et al., 2019; Kingstone et al., 2008; Land & Hayhoe, 2001; Pelz et al., 2001; Risko et al., 2016). For example, Foulsham et al. (2011) asked participants to watch first-person video clips of someone walking across campus. While there was some bias to look in the center of the video, their gaze (i.e., the direction of their eyes) was spread over the whole scene, looking at objects and the people that the walker encountered. When the same participants physically walked across campus with a mobile eye tracker, they often focused on the path, and the eyes remained relatively centered in the visual field as defined by their head orientation (Foulsham & Kingstone, 2017). In other words, when the head was free to move, people tended to move their head in order to redirect their gaze to objects and people. In the lab, Solman and colleagues (2017) have shown that when participants are required to look at a scene through an asymmetric window that is yoked to their eyes, eye movements target regions within the window. However, when the window is yoked to an observer's head movements, the head moves to reveal new information outside the window, presumably so that the eyes can then examine visual information within the new window. These studies suggest that when people are allowed to move their head, they do so, and that the head acts to reveal new information for the eyes to exploit.

In the analysis of head movements, we have learned that the relative timing of eye and head movements may suggest whether attentional selection is reflexive or volitional (Doshi & Trivedi, 2012; Freedman, 2008; Zangemeister & Stark, 1982). For instance, when the eyes move before the head, these are unplanned, reflexive movements (usually to a suddenly presented stimulus such as a flash of lightning) and involve shifts of less than 45° (Barnes, 1979; Flindall et al., 2021). When the head leads the eyes, however, these are thought to be large, planned, purposeful movements, often to a known target location. These conclusions are mainly based on experiments where participants respond to simple light displays or targets on a screen, but it has also been shown in more naturalistic settings, where, for example, when we prepare to cross a street or when we prepare a lane change while driving (Doshi & Trivedi, 2012).

When people are asked to view 360° panoramic scenes in VR, the attention system must coordinate eye movements with head and body movements to explore the full range of the visual field. When looking at 360° scenes, observers spend the majority of the time exploring along the horizon (Rai et al., 2017; Sitzmann et al., 2018), using their head and other body movements to extend the field of view for the eyes (Bischof et al., 2020). This converges with other evidence suggesting that the head acts in service of the eyes to extend the range of possible viewable locations (Solman et al., 2017; Solman & Kingstone, 2014).

On the other hand, studies in VR have also taught us that the eyes, head, and body may move in ways that diverge from what we might expect. When observers are asked to search 3D environments, the effective field of view, or visual span, is much larger than reported in studies using smaller images on a computer monitor (David et al., 2021). In a study where observers viewed large, flat (i.e., non-panoramic) landscape and fractal scenes, it was found that the head and eyes responded differently to scene rotations, where the eyes (e.g., saccade directions) seemed to be more sensitive to scene rotation than the head (e.g., head shift directions, Anderson et al., 2020), with the eyes tending to move along the horizon more than the head. In other work with panoramic scenes, it has been shown that the head is less affected by the masking of central or peripheral scene information than the eyes (David et al., 2022). In addition, the extent to which participants move their head (and body) varies considerably between individuals (Anderson et al., 2020; Bischof et al., 2020; Jacobs et al., 2020). This variation mirrors earlier observations by researchers who looked at the extent that observers prefer to recruit their head in response to peripheral target acquisition – resulting in some observers being dubbed "movers" and others "non-movers" (e.g., Delreux et al., 1991; Fuller, 1992a, 1992b; Goldring et al., 1996; Stahl, 2001). Taken together, this research provides varying degrees of evidence that head and eye movements may diverge in their control strategies, leading researchers to speculate that the head may be under more deliberate, cognitive control (David et al., 2022) or utilize different spatial reference frames (Anderson et al., 2020). Importantly, these and other studies in VR have uncovered interesting findings and generated novel questions about the complex dynamics between eye, head, and other movements in fields of view that extend beyond the standard computer monitor.

With the information provided in this chapter, readers will be able to understand

- The basic hardware and software used for analyzing eye movements in VR
- The fundamental eye tracking concepts relevant to VR
- The methods used for analyzing panoramic virtual environments
- The methods used for analyzing general 3D virtual environments
- The scope of applications in the area of eye tracking in VR

2 Historical Annotations

The concept of virtual reality (VR) can be traced back to several key developments in the mid 1950s to early 1960s. One of the pioneers was Morton Heilig who created the first VR head-mounted display (patented 1960), albeit for a non-interactive film medium without any motion tracking and the Sensorama (patented 1962), a VR machine that combined visual, audio, and sensory effects to immerse users in a simulated environment (Heilig, 1992; Brockwell, 2016). In 1965, Ivan Sutherland described the concept of the “Ultimate Display”, envisioning a virtual world indistinguishable from reality, which included a virtual world viewed through an HMD augmented by 3D sound and tactile feedback. In 1968, Sutherland and Sproull created the first HMD (“Sword of Damocles”) that was connected to a computer and not a camera (Sutherland, 1968). In 1972, General Electric produced the first computerized flight simulator with screens arranged in a 180° configuration that simulated training cockpits and gave pilots the feeling of true immersion. In the 1980s, Jaron Lanier popularized the term “virtual reality” and developed the DataGlove, allowing users to interact with virtual environments through hand movements (Zimmermann et al., 1986). Consumer-grade VR devices began to emerge in the 1990s, although they faced limitations due to technology constraints. Starting in the early 2000s, a multitude of HMDs were developed including, for example, those by Google and Oculus. About 2015, many companies started to develop HMDs, many with commercial success to make them very popular.

The first HMDs with integrated (built-in) eye tracking emerged in the early to mid 2010s, mainly from research labs and a few pioneering companies. The Fove0 HMD, announced in 2015 and shipped in 2017, is widely considered the first consumer VR headset with built-in eye tracking. It used infrared sensors inside the headset and introduced concepts like foveated rendering and gaze-based interaction. Other companies followed soon. SensoMotoric Instruments (SMI) introduced an Eye Tracking HMD around 2014–2016, which was based on a modified version of early Oculus Rift and featured fully integrated eye tracking inside the headset. Tobii developed the VR Development Kit around 2015–2017. The HTC Corporation and Valve collaborated in the development of the HTC Vive Pro, which was released in 2019, and in 2020, the Pico Nero 2 Eye with a built-in Tobii tracker was released. Early versions of HMD eye trackers were add-ons, but quickly evolved into embedded integrations with headset manufacturers.

The introduction of eye tracking into HMDs was motivated by several advantages over simple HMDs: First, eye tracking permits the development of foveated rendering, with rendering of high detail only in the region the user is looking at and low details in peripheral areas, leading to huge savings in computational efforts. Second, users can select objects by looking at them and combine gaze with gestures, voice or buttons for a more intuitive and faster interaction. Third, researchers can track where users look, what captures their attention, and how they behave visually, which is useful, for example, for user interface design or training simulations.

3 Fundamentals

In this section we review major hardware and software used for tracking the eyes, the head and the body in VR environments. We also review major concepts fundamental to the analysis of eye movements.

3.1 Hardware

We first present the hardware used for presenting virtual environments, followed by the presentation of eye tracker and head / body tracker hardware.

3.1.1 Presentation Hardware

VR stimuli can be presented in one of two major configurations, in a Cave system or with a head-mounted display. In *Cave systems,* first developed in 1992 (Cruz-Neira et al., 1992), stimuli are projected onto 3 - 6 walls with the observer(s) standing in the middle (see Fig. 1). Most traditional Cave systems do not directly measure eye movements, but in modern Caves, users either wear stereo shutter glasses with an integrated eye tracker or mobile eye-tracking glasses. In both cases, head movements need to be tracked to convert gaze direction in head-centered coordinates into a stimulus-centered (i.e., allocentric) framework. Alternatively, gaze direction can be converted to allocentric coordinates using markers attached to the walls of the Cave.

Fig. 1 View of a three-walled Cave (Reproduced with permission from Visbox Inc., https://www.visbox.com)

In a *head-mounted display* (HMD), stimuli are presented using a head-mounted viewer with two displays, one for each eye (see Fig. 2). An eye tracker can be mounted inside the viewer, and the position and orientation of the HMD can be tracked with multiple methods, dependent on the headset used.

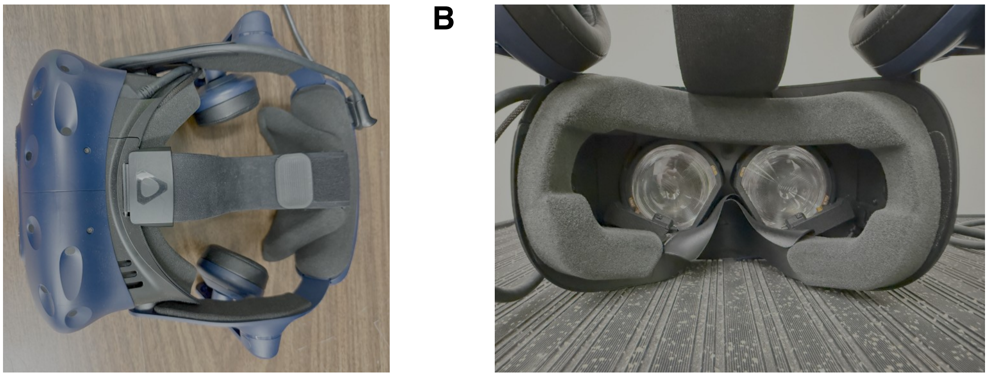


Fig. 2 **A** Outside view of Vive Focus Vision HMD. **B** Inside view with the eye tracker. Photos taken by Cam Hickling.

A comparison between Cave systems and head-mounted displays reveals important differences between the two: Cave systems provide a large field of view (about 170° horizontally for the three-walled Cave in Fig. 1), have a high resolution (2 – 24 megapixels per screen), accommodate multiple observers, and are comfortable with extended use, but are very expensive. In contrast, HMD systems are more limited with respect to the field of view (about 100°), have a lower resolution (about 1 - 2 megapixels per eye), are used by a single observer, are less comfortable with extended use, and are more likely to lead to simulator sickness, but are much cheaper. It is in fact the very low cost of HMD systems that has contributed to the popularity of VR applications in games, research, and many other application areas. For this reason, we focus exclusively on HMD systems in the presentations below.

When setting up a VR eye tracking lab, some consideration needs to be taken for how participants physically interact with the headset. Most headsets with built-in eye tracking are (to date) still tethered to a computer via a cable, as, for example, the HTC Vive. In our experience, participants are aware of any cables from the headset or response apparatus (i.e., keyboard) that are attached to the computer. This may affect their ability or tendency to move freely in the VR environment (if this is desired). For example, in recently published and unpublished work (Bischof et al., 2020; Jacobs et al., 2020) where participants responded on a tethered keyboard while looking at 360° scenes, participants rarely, if ever, rotated the full 360°, despite being seated in a swivel chair. One way to mitigate this is to mount the cables to the ceiling and use VR controllers or wireless keyboards for manual responses.

3.1.2 Eye Tracking Hardware

The goal of eye tracking is to determine the direction of gaze in the coordinate system of the presented virtual world. In the case of HMDs, gaze determination involves two parts: determining gaze direction in head-centered coordinates and determining the position and orientation of the head using a head tracker (see Section 3.1.3) or using scene camera motion analysis (Ohayon & Rivlin, 2006).

Two methods for tracking the eyes are presented here, the corneal reflection method and the pupil tracking method. Other gaze tracking methods are discussed by Jin et al. (2024). In the corneal reflection method, the eye is illuminated with an infrared light source which shows as a bright reflection on the cornea. Using a camera, the pupil and the reflection point are recorded, and from their relative position, eye rotation and eye direction can be inferred. The corneal reflection method is non-invasive and has a high precision, but it is sensitive to lighting conditions, the calibration process is fairly complicated, and it puts fairly heavy demands on the computing resources. In the pupil tracking method, the eye is also illuminated with an infrared light source, but the inference of the eye direction is based on analyzing the location of the pupil's position and movement. This method does not require the detection of the corneal reflection points, and it thus is less sensitive to lighting conditions.

Popular HMDs with built-in eye trackers include the Vive Pro Eye and Vive Focus Vision (https://www.vive.com/), Varjo (https://varjo.com/), and the Play Station VR2 (https://www.playstation.com/). Eye trackers can be added into HMDs, including those by Tobii (https://www.tobii.com/) and Pupil Labs (https://pupil-labs.com/). Note that due to processing limitations, recording eye movements usually requires that the headset remain tethered to a PC.

In addition, there are several other methods for eye tracking, including electrophysiological recordings, electroencephalography, video recordings, and laser scanning. The interested reader is referred to Jin et al. (2024) for a review of all eye tracking methods.

3.1.3 Head and Body Tracking Hardware

Several systems are available for tracking the head position and direction. In many commercial VR systems, the motion tracker is an integral component of the HMD. In the HTC Vive HMD, for example, head movements are tracked using an inertial tracker (see below) that is calibrated using two infrared sensors (base stations) affixed at opposite corners of the room at a height of about eight feet. According to Niehorster et al. (2017), the orientation precision of the Vive headset is on the order of 0.01° for a non-moving observer. Alternatively, head movement can be inferred from scene camera motion analytics (Ohayon & Rivlin, 2006).

While the above systems are used for tracking the position and orientation of the head, more general systems are needed for capturing the motion of torso, limbs, and joints. We briefly present several types of motion capture systems, namely optical motion capture, inertial motion capture, and magnetic motion capture:

- Optical motion capture systems use multiple infrared cameras to track reflective markers placed on the body. They provide sub-millimeter accuracy, making them very popular, but they require a fixed setup in the recording area and are prone to marker occlusion, and tracking errors increase over time. The best systems include Vicon (https://www.vicon.com/), OptiTrack (https://optitrack.com/) and Motion Analysis (https://www.motionanalysis.com/). Some optical motion capture systems are markerless, i.e., they do not require markers on the body and use computer vision and depth sensors to track the body. Although they are typically less precise, the performance of current systems has improved quickly. These systems include Qualisys (https://www.qualisys.com/) and Dari (https://darimotion.com/).
- Inertial motion capture systems use wearable sensors composed of accelerometers and gyroscopes to measure movement without cameras. They work well without a special setup, have no problems with occlusions like the optical systems, but accuracy degrades over time due to sensor drift. The best systems include Xsens (https://www.xsens.com/) and Rokoko Smartsuit Pro (https://www.rokoko.com/products/smartsuit-pro).

- Magnetic motion capture systems track movements using electromagnetic fields, have a high accuracy and have no occlusion problems, but the systems are very sensitive to metal objects and electronic interference. One of the companies offering such systems is Polhemus (https://polhemus.com).

3.2 Eye Tracking

In this section, we present several concepts that are important for the analysis of eye movements in VR: the analysis of fixations and saccades, calibration methods, reference frames, data quality, and visualization methods (see also Clay et al., 2019).

3.2.1 Fixations and Saccades

Typically, many eye movement types are recognized, such as fixations, saccades, smooth pursuit, and micro-saccades (Alexander & Martinez-Conde, Chapter 3 of this volume; Andersson et al., 2017; Holmqvist & Andersson, 2017; Hooge et al., 2019; Lencer et al., Chapter 4 of this volume; Pierce et al., Chapter 2 of this volume), but in this chapter, we restrict our analyses to fixations and saccades. In the studies presented in this chapter, we use the IDT method for extracting fixations from raw eye movement data (Komogortsev et al., 2010; Salvucci & Goldberg, 2000). In Chapter 13 of this volume, we present methods for the detection of fixations and saccades in detail. The reader is encouraged to consult that chapter as well as further sources of information on event detection in eye movement analysis, for example, those by Andersson et al. (2017), Holmqvist and Andersson (2017), Hooge et al. (2018) and Nyström and Holmqvist (2010).

3.2.2 Calibration

A gaze point can be defined as the point in the stimulus where a person is looking at. To obtain a mapping between eye tracker output and a gaze point, a calibration procedure is required (Harezlak et al., 2014). Many simple calibration procedures are in use: Participants are, for example, required to fixate each point of a matrix with 9 to 25 points and wait until the system has recorded the measurement at each point, or they are required to follow a slowly moving dot (Clay et al., 2019). At the beginning of each trial, participants may be required to fixate a point to ensure continued tracking accuracy.

