ORIGINAL MANUSCRIPT

A tutorial: Analyzing eye and head movements in virtual reality

Walter F. Bischof[1] · Nicola C. Anderson[1] · Alan Kingstone[1]

Abstract

This tutorial provides instruction on how to use the eye tracking technology built into virtual reality (VR) headsets, emphasizing the analysis of head and eye movement data when an observer is situated in the center of an omnidirectional environment. We begin with a brief description of how VR eye movement research differs from previous forms of eye movement research, as well as identifying some outstanding gaps in the current literature. We then introduce the basic methodology used to collect VR eye movement data both in general and with regard to the specific data that we collected to illustrate different analytical approaches. We continue with an introduction of the foundational ideas regarding data analysis in VR, including frames of reference, how to map eye and head position, and event detection. In the next part, we introduce core head and eye data analyses focusing on determining where the head and eyes are directed. We then expand on what has been presented, introducing several novel spatial, spatio-temporal, and temporal head–eye data analysis techniques. We conclude with a reflection on what has been presented, and how the techniques introduced in this tutorial provide the scaffolding for extensions to more complex and dynamic VR environments.

Keywords Eye movements · Head movements · Virtual reality · Spatial analysis · Temporal analysis · Head–eyes relationship

Introduction

Historically, conducting eye-tracking research was a specialized, expensive, and solitary affair. Investment in an eye tracker could easily cost a researcher tens of thousands of dollars, and the participants would be required to sit alone, often in the dark, looking at simple images displayed on an oscilloscope. Fast forward to today, where the quality of cameras is exceptional, their availability widespread, and the cost affordable. While there are still specialized high-end desktop eye trackers, the opportunity to track the eyes of observers is ubiquitous. This is especially noticeable in virtual reality (VR) headsets equipped with built-in eye tracking hardware. These headsets give both beginners and experts alike the opportunity to track the head and eyes of observers as they explore environments that are ecologically valid or otherworldly.

A key element to making the best of these opportunities is to know how to analyze the data one collects. The complexity of this issue is compounded by the fact that the eye data collected in virtual systems requires one to factor in the movement of an observer's head resulting from their visual exploration of scenes as well as the movement of other parts of their body, such as the torso and feet. This presents a data-handling problem that has rarely been encountered because historically eye movements were collected from observers who were seated and instructed to keep their head and body perfectly still, often with the aid of a chinrest or some form of head restraint.

The aim of the present tutorial is to introduce different ways to analyze the head and eyes of observers in VR. To assist the many researchers who are interested in – but do not know how to use – the eye tracking technology built into modern VR headsets, we collect and analyze data that connects to and bridges naturally between (a) the desktop head-fixed eye-tracking studies that are relatively commonplace in the literature, and (b) the immersive omnidirectional VR

✉ Walter F. Bischof
wfb@ualberta.ca

Nicola C. Anderson
nccanderson@gmail.com

Alan Kingstone
alan.kingstone@ubc.ca

1 Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC V6T 1Z4, Canada

environment that allows the head as well as the eyes to move freely. For readers interested in learning how to program studies and extract VR data, a recent resource is Anderson et al. (2023). We hope and expect that the approaches to data analysis that we present here will arm users with the skills to capitalize on the many exciting research opportunities in VR that will emerge in the future.

This tutorial is broken down into six distinct sections. In the first section, we briefly describe how VR eye movement research differs from previous forms of eye movement research. The second section introduces the basic methodology used to collect VR eye movement data both in general and with regard to the specific data that we collected to illustrate the different analytical approaches we apply later in the tutorial. The third section introduces the foundational ideas regarding data analysis in VR, including frames of reference, how to map eye and head position, and event detection. The fourth section introduces core head and eye data analyses focusing on determining where the head and eyes are directed. The fifth section expands on what has been presented, introducing several novel spatial, spatio-temporal, and temporal head-eye data analysis techniques. In the final section, we briefly reflect on what has been presented, and how the techniques can be employed to research questions across a broad range of situations.

A brief introduction to VR eye movement research

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 held steady. There is growing recognition, however, that normal looking behavior involves the movement of the head (and body) as well as the eyes (e.g., Backhaus et al., 2020; Foulsham et al. 2011; Foulsham & Kingstone, 2017; Hooge et al., 2019; Kingstone et al., 2008; Land & Hayhoe, 2001; Land & Tatler, 2009; Risko et al., 2016; 't Hart et al., 2009). While these and other studies have addressed foundational questions about the complex dynamics between eyes, head, and body in fields of view that extend beyond the standard computer monitor (see Anderson et al., 2023, for a recent review), this previous research has largely been concerned either with: (a) the spatio-temporal coordination of eyes and head during simple gaze shifts (e.g., Freedman, 2008; Land, 2004; Sidenmark & Gellersen, 2019) or, at the other extreme (b) a consideration of the eyes positioned within the head during active navigation in complex real-world settings (e.g., Einhäuser et al., 2009; Foulsham et al., 2011; Pelz et al., 2001). In order to begin to fill the gap, researchers have begun to focus on detailed analyses of the spatio-temporal coordination of head and

eyes in the exploration of 360° panoramas within a visually complex VR environment. VR is an excellent design choice for behavioral experiments in general (e.g., Rizzo et al., 2021), as tracking the eyes in VR provides several key advantages over traditional computer-based or mobile eye tracking research (Anderson et al., 2023; Clay et al., 2019; David et al., 2021, 2022). Particularly noteworthy is that VR enables the simultaneous tracking of both the eye and head movement with regard to a common reference frame. This capability allows for the precise calculation and differentiation of how different movement systems contribute to overall attentional control, allowing, for example, the comparison between head and eye movements throughout a carefully controlled space.

In VR, when people are asked to view scenes in 360°, the attention system must coordinate eye movements with other head and body movements to explore the full range of the visual field. When looking at 360° panoramic scenes, observers spend the majority of the time exploring along the horizon (e.g., Sitzmann et al., 2018), using their head and other body movements to extend the field of view for the eyes (Bischof et al., 2020). Similarly, when viewing landscape and fractal panoramic scenes that are rotated (for example, 45° clockwise), the head tends to roll in a similar manner in order to bring the scenes closer to their canonical upright position for the eyes (Bischof et al., 2020; Anderson et al., 2020), converging 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).

Studies in VR have also taught us that the eyes, head, and body may move in ways that diverge from what one might expect. 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). Furthermore, in many of these works, large individual differences in the extent to which participants move their head and body have been observed (Anderson et al., 2020; Bischof et al., 2020; Jacobs et al., 2020). Taken together, these works provide 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).

To date, however, the spatio-temporal analyses of the eye and head movements have tended to be analyzed separately with only a moderate emphasis on how closely the two are functionally related (e.g., Solman et al., 2017). When their coordination has been considered, it has been based on relatively simple, fixation-based analyses. In the present tutorial, we will first introduce the reader to these fixation-based analyses. We will then advance to novel ways of examining

how the head and eyes work together, across both space and time. In the next section, we cover the basic methodologies of VR research and the essential procedural details regarding the VR data that we collected for the purpose of this tutorial.

Methods

Virtual reality (VR) immerses users in a computer-generated environment. In modern VR systems, the environment is usually presented with a head-mounted viewer, and the user can interact with the virtual world using multiple devices, for example, controllers built into the VR system (see Anderson et al., 2023), hand tracking via external cameras, or devices for tracking the user's hands (e.g., Ozioko & Dahiya, 2022).

Stimuli were presented in a head-mounted viewer with two displays, one for each eye. In the present tutorial, we collected data using an HTC Vive headset equipped with an SMI eye tracker. The headset has a 110° × 113° field of view, a resolution of 1080 × 1200 pixels, and a refresh rate of 90 Hz. The position and orientation of the headset is measured using an inertial measurement unit (IMU) inside the headset that is calibrated via two infrared base stations located on opposite corners of the space the user is sitting or standing in (Niehorster et al., 2017). Head tracking permits the observer to move and look around in the virtual world and allows one to study the interaction between visual perception and head movements. It should be pointed out that the near-eye displays of many head-mounted viewers may lead to several problems affecting the fidelity of the stimulus perception, including screen-door effects (Lang, 2018) due to the low display resolution (e.g., 11 pixels/degree; Boger, 2017), chromatic aberration (Zhan et al., 2020), especially in the periphery of the displays, and others (Zhao et al., 2022). One can be quite confident, however, that these limitations will continue to decline as the technology continues to improve.

The SMI eye tracker has a sampling rate of 250 Hz, and it was controlled using an SMI-designed Unity plugin (Unity Technologies, 2017). Tracking accuracy was maintained by performing a calibration every 20 trials. This calibration consisted of following with the eyes a moving white circle with a red dot in the middle to five different regions, with the eye tracker reporting only pass or fail after the calibration. We recommend, however, that a validation phase (as defined, for example, in the open implementation for Pupil Labs; Lange, 2019) be added after the calibration to improve the quality of the eye movement data. One thing to note, however, is that the calibration/validation procedure chosen should take into account and be able to accommodate for the gradual slippage of the headset on the head. While VR devices are fixed relatively firmly to the head, in our experience they tend to slide down over time and need adjustments

periodically. This is why we chose to calibrate every 20 trials (roughly every 3–5 min).

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. 1). The sphere radius (15 m) was large enough to reach vergence infinity so that the viewers did not go cross-eyed. The scenes consisted of a balanced mix of indoor and outdoor environments taken from the SUN360 Panorama Database (Xiao et al., 2010), with a resolution of 4096 x 2048 pixels. Binocular and motion depth cues were absent.

We collected data from 25 (six male, 19 female, aged 18–26 years) undergraduate students at the University of British Columbia who participated for course credit. All observers reported normal or corrected-to-normal vision and provided informed consent prior to participation. This data collection was approved by the ethics board of the University of British Columbia (H10-00527).

Observers were instructed that there would be two phases to the data collection, an image encoding phase and an image recognition phase. The latter recognition phase was actually of no interest to us, as it was introduced merely to encourage individuals to visually explore the scenes. We conducted a nominal number of 20 old/new scene recognition trials at the end of data collection but excluded them from analysis (see Bischof et al., 2023).

Before collecting any data, we asked each observer to sit down in a non-swivel chair and familiarize themselves with the VR equipment. They were then fitted with the headset and informed that they would be presented with a total of 80 images of indoor and outdoor scenes, which we would like them to visually explore before their memory for the images was assessed. Each trial began with a uniform gray sphere with a fixation cross (at coordinates [0°, 0°] of the panoramas directly in front of the participant's chair). After fixating on the cross, the participants pressed the 'Spacebar' on a keyboard to indicate that they were ready for a trial to begin. Each scene was presented for 10 s. There were 80 trials. In 40 trials, participants were allowed to move the head (head-free condition) whereas in the other 40 trials, in keeping with classic eye movement methodology, they placed their head on a chinrest to restrict head movements (head-fixed condition). The order of these two conditions was counterbalanced across participants.

It is important to note that our selection of the task was strategic in nature to provide a methodological bridge between (a) a research situation that most researchers will be familiar, specifically, a head-fixed eye-tracking condition; and (b) a novel VR head-free condition. By keeping all the other aspects of the two conditions the same, our tutorial draws a clear line from analyzing eye movement data collected in VR when the head is fixed to when the head is free.
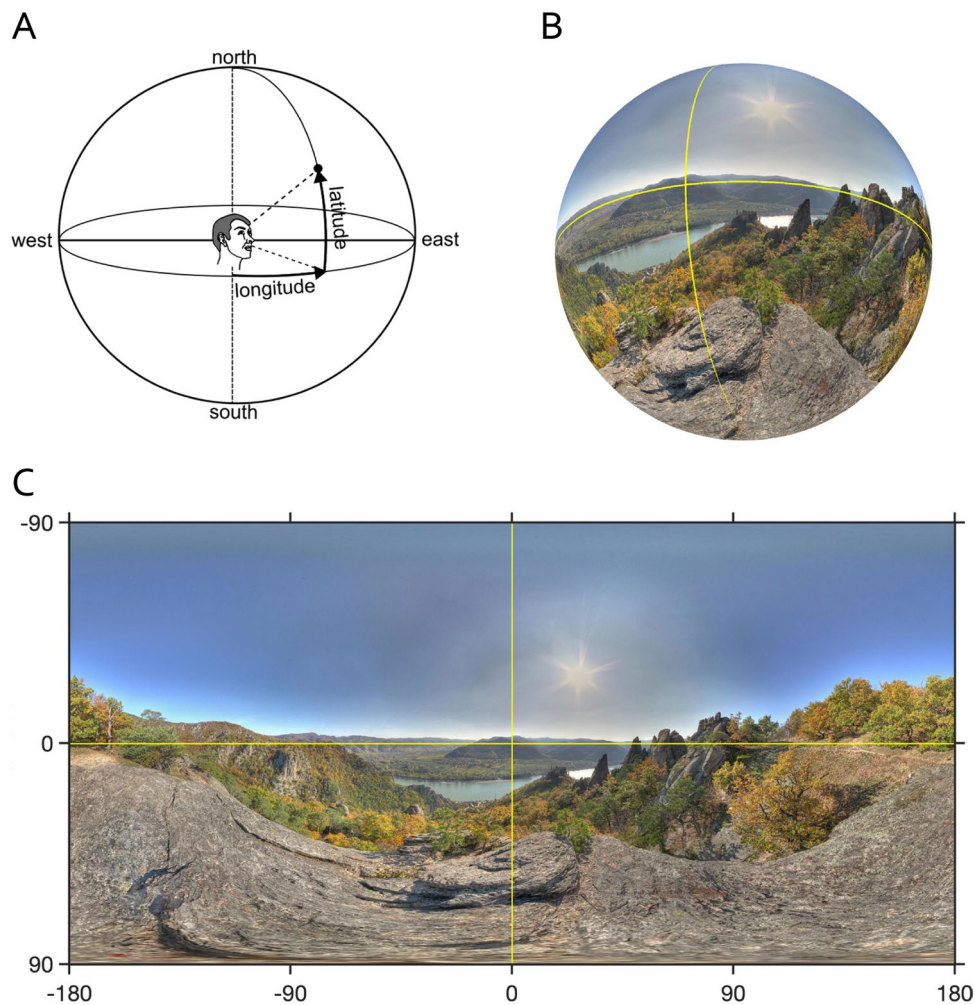
Fig. 1 Stimulus sphere and equirectangular map. *Note.* **A** Sketch of a participant viewing the spherical panorama from the center of the virtual sphere. An eye point (*black circle*) is defined by the intersection of the eye direction vector with the virtual sphere and is described by longitude and latitude, also referred to as azimuth and elevation. **B** Spherical stimulus viewed from outside and from below the equa-

tor, with the yellow lines indicating the equator and the meridian of the initial fixation cross. **C** Equidistant (equirectangular) map of the stimulus corresponding to the spherical stimulus in **B**. The map wraps around at the west meridian (− 180°) and the east meridian (+ 180°). Note the distortions near the north and south poles. They must be taken into account when analyzing fixation patterns

Alternative representations of panoramic spheres

A straightforward approach restricts the eyes-in-space analysis to the viewport defined by the head orientation. This can be handled with simple Euclidean geometry within the viewport-defined screen, and it builds the basis for efficient solutions for streaming VR content. The integration of information across different viewports defined by head orientations and a common representation are, however, not straightforward (see, e.g., Hartley & Zisserman, 2004). The present tutorial focuses on the integration of head and eye movements in VR, and a viewport-defined analysis would thus be definitely sub-optimal.

