

GlanXR: A Hands-Free Fast Switching System for Virtual Screens

Guanghan Zhao*
Osaka University

Jason Orlosky
Augusta University

Kiyoshi Kiyokawa
Nara Institute of Science and Technology

Yuki Uranishi
Osaka University

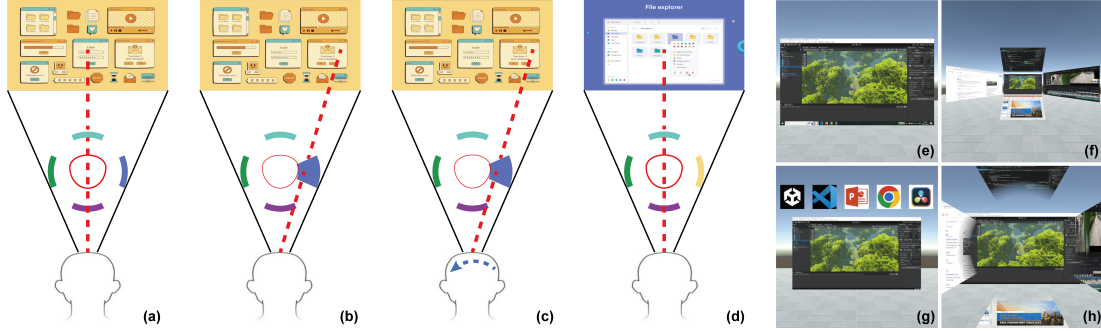


Figure 1: Images showing the procedure to trigger window switching using GlanXR: (a) the user looks at a virtual screen with the eye-head position within the adaptive threshold range, (b) the user selects a peripheral thumbnail by surpassing this range, which displays one of the selectable screens shown in (h), (c) the user sustains focus on the thumbnail and makes a slight head rotation in the opposite direction, and (d) the switch to the newly selected screen. We also show screenshots of the four methods tested in our experiment: (e) traditional taskbar with icons (located at the bottom of the screen), (f) an expansive multi-screen layout, (g) a gaze-selected icon interface, and (h) GlanXR.

ABSTRACT

To date, virtual and augmented reality technologies enable users to view multiple, large virtual screens in their workspaces. However, users must frequently rotate their heads to shift focus among these screens. This paper presents GlanXR, a fast and robust hands-free approach for screen switching in virtual reality. GlanXR incorporates a peripheral interface that remains fixed within the user's view, in which screens can be dynamically selected based on the user's eye-head position beyond an adaptive range. Additionally, the user triggers the switch to the screen chosen by making an opposing head rotation in the direction of the eye-head position to minimize false triggers. We conducted an experiment including a fast-switching scenario and a working simulation scenario with 24 participants to assess the effectiveness of GlanXR as compared to a baseline (taskbar), an expansive multi-screen setup, and a gaze-based screen selection method. The results indicate that GlanXR facilitates precise screen-switching, minimizes the necessity for head rotation, and allows users to maintain a neutral head position.

Index Terms: Gaze, Eye-Head Coordination, Hands-Free Interaction, Screen Switching

1 INTRODUCTION

With recent advancements in virtual reality (VR) and augmented reality (AR) technologies, users can effortlessly view multiple large virtual screens using a single portable head-mounted display (HMD). However, it's still unclear how to efficiently navigate and operate across these screens for information retrieval and manipulation within diverse usage scenarios.

Current popular applications, such as Immersed [2] and Virtual Desktop [1], employ two commonly used methods that align with everyday physical screen usage: the traditional icon-clicking approach and the state-of-the-art expansive multi-screen approach.

However, both approaches have inherent limitations, especially when used in VR/AR environments, where these drawbacks are magnified. For the icon-clicking approach, users must move the cursor down to the taskbar and click on a small icon. This interrupts workflow, costing time for cursor navigation between the taskbar and the interface or requiring users to move their hands away from and back to the keyboard while typing. For the multi-screen approach, although it offers a hands-free experience, users must rotate their heads frequently, often maintaining uncomfortable positions for extended periods. This is likely to induce head and neck fatigue, particularly when considering the additional Head-Mounted Display (HMD) weight [33, 7, 20, 21]. Therefore, virtual screens are typically set to larger sizes to ensure acceptable resolution, amplifying the required amplitude of head rotation.

In light of the limitations associated with current methods, screen-switching systems utilizing hands-free interactions present a potential solution. With gaze sensors in many modern HMDs, gaze-based methods are becoming increasingly popular as a form of hands-free interaction. They are frequently utilized when hands are occupied, such as during driving [19] and cycling [41]. However, accuracy can be compromised by internal and external factors [11, 22]. To compensate for the loss of accuracy, these methods often incorporate additional mechanisms (e.g., dwelling or pursuit) to ensure accurate selection, but this can reduce task performance.