3.2.3 Reference Frame

Before diving into the specifics of eye, head, and gaze analysis in VR, it is worth clarifying the reference frame we are dealing with (see Hessels et al., 2018). Most readers may be familiar with the popular desktop-based eye tracking technology, where observers are required to use a chin rest at a set distance away from a computer monitor. In this situation, the head is fixed, and eye movements are reported in coordinates of the computer monitor, usually in pixels or in degrees of visual angle. In other words, the reference frame for eye movements is the screen, which typically encompasses approximately 30-50° visual angle.

In VR experiments, the eyes are tracked by cameras mounted within the HMD. The head and body are usually free to move naturally, although the weight of the headset, participants' pose (e.g., standing vs. sitting), and potential tethering can influence participants' range of motion. One significant advantage of this setup is that eye and head movements can be tracked with respect to a common reference frame. This means that measures like eye eccentricity in the head coordinate system, as well as the contributions of head movements to gaze position can be calculated precisely (see below). Note however, that without special equipment, neck movements cannot be differentiated from chair and torso movements. For the sake of simplicity, we use the position of the HMD in space as a proxy for the head position.

3.2.4 Data Quality

Eye tracking data quality refers to several eye tracker performance measures, including spatial precision, spatial accuracy and latency (see also Adhanom et al., 2020; Hutton, Chapter 11 of this volume). Spatial accuracy is the average deviation between the true gaze position and the gaze position measured by an eye tracker (Holmqvist & Andersson, 2017). It is measured as the distance between a target point and measurements provided by the eye tracker. It is important to compare the spatial accuracy of the tracker with the stimulus resolution one wants to interpret.

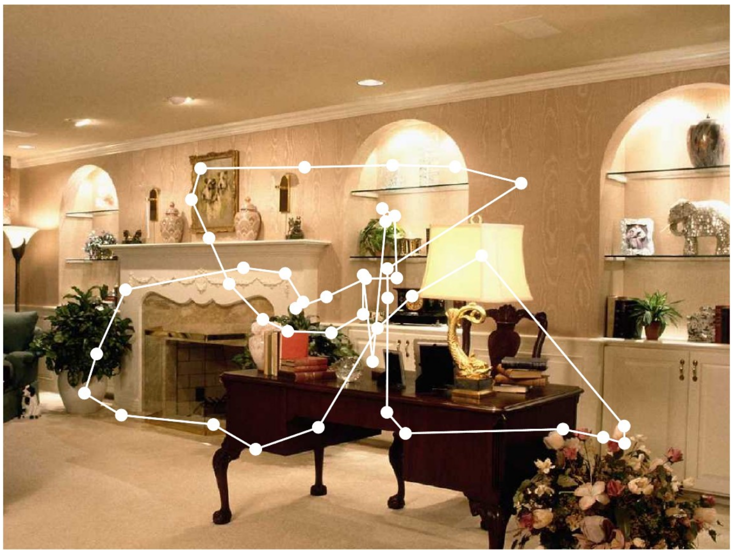
Spatial precision refers to the ability of an eye tracker to reproduce measurements accurately over time (Holmqvist & Andersson, 2017) and is usually measured as the root mean square angular displacement. The precision required for an eye tracking application depends on the resolution of the stimulus that needs to be interpreted, but should be below 0.03° (Andersson et al., 2010).

Latency refers to the delay between eye movement events and the time these events are recorded in the system. Latency is affected by multiple factors, including sampling rate, the time to compute gaze positions, the time to filter the data, by delays introduced in the transmission network, as well as other processing delays. These factors lead to a lag between the occurrence of eye movement events and their recording. It is often assumed that eye movement recordings are occurring at fixed intervals when in fact the lags can be distributed over relatively long intervals. This is particularly evident in gaze-contingent displays or in interactive applications.

3.2.5 Visualization

Fixation patterns for a single participant and a single stimulus can be visualized using a point-based eye movement path. This is illustrated in Fig. 3A, which shows a scene overlaid with an eye movement path and fixations shown as white circles. The radius of the fixation circles can be made proportional to the fixation duration.

A



B

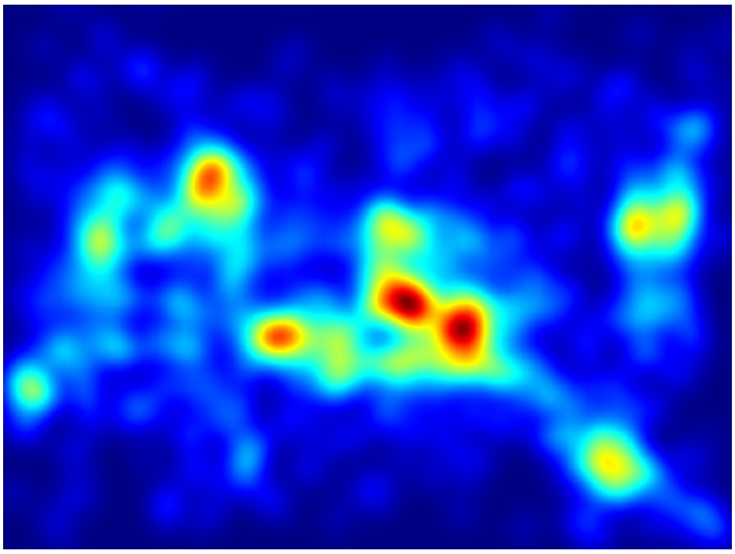


Fig. 3 A Fixation pattern of a single participant with white circles indicating fixations. **B** Fixation patterns of many participants represented in a heatmap.

Fixations of multiple observers can be visualized using a smooth heatmap. In this method, each fixation is replaced by a 2D-Gaussian with a pre-defined standard deviation, and these Gaussians are added together, resulting in a smooth fixation map, which is often referred to as a heatmap (e.g., Duchowski, 2017). Heatmaps can be presented in gray-level, with black representing the minimum value and white representing the maximum value. Alternatively, they can be pseudo-colored with blue representing the minimum value and red representing the maximum value (see Fig. 3B). The standard deviation of the Gaussian in Fig. 3B was 20 pixels with an image size of 1024 by 768 pixels. There are, however, no hard and fast guidelines regarding the best choice of the standard deviation. This can often make it difficult to compare heatmaps across different experimental conditions. The heatmap can also be visualized by assigning a Gaussian standard deviation based on fixation duration. Again, the heatmap makes it clear which image areas were fixated the most. The heatmap can also be overlaid over the original image, but only the peaks are shown in color with the color proportional to the height of the heatmap, whereas in regions with low values, the original image is shown (see Anderson et al., 2013, Figure 10).

Heatmaps are good visualization tools and can give a quick overview of fixation patterns across many participants and highlight fixation hotspots (see Fig. 3B). For several reasons, however, the analysis and the interpretation of heatmaps are difficult. First, heatmaps are specific to a particular stimulus layout, and hence can be used only to compare heatmaps of fixation patterns obtained for the same images (or images with identical spatial layouts). Second, fixation hotspots may not be due to participants fixating a particular scene point, but could be due to the fact that participants tend to fixate the center area of stimuli more frequently than peripheral areas, independent of particular scene contents (Tatler, 2007). In Textbox 1, we provide further details on the statistical comparison of heatmaps. Finally, fixations can also be visualized using areas of interest (AOI). This is further described by Bischof et al. (Chapter 13 in this volume).

Textbox 1: Comparison of Heatmaps

Heatmaps must be analyzed statistically to establish differences and similarities obtained under different experimental conditions, as well as to establish whether certain fixation hotspots are significantly higher than the rest of the heatmap (for a detailed overview of heatmap comparisons, see Le Meur & Baccino, 2013). The similarity of heatmaps can, for example, be measured using a simple correlation between two heatmaps measured over all image positions. Second, one can compare the two heatmaps using the Kullback-Leibler divergence, a non-symmetric measure of the difference between two probability distributions (Kullback & Leibler, 1951). Third, the heatmaps can be compared using ROC analysis (Green & Swets, 1966) where the ROC curve expresses the correspondence between the two maps (for more details see Judd et al., 2009; Le Meur & Baccino, 2013).

For smooth heatmaps, the comparison is more complicated due to the fact that independence between nearby positions is lost due to the spatial smoothing with the Gaussian filter. This is a situation similar to the one in the analysis of fMRI results, where a smoothing is applied to the raw data (Friston et al., 1994, 1995). For eye movements, Caldara (2011) has proposed a pixelwise comparison of heatmaps that takes this dependence into account. In general, it is best to rely on bootstrapping methods (Efron & Tibshirani, 1993) for the statistical analysis of heatmaps.

3.3 Head Tracking

Head movements are inferred from the movements of the VR headset (see also Hu et al., 2023). As explained earlier, the movement of the headset is taken as a proxy of the head movement. Head movements are defined as head yaw, which is achieved through lateral neck rotation, by head pitch, which is achieved through neck extension and flexion, and head roll, which is achieved through lateral bending of the neck. Figure 4 illustrates the range of head orientations. Assuming an upright upper body orientation of the participants, the lateral (i.e., left/right) rotation of the head (yaw) is in the range of ±80°, and the extension / flexion (i.e., up/down) of the head (pitch) is in the range of ±45°. In the studies reported below, lateral bending of the head (roll) in the range of ±45° is ignored.

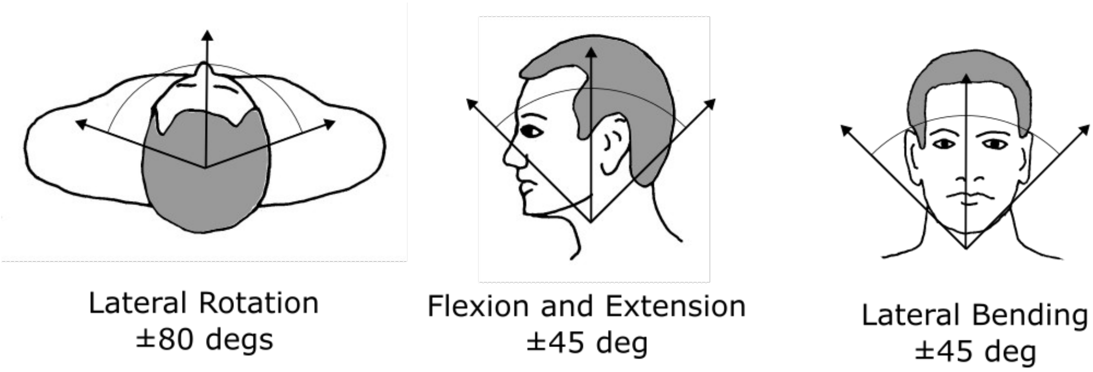


Fig. 4 Axes of head orientation. Left panel: Lateral rotation of the head has a range of ±80 degrees. Middle panel: Flexion-extension of the head has a range of ±45 degrees. Right panel: Lateral bending has a range of ±45 degrees. From (Bischof et al., 2024)

3.4 Data Processing

In this section, we present software systems for the creation and presentation of VR worlds, discuss the synchronization of eye movement data with other sensors, and present the steps in a processing pipeline for the analysis of eye movement data.

3.4.1 Software

There are at least three major software systems for creating VR worlds and recording behavior: Vizard, Unity, and Unreal Engine. Not every VR eye tracker is compatible with each of the above systems, so it is worth making note of the availability and support for different software prior to purchasing any VR eye tracking hardware. No-code or low-code options from MindPort (https://www.mindport.co/vr-builder), ShapesXR (https://www.shapesxr.com/), and Zappar (https://www.zappar.com/) may also be good options.

3.4.1.1 Vizard

Vizard is a VR software development platform developed for academic research by Worldviz (http://worldviz.com). Its architecture provides precise, millisecond-level timing for all events within an immersive virtual environment, which is critical in experiments that involve the rapid presentation of stimuli or require accurate measurement of response times. Vizard is scripted using Python and comes with a large library of example scenes and sample scripts, which can be used as building blocks for experiments. In addition, Worldviz has developed SightLab VR, an extension to Vizard, which offers a complete experiment structure with integrated data collection for eye tracking and a drop-down menu for configuring VR experiments while retaining access to the scripting level.

With the focus on research and on the implementation of experiments, Vizard differs from the general game engines presented in the next section. Vizard offers native support for a large array of VR hardware, including headsets, VR projection systems, motion capture systems, eye trackers, biofeedback, haptic devices, and more. It supports a diverse array of VR headsets with integrated eye-tracking devices, making the integration of eye tracking into experiments easy and requiring only a few lines of script.

3.4.1.2 Unity

Unity is a game engine developed by Unity Technologies in 2005. The engine has been extended to support a large range of platforms, including desktops, mobile devices, and VR platforms. It is very popular for mobile game development and was used for games such as Pokémon Go or Call of Duty. It is considered easy to use for beginners and can be used for creating 3D and 2D worlds. Creating a virtual world consists of setting the world up using an interactive development environment and expanding and controlling it using C# or Java as a scripting language. Experimental control and timing control are, however, more difficult than in Vizard.

There are many toolboxes and packages that help in the development of Unity-based experiments, such as Unity Studio (a web-based, no-code 3D editor), the Terrain Toolbox (for creating complex landscapes), as well as many packages for experiment control, including UXF (Brookes et al., 2020), USE (Watson et al., 2019), bmlTUX (Bebko & Troje, 2020), and VREX (Vasser et al., 2017). The interested reader is referred to Anderson et al. (2023) for an extended and very detailed presentation of the Unity system, programming and applications.

3.4.1.3 Unreal Engine

Unreal Engine is a game engine developed by Epic Games in 1998. The engine was first developed for first-person shooter games, but has since been used in an expanding range of 3D games and has been adopted by the film and television industry. The Unreal Engine is scripted in C++ and supports a wide range of desktop, mobile, console, and virtual reality platforms. For this engine, there are also many toolboxes and packages that help in the development of Unreal-based experiments, including the inventory grid, the mining grid, and other advanced gaming tools. Both Unity and Unreal Engine focus on the development of games using impressive graphics, but are less useful for the design of experiments.

3.4.2 Recording and synchronizing data with other sensors

A key complication is that these different data streams are often recorded at different rates. For example, in our own work using the SMI eye tracker with Unity (Anderson et al., 2023), the eye tracker sampled gaze direction at approximately 250 Hz, whereas Unity's Update() function (which records headset position and orientation) was called at only 50 to 70 Hz. To avoid losing eye tracking samples, we ran the eye tracker independently of the Unity system via multithreading, collecting timestamped gaze samples in a dedicated thread at the tracker's native rate. At the end of each trial, we combined the eye and head data streams in a single output file, using Unix timestamps and linear interpolation to synchronize the two signals. A detailed description of this approach, along with example C# code, is provided by Anderson et al. (2023), and additional discussion of the data handling considerations that arise in VR eye tracking can be found in Bischof et al. (2024).

Although custom scripting offers flexibility, it also requires considerable programming effort and careful attention to timing. Commercial platforms such as Cognitive3D (https://cognitive3d.com) offer an alternative approach, providing pre-built SDKs for Unity, Unreal Engine, and other platforms that handle session recording, gaze and head tracking, object interactions, and the synchronization of external sensors (e.g., heart rate, galvanic skin response, EEG) within a unified data pipeline. For researchers who prefer not to implement their own recording and synchronization code, such tools may substantially reduce the technical overhead involved in collecting well-organized, multi-stream VR data. The complete processing pipeline for the analysis of eye movement data is detailed in Textbox 2.

Textbox 2: Processing Pipeline for the Analysis of Eye Movement Data

The following steps describe the processing of eye and head movements as well as other sensors.

Step 1: *Stimulus presentation and data recording*

The virtual environments are presented in multiple trials and data from the eye tracker, head tracker, and other sensors need to be recorded. Before recording data formally, check the coordinate frame of all data points (local vs. global) and ensure all data output can be synchronized to the same spatial reference frame (Step 6).This has been described in detail in Section 3.4.1.