Another popular representation of VR environment relies on cube mapping in which the environment is mapped onto

the sides of a cube, with the viewer at the center of the cube (Bourke, 2020; Greene, 1986). On the one hand, the cube mapping can lead to efficient rendering, in particular in the context of video streaming (Jeong et al., 2020). On the other hand, these mappings between a sphere surface and a planar surface exhibit either area or angular distortions, and sometimes both: Equal-area mappings preserve area ratios at the cost of large angular distortions, and angular mappings preserve angles locally at the cost of large area distortions (Lambers, 2020).

There is extensive literature on the representation of spherical environments, in particular for geographic applications (see, e.g., Grafarend et al., 2014; Lapaine & Usery, 2017). None of these alternative representations, have, however, clear advantages over the equidistant maps used in the

present tutorial. Moreover, as we also make clear in the discussion, the same analyses we introduce for an observer seated in the center of a virtual sphere can be applied to viewers in a swivel chair and to those that are standing, and reference recent work revealing that the head–eye relationship is comparable when a person is sitting in a stationary chair, a swivel chair, or standing (Mehrotra et al., 2024).

Data and basic analysis methods

In the supplementary material presented on OSF (https://doi.org/10.17605/OSF.IO/THR89), we make available the data used in this tutorial. In addition, while we give general descriptions of the basic analysis methods in the main text, we also present MATLAB programs (The MathWorks Inc., 2023) to ensure that the reader can reproduce the main analyses presented here.

Data handling: Frames of reference and mapping eye and head positions

Frames of reference

Before diving into the specifics of analyzing head and eyes movements, it is important to clarify the reference frames we are dealing with (see e.g., Hessels et al., 2018). Most readers may be familiar with desktop-based eye tracking technology, where observers are required to sit in a chin rest at a set distance away from a computer monitor. In this situation, the head is in a head-fixed condition, and eye movements are reported with respect to the computer monitor, usually with pixel location or degrees of visual angle from the screen center. In other words, the reference frame for eye movements is the screen, which typically encompasses approximately 30–50° visual angle, depending on the particular setup.

In VR experiments, the eyes are tracked by cameras mounted within the HMD. One significant advantage of this setup is that eye and head movements can be tracked with respect to a common reference frame. Usually, this is the 3D virtual world. For example, eye and head movements can be tracked with respect to a 360° scene, or they can be returned as positions on 3D objects in a scene. In addition, VR allows for great flexibility in reference frames as well, where eye eccentricity can be returned as positions in the head coordinate system, such that the contributions of head movements to eye position can be calculated precisely.

In the present paper, we analyze head movements, eye movements in the head coordinate system (denoted "eyes-in-head"; David et al., 2022) and the combination of both in the scene coordinate system (denoted "eyes-in-space" or "gaze"; David et al., 2022). Head rotation is represented with

a quaternion (Goldman, 2022; Quaternion and Spatial Rotation, n.d.; Vince, 2021), and the eyes-in-space direction is obtained by multiplying the eyes-in-head direction vector with the head quaternion.

Mapping of the panoramic sphere to a planar map

As noted before, participants viewed the panoramas from the center inside the sphere on which the panoramas are projected (see Fig. 1A). An eye point is defined as the intersection of the eye direction vector with the panorama sphere. It is described with coordinates longitude in the range [– 180, 180] degrees and latitude in the range [– 90, 90] degrees. Similarly, we define the head point as the intersection of the vector pointing forward from the face with the virtual sphere, and it is also defined in world coordinates (which are, given our restriction to omnidirectional panoramas, given by longitude and latitude on the stimulus sphere).

Figure 1B shows one of the panoramas used in the study, 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. 1C).

When analyzing 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) must 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.

Event detection

In the analysis of eye behavior, we focus on the detection of fixations and saccades, while other ocular events, such as smooth pursuit, micro-saccades, or blinks are ignored (see for example, Andersson et al., 2017; Holmqvist & Andersson, 2017). There are two fundamentally different approaches to gaze analysis. The first approach starts with the detection of fixations, and saccades are defined as differences between successive fixations, whereas the second

approach starts with the detection of saccades, and fixations are defined as stable points between saccades. A popular method for the detection of fixations is the Dispersion-Threshold (IDT) algorithm (Komogortsev et al., 2010; Salvucci & Goldberg, 2000), which assumes that the dispersion of gaze points within a fixation is relatively small (in our studies typically 2.5–3°) and that the duration of fixations exceeds a minimum duration (in our studies typically 80 ms). Specifically, the IDT algorithm proceeds as follows:

1. Initialize a window of gaze points to cover duration threshold. Drop the data points if the dispersion of the gaze points exceeds the dispersion threshold.
2. Add further gaze points to the window as long as the dispersion of the gaze points does not exceed the dispersion threshold.
3. Define the fixation position as the centroid of gaze points.
4. Remove the gaze points of the fixation and start again from step 1.

The MATLAB program of the IDT analysis is presented in the supplementary material. The result of the IDT approach is illustrated in Fig. 2, where black circles show raw eyes-in-space points, red circles show fixations extracted by the IDT algorithm, and green circles show the raw eyes-in-space points that are used in computing the (nearby) fixation. Alternatively, fixation detection relies on gaze vector velocities, where in step 2, gaze points are added to the window as long as the velocity of successive gaze points does not exceed the velocity threshold. For both the IDT algorithm and the velocity algorithm, saccades are defined as differences between successive fixations.

The second approach begins with the detection of saccades, and fixations are defined as stable points between saccades. The detection of saccades is based on the assumption that motion above a velocity threshold is assumed to be (part of) a saccade. Specifically, the algorithm proceeds as follows:

1. Calculate the gaze velocities between all successive gaze points.
2. Detect peak velocities (which are assumed to define the middle of a saccade).
3. Add velocities immediately before the peaks and immediately after the peaks as long as they exceed a velocity threshold. Velocities below that threshold are assumed to be part of a fixation.
4. Peak velocities must be below a certain limit to exclude artefacts, such as blinks.
5. Finally, fixations are defined as the relatively stable positions between saccades.

The reader is advised to consult further sources of information on event detection in eye movement analysis, including, for example, those by Andersson et al. (2017), Holmqvist & Andersson (2017), Hooge et al. (2018), and Nyström & Holmqvist (2010).

Given a sequence of fixations, saccades can be defined by the difference between successive fixations. They are described by direction, amplitude and duration, and saccade velocity is defined as the ratio of amplitude and duration. Saccades exceeding a maximum duration (in our study 300 ms) are ignored. Although the head moves smoothly (rather than jumping abruptly like the eyes), one can still analyze sequences of head positions and define head shifts

Fig. 2 Analysis of eyes-in-space points using the IDT method. *Note.* The *black circles* indicate raw eyes-in-space points, the *red filled circles* indicate fixations extracted using the IDT algorithm, and the *green filled circles* indicate raw eyes-in-space points that are used in computing the nearby fixation

as the difference between successive head positions. They are also described by direction, amplitude and duration, and head velocity is defined by the ratio of amplitude and duration.

Standard head and eye movement analyses

While it is relatively simple to understand VR eye tracking in terms of what people are looking at in 3D space, tracking the eyes and head through a virtual space introduces additional complexity. Not only does the researcher now have access to what people are looking at, they may also consider how the eyes move within the head, and how the head moves, independently of the eyes. Given the complexity and richness of these data, it is easy to get lost in analyses, making it even more critical to approach data analysis with a question, or series of questions in mind. This is good advice, generally speaking, but it becomes even more apparent when faced with the plethora of possible analyses and potentially interesting pathways for analyzing eye, head, and body movements.

In this section, we introduce how one goes about analyzing head and eye movement data collected in VR by systematically reporting the data we collected in three stages. First, the head data with regard to where the head is pointed in space. Next, the eye data with respect to where the eyes are directed relative to the head. Finally, the eyes-in-space data, the direction of which demands a consideration of the position and orientation of the head in order to calculate the direction of the eyes in world coordinates.

Head analysis

Head movements are inferred from the movements of the VR headset. The position and orientation of the headset are tracked using IMUs inside the headsets. Recall that the head point refers to the intersection of the vector pointing forward from the face with the virtual sphere, and it can thus be defined in world coordinates.

Figure 3 illustrates the range of head movements, more precisely, 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; in the range of ± 80°) determines the *longitude* of the head point and the extension/flexion (i.e., up/down) of the head (pitch; in the range of ± 45°) determines the *latitude* of the head point. In this study, lateral bending of the head (roll; in the range of ± 45°) is ignored.

Figure 4 shows the heatmaps of head points of the two experimental conditions of the present study, namely the head-free condition, in which participants were allowed to freely move the head, and the head-fixed condition, in which participants kept the head in a chin rest. The heatmaps were obtained from the maps of 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 MATLAB program for plotting these heatmaps (Plot_Heatmaps) is shown in the supplementary material. In the head-free condition (Fig. 4A), the head patterns are concentrated along the equator line of the panoramas. In the head-fixed condition (Fig. 4B), participant rested the head in a chin rest, and consequently, the head positions are concentrated around position [0°, 0°] of the panoramas, confirming that the head indeed remained at the initial fixation position.

The means and standard deviations of the head points are reported in Table 1. These can be examined by a participant-wise analysis which shows that, in the head-free condition, the head points are on average somewhat above the equator line, with mean latitude 2.4°, $t(24) = 2.21$, $p = .037$, but not biased east or west of the starting point, i.e., with mean longitude -1.5°, $t(24) = -0.69$, $p = .497$. Intriguingly, in the head-fixed condition, the head points were also on average somewhat above the equator line, with mean latitude 2.7°, $t(24) = 1.76$, $p = .09$ and biased east, with mean longitude 5.0°, $t(24) = 3.52$, $p = .002$. The distribution of head points in the head-free condition is
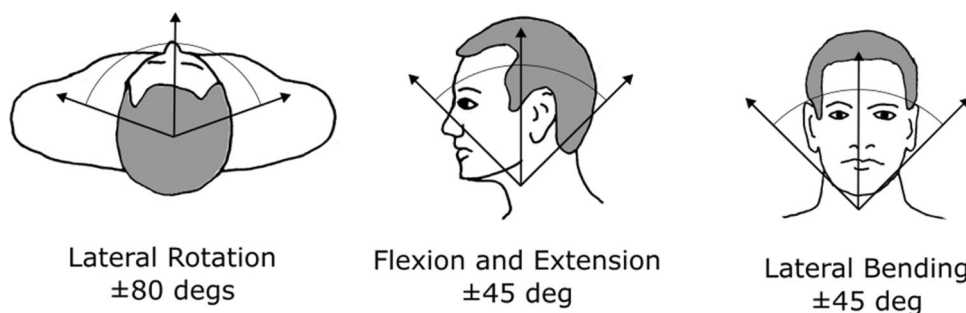

Fig. 3 The three axes of head orientation. *Note*. Left panel: Lateral rotation of the head has a range of ± 80 degrees. *Middle panel*: Flexion-extension of the head has a range of ±45 degrees. *Right panel*: Lateral bending has a range of ±45 degrees
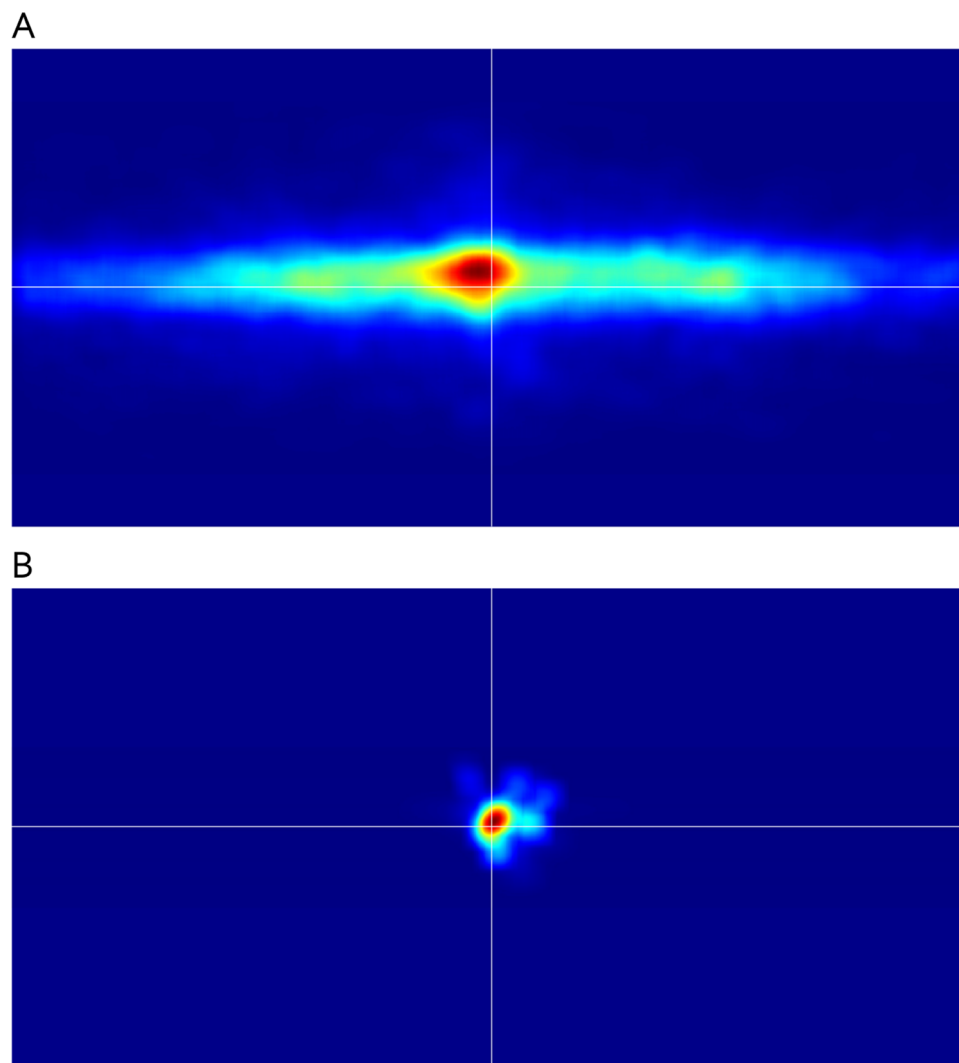
A

B

Fig. 4 Head Heatmap. *Note*. 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. **A** Heatmap for the head-free condition, in which participants were allowed to freely move the head. The head points were concentrated along the equator of the panoramas. **B** Heatmap for the head-fixed condition, in which participants kept the head in a chin rest. The head points were concentrated at coordinates [0°, 0°] of the panoramas

Table 1 Means and standard deviations of effector distributions

	Head-free condition		Head-fixed condition	
	Longitude	Latitude	Longitude	Latitude
Head	− 1.5 ± 79.8	2.4 ± 15.3	5.0 ± 3.4	2.7 ± 2.8
Eyes-in-head	2.2 ± 16.6	− 4.8 ± 10.8	1.7 ± 16.4	− 5.1 ± 12.3
Eyes-in-space	1.9 ± 86.2	− 2.4 ± 20.7	6.8 ± 18.0	− 2.5 ± 12.7

highly anisotropic (longitude STD = 79.8°, latitude *STD* = 15.3°), whereas in the head-fixed condition, the head points are concentrated more or less isotropically around the mean (longitude STD = 3.4°, latitude STD = 2.8°).

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_i (see the program Analyze_IDT in the supplementary material).

Figure 5 shows the direction distributions (panels A and C) and velocity distributions (panels B and D) of the head shifts, for the head-free condition (panels A and B) and the head-fixed condition (panels C and D). The MATLAB program for plotting these distributions (Plot_Saccades) is shown in the supplementary material. In the head-free
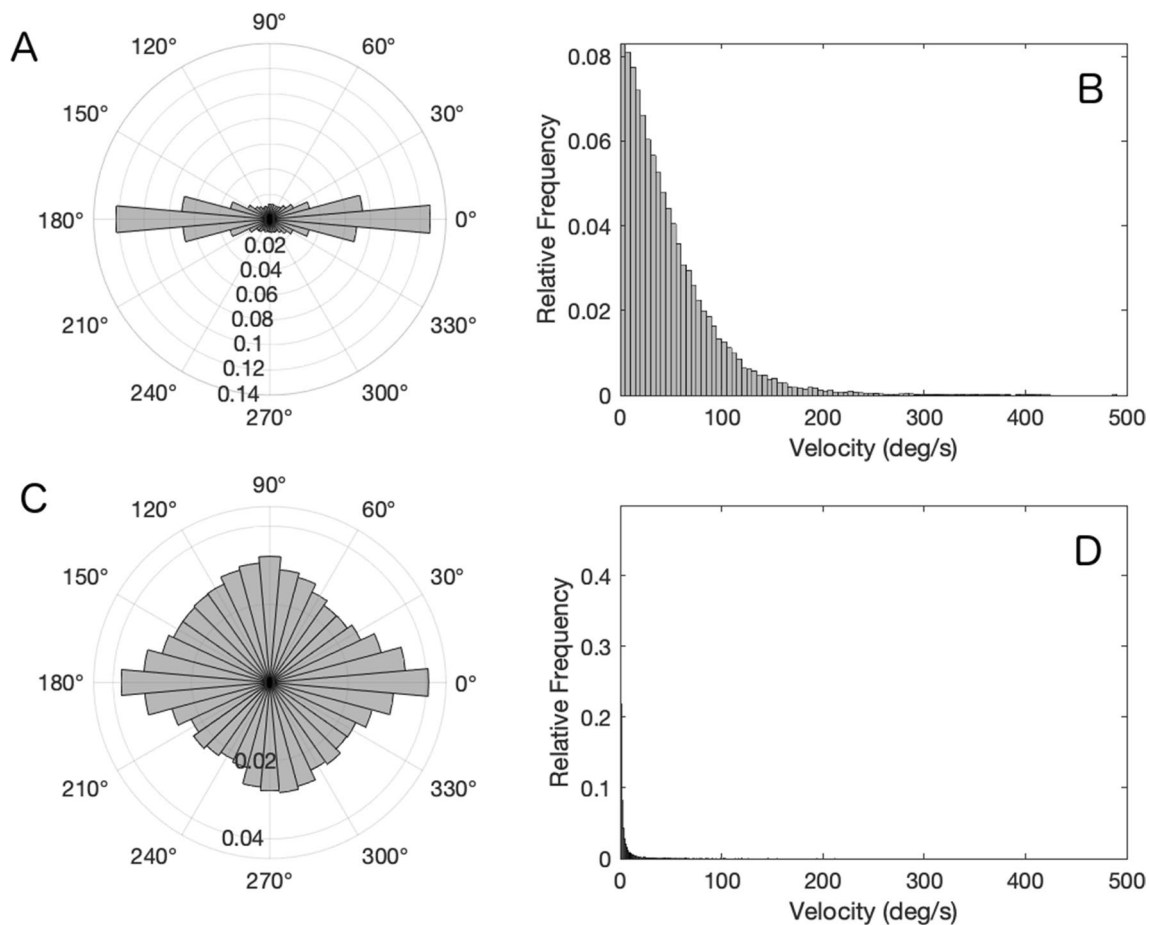
Fig. 5 Head shifts. *Note*. Direction distributions (**A** and **C**) and velocity distributions (**B** and **D**) of the head shifts, for the head-free condition (**A** and **B**) and the head-fixed condition (**C** and **D**)

condition, 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. In the head-fixed condition, there is, for obvious reasons, minimal head movement in the horizontal and vertical directions. Accordingly, the head shift velocity is close to zero (panel D) and the direction of head shift directions are distributed isotropically (panel C).

In the head-free condition, the was 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. This result is known as the equator bias and is consistent with results obtained previously 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). Second, if participants explore the panoramas with neck extension/flexion in a resting state, then this distribution of head positions is most likely to occur. In the

head-fixed condition, participants rested the head in a chin rest and head movements were thus minimal.

Eyes-in-head analysis

While the previous section focused on analyzing head direction in world coordinates, the eye tracker built into VR headsets provides eye directions in head-centered coordinates. Figure 6 shows the heatmaps of the eyes in head-centered coordinates, again for the two experimental conditions, the head-free condition, and the head-fixed condition. The heatmaps were obtained from the maps of eyes-in-head points using a Gaussian filter with σ = 4° angular distance. Note that we use the same filter size for all heatmaps to allow a direct comparison. The eye-in-head heatmap for the head-free condition are shown in Fig. 6A and for the head-fixed condition in Fig. 6B. The analysis of eye-in-head latitudes shows that the eyes are directed below the head-defined center of the visual field, both in the head-free condition, with mean latitude $-4.8°$, $t(24) = -4.45$, $p < .001$, and in the head-fixed condition, with mean latitude $-5.1°$, $t(24) =$
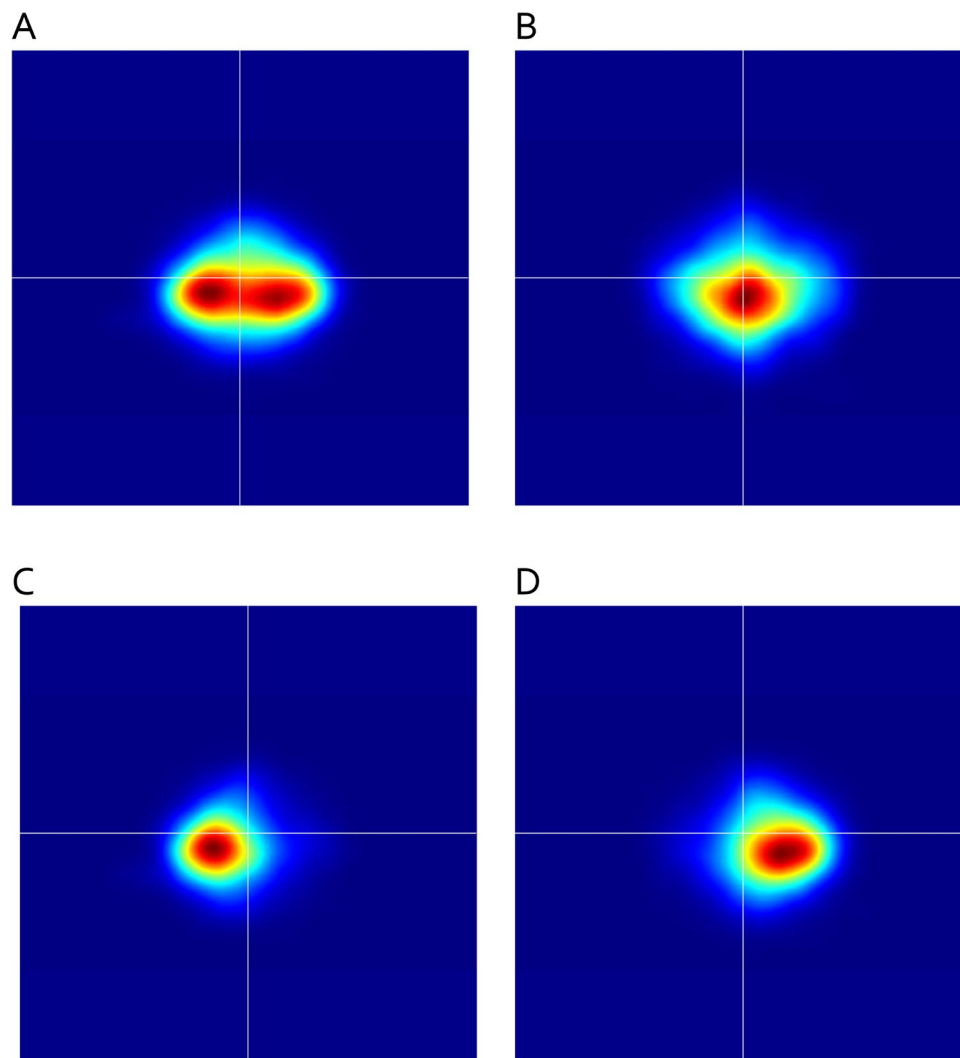
A

B

C

D

Fig. 6 Eyes-in-head heatmaps. *Note*. Heatmaps of eyes-in-head points in the longitude range [– 90°, 90°] and the latitude range [– 90°, 90°], with the white lines showing 0° longitude and latitude. Heatmaps of eyes-in-head points are shown for the head-free condition (**A**) and the head-fixed condition (**B**). In the head-free condition, there are two peaks, one to the left and the other to the right of the head-defined center. Panels **C** and **D** show how these double peaks arise: if the head is moving to the left the eyes-in-head points are left of the head-defined center (**C**) and if the head is moving to the right the eyes-in-head points are to the right of the head-defined center (**D**)

– 3.97, $p < .001$. In other words, in both experimental conditions, the eyes are directed below the head-defined center of the visual field. Regarding the horizontal bias, the eyes were directed east of the 0° meridian in the head-fixed condition, with mean longitude 2.2°, $t(24) = 3.78$, $p < .001$, and in the head-fixed condition, with mean longitude 1.7°, $t(24) = 2.68$, $p = .013$.

A more detailed analysis of eyes-in-head longitudes of the head-free condition shows a somewhat unexpected result: There is evidence of 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. This double peak of the eyes-in-head heatmap has been found before (e.g., Sidenmark & Gellersen, 2019, Fig. 3) but has not been commented on. A

closer analysis reveals the following: As will be explained in the later sections that consider the temporal relation between eyes and head, eyes lead the head in almost all cases of panoramic viewing. Thus, if the head moves towards the left, the eyes tend to be to the left of the head position (Fig. 6C), and conversely, if the head moves towards the right, the eyes tend to be to the right of the head positions (Fig. 6D). If head movement is ignored, the double peak of the eye-in-head heatmap is obtained (Fig. 6A). More precisely, if the head is moving to the left, i.e., $\text{longitude}_{\text{head}}(t + \Delta t) < \text{longitude}_{\text{head}}(t)$, then the peak of the heatmap at longitude – 15.9° is significantly to the left of the 0° meridian, $t(24) = -5.87$, $p < .001$. Conversely, if the head is moving to the right, i.e., $\text{longitude}_{\text{head}}(t + \Delta t) > \text{longitude}_{\text{head}}(t)$, then the

peak of the heatmap at longitude 11.7° is significantly to the right of the 0° meridian, $t(24) = 192.4$, $p < .001$. Finally, in the head-fixed condition (Fig. 6B) with no head movements, the eye positions are concentrated somewhat to the right of the 0°meridian, with longitude 2.2°, $t(24) = 2.68$, $p = .013$.

In the head-fixed condition, the standard deviation of the eyes-in-head longitudes was 16.4° and of the latitudes was 12.3°. In the head-free condition, 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), although it is noteworthy that these values do not seem to vary meaningfully whether one is moving or not moving the head.

We define eyes-in-head as follows: Given two fixations f_i and f_j and the corresponding head positions h_i and h_j, we define eyes-in-head coordinates e_i and e_j by the great circles h_i to f_i and h_j to f_j, and the corresponding eyes-in-head

saccade is defined by the great circle 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_i (see the program Analyze_IDT in the supplementary material presented on OSF (https://doi.org/10.17605/OSF.IO/THR89).

Figure 7 shows the direction distributions (panels A and C) and velocity distributions (panels B and D) of the eyes-in-head saccades, for the head-free condition (panels A and B) and the head-fixed condition (panels C and D). In both conditions, the saccade directions and the saccade velocities are similar, indicating again an independence of eye saccades from head positions.