Step 2: *Data preparation and event detection*

Data from all sensors may have to be transformed into a common format. For example, the eye tracker might record the position and quaternions of both eyes, which then might have to be converted to mid-eye position and monocular viewing direction. Similarly, the position and quaternions of the head-mounted display might have to be transformed into head position and the head forward vector. In addition, it must be ensured that all data sets are stored in a uniform format.

Step 3: *Data cleaning*

Several steps have to be followed to ensure the integrity of the data sets (see, e.g., Miller et al., 2021; Landes et al., 2024). First, each data set has to be limited to the relevant time interval, and leading or trailing measurements should be removed. Furthermore, missing data has to be treated in a uniform manner. This includes the elimination of eye movement data during blinks or head movement errors due to tracker failure or slippage (Niehorster et al., 2020). Second, data outliers can affect data integrity and should be removed. These outliers can arise for multiple reasons, including measurement errors or participant distractions. Their treatment includes, for example, the elimination of very long and very short fixation durations or the elimination of excessively large saccades. Methods for identifying outliers have been presented in detail by Chen et al. (2010).

Step 4: *Filtering, smoothing, and Event Detection*

To reduce noise and increase fixation stability, data can be interpolated and smoothed using smoothing filters, which may include Fourier transform methods or filters like

Savitzky-Golay. This permits the elimination of unwanted frequency components of the eye movement data and those of the other sensors. This step should be followed by event detection, e.g., the detection of eye fixations and saccades, as described in Section 3.2.1.

Step 5: *Temporal Synchronization of Data*

Alignment of data from eye tracker, head tracker, body tracker, and other sensors could be synchronized by aligning the start points, but due to different recording rates temporal alignment using timestamps (with a resolution of one ms or better) and temporal interpolation of data sets is recommended.

Step 6: *Spatial Synchronization of Data*

Data should be processed in a common spatial coordinate system. This is achieved by representing data in a common reference frame, as discussed in Section 3.2.3.

Step 7: *Statistical Analysis*

All resulting tracker data should be analyzed using statistical methods, which depend on the purpose of the eye tracking study. It is important to note that reliance on visualizations alone is not sufficient for presenting the results of a study.

The steps above are not subject to a strict temporal order, but they simply emphasize different aspects of the data analysis. At each step of the analysis, quality checks and visualizations can help to eliminate data processing errors. The interested reader should consult the very detailed presentation of eye movement data processing by Zhang and Jonides (2026).

3.5 Summary

This section provided an overview of a diverse set of topics fundamental to the analysis of eye movements in VR: methods of eye movement analysis, eye tracker and head tracker hardware, software for the presentation of VR and the recording of eye movements, synchronization of data streams, and in a textbox, a description of the processing pipeline for the analysis of eye movement data. The interested reader is encouraged to consult Holmqvist and Anderson (2017), Anderson et al. (2023), and Bischof et al. (2024) for further information.

4 Panoramic Virtual Environments

In this section, we present some of our work on eye and head movements in panoramic VR scenes, i.e., in a static 360° spherical environment without any moving objects (Bischof et al., 2020, 2023, 2024; Mehrotra et al., 2025). The stimulus sphere was centered to the participants' head to ensure that it remained unaffected by the participants' body movements. They were asked to move their eyes and head to look at a random selection of indoor and outdoor 360° panoramic scenes. Participants looked at the scenes using an HMD, and each scene was displayed for 10 seconds. A uniform gray scene with a black fixation cross presented directly ahead of the participant was displayed between images, and participants were instructed to look at this fixation cross and press a key to initiate the next trial.

Due to this setup, only a limited number of depth cues were available to the participants: Perspective distortion (parallel lines converging in the distance allow us to reconstruct the relative distance of two parts of an object), relative size (the relative size of known objects can provide information about their relative depth), interposition (near surfaces overlap far surfaces and are perceived as being closer), and texture (fine details on nearby objects can be seen clearly, whereas they are not visible on faraway objects). In the case where the 360° panoramic scenes are produced using stereoscopic videos, participants can also use disparity information (where each eye views a slightly different angle of an object seen) as depth cues. Most videos available in public databases are, however, monoscopic and thus lack disparity information (see also Hornsey & Hibbard, 2021).

In the studies reported here, participants were exploring the environment using passive observation with the task of remembering the scenes for a later memory test. Alternatively, participants can be instructed to actively observe the environment, for example, by searching for particular objects in the scenes (see also Haskins et al., 2020).

4.1 Panoramic Analysis

The virtual space consisted of a sphere around an observer, onto which different omni-directional panoramic scenes were projected, effectively immersing an observer in the scenes, with the head positioned in the middle (see Fig. 5A). Figure 5B shows one of the panoramas projected on the sphere, with the yellow lines indicating the equator and the meridian of the initial fixation. One way to visualize and analyze eye points, or sets of eye points, is to project the panoramas onto a flat map, for example, an equirectangular (or equidistant) projection map (Equirectangular, n.d.; Grafarend et al., 2014; Lapaine & Usery, 2017). This projection maps meridians into vertical straight lines of constant spacing, introducing distortions near the poles compared to the equator (see Fig. 5C). Map positions are defined by longitude (in the range ±180°) and latitude (in the range ±90°).

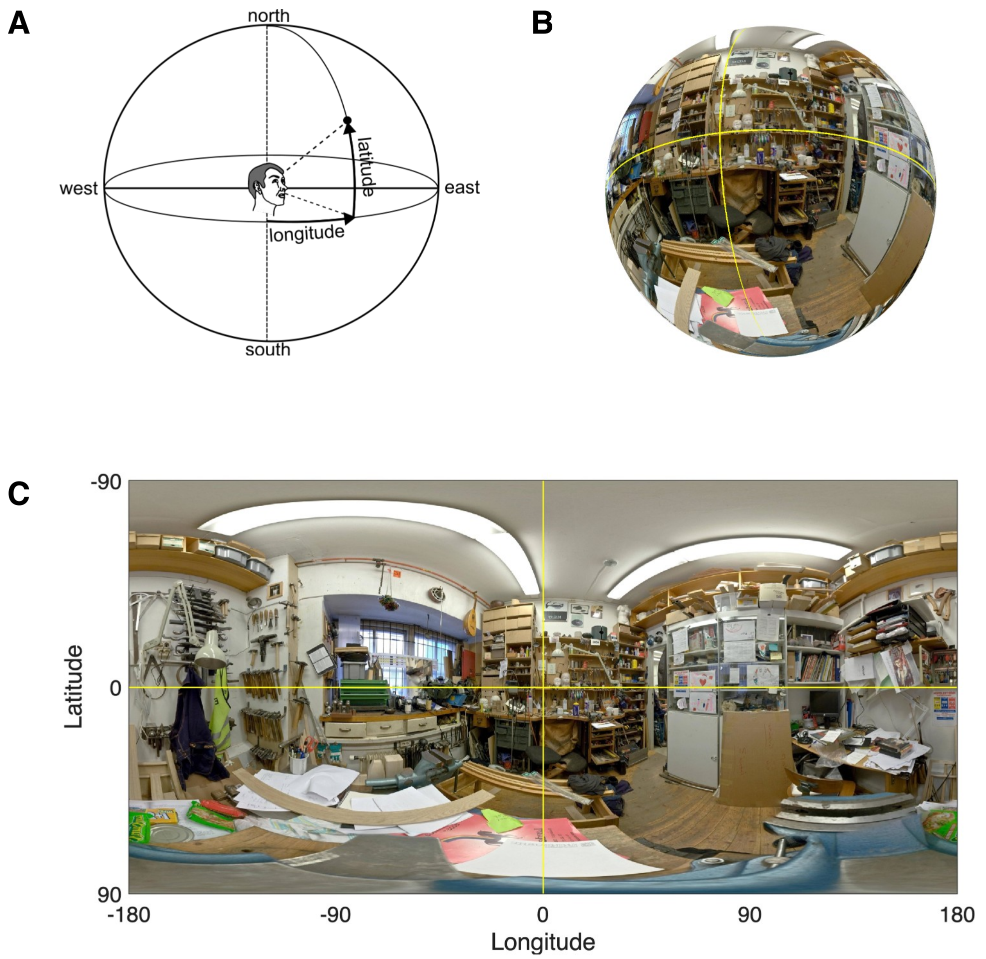


Fig. 5 A Sketch of a participant viewing the spherical panorama from the center of the virtual sphere. An eye point (black circle) is defined by the intersection of the eye direction vector with the virtual sphere and is described by longitude (azimuth) and latitude (elevation). **B** Spherical stimulus viewed from outside and from below the equator, with the yellow lines indicating the equator and the meridian of the initial fixation cross. **C** Equirectangular (equidistant) map of the stimulus. The yellow lines indicate the equator and the center meridian. Note the distortions near the north and south poles. In analyzing fixation patterns, these distortions must be taken into account. From Bischof et al., (2024).

In the analysis of eye movements in a fully immersive VR environment, wrap-around of the coordinate system at longitudes ±180° and distortions near the poles must be taken into account, hence circular (Batschelet, 1981; Bischof et al., 2020 Appendix 1; Mardia & Jupp, 2000) or spherical statistics (Fisher et al., 2010) should be used. Particular attention must be paid to how distances between successive eye positions are calculated. This has implications, for example, for the dispersion threshold of fixations (usually represented in degrees visual angle) as well as saccade amplitudes (distances between fixations). In a fully immersive 360° world, these distances must be calculated using the orthodromic (great circle) distance, which is defined as the shortest distance between two points on a sphere.

4.2 Basic Terms

As indicated before, participants viewed the panoramas from the center inside the sphere on which the panoramas are projected (see Fig. 5A). We now define three basic terms for our analyses: A *head point* is defined as the intersection of the vector pointing forward from the face with the panorama sphere. Azimuth and elevation of this point are described with longitude coordinates ranging from -180 to 180 degrees and latitude coordinates ranging from -90 to 90 degrees. An *eyes-in-space point* (also referred to as gaze point) is defined as the intersection of the eye direction vector with the panorama sphere with the same azimuth and elevation range as head points. Finally, an *eyes-in-head point* is defined as the intersection of the eye direction with the panorama vector, and is described in the head coordinate system with longitudes ranging from -110 to 110 degrees and latitudes ranging from -80 to 70 degrees. This more restrictive range of longitudes and latitudes derives from the limited field of view of the eyes (Strasburger et al., 2011).

In the sections below, we introduce how one goes about analysing head and eye movement data collected in VR by systematically reporting the data: 1) eyes-in-head data, i.e., the direction of the eyes relative to the head, as measured by the eye tracker; 2) head data, i.e., the direction of the head in world coordinates (coordinates on the sphere), as measured by the head tracker; and 3) the eyes-in-space data, i.e., the direction of the eyes in world coordinates (coordinates in the sphere) as obtained by combining the eyes-in-head coordinates with head coordinates.

4.3 Eyes-in-Head Analysis

Figure 6 shows the heatmaps of the eyes in head-centered coordinates. The heatmaps were obtained from the maps of eyes-in-head points using a Gaussian filter with $\sigma = 4°$ angular distance. The analysis of the eyes-in-head coordinates shows that the eyes are directed below the head-defined center of the visual field and with a bimodal distribution, with one peak to the left (west of) the 0° meridian and the other to the right (east of) the 0° meridian (Fig. 6A). A closer analysis reveals the following: As explained by Bischof et al. (Chapter 13 of this volume), the eyes lead the head in almost all cases. Thus, if the head moves towards the left, the eyes tend to be to the left of the head position (Fig. 6B), and conversely, if the head moves towards the right, the eyes tend to be to the right of the head position (Fig. 6C). If head movement is ignored, the double peak of the eye-in-head heatmap is obtained (Fig. 6A).

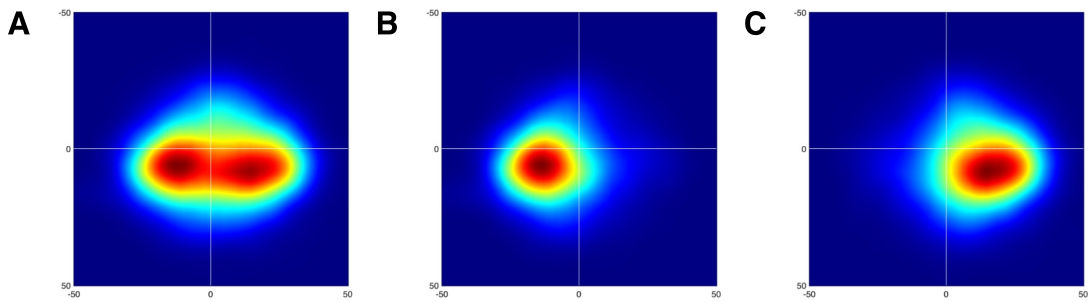


Fig. 6 Heatmaps of eyes-in-head points in the longitude range [-50°, 50°] and the latitude range [-50°, 50°], with the white lines showing 0° longitude and latitude. **A** Heatmaps of eyes-in-head points. There are two peaks, one to the left and the other to the right of the head-defined center. Panels B and C show how these double peaks arise: **B** if the head is moving to the left the eyes-in-head points are left of the head-defined center. **C** if the head is moving to the right the eyes-in-head points are to the right of the head-defined center. Adapted from (Bischof et al., 2024)

The standard deviation of the eyes-in-head longitudes was 16.6° and of the latitudes was 10.8° (taking head movements to the left and right into account). These results are consistent with other results in the literature indicating that eye positions are limited to about 15 - 25° of the head positions (David et al., 2022; Freedman & Sparks, 1997).

Given two fixations f_i and f_j with eyes-in-head coordinates e_i and e_j, the eyes-in-head saccade is defined by the great circle from e_i to e_j, giving the direction and the amplitude of the eyes-in-head saccade. The velocity of the eyes-in-head saccades is defined by dividing the amplitude of the eyes-in-head saccades by the duration of the saccade from f_i to f_j. Figure 7 shows the direction and velocity distributions of the eyes-in-head saccades. The saccade velocity distribution is bimodal. A closer analysis suggests that this is the result of two different types of saccades, slow saccades and fast saccades. The slow saccades have an isotropic direction distribution, and the fast saccades are almost exclusively in the horizontal direction. The latter are likely related to a strategy of revealing new regions of the panoramic stimulus with horizontal saccades.

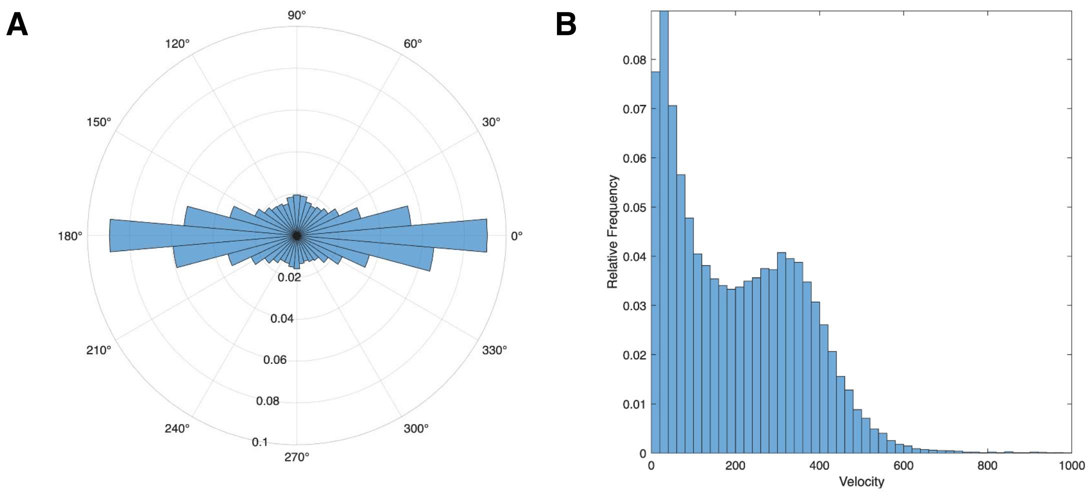


Fig. 7 **A** Direction distribution of the eyes-in-head saccades. **B** Velocity distributions of the eyes-in-head saccades. Adapted from (Bischof et al., 2024)

In summary, the results show the direction of the eyes in head coordinate, and how fast they move. Mapping where people look yields a bimodal distribution which results from the spatio-temporal dynamics of eye and head movements.