In summary, this initial introduction to analyzing VR eye movement data shows how one can examine eye movement data from a number of different spatial measurements: where the eyes are positioned in the head, how much they move in the head, and how fast they move in the head. For the head-head-fixed condition data, the findings
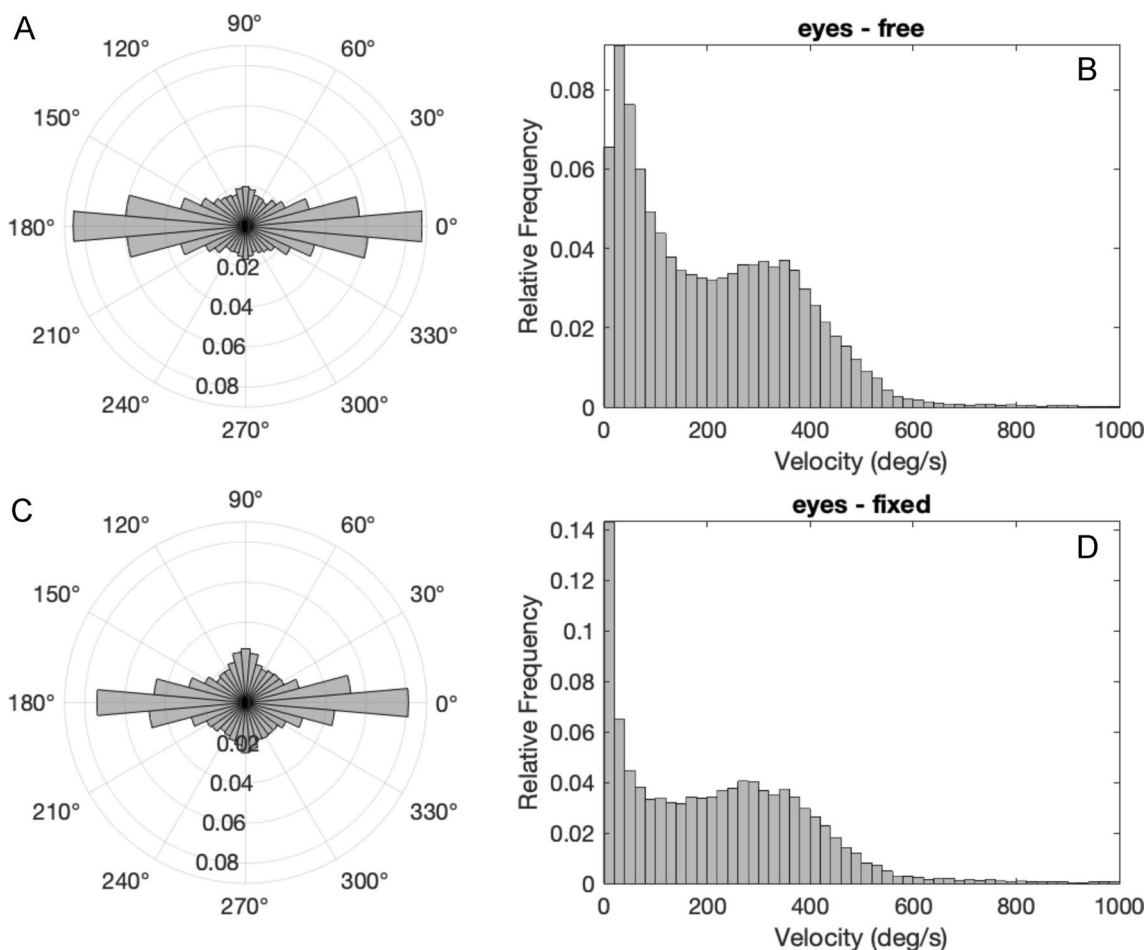

Fig. 7 Eyes-in-head saccades. *Note.* Direction distributions (**A** and **C**) and velocity distributions (**B** and **D**) of the eyes-in-head saccades, for the head-free condition (**A** and **B**), and the head-fixed condition (**C** and **D**)

replicate much of what has previously been reported in the literature, consistent with the fact that eye movement data with the head constrained makes up much of what is known in the field. However, when the head is free to move, we are already making some new and intriguing observations. For instance, mapping where people look can yield bimodal distributions which, as we have foreshadowed, results from the spatio-temporal dynamics of eye and head movements (i.e., the eyes tend to lead head movements). It is also interesting to observe that the direction, velocity, and angular variance of the eyes relative to the head seems to be remarkably stable when the head is moving versus when it is head-fixed condition, suggesting it is the visual input to the eyes that drives these selections

independent of head kinematics. This is very good news for researchers worried about the generalizability of more traditional head-fixed eye tracking studies.

In the next section, we consider where the eyes are directed with regard to the virtual environment (eyes-in-space or 'gaze'). Note that in VR 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 general issue).
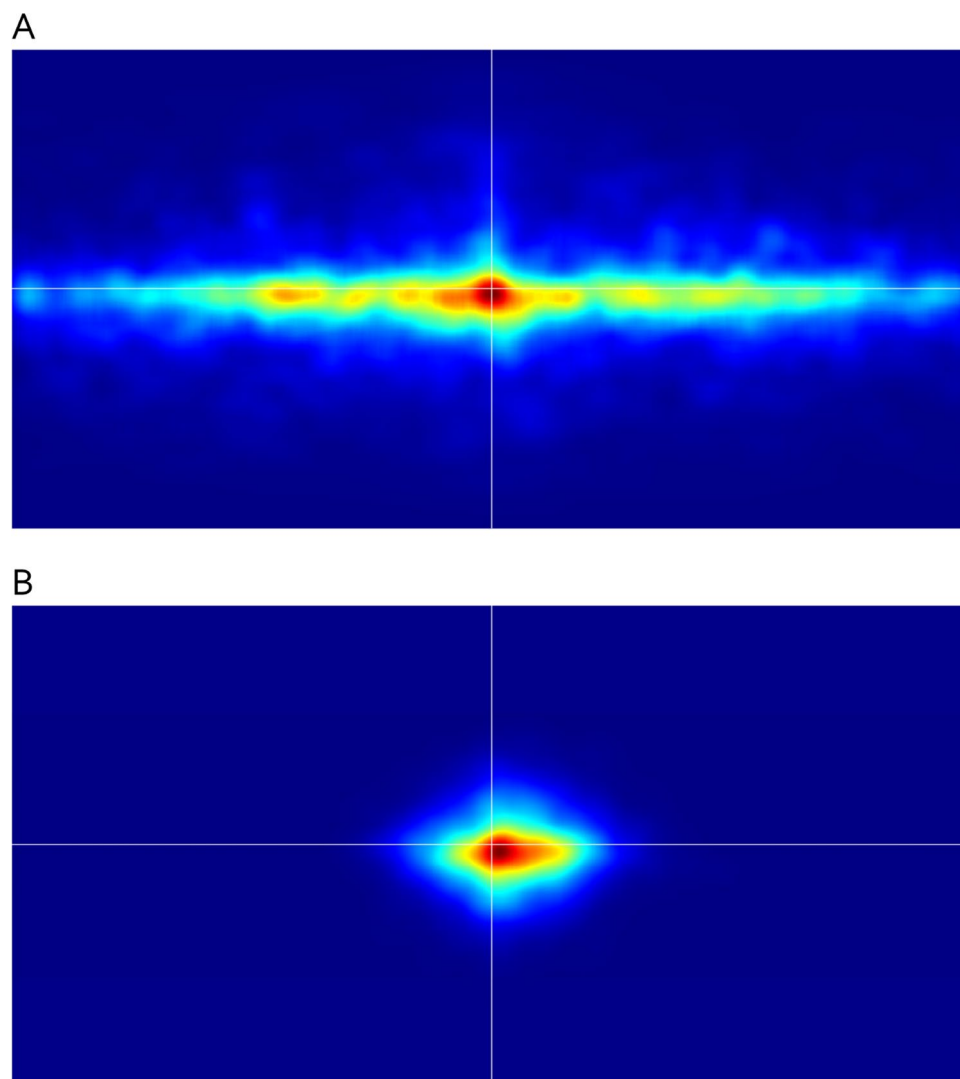

Fig. 8 Eyes-in-space heatmaps. *Note*. Heatmaps of eyes-in-space points in the longitude range [– 180°, 180°] and the latitude range [– 90°, 90°], with the *white lines* showing 0° longitude and latitude. **A** Heatmap for the head-free condition, in which participants were allowed to freely move their head. **B** Heatmap for the head-fixed condition, in which participants were instructed to keep the head in a chin rest. The distribution of eyes-in-space points is determined by the distribution of the head positions and the distribution of eyes-in-head points, as described in the text

Eyes-in-space analysis

Heatmaps of the eyes-in-space patterns for head free and head-fixed condition are illustrated in Fig. 8. In the head-free condition (Fig. 8A), the eyes-in-space fixation patterns are concentrated along the equator line of the panoramas, and in the head-fixed condition (Fig. 8B), they are concentrated around the head-fixed condition head position at longitude 0° and latitude 0° of the panorama. These heatmaps are best understood as the combination of the head heatmaps (Fig. 4) and eye-in-head heatmaps (Fig. 6). A comparison of Fig. 8A and B illustrates the strong effect of head movements on the practical field of fixation, that is, the idea that freeing the head enlarges the space within which a fixation can be made (David et al., 2020, 2022; Rötth, 1925; von Noorden & Campos, 2002, p. 79).

A participant-wise analysis showed that, in the head-free condition, the eyes-in-space points were neither biased above or below the equator line, with latitude – 2.4°, $t(24)$ = – 1.84, p = .078, nor were they biased east or west of the 0° meridian line, with longitude 1.9°, $t(24)$ = 1.12, p = .272.

In contrast, eyes-in-space points in the head-fixed condition were biased in the vertical direction, with latitude – 2.5°, $t(24)$ = – 2.53, p = .019, and in the horizontal direction, with longitude 6.8° line, $t(24)$ = 4.02, p < .001 (see Table 1). The distribution of eyes-in-space points in the head-free condition is anisotropic (longitude STD = 86.2°, latitude STD = 20.7°), whereas in the head-fixed condition, they are distributed closer to an isotropic distribution (longitude STD = 18.0°, latitude STD = 12.7°). The eyes-in-space distributions are similar to the head distributions, albeit somewhat less extreme.

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 (see the program Analyze_IDT in the supplementary material).

Figure 9 shows the direction distributions (panels A and C) and velocity distributions (panels B and D) of the eyes-in-space saccades, for the head-free condition (panels A and
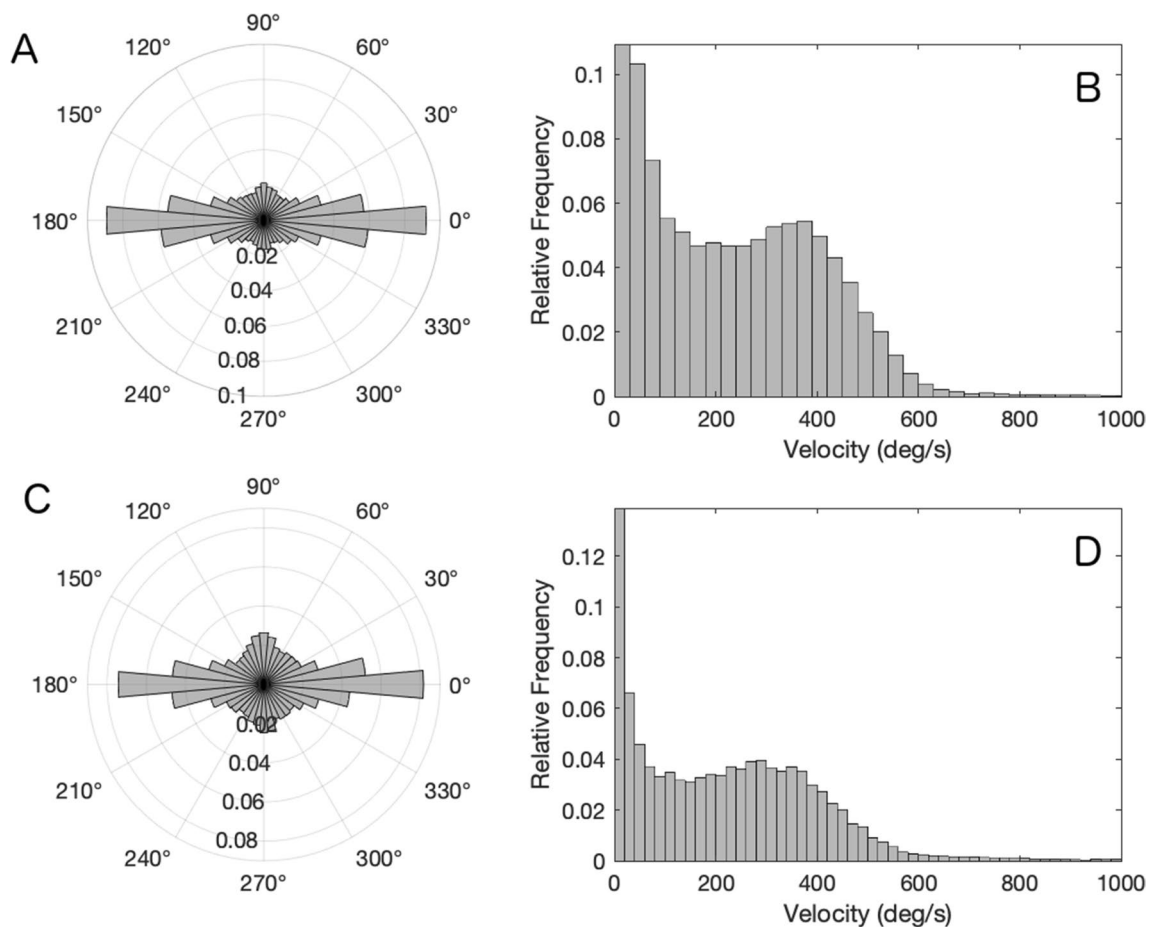

Fig. 9 Eyes-in-space saccades. *Note.* Direction distributions (**A** and **C**) and velocity distributions (**B** and **D**) of the eye-in-space saccades, for the head-free condition (**A** and **B**) and the head-fixed condition (**C** and **D**)

B) and the head-fixed condition (panels C and D). The distributions of the directions and velocities of the eyes-in-space saccades were almost identical to those of the eyes-in-head saccades.

The eyes-in-space analysis confirms the equator bias obtained earlier by Bischof et al. (2020) and Sitzmann et al. (2018). This result may reflect several potential variables. First, if participants inspect the panorama with neck extension and flexion in a resting state and the eyes are centered within the head coordinate system then there is a natural preference for fixations along the horizon. Second, an analysis of typical panorama images shows that on average, edge density is strongest along the equator (as shown below), which may be due to the fact that there is simply more content along the horizon in typical panoramic scenes (see e.g., Solman et al., 2017 or Torralba et al., 2006).

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. This is analyzed in the following paragraphs.

All panoramas were grouped into either indoor or outdoor panoramas, and in the study reported here, participants were presented with 40 panoramas of each group. Figure 10 shows the distribution of edges of each group, obtained with
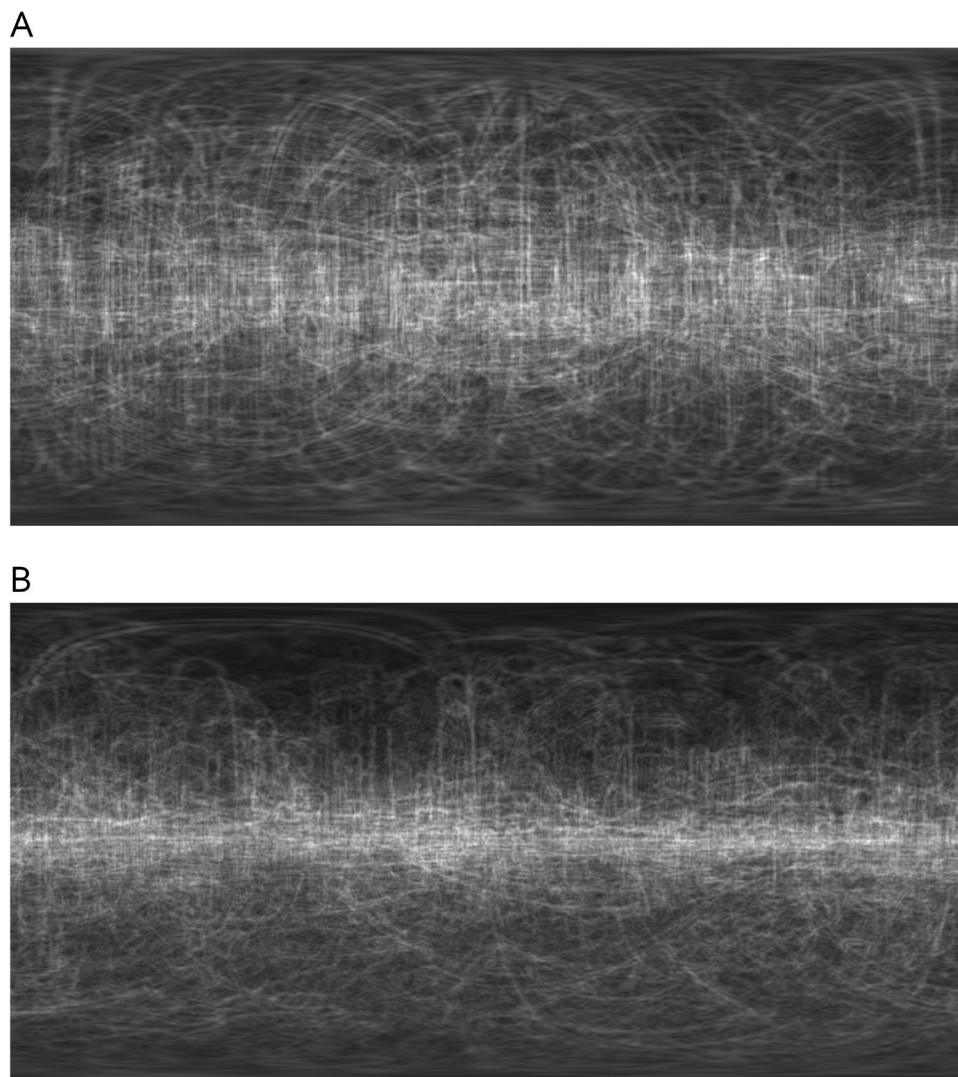

Fig. 10 Edge map of indoor and outdoor panoramas. *Note.* **A** Sum of edge maps over all indoor scenes; **B** Sum of edge maps over all outdoor scenes. The edge maps were obtained using an edge detector combined with a Gaussian filter of size σ = 0.5° at the equator and increasing horizontally by sec(latitude) towards the north and south. The edge maps show that the edges of indoor scenes are distributed over a wider vertical range than the edges of the outdoor scenes

an edge detector combined with a Gaussian filter of σ = 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.

Figure 11 shows the eyes-in-space heatmaps for the indoor and outdoor scenes and the two experimental conditions, the head-fixed condition, and the head-free condition. In the head-free condition, the standard deviation of the latitudes was larger for the indoor panoramas (23.6°) than for the outdoor panoramas (19.0°), Levene $F(1, 38846) = 831.3$, $p < .001$, and the standard deviation of the longitudes was somewhat smaller for the indoor panoramas (85.9°) than for the outdoor panoramas (87.6°), Levene $F(1, 38846) = 19.8$, $p < .001$. In the head-fixed condition, the standard deviation of the latitudes was larger for the indoor panoramas (14.8°) than for the outdoor panoramas (12.4°), Levene $F(1, 30168) = 426.5$, $p < .001$, but the standard deviations of the longitudes were almost the same (20.7° and 20.3°), Levene $F(1, 30168) = 0.014$, $p = .906$. The results show that the distribution of latitudes are consistent with the expectations from the edge analysis and, more generally, that eyes-in-space is affected by the spatial properties of the panoramas.

Spatio-temporal head-eye data analyses

While the above sections have applied analyses that are aligned with the most recent VR investigations, in the following sections we present several novel methods for analyzing the relation between head, eyes-in-head, and eyes-in-space. These analyses capitalize on the fact that the eyes and head are interrelated and nested effectors in a dynamic system. In the analysis our intent is to both arm the reader with new tools to analyze the data that they collect, as well as to demonstrate to others that one can be creative with how they engage with the data in order to address and answer a range of questions. Naturally, the following analyses include data from the head-free condition only.

To illustrate the close connection between eyes-in-space and head, Fig. 12 shows a panorama overlayed with eye points and head points. The red circles indicate eye fixations, the black line shows the head positions, the blue circles indicate head averages during the fixations, and the blue lines connect fixations with the corresponding head averages. These lines thus represent eyes-in-head.

Spatial relation between eyes and head

A comparison of the latitude means in Table 1 shows that, in the head-free condition, the mean of the head points tended to be above the equator lines (with latitude + 2.4°),

but eyes-in-head means were below the equator lines (with latitude – 4.8°) with the result that the mean of the eyes-in-space points tended to be below the equator line (with latitude – 2.4°). The fact that the eyes-in-head means were substantially below the equator line is consistent with the results obtained by Lee et al. (2019).

The analysis of head and eyes showed that the spread of eye points is systematically larger than the spread of head points (see Table 1). A more detailed analysis of eyes-in-head reveals how this difference arises. Figure 13 shows eyes-in-head directions and amplitudes, averaged at regular head-point intervals of 10° longitude and 10° latitude and averaged over all scenes and participants in the head-free condition. Arrows start at the head points and end at the eye points, with zero length indicating that, on average, head points and eye points coincide. Above latitude 30°, most arrows are pointing upwards, below latitude 30°, most are pointing downwards, west of longitude – 10° the arrows are pointing west, and east of longitude – 10°, they are pointing east. Near the poles, there are only few measurements (see Figs. 4A and 8A), and consequently, the eyes-in-head directions and amplitudes have higher sampling errors. Together, this analysis shows that the eyes are expanding the range of locations defined by the head. This can be explained by the fact that the physical effort to move they eyes is much smaller than the effort to move the head (Solman et al., 2017).

### Spatio-temporal relation between eyes and head

One critical type of information that can be extracted from VR data, in addition to the spatial relationship between the eyes and the head (as reported above), is the spatio-temporal relationship between the eyes and head. This latter information is key to understanding the complex temporal dynamics between gaze, the head, and (in future work) other body movements. One question often asked is whether eye movements lead or follow head movements, as this temporal relationship is thought to have cognitive consequences [i.e., the relative timing of eye and head movements indicates whether attentional selection is reflexive or volitional – see Solman & Kingstone (2014); Solman et al. (2017); Zangemeister & Stark (1982); Doshi & Trivedi (2012); Freedman (2008)]. We have developed several original methods for determining the spatio-temporal relation between eyes and head, namely the minimum distance method (see Bischof et al., 2020), the direction method, and the distance method. In addition, we report the average distance method proposed by Kangas et al. (2022). These are described in the following paragraphs.

Minimum-distance method For this method, the aim is to determine if the eyes or the head moves first when a person
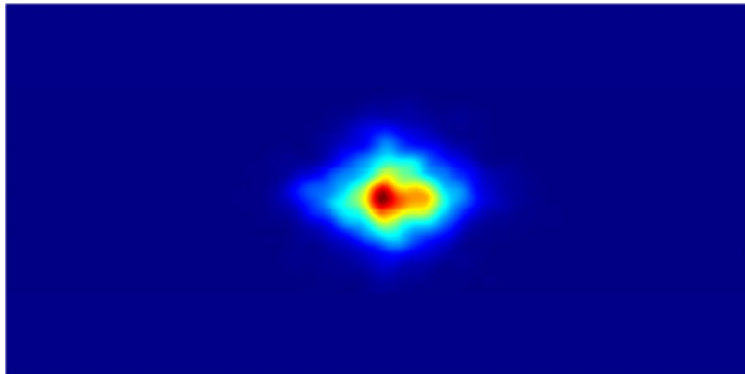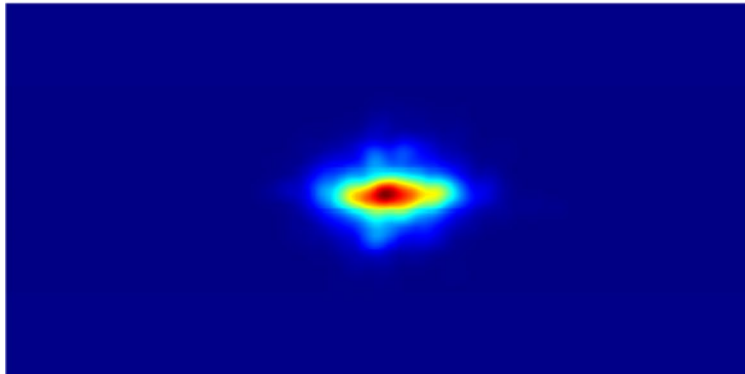
A
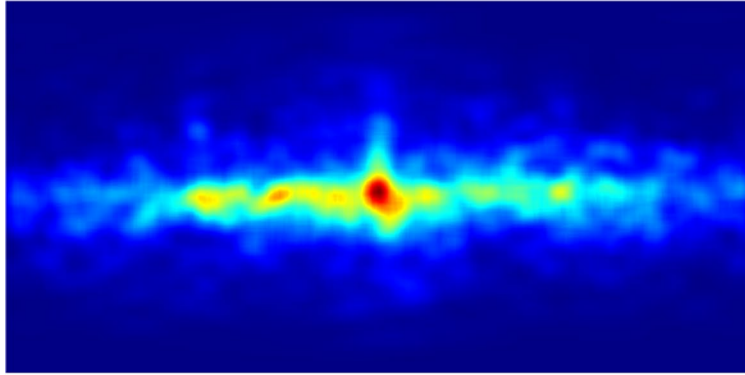
B
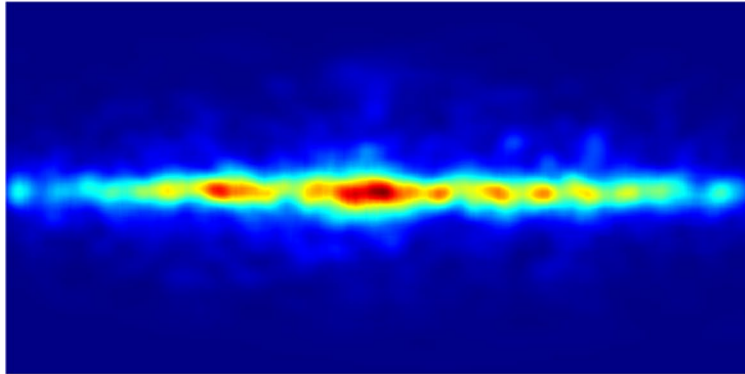
C

D

◀**Fig. 11** Eyes-in-space heatmaps of indoor and outdoor panoramas in the head-free and head-fixed conditions. *Note.* **A** Heatmap of indoor panoramas in the head-free condition; **B** Heatmap of the outdoor panoramas in the head-free condition; **C** Heatmap of the indoor panoramas in the head-fixed condition; **D** Heatmap of the outdoor panoramas in the head-fixed condition

looks at a scene. To do this, we compare where their eyes are focused (fixations) with where their head is on average (average head positions). Specifically, for each fixation (let us call it g_i), we look at the head positions (h_j) both before and after that fixation. We then find the head position (h_{min}) that is closest to where the person is looking (g_i). If this closest head position (h_{min}) happens *after* the person starts looking (a positive lag), it means their eyes moved first, and then their head followed. However, if it happens before the person starts looking, it means their head moved first, and then their eyes followed.

One can see the details of how this analysis is done and what it shows in Fig. 14A. The black and grey points indicate fixations, the dashed line indicates the smooth movement of head position, the white points indicate average head positions during each fixation, and the black lines show the distance between fixations and average head positions. As noted, the minimum distance method determines, for each fixation g_i, the closest average head position h. In this example, h_{i+1} is closest to g_i, indicating that the eyes are leading head by a lag of about one fixation. The histogram of eyes-head lags for the head-free condition (Fig. 14B) shows that most lags (86.2%) are positive, that is, eyes are leading head in most cases. An analysis of negative lags (13.8%) showed that they occur exclusively when the head is moving slowly or varies randomly around a more or less stable position. In summary, the results show that the lag peak is around one fixation, suggesting that the eye leads head by one fixation or approximately 200 ms.

Average-distance method A method related to our minimum-distance method was proposed by Kangas and colleagues. It estimates the eyes-head delay based on the minimum distance between fixations and head, averaged over participants and images (Kangas et al., 2022; Fig. 5 and Eq. 1). The results of this method with the data of our study are shown in Fig. 15A and indicate an average delay of approximately 170 ms between eyes and head. A participant-wise analysis of the average distance method shows a substantial variation of the distance curves over participants (Fig. 15B) and the location of the curve minima, which is also shown by Kangas et al. (their Fig. 6). Nonetheless, the results confirm the positive lag between eyes and head.

Direction method This method determines the relation between eyes and head based on directional information,

as illustrated in Fig. 16, which shows a series of fixations in black and a series of head positions in white, that is, the black dots g_1, g_2, … show the position of successive fixations and the white dots h_1, h_2, … show the head averages during the fixations. The grey angles indicate the angle between the lines g_i - h_i and the lines connecting successive head positions $h_i - h_{i+1}$, and the dashed line indicates the smooth movement of the head positions. The upper panel of Fig. 16A shows examples of the eye leading head, with the distribution of angles centered around 0°, as shown in the upper panel of Fig. 16B. In contrast, the lower panel of Fig. 16A shows examples of the eye trailing head, with the distribution of angles centered around ± 180°, as shown in the lower panel of Fig. 16B. In summary, the angles between lines g_i - h_i and $h_i - h_{i+1}$ are centered around 0° if the eyes are leading head and centered around ± 180° if the eyes are lagging behind head. The empirical histogram of the angles is shown in Fig. 16C, with a majority of cases (85.7%) centered around 0°, indicating that the eyes are leading head, and in the other cases (14.3%), head is trailing the eyes. It should be added that the direction method cannot produce a temporal estimate of the eyes–head delay and only indicates whether the eyes are leading or lagging the head.

Tracking method The fourth method is based on estimating how long it takes the head to move to the head position closest to a fixation. In Fig. 17A, the black dots g_1, g_2, … show the position of successive fixations and the white dots h_1, h_2, … show the head averages during the fixations. The dashed line indicates the smooth movement of the head positions. The black lines are obtained by projecting the lines g_i - h_i onto the continuous lines of head positions, and one can measure how long (in ms) it takes the head to reach the end position of the projected line. In this example, g_1, g_2, and g_4 are leading head while g_3 is trailing head. Figure 17B shows the histogram of eyes–head lags (in ms). Most of the values (86.0%) are positive, indicating that eyes are leading head, and only a small proportion of cases (14.0%) indicate that the eyes are trailing head.