4.4 Head analysis

Figure 8 shows the head heatmap, which was obtained from all head points using a Gaussian filter with σ = 4° angular distance at the equator and horizontally increasing by sec(latitude) towards north and south. The head points are concentrated along the equator line of the panoramas. This is known as the equator bias and is consistent with results obtained by Bischof et al. (2020) and Sitzmann et al. (2018). There are multiple possible reasons for the equator bias. First, the equator area of the panoramas corresponds to the area with the highest statistical information density (e.g., Anderson et al., 2023; Solman et al., 2017; Torralba et al., 2006). Second, if participants explore the panoramas with neck extension / flexion in a resting state, then there is a natural preference for fixations along the equator. The head points are biased neither east nor west of the midpoint (with longitude and latitude 0°), which is most likely due to the fact that participants started each trial head and eyes directed at the same position. The distribution of head points is highly anisotropic (longitude *STD* = 79.8°, latitude *STD* = 15.3°).

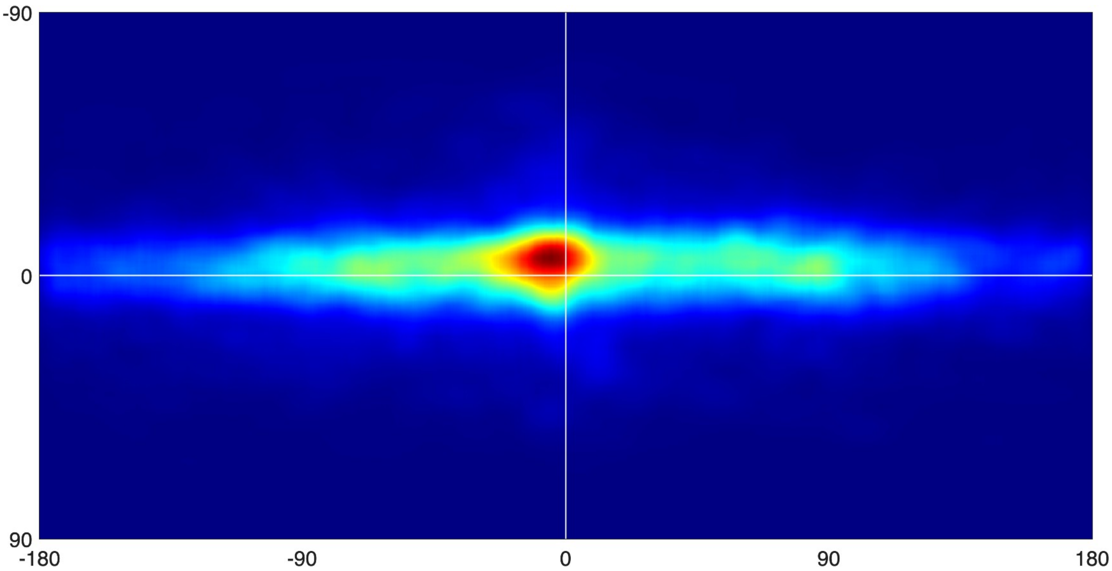


Fig. 8 Heatmaps of head points in the longitude range [-180°, 180°] and the latitude range [-90°, 90°], with the white lines showing 0° longitude and latitude. The head points are concentrated along the equator of the panoramas. From (Bischof et al., 2024)

We define fixation-based head shifts as follows: Given two fixations f_i and f_j, we define the head positions h_i and h_j by the mean longitudes and mean latitudes of the head during f_i and f_j. The head shift from h_i to h_j is defined by the great circle from h_i to h_j, which defines the direction and the amplitude of the fixation-based head shift. The velocity of the head shift is defined by dividing the amplitude of the head shift by the duration of the saccade from f_i to f_j. Figure 9 shows the direction and velocity distributions of the head shifts. The head shifts are primarily distributed along the horizontal direction of the panoramas (*STD* = 79.8°, with relatively little variation along the vertical direction (*STD* = 15.3°), consistent with the fact that head positions are concentrated along the equator. There is a wide range of head movements along the equator of the panoramas, i.e., along latitude 0°, and a relatively small range of head movements in the vertical direction.

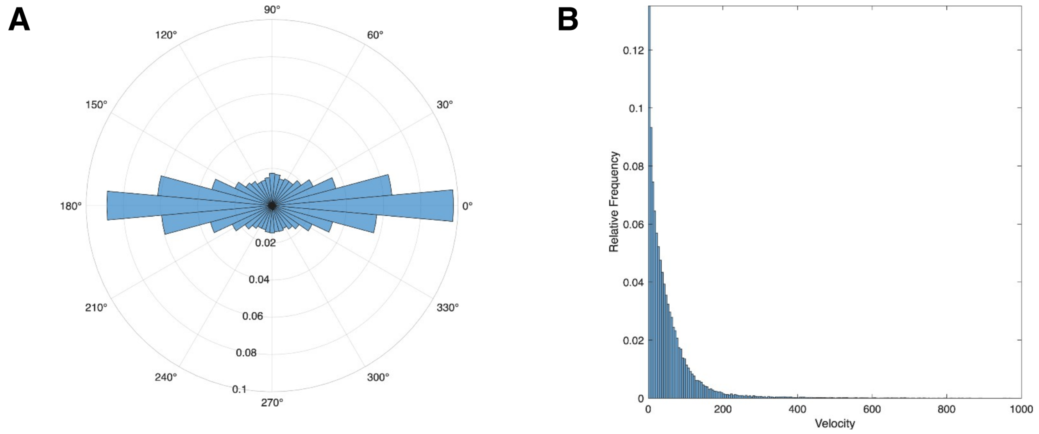


Fig. 9 **A** Direction distribution of head shifts; **B** Velocity distribution of head shifts. Adapted from (Bischof et al., 2024)

In the next section we consider where the eyes are directed with regard to the virtual environment (eyes-in-space or 'gaze'). Determining where the eyes are directed in space involves combining the position of the eyes in the head and the direction of the head in space (see Valenti et al., 2012 for a thorough discussion of this issue).

4.5 Eyes-in-Space Analysis

The eyes-in-space heatmap is illustrated in Fig. 10, which indicates that the eyes-in-space fixations are concentrated along the equator of the panoramas. This heatmap is best understood as the combination of the head heatmap (Fig. 8) and eye-in-head heatmap (Fig. 6, see also Valenti et al., 2011). Figure 10 illustrates the strong effect of head movements on the practical field of fixation, that is, the idea that the head enlarges the space within which a fixation can be made (David et al., 2020, 2022; v. Rötth, 1925; von Noorden et al., 2002, p. 79).

A detailed analysis shows that the eyes-in-space points are not biased above or below the equator line, nor are they biased east or west of the 0° meridian line. Again, this is most likely due to the fact that participants started each trial looking at the same position on the stimulus sphere. The distribution of eyes-in-space points is anisotropic (longitude *STD* = 86.2°, latitude STD = 20.7°). The eyes-in-space distributions are similar to the head distributions, albeit somewhat less extreme.

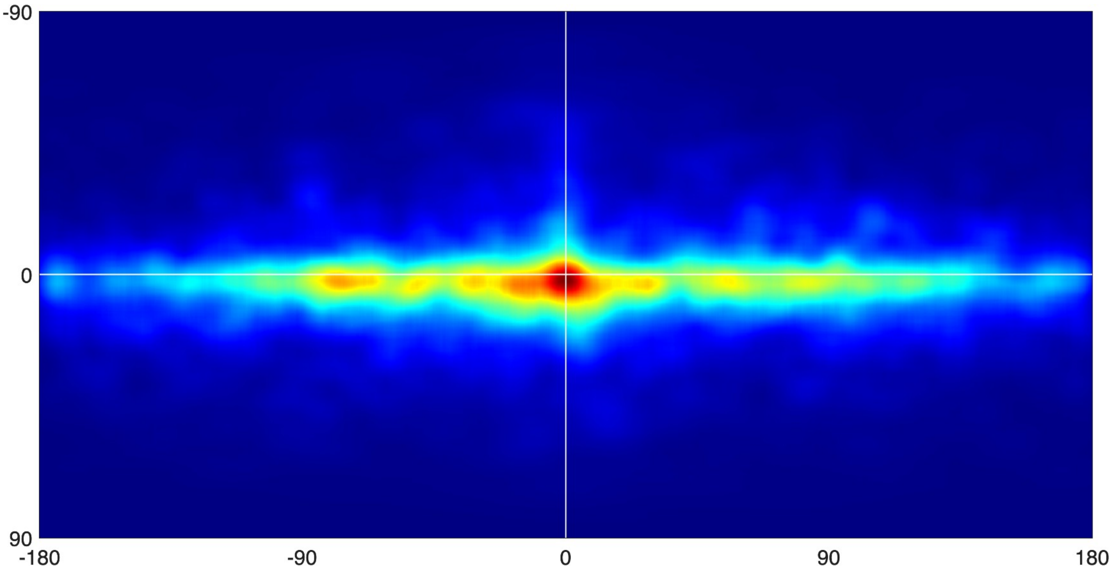


Fig. 10 Heatmap of eyes-in-space points in the longitude range [-180°, 180°] and the latitude range [-90°, 90°], with the white lines showing 0° longitude and latitude. The distribution of eyes-in-space points is determined by the distribution of the head positions and the distribution of eyes-in-head points. From (Bischof et al., 2024)

Given two eyes-in-space fixations f_i and f_j, we define eyes-in-space saccades by the great circle $f_i - f_j$, giving the direction and the amplitude of the eyes-in-space saccades. The velocity of the eyes-in-space saccades is defined by dividing the amplitude of the eyes-in-space saccades by the duration of the saccade from f_i to f_j. Figure 11 shows the direction distribution and velocity distributions of the eyes-in-space saccades. The distributions of the directions and velocities of the eyes-in-space saccades are almost identical to those of the eyes-in-head saccades. The bimodal saccade velocity distribution again suggests that this is due to two different classes of saccades, slow saccades and fast saccades with the latter likely related to a strategy of revealing new regions of the panoramic stimulus with horizontal saccades.

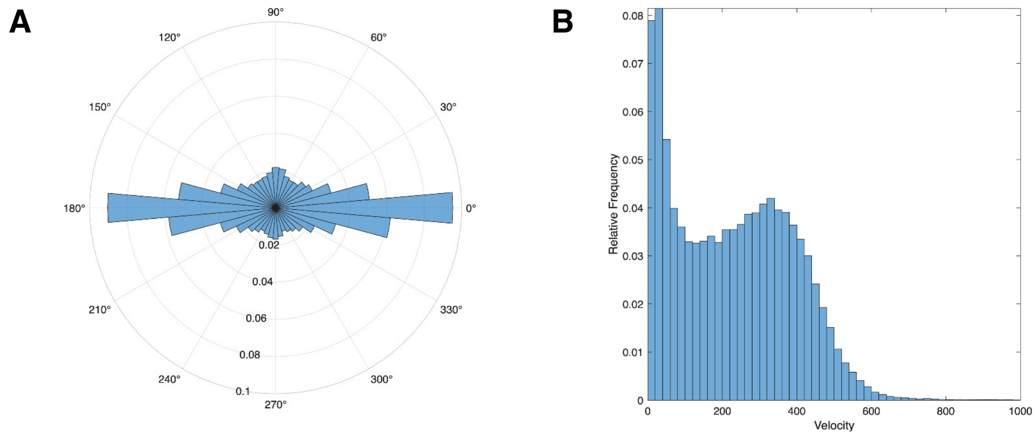
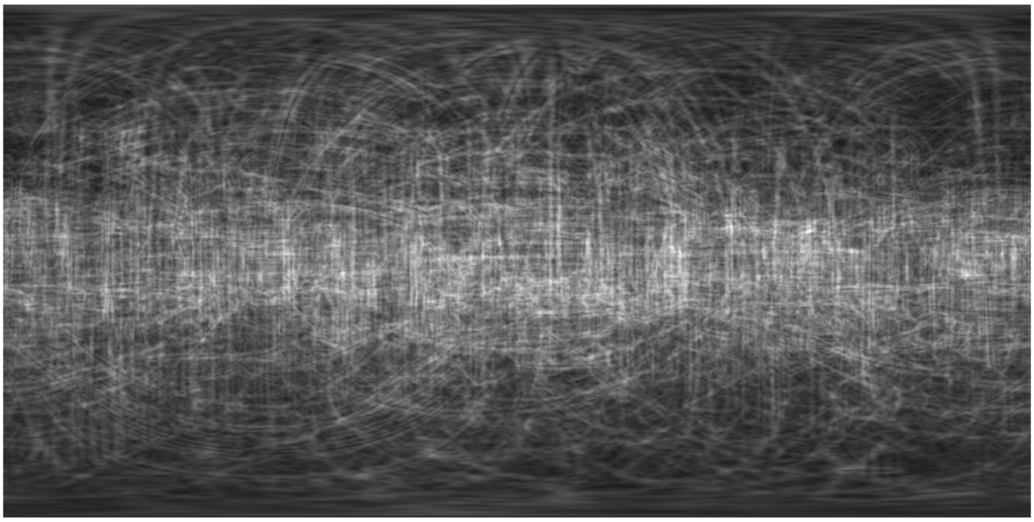


Fig. 11 **A** Direction distribution of eyes-in-space. **B** Velocity distribution of eyes-in-space. Adapted from (Bischof et al., 2024)

The panoramas differ with respect to the distribution of information in the vertical dimension. For example, panoramas of indoor scenes tend to have information spread further along the vertical dimension than panoramas of outdoor (landscape) scenes. It is plausible that this may have an effect on all effectors, but on the distribution of eyes-in-space in particular. All panoramas were grouped into either indoor or outdoor panoramas. Figure 12 shows the distribution of edges of each group, obtained with an edge detector combined with a Gaussian filter of $\sigma = 0.5°$ at the equator and increasing horizontally by sec(latitude) towards the north and south. The filter size was chosen to visualize all major edges of the panoramas. The edge maps show that the indoor scenes contain more information in the vertical dimension than the outdoor scenes.

A



B

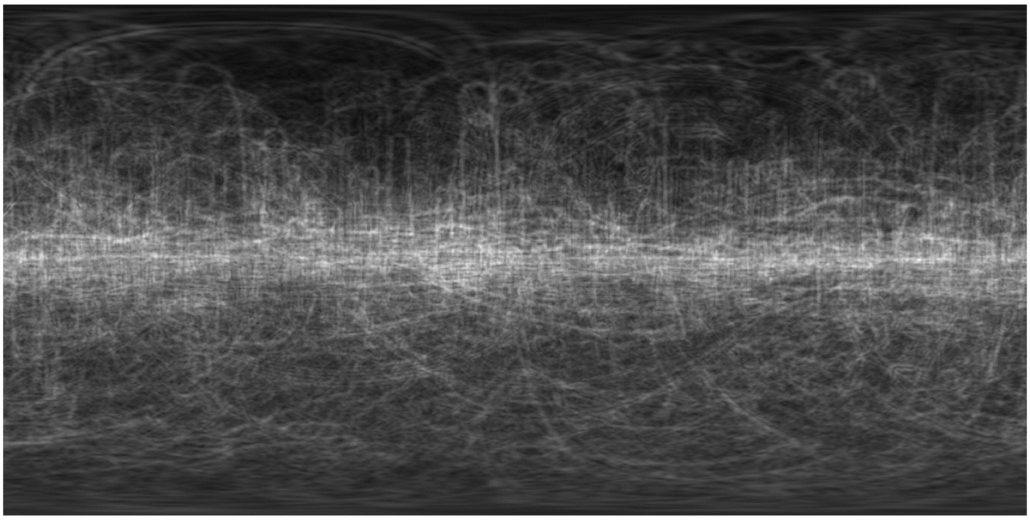


Fig 12. Edge distribution of **A** indoor images and **B** outdoor images. Note the much larger spread of edges in the vertical direction for indoor images compared to outdoor images. Adapted from (Bischof et al., 2024)

Figure 13 shows the eyes-in-space histograms for the indoor and outdoor scenes. The standard deviation of the latitudes is signiticantly larger for the indoor panoramas (23.6°) than for the outdoor panoramas, and the standard deviation of the longitudes is marginally smaller for the indoor panoramas (85.9°) than for the outdoor panoramas (87.6°). The results show that the distribution of latitudes is consistent with the expectations from the edge analysis and, more generally, that eyes-in-space are affected by the spatial properties of the panoramas.

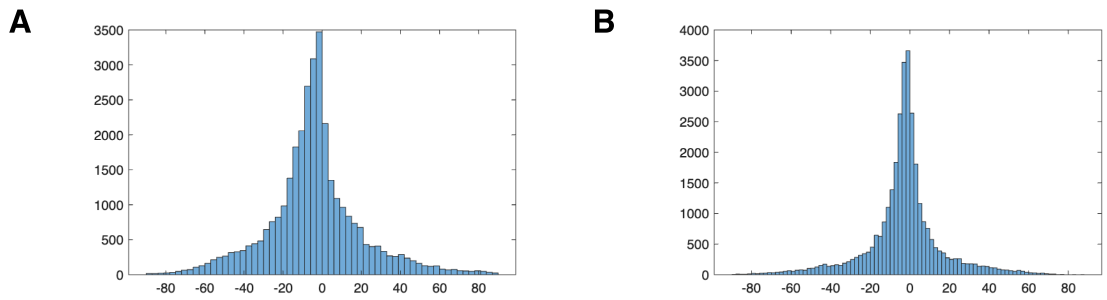


Fig. 13 A Histogram of eyes-in-space latitudes for indoor scenes. **B** Histogram of eyes-in-space latitudes for outdoor scenes. From (Bischof et al., 2024)

4.6 Summary

In this section, we introduced panoramic displays, a restricted type of VR environments where the observer cannot move within the environment. We discussed basic concepts and terms of eye movement analyses in panoramic displays and then presented results on eyes-in-head, head and eyes-in-space analyses. The interested reader is referred to Bischof et al. (2020, 2024) for more detailed information on these topics.

5 3D Virtual Environments

In the previous section, we presented analysis methods and results of eye tracking studies in a restricted virtual environment: In panoramic displays, the stimuli are projected onto a sphere centered at the head of the participant (see Fig. 5A). For this reason, head and body position remain irrelevant since they do not affect the view of the stimulus. In this section, we present VR experiments in a 3D world, where participants are allowed to move through and view the world from different positions, viewing angles and distances.

Locomotion through the VR world can be achieved using virtual locomotion techniques. Participants can walk virtually, either on a treadmill (Bashir et al., 2023) or by walking in place while the body motions are recorded using a motion capture system (e.g., Shin et al., 2024). Alternatively, locomotion can be achieved using a controller or joystick or using teleporting (see, e.g., Boletsis & Chasanidou, 2022). Finally, in gaze-based steering, a 2D gaze interface is superimposed on the virtual environment, and participants can steer through the environment by fixating different direction arrows in the gaze interface (Stellmach & Dachselt, 2012; Araujo et al., 2020). In the point and fly technique, participants can fly to a position in the virtual environment by gazing at it (Qian & Teather, 2018). For further information, the reader is advised to consult the reviews by Adhanom et al. (2023) and by Boletsis and Chasanidou (2022).

In panoramic displays, participants either passively observe the spherical environment or can be asked to actively search for certain objects. In 3D virtual environments, more general tasks can be implemented, including active navigation through the environment and interacting with objects by selecting or manipulating them, e.g., by turning on a switch, opening a door, and more.

Finally, 3D virtual environments provide distance cues in addition to those available in panoramic displays (see Section 4). Namely binocular distance cues (by interpreting the disparity between the retinal images in terms of depth), motion parallax (the apparent relative motion of several stationary objects against a background, providing information about their relative distance), and in the case of moving objects in the virtual environment, the kinetic depth effect (the movement of object parts or silhouettes, providing information for the perception of depth).

In the sections below, we show pilot results from a study in which participants walked through a 3D virtual environment while their eye and head movements were recorded. The environment was a medieval castle, built in the Unity game engine, in which participants completed a goal-directed exploration task ("Open a Treasure Chest"). Participants navigated through a series of interconnected rooms, searching for a key and using it to unlock a treasure chest. Along the way, they could interact with objects in the environment (e.g., picking up and using a sword). Locomotion was achieved primarily through virtual locomotion using the controller joysticks; teleportation was not available. Figure 14 shows the layout of the castle environment with user position overlaid.

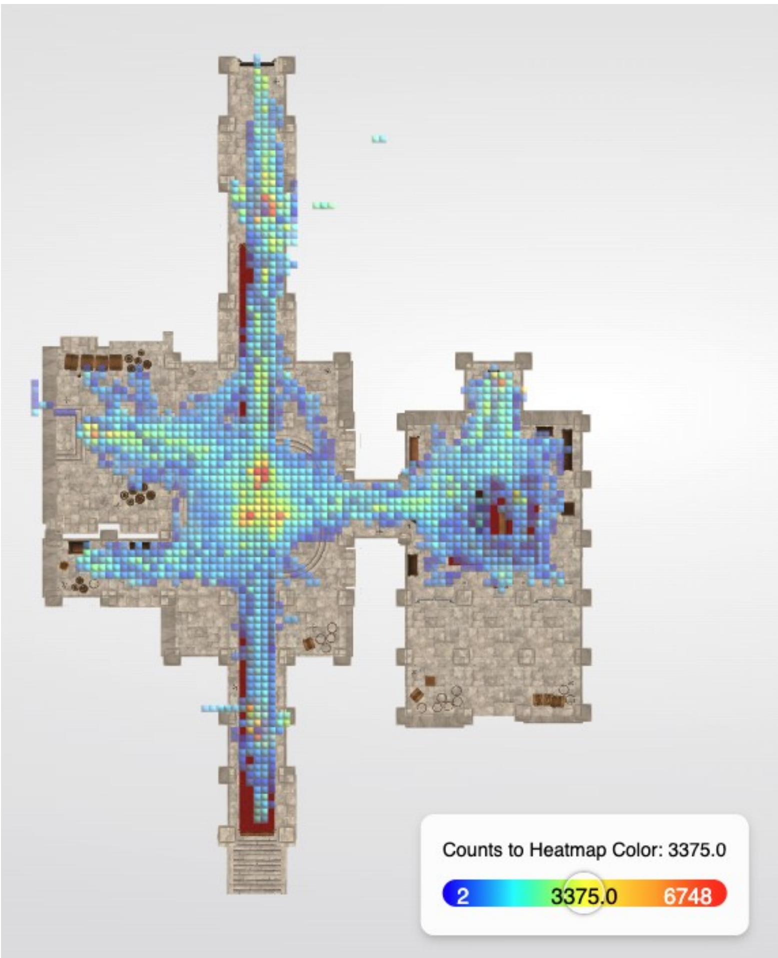


Fig. 14 Top-down view of the medieval castle environment with a heatmap of user position aggregated across participants and projected onto the scene geometry. The heatmap was generated by the Cognitive3D analytics platform, which reconstructs the 3D scene and maps

recorded head position to world coordinates. This view shows only the bottom 10% of the scene geometry (cutting off the roof of the building). Color indicates the number of position samples at each location, ranging from blue (low) through green and yellow (medium) to red (high), as shown in the legend. The heatmap reveals common navigation paths, with high-density areas visible along the main hallway and at key interaction locations where participants paused to complete task objectives (e.g., finding the key, opening the treasure chest).

Eye tracking data were collected using the HTC Vive Focus Vision head-mounted display with integrated eye tracking. Head position and rotation were recorded at 10 Hz, and eye tracking data were recorded at 60 Hz. Eye and head data streams were synchronized using timestamps within a unified recording pipeline (see Section 3.4.2). In the analyses that follow, we present representative results from a random subset of these data.

In the following sections, we show 1) eyes-in-head data, i.e., the direction of the eyes relative to the head; 2) head data, i.e., the direction of the head in world coordinates; and 3) eyes-in-space data, i.e., the direction of the eyes in world coordinates as obtained by combining the eyes-in-head coordinates with head coordinates. These analyses follow very closely those for the panoramic virtual environment. The we discuss gaze visualization in 3D environments and present our own visualization of fixation patterns in 3D space.

5.1 Eyes-in-Head Analysis

Figure 15 shows the heatmap of the eyes in head-centered coordinates. The heatmaps were obtained from the maps of eyes-in-head points using a Gaussian filter with $\sigma = 4°$ angular distance. The analysis of eye-in-head coordinates shows that the eyes are directed below the head-defined center of the visual field, but centered along the 0° meridian. In contrast to the heatmap in Fig. 6, the heatmap in Fig. 15 does not show a double peak that depends on the head movement. Most likely, this is due to a less systematic pattern of left-right head movements in the 3D virtual environments compared to panoramic environments.

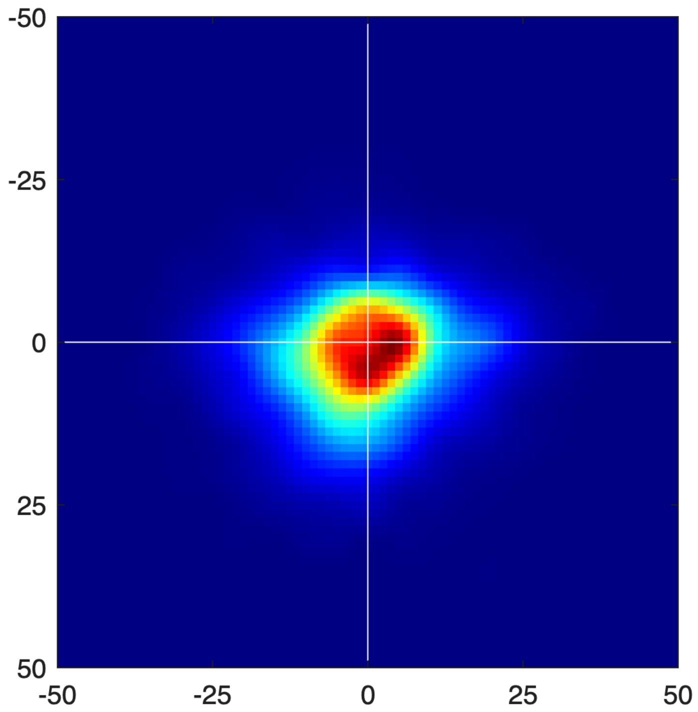


Fig. 15 Heatmaps of eyes-in-head points in the longitude range [-50°, 50°] and the latitude range [-50°, 50°], with the white lines showing 0° longitude and latitude.

5.2 Head and Eyes-in-Space Analysis

Figure 16A shows the heatmap of head points, and Figure 16B shows the heatmap of the eyes-in-head points. Both heatmaps were obtained from the maps of all head points and eyes-in-head points using a Gaussian filter with $\sigma = 4°$ angular distance at the equator and horizontally increasing by sec(latitude) towards north and south. Both heatmaps are biased below the equator line and biased east of the 0° meridian line, with hotspots near 0°, ±90° and ±180°.

A

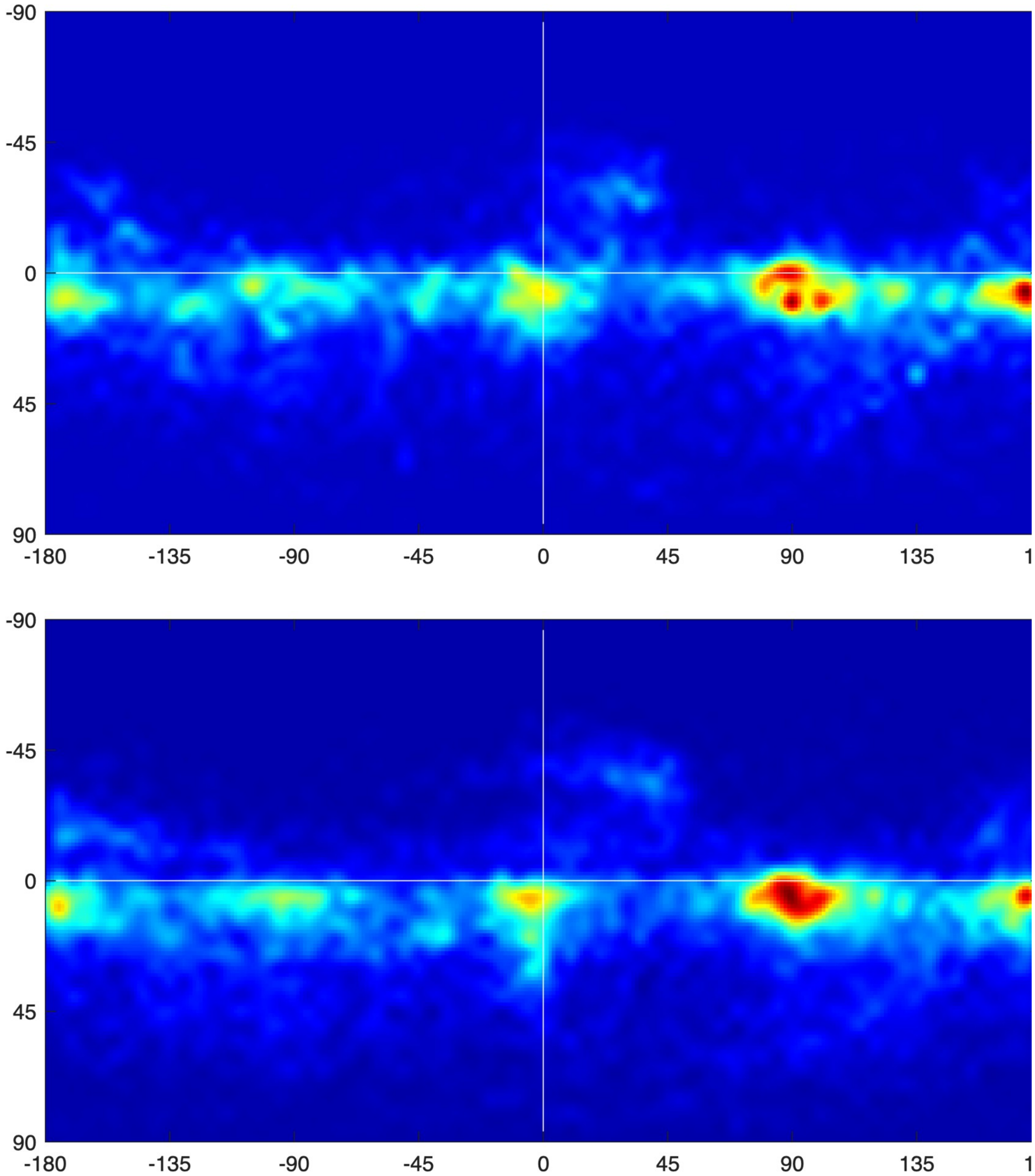


Fig. 16A Heatmap of head points in the longitude range [-180°, 180°] and the latitude range [-90°, 90°], with the white lines showing 0° longitude and latitude. **B** Heatmap of eyes-in-space points in the longitude range [-180°, 180°] and the latitude range [-90°, 90°], with the white lines showing 0° longitude and latitude.

A comparison of the head heatmaps for the panoramic environments (Figs. 8 and 10) and those for the 3D VR environments (Figs. 16A and 16B) reveals important differences between the two environments. With respect to latitudes, the heatmaps of the panoramic environments are close to the equator whereas those of the 3D VR environment are clearly below the equator. This may be attributed to the fact that in the 3D VR environment, participants frequently looked at the ground and interacted with objects in the lower visual field, in contrast to the panoramic environments, which were viewed passively. With respect to longitudes, there is a single hotspot near the 0° meridian for the panoramic heatmap, where participants started each trial. In contrast, the heatmaps of the 3D environments show hotspots near 0°, ±90° and ±180°. This is most likely attributable to the architecture of the 3D environment, which consisted primarily of long walls in the east-west and in the north-south direction (see Fig. 14). Despite these simple characteristics, the heatmaps of the 3D VR environment are difficult to interpret: Although the viewing direction is defined precisely, the maps do not take into account the position within a virtual environment from where the view is taken. In other words, while viewing direction is well defined, the viewer's position in the environment is not.