Note that all four methods indicate that eyes were leading head in a substantial majority of cases (86.2% for the minimum-distance method, 85.7% for the direction method, and 86.0% for the tracking method). An analysis of the cases where the head lead the eyes showed that they occurred mostly when the head is either moving slowly or varies randomly around a more or less stable position. It is noteworthy, however, that the present study involved passive viewing of a static image. Paradigms that include situations where a person is reacting to a flashing light in the periphery, or other dynamic changes, could produce different results.

Previous work on the coordination between eye and head movements suggests that, for small eye movements (< 45°), the eye leads the head, and for larger shifts (> 60°)
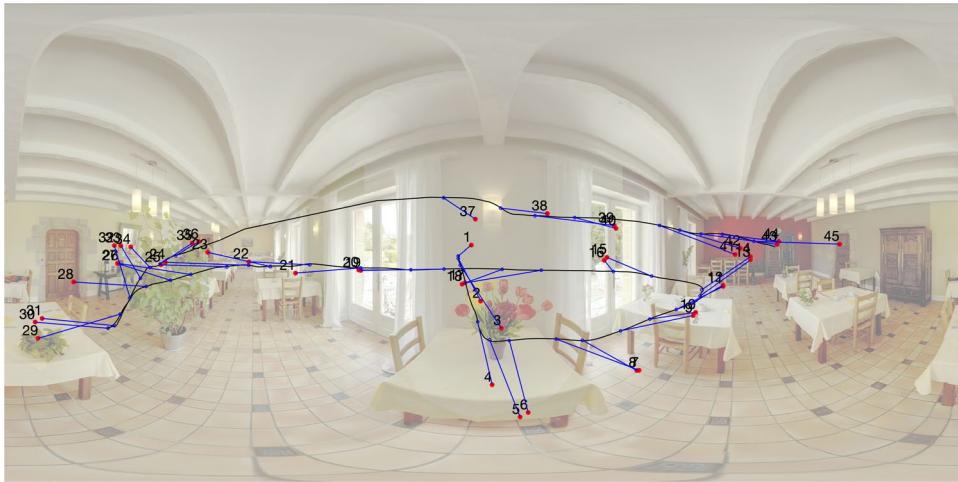
Fig. 12 Example panorama map with eye and head points. *Note.* Example map with eye and head positions. The *red circles* indicate eye fixations, the *black line* shows the head positions, the *blue circles* indicate head averages during the fixations, and the *blue lines* con- nect fixations with the corresponding head averages. These lines thus represent eyes-in-head, i.e., the direction and distance between eye points and head points
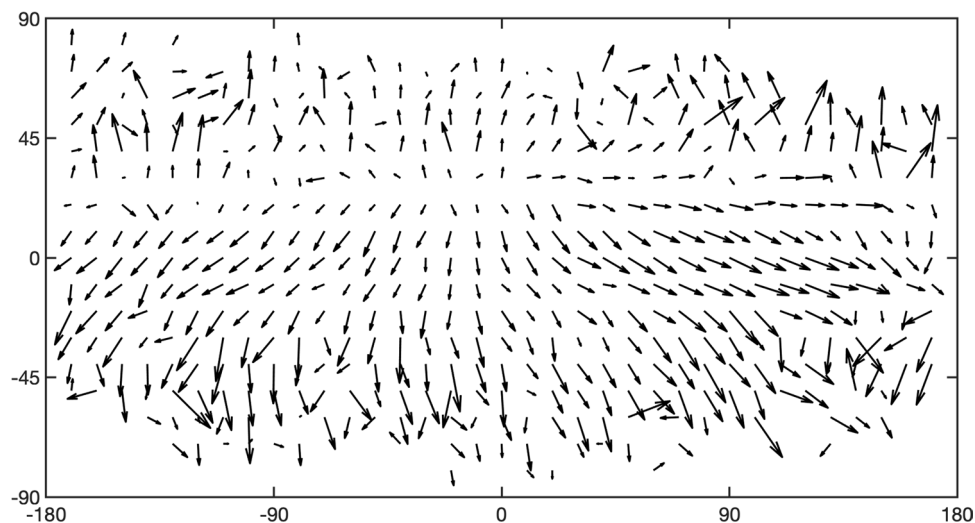

Fig. 13 Eyes-in-head directions and amplitudes. *Note.* Eyes-in-head directions and amplitudes, averaged at regular head-point intervals of 10° longitude and 10° latitude and averaged over all scenes and participants in the head-free condition. *Arrows* start at the head points and end at the eye points, with zero length indicating that, on average, head points and eye points coincide. Above latitude 30°, most arrows are pointing up, below latitude 30°, most are pointing down, west of longitude – 10° most are pointing west, and east of longitude – 10°, most are pointing east. This indicates that the eyes are expanding the range of positions defined by the head

the initiation of the two tends to be more synchronous (e.g., Barnes, 1979). In both cases, however, the eyes terminate in advance of the slower head movements owing to longer contraction times for the neck muscles and the greater inertial forces acting on the head compared to the eye (Bizzi et al., 1971; Freedman, 2008; Gilchrist et al., 1998). Interestingly, the conditions that result in the head leading the eyes are relatively few, including, for example, preparation for a specific task-oriented event, such as shoulder checking in a car before changing lanes (Doshi & Trivedi, 2012) or choosing to move the eyes into space that is outside visible range, such as when looking at the world through binoculars (Sidenmark & Gellersen, 2019; Solman et al., 2017).
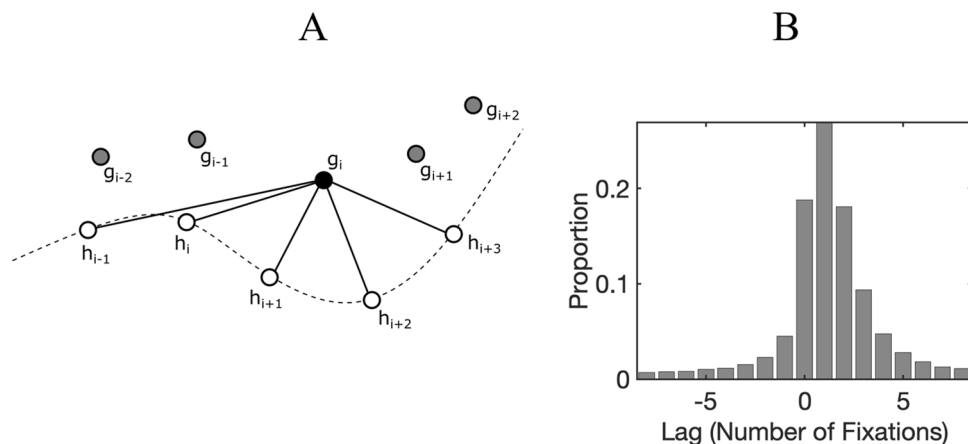
A

B

Fig. 14 Minimum distance method for analyzing the temporal relationship between eyes and head. *Note.* **A** The *black dot* shows the position of a fixation g_i, the *grey dots* g_{i-2}, g_{i-1}, g_{i+1}, g_{i+2} show the position of fixations preceding and following g_i, and the *white dots* h_{i-1}, h_i, … show the head averages during the fixations. The *dashed line* indicates the smooth movement of the head positions. The *black lines* show the distances between the fixations g_i and the head posi-

tions. In this case, the distance $g_i - h_{i+1}$ is minimal, indicating the gaze is leading head at this time point. **B** Histogram of gaze-head lags with the lag expressed in number of fixations. If gaze has a positive lag, then gaze is leading head, and if it has a negative lag, then gaze is trailing head. The results show that the eyes are leading head in most cases

Temporal relation between eyes and head

In addition to examining the spatio-temporal relationship between head and eyes, one can also ask, independent of where in space an eye movement is going, what the temporal relationship between the head and eyes is (e.g., how long after an eye movement does a head movement begin)? Freedman (2008) has already provided an extensive discussion of the coordination of eyes and head during visual orienting, both with head-restrained saccades (corresponding to our head-fixed condition) and head-free saccades. In the present tutorial, we have explored similar issues with longer-duration (10 s) explorations of omni-directional panoramas in the head-free condition.

Figure 18 shows that eyes-in-head eccentricity varies systematically in temporal relation to saccade initiation. Before the start of a saccade, that is during fixations, eyes-in-head eccentricity diminishes gradually as the trailing head catches up to the eye, reaching a minimum of about 17.5° at the beginning of a saccade, followed by a rapid increase. An analysis by participants revealed that all participants showed the same effect, with the participant-wise minimum eccentricity (in a range of approximately 11° to 22.5°) coinciding with saccade starts. The temporal eccentricity effect can be explained as follows: Head movements are smooth and follow the eyes with a lag of about one fixation. The eyes jump in saccades and stay more or less locked in place during fixations. Hence, the head has time to catch up somewhat with the eyes during a fixation, until the eyes jump in a saccade to the next position.

Note that the finding presented in Fig. 18 should not be confused with the vestibulo-ocular reflex (VOR) or the opto-kinetic nystagmus. These compensatory mechanisms move the eyes in the direction opposite to the head movement in order to stabilize the perceptual input (see e.g., Barnes, 1979; Carpenter, 1988; Laurutis & Robinson 1986; Leigh & Zee, 2015). Figure 19 shows the angle between head direction and eyes-in-head direction for different ranges of head velocities. For head velocities < 200 °/s, most head directions and eyes-in-head directions are aligned, but for head velocities > 200 °/s, eyes-in-head directions are opposite to the head direction. We attribute the latter effect to the VOR, and the results show that in our study the VOR occurs only for relatively large head velocities.

Summary and future directions

In the present tutorial, we began by identifying a meaningful gap in the scientific literature that can be addressed by allowing observers to move their head and eyes freely in VR (Section "Introduction"). We then reviewed the method for collecting such data, detailing the manner that the data for this tutorial was collected (Section "A brief introduction to VR eye movement research"). In Section "Method" we introduced the different frames of reference that are essential to handling the data and isolating different head and eye movement events. We then considered foundational head and eye movement analyses that are currently used in the field, bench-marking these against the classic eye movement method of immobilizing the head (Section "Alternative
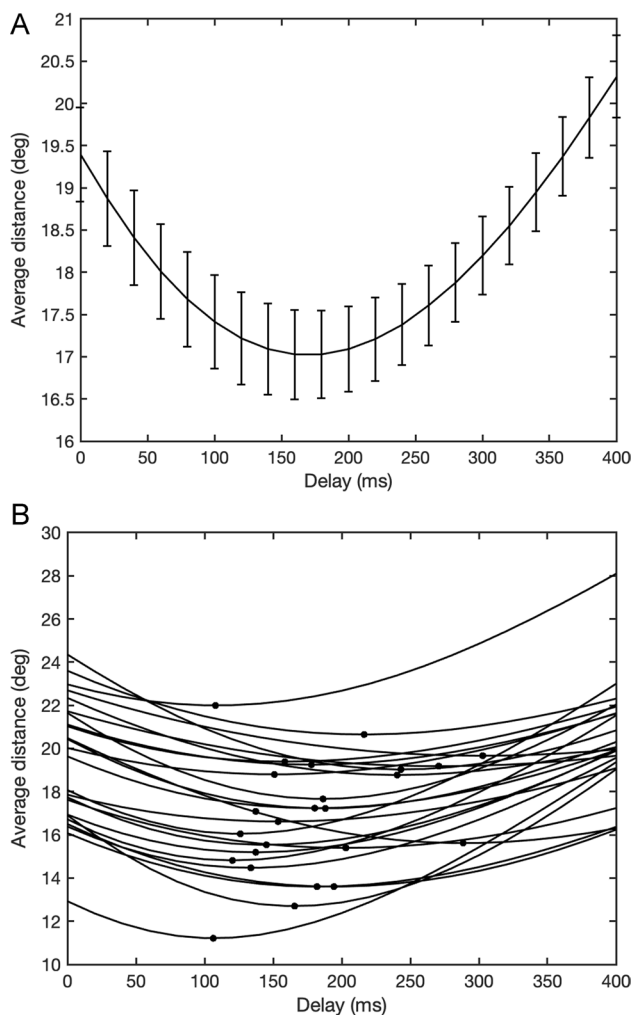
Fig. 15 Average distance method for analyzing the temporal relationship between eyes and head. *Note.* **A** The *curve* shows the Kangas et al. (2022) distance curve for our study, averaged over participants and scenes, with the *bars* indicating ± 1 standard error. The minimum is reached at a delay of about 170 ms. **B** Individual curves for each participant in our study, averaged over all scenes, with the *dots* indicating the minimum of each curve. The results show a substantial variation of the distance curves between participants, with the minima ranging between about 100 and 300 ms

representations of panoramic spheres"). In Section "Data and basic analysis methods" we extended the data analyses to novel approaches that consider the interplay between head and eyes in space, time, and space-and-time.

Below we provide a brief high-level summary of what these analyses revealed regarding the similarities and differences between the eyes and head, and what they suggest regarding the role of head and eyes in the perception of the environment. We close by considering how the analytical techniques we introduced for panoramic scene viewing can be extended to examine head and eye movements in other VR environments.

Similarities between eyes and head

The analyses revealed a close spatial relationship between eyes and head. Eye points remain in close spatial proximity to head points at all times, with a standard deviation of eyes-in-head eccentricity below 15°. This result is consistent with earlier results in the literature that eye positions are limited to a range of about 15–25° of the head positions (David et al., 2022; Freedman & Sparks, 1997). In the exploration of panoramic scenes, this spatial proximity is the result of the head following the eyes continuously. Given the rapid displacement of the eyes during saccades, the discrepancy between eyes and head is reduced during fixations.

In classical eye movement studies with an immobilized head, the eye movements are alone responsible for directing attention to different regions of the stimulus, while in studies that permit free head movements, this is achieved through a combination of eye and head movements. As a result, the spatial distributions of head points and eye points are similar, as illustrated in Fig. 4, which shows the heatmaps of the head points (for conditions free and head-fixed condition) and in Fig. 6, which shows the heatmaps of the eye points (again for the two conditions).