5.3 Visualization of Gaze in 3D VR environments

To overcome the problems described in the previous section, representations of gaze must be used that take the position of the viewer within the 3D VR environment into account. Crucially, these visualizations are defined with respect to the virtual environment rather than the viewer. Several methods have been developed to visualize gaze in 3D VR environments.

In the gaze point visualization method, fixation locations are defined as gaze points in the 3D environments, for example using spheres with the sphere centers at the fixation locations (e.g., Breen et al., 2021; Clay et al., 2019; Duchowski et al., 2002; Maurus et al., 2014; Shi et al., 2020). This visualization is independent of a particular point of view and can thus be viewed from any viewpoint. In addition, the sphere radius can be chosen to be proportional to the fixation duration.

Gaze patterns in 3D environments can also be visualized using heatmaps: For VR scenes with a single or very few objects, the heatmaps can be related directly to those objects. For example, Stellmach et al. (2010) overlaid heatmaps on top views of a set of objects, or projected them directly onto the objects' surfaces. Similarly, Bronstein et al. (2008) wrapped heatmaps onto single objects. A similar approach can be used in the case of complex virtual environments: Hagihara et al. (2018) had participants view a set of objects from above, and then pseudo-colored the objects depending on the fixation frequency. Maurus et al. (2014)

used spherical gaze fixations in scenes, and they projected fixation heatmaps directly onto the surfaces of the objects. Singh et al. (2018) also projected heatmaps onto sets of objects and complemented this visualization with quantitative analyses.

One can also use AOI-based methods (See Bischof et al., Chapter 13 in this volume): One can define particular areas of interest with respect to the 3D VR environment and describe the temporal fixation patterns of each AOI (see e.g., Richardson & Dale, 2005). For all visualization methods, the interested reader is referred to the reviews by Blascheck et al. (2017) and by Sundstedt and Garro (2022) who discuss in detail visualization techniques for 3D environments.

Below, we illustrate several of these visualization methods using the Cognitive3D analytics platform (https://cognitive3d.com; see also Section 3.4.2) applied to data collected in the castle environment described above. In this approach, the eye tracker's gaze direction is combined with head position and orientation to compute a gaze ray, which is cast into the 3D scene to determine the point on a surface where the participant was looking (a "gaze hit point"). Because each hit point is defined in world coordinates, it is inherently tied to the location in the virtual environment that was fixated, regardless of where the observer was physically standing. These hit points can then be aggregated across participants and time to generate surface-projected heatmaps, with color indicating fixation density at each location.

Figure 17 illustrates this approach. Gaze hit points are aggregated across participants and projected onto the surfaces of the environment, with color intensity indicating the number of fixations at each location. Because the heatmap is rendered within the 3D scene geometry, it can be rotated and inspected from any viewpoint, providing a spatial overview of where participants directed their attention.

A

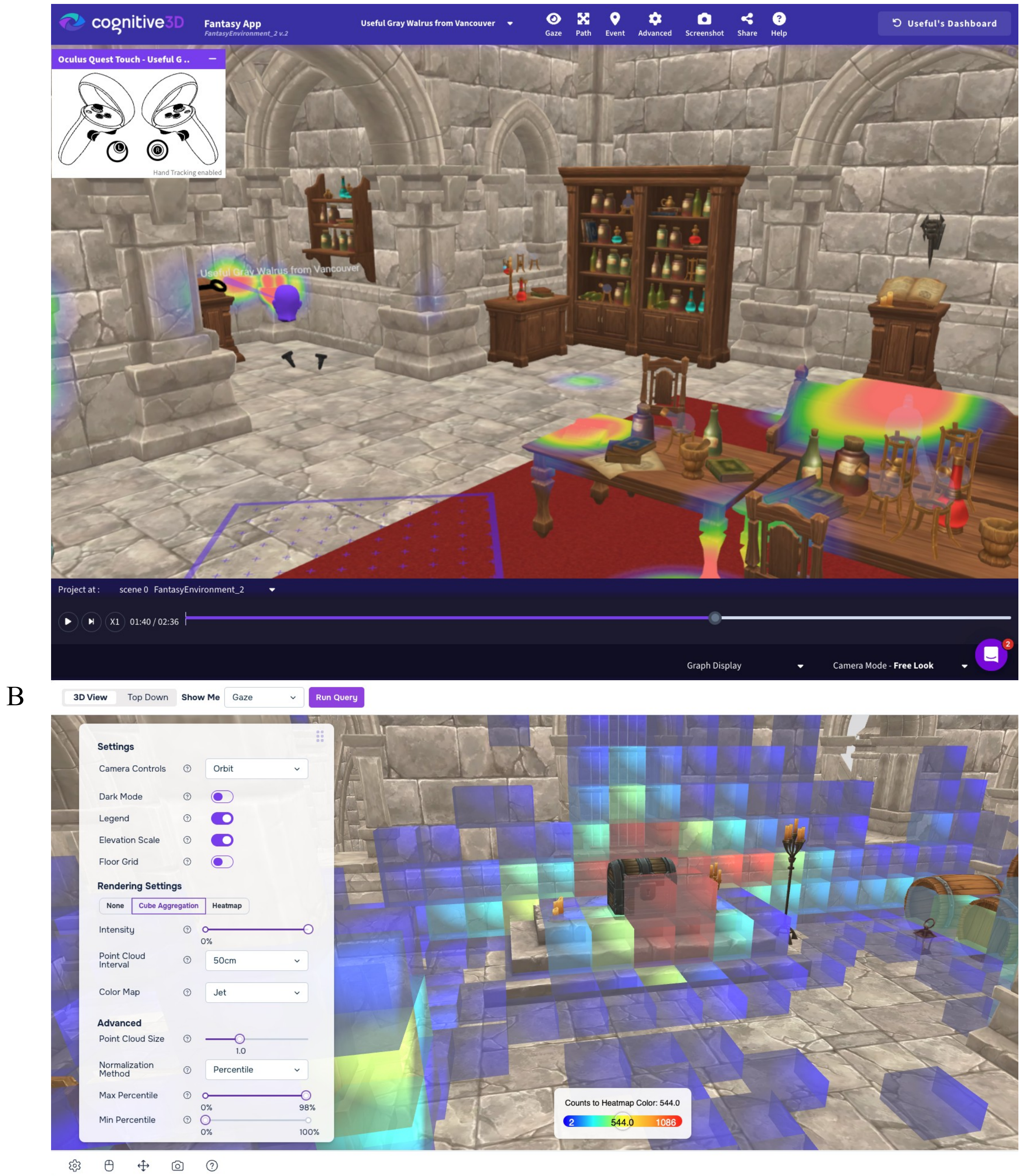


Fig. 17A Session replay of a single participant in the castle environment, viewed from a free-look camera perspective within the Cognitive3D analytics platform. The participant's head position is represented by the avatar, with gaze fixations visualized as a surface-projected heatmap (color indicating fixation density). Controller button presses are shown in the upper left. The timeline allows the session to be reviewed at any point, and the camera can be positioned freely within the 3D scene. **B** Aggregate gaze visualization across multiple participants in the same environment, using volumetric cube aggregation. Gaze hit points are binned into discrete spatial volumes (50 cm intervals) and displayed as colored cubes, with color indicating the number of fixations at each location, ranging from blue (low, 2) through yellow (medium) to red (high, 1086). The treasure chest, a key task objective, is visible at the center of a high-density cluster.

These tools complement the analytical approaches described in the preceding sections by providing an interactive, spatial representation of gaze data that can be shared among collaborators and inspected without requiring custom analysis scripts. For a comprehensive review of gaze visualization methods in 3D environments, the reader is again referred to Blascheck et al. (2017) and Sundstedt and Garro (2022).

5.4 Analysis of Eyes-in-Space in 3D VR environments

We return to the analysis of head and eyes-in-space presented in Section 5.2. Here, we present a top view of the 3D environment, and the map indicates all the locations the participants fixated in this environment (about 33,000 fixations). Locations with low saturation indicate one or very few fixations of that point, whereas locations with high saturation indicate that a point was fixated very often (Fig. 17A). This was achieved by plotting a scatter plot with each fixation shown using a semi-transparent disc.

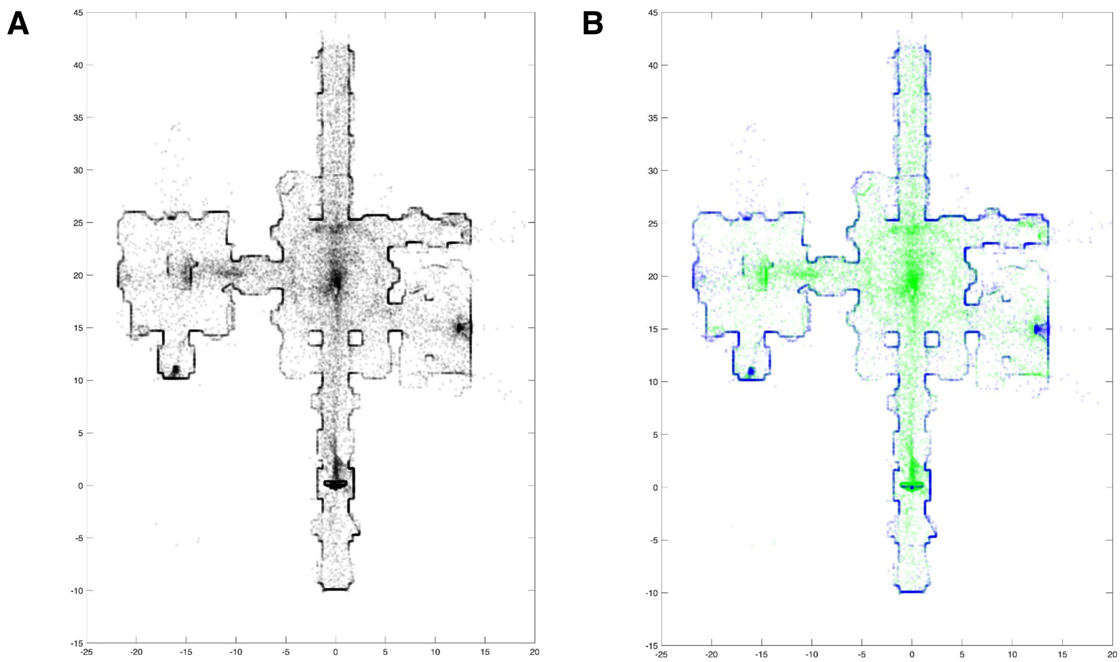


Fig. 18A Top view of a virtual 3D environment. Each dot indicates a fixation, with saturation indicating how often that location was fixated: Low saturation indicates one or very few fixations of that location, high saturation indicates that a location was fixated very often. **B** Same as A, but color of fixations depends on fixation direction: Green dots indicate fixations in the downward direction, i.e., on the floor and blue dots indicate fixations above ground, i.e., at the wall or at objects in the environment.

Fig. 18B shows a colored version with fixations on the ground in green and fixations above ground, e.g., walls and objects in blue. The blue contours indicate that participants fixated the walls of the environment very frequently, although there are obvious gaps, indicating that not every wall location in the environment was fixated. The green clusters indicate fixated locations on the ground. Fixations outside the 3D environment are due to eye tracking errors. Inside the walls, there are several interesting fixation clusters: The cluster at coordinates (0, 0) is the location where the participant spawned (i.e., their starting position in virtual space), the cluster at coordinates (0, 20) was a popular stopping location directly down the hall from the spawn point, where participants pulled a sword embedded in a stone. Other clusters are associated with key story points such as picking up a key used to unlock a chest.

The visualization presented in this section is somewhat similar to the one by Hagihara et al. (2018) presented in the previous section except that they applied it to a small number of objects. In contrast to the 3D visualizations discussed in the previous section, this visualization permits the construction of complete fixation maps in a top-down view or in fact in any other projection plane.

5.5 Summary

In this section, we introduced the measurement of eye movements in full 3D environments, where the observer is free to move around. A first analysis showed that the methods developed in Section 4 have limited applicability in 3D VR environments. We then reviewed techniques for visualizing eye movements in 3D environments and introduced our own visualization techniques. The interested reader is advised to consult reviews by Blascheck et al. (2017) and by Sundstedt and Garro (2022) who review visualization techniques for 3D environments.

6 Applications

VR applications are used across a diverse range of domains, e.g., attention research, neuroscience, clinical applications, engineering, architecture, and games. Below we discuss some of these applications that use eye tracking.

6.1 Attention Research

VR is an excellent design choice for behavioral experiments in general (e.g., Rizzo et al., 2021). Tracking the eyes in VR provides key advantages over traditional computer-based or mobile eye tracking research: Virtual environments are under the experimenter's control, yet ensure that the participants are fully immersed in the environment (Anderson et al., 2023; Clay et al., 2019; David et al., 2021, 2022). This includes acoustic signals, often spatially localized, which are delivered through the headset or via headphones. It is noteworthy that VR enables the simultaneous tracking of both the eye and head movement with regard to a common reference frame, allowing one to calculate and differentiate how different movement systems contribute to overall attentional control.

6.2 Clinical Applications

It is well known that eye movements can help in identifying psychiatric, neurological or ophthalmological disorders (e.g., Müri et al., Chapter 21 of this volume; Klein & Smyrnis, Chapter 20 of this volume; Trillenberg et al., 2004). It is also known that they can be used for developing complete diagnostic systems: Eye movements in VR can be used for the diagnosis of neurological disorders such as Parkinson's disease or Alzheimer's disease, given that abnormalities in eye movements can be detected long before other clinical symptoms are detected. Orlosky et al. (2017) achieved this by designing tasks that can elicit abnormal eye movements associated with these diseases. Specifically, they construct a 3D virtual environment designed to emulate common tasks used to evaluate patients, such as fixating a point, conducting smooth pursuit of an object, or executing saccades. These virtual tasks are designed to elicit eye movements commonly associated with neurodegenerative disease, such as abnormal saccades, square wave jerks, and ocular tremor. Tao et al. (2020) explored eye tracking in neurological disorders that are commonly associated with cognitive deficits, specifically, amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, multiple sclerosis, and epilepsy. They focused on the application of ocular measures in these disorders, with the goal of understanding how eye tracking technology can be used in the clinical setting.

VR applications can also be used to diagnose ophthalmological problems: In low vision patients, the assessment of contrast sensitivity is an essential tool for determining the stage of visual impairment. Traditional contrast sensitivity tests rely on verbal feedback, and the expertise of the examiner. In contrast, Tatiyosyan et al. (2020) developed a fast, OKN-based virtual diagnosis tool for estimating contrast sensitivity automatically without active cooperation of the patient or the practitioner. Strabismus is an eye movement disorder characterized by abnormal ocular deviation. Miao et al. (2020) developed an automated technique for measurement of ocular deviation using a VR device, in which the screens presenting the fixation target are alternately switched on and off to simulate standard strabismus diagnosis steps. Their results indicate that these techniques could identify ocular deviation with high accuracy and efficiency.

Finally, VR applications using eye movements are also useful for the diagnosis, as well as the treatment of social phobias. Grillon et al. (2006) presented a study with social phobic subjects, whose aim was to evaluate the efficiency and flexibility of virtual reality as a therapeutic tool in a social phobia behavioral therapeutic program. Their results show that the virtual reality exposure therapy could be successfully applied to social therapy

6.3 Consumer Research

Substantial financial resources are being put into research on advertising and consumer research using VR combined with eye tracking (see e.g., Rothensee & Reiter, Chapter 23 of this volume). For example, Meißner et al. (2019) designed a virtual supermarket with the goal of creating a realistic shopping experience. The project was aimed at designing and testing an assistance system that provided additional information in real time based on consumers' gaze behavior. In addition, they tested whether recommending products based on previous purchases influences consumers' choice behavior. The advantage of a VR study for this task was that shop and shelf layouts could be designed and changed easily, keeping other influences under control, in contrast to studies in a real supermarket, which require major efforts for redesigning layouts.