Differences between eyes and head

Despite the similarities between eyes and head, the analyses revealed that there are systematic differences between the two, both with respect to spatial and temporal characteristics. The spread of eye points was larger than the spread of head points (see Table 1; and Bischof et al., 2019). The analysis of the spatial relation between eyes and head (Fig. 13) showed why this difference arises. Above latitude 30°, most eyes-in-head arrows are pointing upward, and below, most are pointing downwards. Along the horizontal direction, most arrows west of the − 10° meridian are pointing westward, and east of the meridian, they are pointing eastward. Together, this shows that the eyes are expanding the range of locations defined by the head, leading to a larger eye spread. This result can be understood in terms of the differential effort to move head and eyes, with the heatmaps of head and eyes characterizing the respective distributions. One potential limitation of this finding is how much the weight of the headset itself may impact the effort required to move the head, and therefore, the propensity to do so.

We also found that eyes-in-head eccentricity varies systematically in temporal relation to fixations and saccades (Fig. 18). Before the start of a saccade, i.e., during a fixation, eyes-in-head eccentricity diminished slowly, reaching a minimum at the beginning of a saccade. In other words, during fixations, the head movements lead to the head direction being more closely aligned with eye direction. After the initiation of a saccade, there is a rapid increase of the
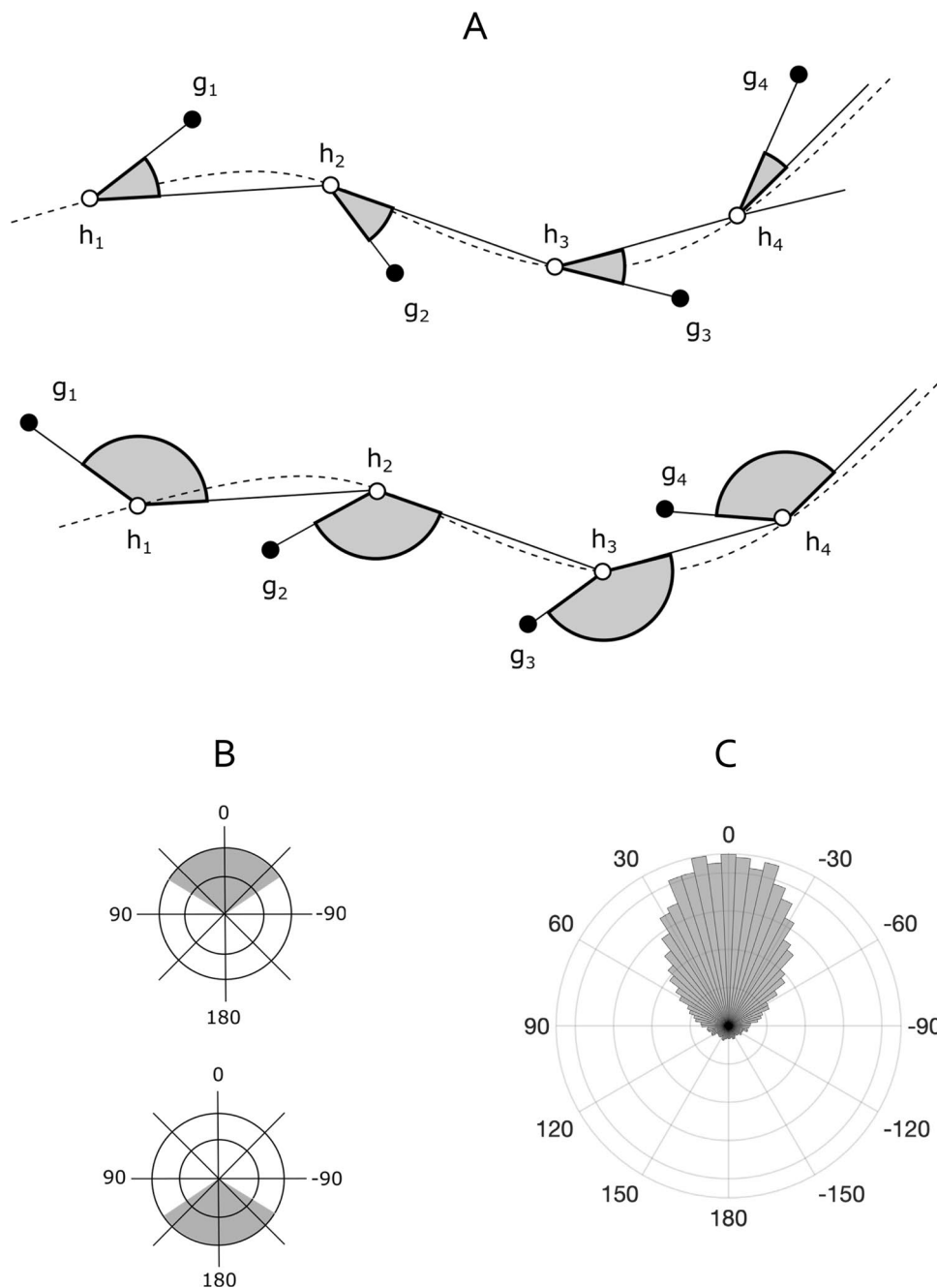
Fig. 16 Direction method for analyzing the temporal relationship between eyes and head. *Note.* **A** The *black dots* g_1, g_2, … show the position of successive fixations and the *white dots* h_1, h_2, … show the head averages during each fixation. The *grey angles* indicate the angle between the lines g_i-h_i and the lines connecting successive head positions $h_i - h_{i+1}$, and the dashed line indicates the smooth movement of the head positions. The *small angles* in the upper sketch indicate that eye is leading head, whereas the *large angles* in the lower sketch indicate that eyes are trailing head. **B** If eyes are leading head, then the gray angles are on average small, as indicated by the upper sketch of a polar histogram. If eyes are lagging behind head then the grey angles are on average large., as indicated by the lower sketch of a polar histogram. **C** Empirical polar histogram

eyes-in-head eccentricity, lasting for the duration of the saccade, after which the cycle repeats.

Regarding the temporal relation between eyes and head, it was found consistently that the eyes are leading head in most cases by about 100–200 ms (or one fixation). While this is true for the passive exploration of panoramas investigated in the present study, it may be different in cases of active interaction with objects in the environment or active navigation through an environment. A fruitful avenue of future research would be to understand how control mechanisms (top-down/
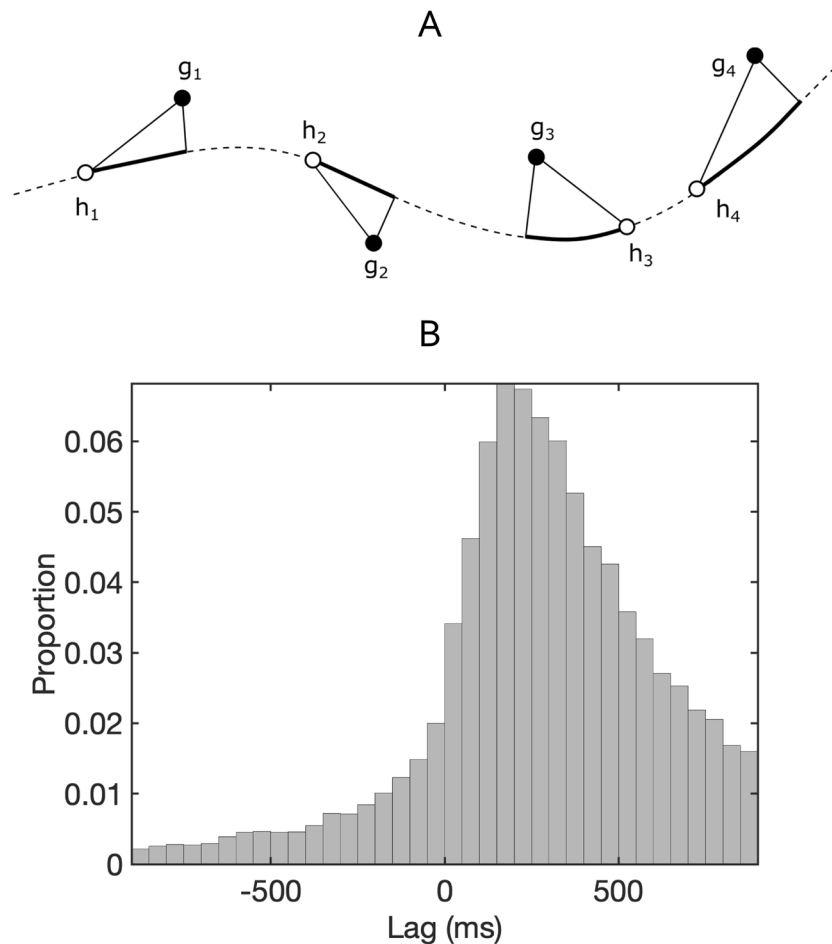
Fig. 17 Tracking method for analyzing the temporal relationship between eyes and head. *Note.* **A** The *black dots* g_1, g_2, … show the position of successive fixations and the *white dots* h_1, h_2, … show the head averages during each fixation. The *dashed line* indicates the smooth movement of the head positions. The *black lines* are obtained by projecting the lines g_i-h_i onto the continuous lines of head positions. In this example, g_1, g_2, and g_4 are leading head while g_3 is trailing head. **B** Histogram of eye-head lags (in ms). Most of the values are positive, indicating that the eyes are leading head

bottom-up) and high-level tasks may affect these very consistent spatio-temporal relationships encountered here.

Taken together, the results suggest that the functional relationship between head and gaze appears to be complementary: First, the eyes appear to compensate for the limited range of head positions by expanding the range of eye positions (Fig. 13). Second, while the eyes can be adjusted rapidly, the head can slowly catch up with the eye direction during fixations (Fig. 18). Third, the eyes are responsible for stabilizing the retinal image during rapid head movements (Fig. 19).

The role of head and eyes in visual perception

The majority of past studies measured eye movements when the observer's head was immobilized. Eye behavior without head movements may, however, not reflect the dynamics of gaze selection in the real world where large scanning eye movements are supported by head movements. First, when studying gaze behavior without head movements, visual information is pre-selected by the experimenter, whereas in studies with free head movements, visual information is selected by the observer. Second, when head movements are permitted, eyes and head jointly control the direction of gaze, and the eyes enable the stability of the retinal images during large head movements (Fig. 17; Einhäuser et al., 2009). Third, looking behavior changes substantially from head-fixed passive viewing to viewing with free head movements.

The present paper reported a direct comparison between visual exploration with the head-fixed condition and visual exploration with the head freely moving. Most notably, the direction, velocity, and angular variance of the eyes relative to the head seemed to be remarkably stable regardless of whether the head was restrained. This is an encouraging finding for the generalizability of head-fixed eye tracking
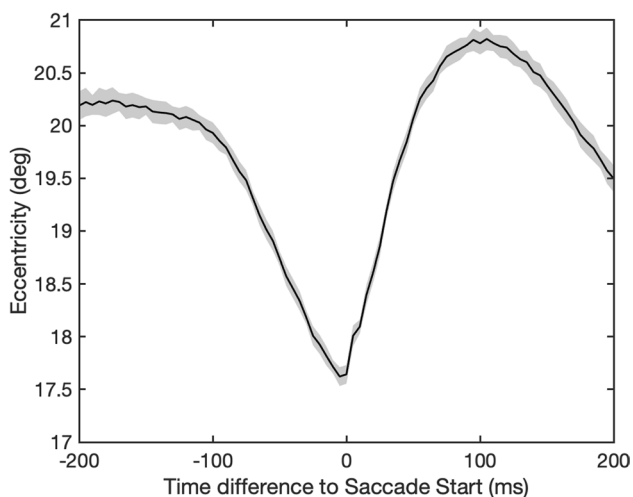
Fig. 18 Eyes-in-head eccentricity in temporal relation to saccade starts. *Note*. Average eyes-in-head eccentricity in relation to the saccade starts, with the *grey band* indicating its 95% confidence interval. It diminishes during fixations, reaching a minimum just before a new saccade is initiated (at time difference 0), followed by a rapid increase in eccentricity

experiments. In addition, the contributions of head and eyes with free head movements also depend on the observer's task. When a participant has to focus on additional tasks, e.g., locating a target in the environment, actively interact with object in the environment (Land & Hayhoe, 2001; Land & Tatler, 2009), or navigate through an environment (e.g., Foulsham et al., 2011), then the head may be under more deliberate cognitive control.

The present tutorial was concerned exclusively with the exploration of static 360° panoramas. Although observers were seated in a stationary chair, recent work indicates that the key analyses and findings extend to when a person is seated in a swivel chair or is standing (Mehrotra et al., 2024) although, overall, more of the panorama is generally explored. Similarly, we expect that with some additions, the methods described here can be extended to dynamic panoramas (i.e., panoramas that change over time). Such stimuli can be produced by projecting 360° movies rather than static images onto the virtual sphere surrounding the participants (see for example, 360cities, n.d.). To deal with the dynamic stimulation and addition of smooth-pursuit eye movements,
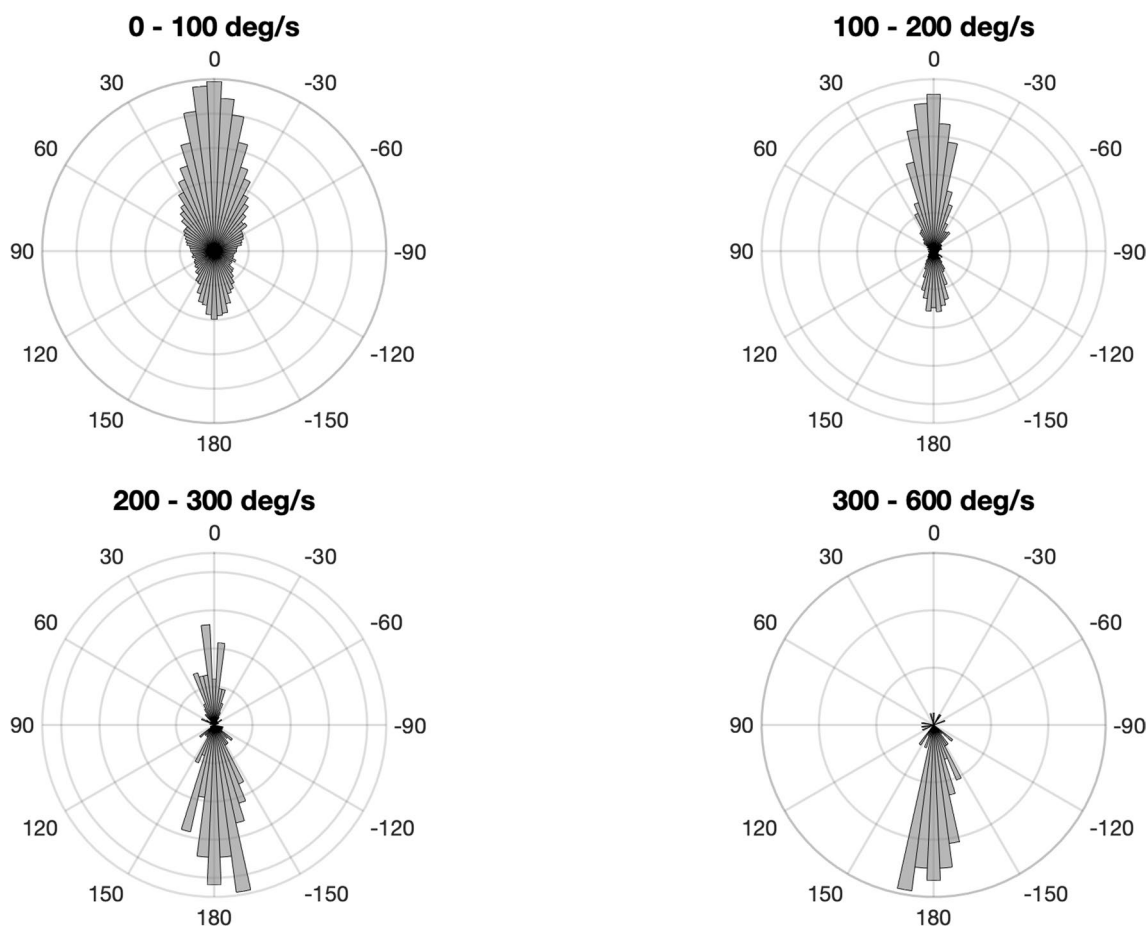