6.4 Architecture

Another obvious area of application is in engineering and architecture, where VR was applied to diverse areas from risk assessment to situational awareness in architectural design. Kang et al. (2020) simulated an environment where VR enabled them to evaluate trainees' and operators' situation awareness in a non-hazardous environment using eye movements, specifically visual scan paths. Their results show that the trainees, having lower situational awareness, tended to create random visual scan paths whereas the trainees with higher situational awareness tended to create concentrated and refined visual scan paths. In architectural design, Zhang et al. (2019) proposed an integrated approach to protect and design cityscapes based on virtual reality (VR) and eye tracking. Their research revealed the mechanism for identifying cityscape features, and discovered the differences in the perception of the cityscape features by different people, thus proposing a strategy for protecting cityscape. This research is relevant for building a human-oriented scientific planning and protection method and promoting the application of digital technology in the field of smart city governance.

6.5 Summary

In this section, we presented applications of eye tracking in four areas, attention research, clinical applications, consumer research and architecture. The interested reader is referred to the reviews by Adhanom et al. (2023) and Moreno-Arjonilla et al. (2024) for further applications of eye-tracking in VR.

7 General Summary

This chapter is aimed at introducing the reader to the use of eye tracking in virtual reality (VR) environments.

In Sections 1-3, we presented an overview of a diverse set of topics fundamental to the analysis of eye movements in VR: fundamentals of eye movement analysis, eye tracker hardware, head tracker hardware, software for the presentation of VR and the recording of eye movements, and finally, a description of the processing pipeline for the analysis of eye movement data.

In Section 4, we introduced panoramic displays, a restricted type of VR situation where the observer cannot move within the environment. We discussed basic concepts and terms of eye movement analyses in panoramic displays and then presented results on eyes-in-head, head, and eyes-in-space analyses.

In Section 5, we introduced the measurement of eye movements in full 3D environments, where the observer can move freely. The analytical methods presented in Section 4 were revealed to have limited utility in 3D VR environments. We then reviewed techniques for visualizing eye movements in 3D environments and introduced our own visualization technique.

In Section 6, we presented applications of VR eye tracking in four areas: attention research, clinical applications, consumer research, and architecture.

Overall, this chapter highlights the promise and complexity of eye tracking in VR environments. Future work will benefit from continued advances in hardware, analytical methods, and experimental design, enabling increasingly precise and ecologically valid investigations of human behavior.

8 Suggested Readings

Adhanom, I. B., MacNeilage, P., & Folmer, E. (2023). Eye Tracking in Virtual Reality: A Broad Review of Applications and Challenges. *Virtual Reality, 27*(2), 1481–1505.

This paper reviews recent publications dealing with eye tracking in VR and highlights their areas of application, challenges, and limitations.

Anderson, N. C., Bischof, W. F., & Kingstone, A. (2023). Eye Tracking in Virtual Reality. In: Maymon, C., Grimshaw, G., Wu, Y.C. (Eds) Virtual Reality in Behavioral Neuroscience: New Insights and Methods. Current Topics in Behavioral Neurosciences, Vol 65. Springer, Cham.

This paper presents research on eye tracking in VR, with sections on how to implement an experiment with a panoramic display in Unity, foundational ideas regarding data analysis, and techniques for analyzing VR eye tracking data.

Bischof, W. F., Anderson, N. C., & Kingstone, A. (2024). A Tutorial: Analysing Eye and Head Movements in Virtual Reality. Behavior Research Methods, 56, 8396-8421.

This tutorial discusses eye tracking technology, beginning with the analysis of head and eye movement data when an observer is situated in the center of a panoramic display, and then extending to spatio-temporal head-eye analysis techniques.

Sundstedt, V., & Garro, V. (2022). A Systematic Review of Visualization Techniques and Analysis Tools for Eye-Tracking in 3D Environments. *Frontiers in Neuroergonomics, 3*, 910019.
This review provides a systematic overview of recent developments in the areas of visualization techniques and analysis tools for eye movement data in 3D environments.

WorldViz (https://worldviz.com) has published several detailed guides on how to set up a VR lab, on budgeting guidelines, on how to write a grant application for a VR lab, as well as examples of many experiments in VR research. Anyone who is considering starting a VR study is advised to read their very helpful documents.

9 Questions to Students

1. Under what circumstances would smooth pursuit movements in VR have to be considered?
2. In what ways are temporal characteristics of eye movements important in VR applications?
3. Describe the key steps in a processing pipeline for VR eye movement data and explain why temporal and spatial synchronization are necessary.
4. How do eye and head movements relate in VR applications?
5. Why is it difficult to produce stereoscopic stimuli for panoramic virtual environments and how do 3D virtual environments address this limitation?
6. What depth cues are available in panoramic and in 3D virtual environments, and which cues are unique to each one?
7. Why do panoramic analysis methods not transfer directly to 3D virtual environments?
8. Choose one application area discussed in the chapter and explain how VR-based eye tracking provides advantages over traditional eye tracking methods in that domain.

Acknowledgements This work was supported by a Canadian Mitacs Accelerate grant awarded in partnership with Cognitive3D and the Brain, Attention, and Reality Lab at the University of British Columbia, as well as by an NSERC Discovery Grant to AK.

References

Adhanom, I. B., Lee, S. C., Folmer, E., & MacNeilage, P. (2020). GazeMetrics: An Open-Source Tool for Measuring the Data Quality of HMD-based Eye Trackers. *Proceedings. Eye Tracking Research & Applications Symposium*, *2020*, 19. https://doi.org/10.1145/3379156.3391374

Adhanom, I. B., MacNeilage, P., & Folmer, E. (2023). Eye Tracking in Virtual Reality: A Broad Review of Applications and Challenges. *Virtual Reality, 27*(2), 1481–1505. https://doi.org/10.1007/s10055-022-00738-z

Anderson, N. C., Bischof, W. F., & Kingstone, A. (2023). Eye Tracking in Virtual Reality. In: Maymon, C., Grimshaw, G., Wu, Y.C. (Eds) *Virtual Reality in Behavioral Neuroscience: New Insights and Methods. Current Topics in Behavioral Neurosciences*, *Vol. 65*. Springer.

Araujo, J. M., Zhang, G., Hansen, J. P. P., & Puthusserypady, S. (2020). Exploring eye-gaze wheelchair control. In: *Symposium on Eye Tracking Research and Applications*, Association for Computing Machinery, pp 1–8. https://doi.org/10.1145/3379157.3388933

Arnaldi, B., Guitton, P., & Moreau, G. (Eds.) (2018). *Virtual and Augmented Reality: Myths and Realities*, Wiley. http://doi.org/10.1002/9781119341031

Backhaus, D., Engbert, R., Rothkegel, L. O. M., & Trukenbrod, H. A. (2020). Task-dependence in scene perception: Head unrestrained viewing using mobile eye-tracking. *Journal of Vision*, 20(5), 3. https://doi.org/10.1167/jov.20.5.3

Barnes, G. R. (1979). Vestibulo-ocular function during co-ordinated head and eye movements to acquire visual targets. *Journal of Physiology*, 287, 127–147. https://doi.org/10.1113/jphysiol.1979.sp012650

Bashir, A., De Regt, T., & Jones, C.M. (2023). Comparing a friction-based uni-directional treadmill and a slip-style omni-directional treadmill on first-time HMD-VR user task performance, cybersickness, postural sway, posture angle, ease of use, enjoyment, and effort. *International Journal of Human-Computer Studies, 179*, 103101. https://doi.org/10.1016/j.ijhcs.2023.103101

Batschelet, E. (1981). Circular statistics in biology. Academic Press, 111 Fifth Ave., New York, NY 10003, 1981, 388.

Bebko, A. O., & Troje, N. (2020). bmlTUX: Design and control of experiments in virtual reality and beyond. PsyArXiv. https://doi.org/10.31234/osf.io/arvkf

Bischof, W. F., Anderson, N. C., Doswell, M. T., & Kingstone, A. (2020). Visual exploration of omnidirectional panoramic scenes. Journal of Vision, 20(7), 23. https://doi.org/10.1167/jov.20.7.23

Bischof, W. F., Anderson, N. C., & Kingstone, A. (2023). Eye and Head Movements while Encoding and Recognizing Panoramic Scenes in Virtual Reality. *PLOS One,* 18(2), e0282030. https://doi.org/10.1371/journal.pone.0282030

Bischof, W. F., Anderson, N. C., & Kingstone, A. (2024). A Tutorial: Analysing Eye and Head Movements in Virtual Reality. Behavior Research Methods, 56, 8396-8421. https://doi.org/10.3758/s13428-024-02482-5

Blascheck, T., Kurzhals, K., Raschke, M., Burch, M., Weiskopf, D., & Ertl, T. (2017). Visualization of eye tracking data: a taxonomy and survey. *Computer Graphics Forum, 36*, 260–284. https://doi.org/10.1111/cgf.13079

Boletsis, C., & Chasanidou, D. (2022). A typology of virtual reality locomotion techniques. *Multimodal Technologies and Interaction, 6*, 72. https://doi.org/10.3390/mti6090072

Breen, M., McClarty, J., Langley, C., Farzidayeri, J., Trevethan, K., Swenson, B., … & Sarkar, N. (2021). 2D and 3D visualization of eye gaze patterns in a VR-based job interview simulator: application in educating employers on the gaze patterns of autistic candidates. In Antona, M. & Stephanidis, C. (Eds.), *Universal Access in Human-Computer Interaction. Design Methods and User Experience, HCII 2021*. Lecture Notes in Computer Science, vol 12768, Springer, pp. 533–544. https://doi.org/10.1007/978-3-030-78092-0_36

Brockwell, H. (2016, April 3). Forgotten genius: the man who made a working VR machine in 1957. Techradar. https://www.techradar.com/news/wearables/forgotten-genius-the-man-who-made-a-working-vr-machine-in-1957-1318253

Bronstein, A., Bronstein, M., & Kimmel, R. (2008). *Numerical Geometry of Non-Rigid Shapes*, Springer. https://doi.org/10.1007/978-0-387-73301-2_1

Brookes, J., Warburton, M., Alghadier, M., Mon-Williams, M., & Mushtaq, F. (2020). Studying human behavior with virtual reality: The Unity Experiment Framework.

Behavior Research Methods, 52(2), 455–463. https://doi.org/10.3758/s13428-019-01242-0

Caldara, R., & Miellet, S. (2011). iMap: A Novel Method for Statistical Fixation Mapping of Eye Movement data. *Behavior Research Methods, 43*(3), 864-878. https://doi.org/10.3758/s13428-011-0092-x

Chen, S., Wang, W., & van Zuylen, H. (2010). Comparison of outlier detection algorithms for ITS data, *Expert Systems with Applications, 37*, 1169–1178. https://doi.org/10.1016/j.eswa.2009.06.008

Clay, V., König, P., & König, S. (2019). Eye tracking in virtual reality. *Journal of Eye Movement Research, 12*(1), 3. https://doi.org/10.16910/jemr.12.1.3

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., & Hart, J. C. (1992). The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM, 35*(6), 64-72. https://doi.org/10.1145/129888.129892

David, E. J., Beitner, J., & Võ, M. L.-H. (2020). Effects of Transient Loss of Vision on Head and Eye Movements during Visual Search in a Virtual Environment. *Brain Sciences, 10*(11), 841. https://doi.org/10.3390/brainsci10110841

David, E. J., Beitner, J., & Võ, M. L.-H. (2021). The importance of peripheral vision when searching 3D real-world scenes: A gaze-contingent study in virtual reality. *Journal of Vision, 21*(7), 3–3. https://doi.org/10.1167/jov.21.7.3

David, E. J., Lebranchu, P., Perreira Da Silva, M., & Le Callet, P. (2022). What are the visuo-motor tendencies of omnidirectional scene free-viewing in virtual reality? *Journal of Vision, 22*(4), 12. https://doi.org/10.1167/jov.22.4.12

Delreux, V., Abeele, S. V., Lefevre, P., & Roucoux, A. (1991). Eye-head coordination: Influence of eye position on the control of head movement amplitude. In J. Paillard (Ed.), Brain and Space (pp. 38–48). Oxford University Press.

Doshi, A., & Trivedi, M. M. (2012). Head and eye gaze dynamics during visual attention shifts in complex environments. *Journal of Vision*, 12(2), 9. https://doi.org/10.1167/12.2.9

Duchowski, A. T. (2017). *Eye Tracking Methodology: Theory and Practice* (3rd ed.). Springer. https://doi.org/10.1007/978-3-319-57883-5

Duchowski, A. T., Medlin, E., Cournia, N., Gramopadhye, A., Melloy, B., & Nair, S. (2002). 3D eye movement analysis for VR visual inspection training. In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications, ETRA '02*, Association for Computing Machinery, pp. 103–110. https://doi.org/10.1145/507072.507094

Efron, B., & Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Chapman & Hall. https://doi.org/10.1201/9780429246593

Equirectangular Projection. (n. d.). In *Wikipedia*.

https://en.wikipedia.org/wiki/Equirectangular_projection

Fisher, N. I., Lewis, T., & Embleton, B. J. J. (2010). *Statistical Analysis of Spherical Data*. Cambridge University Press. https://doi.org/10.1017/CBO9780511623059

Flindall, J., Sara, A., & Kingstone, A. (2021). Head and eye movements are each facilitated by the offset of a central fixation point in a virtual gap paradigm. *Experimental Brain Research*, 239, 197–208. https://doi.org/10.1007/s00221-020-05905-9

Foulsham, T., & Kingstone, A. (2017). Are fixations in static natural scenes a useful predictor of attention in the real world? *Canadian Journal of Experimental Psychology*, 71(2), 172–181. https://doi.org/10.1037/cep0000125

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research, 51*(17), 1920–1931. https://doi.org/10.1016/j.visres.2011.07.002

Freedman, E. G. (2008). Coordination of the eyes and head during visual orienting. *Experimental Brain Research*, 190(4), 369–387. https://doi.org/10.1007/s00221-008-1504-8

Freedman, E. G., & Sparks, D. L. (1997). Eye-head coordination during head-unrestrained gaze shifts in rhesus monkeys. *Journal of Neurophysiology*, 77(5), 2328–2348. https://doi.org/10.1152/jn.1997.77.5.2328

Friston, K. J., Jezzard, P., & Turner, R. (1994). Analysis of Functional MRI Time-Series, *Human Brain Mapping, 1*, 153-171.

Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J., & Turner, R. (1995). Analysis of fMRI Time-Series Revisited, *Neuroimage, 2*, 45-53.

Fuller, J. H. (1992). Head movement propensity. *Experimental Brain Research*, 92(1), 152–164. https://doi.org/10.1007/BF00230391

Goldring, J. E., Dorris, M. C., Corneil, B. D., Ballantyne, P. A., & Munoz, D. P. (1996). Combined eye-head gaze shifts to visual and auditory targets in humans. *Experimental Brain Research*, 111(1), 68–78. https://doi.org/10.1007/BF00229557

Grafarend, E. W., You, R.-J., & Syffus, R. (2014). *Map Projections: Cartographic Information Systems* (2nd ed.). Springer. https://doi.org/10.1007/978-3-642-36494-5

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley.

Grillon, H., Riquier, F. F., Herbelin, B., Thalmann, D., & Riquier, F. F. (2006) Virtual reality as a therapeutic tool in the confines of social anxiety disorder treatment. *International Journal on Disability and Human Development, 5*(3), 243–250. https://doi.org/10.1515/IJDHD.2006.5.3.243

Hagihara, K., Taniguchi, K., Abibouraguimane, I., Itoh, Y., Higuchi, K., Otsuka, J., Sugimoto, M., & Sato, Y. (2018). Object-wise 3D gaze mapping in physical workspace. In *Proceedings of the 9th Augmented Human International Conference, AH '18*, Association for Computing Machinery. https://doi.org/10.1145/3174910.3174921

Haskins, A.J., Mentch, J., Botch, T.L., & Robertson, C. E. (2020). Active vision in immersive, 360° real-world environments. *Scientific Reports, 10*, 14304. https://doi.org/10.1038/s41598-020-71125-4

Harezlak, K., Kasprowski, P., & Stasch, M. (2014). Towards accurate eye tracker calibration – methods and procedures, *Procedia Computer Science, 35*, 1073 – 1081. https://doi.org/10.1016/j.procs.2014.08.194

Heilig, M. L. (1992). EL Cine del Futuro: The Cinema of the Future. *Presence: Teleoperators and Virtual Environments, 1*(3), 279–294. Reprinted from Espacios 23-24, 1955. https://doi.org/10.1162/pres.1992.1.3.279

Hessels, R. S., Niehorster, D. C., Nyström, M., Andersson, R., & Hooge, I. T. (2018). Is the eye-movement field confused about fixations and saccades? A survey among 124 researchers. *Royal Society Open Science, 5*(8), 180502. https://doi.org/10.1098/rsos.180502

Holmqvist, K., & Andersson, R. (2017). *Eye tracking: A comprehensive guide to methods. Paradigms and Measures*. Oxford University Press.

Holmqvist, K., Nyström, M., & Mulvey, F. (2012). Eye tracker data quality: What it is and how to measure it. In *Proceedings of the Symposium on Eye Tracking Research and Applications (ETRA '12)* (pp. 45–52). ACM. https://doi.org/10.1145/2168556.2168563

Hooge, I. T., Hessels, R. S., Niehorster, D. C., Diaz, G. J., Duchowski, A. T., & Pelz, J. B. (2019). From lab-based studies to eye-tracking in virtual and real worlds: Conceptual and methodological problems and solutions. *Symposium 4 at the 20th European Conference on Eye Movement Research (ECEM) in Alicante*. https://doi.org/10.16910/jemr.12.7.8

Hornsey, R.L., & Hibbard, P.B. (2021). Contributions of pictorial and binocular cues to the perception of distance in virtual reality. *Virtual Reality, 25*, 1087–1103 (2021). https://doi.org/10.1007/s10055-021-00500-x

Hu, Z., Bulling, A., Li, S., & Wang, G. (2023). EHTask: Recognizing User Tasks from Eye and Head Movements in Immersive Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics, 29*(4), 1992–2004. https://doi.org/10.1109/TVCG.2021.3138902

Jacobs, O., Anderson, N. C., Bischof, W. F., & Kingstone, A. (2020). Into the unknown: Head-based selection is less dependent on peripheral information than gaze-based selection in 360-degree virtual reality scenes. PsyArXiv. https://doi.org/10.31234/osf.io/2qtcw

Jin, X., Chai, S., Tang, J., Zhou, X., & Wang, K. (2024). Eye-tracking in AR/VR: A technological review and future directions. *IEEE Open Journal on Immersive Displays, 1*, 146-154, 2024. https://doi.org/10.1109/OJID.2024.3450657

Judd, T., Ehinger, K., Durand, F., & Torralba, A. (2009). Learning to predict where humans look. Paper presented at the IEEE International Conference on Computer Vision (ICCV), Kyoto, Japan. pp. 2106-2113. https://doi.org/10.1109/ICCV.2009.5459462

Kang, Z., Jeon, J., & Salehi, S. (2020). Eye tracking data analytics in virtual reality training: Application in Deepwater Horizon oil drilling operation. *Proceedings of the Human*

Factors and Ergonomics Society Annual Meeting, *64*(1), 821-825. https://doi.org/10.1177/1071181320641191

Kingstone, A., Smilek, D., & Eastwood, J. D. (2008). Cognitive ethology: A new approach for studying human cognition. *British Journal of Psychology*, *99*(3), 317–340. https://doi.org/10.1348/000712607X251243

Komogortsev, O. V., Gobert, D. V., Jayarathna, S., & Gowda, S. M. (2010). Standardization of automated analyses of oculomotor fixation and saccadic behaviors. *IEEE Transactions on Biomedical Engineering, 57*(11), 2635–2645. https://doi.org/10.1109/TBME.2010.2057429

Kullback, S., & Leibler, R.A. (1951). On information and sufficiency. *Annals of Mathematical Statistics, 22* (1): 79–86. https://doi.org/10.1214/aoms/1177729694

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research*, 41(25–26), 3559–3565. https://doi.org/10.1016/S0042-6989(01)00102-X

Landes, J., Köppl, S., & Klettke, M. (2024). Data Processing Pipeline for Eye-Tracking Analysis. *35th GI-Workshop on Foundations of Databases (Grundlagen von Datenbanken), May 22-24, 2024, Herdecke, Germany.*

Lapaine, M., & Usery, E. L. (2017). *Choosing a map projection*. Lecture notes in geoinformation and Cartography. Springer. https://doi.org/10.1007/978-3-319-51835-0

Le Meur, O., & Baccino, T. (2013). Methods for comparing scanpaths and saliency maps: strengths and weaknesses. *Behavior Research Methods, 45*, p. 251-266. https://doi.org/10.3758/s13428-012-0226-9

Mardia, K. V., & Jupp, P. E. (2000). *Directional Statistics*. Wiley. https://doi.org/10.1002/9780470316979

Maurus, M., Hammer, J. H., & Beyerer, J. (2014). Realistic heatmap visualization for interactive analysis of 3D gaze data. In *Proceedings of the Symposium on Eye Tracking Research and Applications, ETRA '14*, Association for Computing Machinery, pp. 295–298. https://doi.org/10.1145/2578153.2578204

Mehrotra, A., Silver, C., Bischof, W. F., & Kingstone, A. (2025). Sit, stand, and swivel: Posture affects visual exploration of panoramic scenes in virtual reality. *Plos One, 20*(10), e0334182. https://doi.org/10.1371/journal.pone.0334182

Miao, Y., Jeon, J. Y., Park, G., Park, S. W., & Heo, H. (2020) Virtual reality-based measurement of ocular deviation in strabismus. *Computer Methods and Programs in Biomedicine, 185*, 105132. https://doi.org/10.1016/j.cmpb.2019.105132

Meißner, M., Pfeiffer, J., Pfeiffer, T., & Oppewal, H. (2019). Combining virtual reality and mobile eye tracking to provide a naturalistic experimental environment for shopper research. *Journal of Business Research, 100,* 445–458. https://doi.org/10.1016/j.jbusres.2017.09.028

Miller, H. L., Zurutuza, I. R., Fears, N. E., Polat, S. O., & Nielsen, R. D. (2021). Post-processing integration and semi-automated analysis of eye-tracking and motion-capture data obtained in immersive virtual reality environments to measure visuomotor integration. *Proceedings. Eye Tracking Research & Applications Symposium*, *17*. https://doi.org/10.1145/3450341.3458881

Moreno-Arjonilla, J., López-Ruiz, A., Jiménez-Pérez, J. R., Callejas-Aguilera, J. E., & Jurado, J. M. (2024). Eye-tracking on virtual reality: a survey. *Virtual Reality*, *28*(1). https://doi.org/10.1007/S10055-023-00903-Y

Niehorster, D. C., Li, L., & Lappe, M. (2017). The accuracy and precision of position and orientation tracking in the HTC Vive virtual reality system for scientific research. *I-Perception, 8*(3), 1–23. https://doi.org/10.1177/2041669517708205

Niehorster, D. C., Santini, T., Hessels, R. S., Hooge, I. T. C., Kasneci, E., & Nyström, M. (2020). The impact of slippage on the data quality of head-worn eye trackers. *Behavior Research Methods*, *52*(3), 1140–1160. https://doi.org/10.3758/s13428-019-01307-0

Ohayon, S., & Rivlin, E. (2006). Robust 3d head tracking using camera pose estimation. 18th International Conference on Pattern Recognition (ICPR'06), 1, 1063–1066. pp. 1063-1066, https://doi.org/10.1109/ICPR.2006.999

Orlosky, J., Itoh, Y., Ranchet, M., Kiyokawa, K., Morgan, J., & Devos, H. (2017). Emulation of physician tasks in eye-tracked virtual reality for remote diagnosis of neurodegenerative disease. *IEEE Transactions on Visualization and Computer Graphics. 23*(4), 1302–1311. https://doi.org/10.1109/TVCG.2017. 2657018

Pai, Y.S., Dingler, T. & Kunze, K. (2019). Assessing hands-free interactions for VR using eye gaze and electromyography. *Virtual Reality, 23*, 119–131. https://doi.org/10.1007/s10055-018-0371-2

Pelz, J., Hayhoe, M., & Loeber, R. (2001). The coordination of eye, head, and hand movements in a natural task. *Experimental Brain Research, 139*(3), 266–277. https://doi.org/10.1007/s002210100745

Qian, Y. Y., & Teather, R. J. (2018) Look to go: an empirical evaluation of eye-based travel in virtual reality. In *SUI 2018 – proceedings of the symposium on spatial user interaction*, https://doi.org/10.1145/3267782. 3267798.

Rai, Y., Gutiérrez, J., & Le Callet, P. (2017). A dataset of head and eye movements for 360 degree images. In *Proceedings of the 8th ACM Multimedia Systems Conference (MMSys '17)* (pp. 205–210). ACM. https://doi.org/10.1145/3083187.3083218

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science, 29*, 1045–1060. https://doi.org/10.1207/s15516709cog0000_29

Risko, E. F., Richardson, D. C., & Kingstone, A. (2016). Breaking the fourth wall of cognitive science: Real-world social attention and the dual function of gaze. *Current Directions in Psychological Science*, 25(1), 70–74. https://doi.org/10.1177/0963721415617806

Rizzo, A.S., Chen, J.Z., Wang, J., Ma, A., Chang, C-Y., Turnbull, J. & Shao, C._ (2021). Normative Data for a Next Generation Virtual Classroom for Attention Assessment in Children with ADHD and Beyond! *Proceedings of the International Conference on Disability, Virtual Reality, and Associated Technologies*

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*, 71–78.

Shi, Y., Du, J., & Ragan, E. (2020). Review visual attention and spatial memory in building inspection: toward a cognition-driven information system. *Advances in Engineering Informatics, 44*, 101061. https://doi.org/10.1016/j.aei.2020.101061

Shin, R., Choi, B., Choi, S.-M., & Lee, S. (2024). Implementation and Evaluation of Walk-in-Place Using a Low-Cost Motion-Capture Device for Virtual Reality Applications. *Sensors, 24*, 2848. https://doi.org/10.3390/s24092848

Singh, K., Kalash, M., & Bruce, N. (2018). Capturing real-world gaze behaviour: Live and unplugged. In *Proceedings of the 2018 ACM Symposium on EyeTracking Research & Applications, ETRA '18*. Association or Computing Machinery. https://doi.org/10.1145/3204493.3204528

Sitzmann, V., Serrano, A., Pavel, A., Agrawala, M., Gutierrez, D., Masia, B., & Wetzstein, G. (2018). Saliency in VR: How Do People Explore Virtual Environments? *IEEE Transactions on Visualization and Computer Graphics, 24*(4), 1633–1642. https://doi.org/10.1109/TVCG.2018.2793599

Solman, G. J. F., Foulsham, T., & Kingstone, A. (2017). Eye and head movements are complementary in visual selection. *Royal Society Open Science*, 4(1), 160569. https://doi.org/10.1098/rsos.160569

Solman, G. J. F., & Kingstone, A. (2014). Balancing energetic and cognitive resources: Memory use during search depends on the orienting effector. *Cognition*, 132(3), 443–454. https://doi.org/10.1016/j.cognition.2014.05.005

Stahl, J. S. (2001). Eye-head coordination and the variation of eye-movement accuracy with orbital eccentricity. *Experimental Brain Research*, 136(1), 81–100. https://doi.org/10.1007/s002210000593

Stellmach, S., & Dachselt, R. (2012). Designing gaze-based user interfaces for steering in virtual environments. *Proceedings of the symposium on eye tracking research and applications - ETRA'12*, Association for Computing Machinery, p 131. https://doi.org/10.1145/2168556.2168577

Stellmach, S., Nacke, L., & Dachselt, R. (2010). Advanced gaze visualizations for three-dimensional virtual environments. In *Proceedings of the 2010 Symposium on Eye-Tracking Research and Applications, ETRA'10*, Association for Computing Machinery, pp. 109–112. https://doi.org/10.1145/1743666.1743693

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: a review. *Journal of Vision, 11*(5), 13. https://doi.org/10.1167/11.5.13

Sundstedt, V., & Garro, V. (2022). A Systematic Review of Visualization Techniques and Analysis Tools for Eye-Tracking in 3D Environments. *Frontiers in Neuroergonomics, 3*, 910019. https://doi.org/10.3389/fnrgo.2022.910019

Sutherland, I.E. (1968). A head-mounted three dimensional display. *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I on - AFIPS '68*. https://doi.org/10.1145/1476589.1476686

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, 7(14), 4. https://doi.org/10.1167/7.14.4

Tao, L., Wang, Q., Liu, D., Wang, J., Zhu, Z., & Feng, L. (2020) Eye tracking metrics to screen and assess cognitive impairment in patients with neurological disorders. *Neurological Sciences, 41*(7), 1697–1704. https://doi.org/10.1007/s10072-020-04310-y

Tatiyosyan, S. A., Rifai, K., & Wahl, S. (2020). Standalone cooperation-free OKN-based low vision contrast sensitivity estimation in VR - a pilot study. *Restorative Neurology and Neuroscience, 38*(2), 119–129. https://doi.org/10.3233/RNN-190937

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*(4), 766–786. https://doi.org/10.1037/0033-295X.113.4.766

Trillenberg, P., Lencer, R., & Heide, W. (2004). Eye movements and psychiatric disease. *Current Opinion in Neurology, 17*(1), 43–47. https://doi.org/10.1097/00019052-200402000-00008

Valenti, R., Sebe, N., & Gevers, T. (2012). Combining head pose and eye location information for gaze estimation. *IEEE Transactions on Image Processing, 21*(2), 802–815. https://doi.org/10.1109/TIP.2011.2162740

Vasser, M., Kängsepp, M., Magomedkerimov, M., Kilvits, K., Stafinjak, V., Kivisik, T., Vicente, R., & Aru, J. (2017). VREX: An open-source toolbox for creating 3D virtual reality experiments. *BMC Psychology, 5*(1), 4. https://doi.org/10.1186/s40359-017-0173-4 von Noorden, G. K., & Campos, E. C. (2002). *Binocular Vision and Ocular Motility: Theory and Management of Strabismus* (6th ed.). Mosby. https://doi.org/10.1001/archopht.120.5.677

v. Rötth, A. (1925). Über das praktische Blickfeld. *Graefes Archiv für Ophthalmologie, 115*, 314–321. https://doi.org/10.1007/BF02283709

Watson, M. R., Voloh, B., Thomas, C., Hasan, A., & Womelsdorf, T. (2019). USE: An integrative suite for temporally-precise psychophysical experiments in virtual environments for human, nonhuman, and artificially intelligent agents. *Journal of Neuroscience Methods, 326*, 108374. https://doi.org/10.1016/j.jneumeth.2019.108374

Zangemeister, W. H., & Stark, L. (1982). Types of gaze movement: Variable interactions of eye and head movements. *Experimental Neurology, 77*(3), 563–577. https://doi.org/10.1016/0014-4886(82)90228-X

Zhang, H., & Jonides, J. (2026). PupEyes: An interactive Python library for eye movement data processing. *Behavior Research Methods, 58*, 29. https://doi.org/10.3758/s13428-025-02830-z

Zhang, L.-M., Zhang, R.-X., Jeng, T.-S., & Zeng, Z.-Y. (2019). Cityscape protection using VR and eye tracking technology. *Journal of Visual Communication and Image Representation, 64*, 102639. https://doi.org/10.1016/j.jvcir.2019.102639

Zimmermann, T. G., Lanier, J., Blanchard, D., Bryson, S., & Harvil, Y. (1986). A hand gesture interface device. *SIGCHI bulletin, 17, Issue SI.*