Fig. 19 Vestibulo-ocular reflex. *Note*. Angles between head direction and eye direction for different ranges of head velocities. For velocities < 200°/s, most head directions and eye directions are aligned, but for head velocities > 200°/s, most eye directions are opposite to the head direction, which we attribute to the vestibulo-ocular reflex

the algorithm for detecting ocular events (Method) will need to be expanded (e.g., by adapting the algorithms proposed by Dar et al., 2021; Komogortsev & Karpov, 2013), possibly by relating eye and head movements to local motions in the panoramas (see e.g., Chapel & Bouwmans, 2020). If one knows what objects are being introduced into the dynamic environment, then this should be relatively straight-forward. If not, then at this moment, it is not yet possible to recognize and identify objects in a scene automatically, although the technology on this front is making rapid progress. Regardless, the analyses of head, eyes-in-head, and eyes-in-space will be comparable to those presented here. On a similar note, the principles of eye and head analysis can also be extended to VR environments in which the observer is free to locomote within a dynamic environment.

In summary, the analysis of eye movements in dynamic panoramas is an exciting avenue for future research. We hope that this tutorial has provided the reader with an overview of the analytical tools to study head and eye movements, not only in the exploration of 360° panoramas, but also in more general settings and empower them to create their own analyses to answer the unique questions about eye movements that emerge.

Data availability The data underlying the results presented in the study are available in the repository Open Science Framework (https://doi.org/10.17605/OSF.IO/THR89). None of the experiments was preregistered.

Code availability The code of the main analysis programs is available in the repository Open Science Framework (https://doi.org/10.17605/OSF.IO/THR89).

Declarations

Conflict of interest / Competing interests All authors report no conflict of interests

Ethics approval The study was approved by the ethics board of the University of British Columbia (H10-00527). The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

Consent to participate All participants provided informed consent prior to participation.

Consent for publication We confirm that this work is original and has not been published elsewhere, nor is it currently under consideration for publication elsewhere.

References

360cities (n.d.) *360cities*. Retrieved October 10, 2023, from https://360cities.net

Andersson, R., Larsson, L., Holmqvist, K., Stridh, M., & Nyström, M. (2017). One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms. *Behavior Research Methods, 49*, 616–637. https://doi.org/10.3758/s13428-016-0738-9

Anderson, N. C., Bischof, W. F., Foulsham, T., & Kingstone, A. (2020). Turning the (virtual) world around: Patterns in saccade direction vary with picture orientation and shape in virtual reality. *Journal of Vision, 20*(8), 1–19. https://doi.org/10.1167/jov.20.8.21

Anderson, N. C., Bischof, W. F., & Kingstone, A. (2023). Eye Tracking in Virtual Reality. In C. Maymon, G. Grimshaw, & Y. C. Wu (Eds.), *Virtual Reality in Behavioral Neuroscience: New Insights and Methods.* Springer, UK: Current Topics in Behavioral Neurosciences. https://doi.org/10.1007/7854_2022_409

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), 3–3. https://doi.org/10.1167/jov.20.5.3

Barnes, G. R. (1979). Head-eye coordination in normals and in patients with vestibular disorders. *Advances in Oto-Rhino-Laryngology, 25*, 197–201. https://doi.org/10.1159/000402942

Batschelet, E. (1981). *Circular statistics in biology*. Academic Press.

Bischof, W. F., Anderson, N. C., & Kingstone, A. (2019). Temporal Methods for Eye Movement Analysis. In C. Klein & U. Ettinger (Eds.), *Eye Movement Research: An Introduction to its Scientific Foundations and Applications* (pp. 407–448). Springer. https://doi.org/10.1007/978-3-030-20085-5_10

Bischof, W. F., Anderson, N. C., Doswell, M. T., & Kingstone, A. (2020). Visual exploration of omni-directional panoramic scenes. *Journal of Vision, 20*(7), 1–29. 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

Bizzi, E., Kalil, R. E., & Tagliasco, V. (1971). Eye-head coordination in monkeys: Evidence for centrally patterned organization. *Science, 173*, 452–454.

Boger, Y. (2017). *Understanding Pixel Density & Retinal Resolution, and Why It's Important for AR/VR Headsets.* Retrieved October 10, 2023, from https://www.roadtovr.com/understanding-pixel-density-retinal-resolution-and-why-its-important-for-vr-and-ar-headsets

Bourke, P. (2020). *Converting to/from cubemaps.* http://www.paulbourke.net/panorama/cubemaps

Carpenter, R. H. S. (1988). *Movements of the Eyes* (2nd ed.). Pion Limited.

Chapel, M.-N., & Bouwmans, T. (2020). Moving objects detection with a moving camera: A comprehensive review. *Computer Science Reviews, 38*, 100310. https://doi.org/10.1016/j.cosrev.2020.100310

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

Dar, A. H., Wagner, A. S., & Hanke, M. (2021). REMoDNaV: Robust eye-movement classification. *Behavior Research Methods, 53*, 399–414. https://doi.org/10.3758/s13428-020-01428-x

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

Doshi, A., & Trivedi, M. M. (2012). Head and eye dynamics during visual attention shifts in complex environments. *Journal of Vision, 12*(2), 1–16. https://doi.org/10.1167/12.2.9

Einhäuser, W., Moeller, G. U., Schumann, F., Conradt, J., Vockeroth, J., Bartl, K., Schneider, E., & König, P. (2009). Eye-head coordination during free exploration in human and cat. *Annals of the New York Academy of Sciences, 1164*, 353–366. https://doi.org/10.1111/j.1749-6632.2008.03709.x

Equirectangular Projection. (n.d.). In *Wikipedia*. Retrieved December 23, 2019, from 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

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 / Revue canadienne de psychologie expérimentale, 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*, 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

Gilchrist, I. D., Brown, V., Findlay, J. M., & Clarke, M. P. (1998). Using the eye-movement system to control the head. *Proceedings of the Royal Society of London B, 265*, 1831–1836. https://doi.org/10.1098/rspb.1998.0509

Goldman, R. (2022). Rethinking Quaternions. *Springer*. https://doi.org/10.1007/978-3-031-79549-7

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

Greene, N. (1986). Environment mapping and other applications of world projections. *IEEE Computer Graphics and Applications, 6*, 21–29.

Hartley, R. & Zisserman, A. (2004). *Multiple View Geometry in Computer Vision* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511811685

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

Holmqvist, K., & Andersson, R. (2017). *Eye tracking: A comprehensive Guide to Methods*. CreateSpace Independent Publishing Platform.

Hooge, I. T., Niehorster, D. C., Nyström, M., Andersson, R., & Hessels, R. S. (2018). Is human classification by experienced untrained observers a gold standard in fixation detection? *Behavior Research Methods, 50*, 1864–1881. https://doi.org/10.3758/s13428-016-0738-9

Hooge, I., 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, 20.8.2019. *Journal of Eye Movement Research, 12*(7), https://doi.org/10.16910/jemr.12.7.8. https://doi.org/10.16910/jemr.12.7.8

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

Jeong, J.-B., Lee, S., Ryu, I.-W., Le, T. T., & Ryu, E.-S. (2020). Towards Viewport-dependent 6DoF 360 Video Tiled Streaming for Virtual Reality Systems. In: *MM '20: Proceedings of the 28th ACM International Conference on Multimedia*, (pp. 3687–3695). https://doi.org/10.1145/3394171.3413712

Kangas, J., Špakov, O., Raisamo, R., Koskinen, O., Järvenpää, T., & Salmimaa, M. (2022). Head and gaze orienting in hemispheric image viewing. *Frontiers in Virtual Reality, 3*, 822189. https://doi.org/10.3389/frvir.2022.822189

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., & Karpov, A. (2013). Automated classification and scoring of smooth pursuit eye movements in the presence of fixations and saccades. *Behavior Research Methods, 45*(1), 203–215. https://doi.org/10.3758/s13428-012-0234-9

Komogortsev, O. V., Gobert, D. V., Jayarathna, S., Koh, D., & Gowda, S. (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

Lambers, M. (2020). Survey of cube mapping methods in interactive computer graphics. *The Visual Computer, 36*, 1043–1051. https://doi.org/10.1007/s00371-019-01708-4

Land, M. F. (2004). The coordination of rotations of the eyes, head and trunk in saccadic turns produced in natural situations. *Experimental Brain Research, 159*, 151–160. https://doi.org/10.1007/s00221-004-1951-9

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

Land, M. F., & Tatler, B. W. (2009). *Looking and acting: Vision and eye movements in natural behaviour*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198570943.001.0001

Lang, B. (2018). *Understanding the difference between 'Screen Door Effect', 'Mura', & 'Aliasing'*. Retrieved October 10, 2023, from https://www.roadtovr.com/whats-the-difference-between-screen-door-effect-sde-mura-aliasing-vr-headset

Lange, F. (2019). HMD-Eyes. GitHub Repository, https://github.com/pupil-labs/hmd-eyes

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

Laurutis, V., & Robinson, D. (1986). The vestibulo- ocular reflex during human saccadic eye movements. *Journal of Physiology, 373*, 209–33. https://doi.org/10.1113/jphysiol.1986.sp016043

Lee, W. J., Kim, J. H., Shin, Y. U., Hwang, S., & Lim, H. W. (2019). Differences in eye movement range based on age and gaze direction. *Eye, 33*, 1145–1151. https://doi.org/10.1038/s41433-019-0376-4

Leigh, R. J., & Zee, D. S. (2015). *The Neurology of Eye Movements* (5th ed.). Oxford Academic. https://doi.org/10.1093/med/9780199969289.001.0001

Mardia, K. V., & Jupp, P. E. (2000). *Directional Statistics*. Wiley. ISBN: 978-0-471-95333-3.

Mehrotra, A., Silver, C., Jacobs, O., Bischof, W. F., & Kingstone, A. (2024). *Sit, Stand, or Swivel? Posture Affects Visual Exploration*

of Panoramic Scenes in Virtual Reality [Manuscript in preparation]. University of British Columbia.

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

Nyström, M., & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade and glissade detection in eyetracking data. *Behavior Research Methods, 42*, 188–204. https://doi.org/10.3758/BRM.42.1.188

Ozioko, O., & Dahiya, R. (2022). Smart tactile gloves for haptic interaction, communication, and rehabilitation. *Advanced Intelligent Systems, 4*, 2100091. https://doi.org/10.1002/aisy.202100091

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

Quaternions and spatial rotation. (n.d.) In *Wikipedia*. Retrieved October 10, 2023, from https://en.wikipedia.org/wiki/Quaternions_and_spatial_rotation

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 eye. *Current Directions in Psychological Science, 25*(1), 70–74. https://doi.org/10.1177/0963721415617806

Rizzo, A. S., Goodwin, G. J., De Vito, A. N., & Bell, J. D. (2021). Recent advances in virtual reality and psychology: Introduction to the special issue. *Translational Issues in Psychological Science, 7*(3), 213–217. https://doi.org/10.1037/tps0000316

Rötth, A. (1925). Über das praktische Blickfeld [On the practical field of fixation]. *Graefe's Archive for Clinical and Experimental Ophthalmology, 115*(2), 314–321.

Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the Eye-Tracking Research and Applications Symposium* (pp. 71–78). ACM Press. https://doi.org/10.1145/355017.355028

Sidenmark, L, & Gellersen, H. (2019). Eye, head and torso coordination during eye shifts in virtual reality. *ACM Transaction on Computer–Human Interaction, 27*(1), 4:1–4:40. https://doi.org/10.1145/3361218

Sitzmann, V., Serrano, A., Pavel, A., Agrawala, M., Gutiérrez, 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., & 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

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

't Hart, B. M., Vockeroth, J., Schumann, F., Bartl, K., Schneider, E., König, P., & Einhäuser, W. (2009). Eye allocation in natural stimuli: Comparing free exploration to head-head-fixed condition viewing conditions. *Visual Cognition, 17*(6–7), 1132–1158. https://doi.org/10.1080/13506280902812304

The MathWorks Inc. (2023). *MATLAB version: 9.14.0 (R2023a)*. Natick, Massachusetts: The MathWorks Inc. https://www.mathworks.com

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

Unity Technologies. (2017). *Unity Software: Release 2017.4.1*. Retrieved September 15, 2017, from https://unity3d.com

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

Vince, J. (2021). *Quaternions for Computer Graphics*. Springer. https://doi.org/10.1007/978-1-4471-7509-4

von Noorden, G. K., & Campos, E. C. (2002). *Binocular vision and ocular motility: Theory and management of strabismus* (vol. 6). Mosby.

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). Sun database: Large-scale scene recognition from abbey to zoo. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* (pp. 3485–3492). IEEE http://ieeexplore.ieee.org/abstract/document/5539970/

Zangemeister, W. H., & Stark, L. (1982). Types of eye movements: Variable interactions of eye and head movements. *Experimental Neurology, 77*, 563–577.

Zhan, T., Zou, J., Xiong, J., Chen, H., Liu, S., Dong, Y., & Wu, S.-T. (2020). Planar optics enables chromatic aberration correction in immersive near-eye displays. In: B. C. Kress, & C. Peroz (Eds.), *Optical Architectures for Displays and Sensing in Augmented, Virtual, and Mixed Reality (AR, VR, MR), Proceedings SPIE, 11310* (p 1131003). https://doi.org/10.1117/12.2542365

Zhao, C., Kim, A. S., Beams, R., & Badano, A. (2022). Spatiotemporal image quality of virtual reality head mounted displays. *Scientific Reports, 12*, 20235. https://doi.org/10.1038/s41598-022-24345-9