To help address these issues, we introduce a hands-free approach named "GlanXR," which integrates eye gaze, head rotation, and eye-head position. Specifically, the eye gaze refers to the combined gaze using eyes only, and the eye-head position represents the direction in the view where the eye gaze is pointing. In addition, we take advantage of the rapid and less energy-intensive nature of eye movement, complemented by the less jittery and more controlled characteristics observed in head movement [3, 22]. We employ real-time eye-gaze detection through the integrated eye tracker on the HMD. When the user's eye-head position surpasses a dynamically adjusted range, which adapts to eye-gaze movement and is determined by the root mean square method, the system selects the thumbnail associated with the peripheral interface in that direction (Figure 1). To mitigate potential visual distraction, we designed the thumbnails on the peripheral interface with soft-edged

*e-mail: zhao.guanghan@lab.ime.mmc.osaka-u.ac.jp

semi-transparency. When selected, these thumbnails expand to the margin of the dynamically adjusted range and smoothly transition from complete transparency to semi-transparency. This approach prevents flickering induced by random glancing and aids users in understanding the range for selecting the thumbnails, thereby reducing the likelihood of large-amplitude eye movements. Our contributions in this paper are summarized as follows:

- The design and implementation of an adaptive hands-free screen switching method that offers unique advantages in practical scenarios
- An experiment evaluating different hands-free methods against a baseline for metrics such as performance, head rotation, and user fatigue, addressing a gap in previous works focused on efficiency.
- Investigation and analysis of the natural ranges of users' eye-head positions, adding to the body of knowledge of gaze-interactive techniques.

2 RELATED WORK

In pursuit of robust and easy-to-use switching, researchers have directed their attention to the eyes and head—flexible parts of the human body well-suited for interactions beyond hands. Many studies have been conducted to demonstrate natural eye and head behaviors and propose hands-free interactive techniques that take advantage of those behaviors.

2.1 Eye-Head Position and Coordination

The physical motion limitations of human eyes, as demonstrated by Lee et al. [23], are approximately 44.9° in adduction, 44.2° in abduction, 27.9° in elevation, and 47.1° in depression. However, the eyes do not typically reach these limitations during normal viewing. Eye movements primarily govern gaze shifts, but head movements also contribute, especially when acquiring a target outside the current field of view [5, 4, 40, 12].

Recent studies have further explored gaze and head behaviors using modern technologies such as Head-Mounted Displays (HMDs) and have observed similar patterns as earlier studies in virtual environments [15, 30]. These prior studies have provided guidelines for designing new hands-free human-computer interaction techniques that take advantage of gaze navigation and head orientation. However, it is undeniable that the range of eye-head position and how much the head contributes may vary due to individual differences [13, 26, 35, 30]. While gaze and head behaviors have been observed to vary among different individuals, they have also been found to be consistent across different contexts [34, 35].

The work by Sidemark and Gellersen [30] demonstrated that the range of eye-head position is not a complete circular shape but extends slightly further in the downward direction relative to the center of view. Another study by Yi et al. [37] also supports the notion that the range is not a completely circular shape. However, they propose that the shape is an ellipse centered below the center of the view. To further investigate this, we explored the range of eye-head positions using the data gathered from our experiment in this study. We also consider a variety of content layouts for daily usage of screens and provide an adaptive range as the threshold for triggering screen switching through eye movements.

2.2 Gaze-Based Interactive Techniques

Substituting gaze ray hitting position directly for a cursor is considered impractical and potentially irritating in user interaction. The phenomenon referred to as the “Midas Touch” problem highlights the challenges associated with using eye movement for command activation, where every gaze instance triggers a command [17, 3]. Other deficits with tracker accuracy or hardware limitations can make affect the usability of techniques that rely on very precise tracking. In light of the challenges in pure gaze-based interactions,

approaches integrating additional mechanisms such as dwelling and pursuit have been proposed. For instance, Hansen et al. [14] examined gaze interaction with dwell time activation, comparing typing performance using mouse clicks and dwell time on keyboards.

Additional studies have explored solutions such as adjustable dwell time design [25], enlarged buttons [28], glanceable interfaces [24], or virtual copies [27], to reduce the required dwell time and enhance precision. However, these efforts have not made dwell-based gaze interactions satisfactory for daily use. In addition, for gaze pursuit inputs, Vidal et al. [36] introduced a pursuit interaction with moving objects by employing a tracking principle that aligns eye and object movement, eliminating the need for precise gaze calibration. While recognized as faster and more robust than dwell-based input, these methods need specially designed interfaces.

Given that the natural eye-head position range is smaller than the physical limitation of eye rotation, as demonstrated in previous studies, this phenomenon can potentially prevent the Midas Touch problem when interacting with eye-gaze without additional time-consuming mechanisms. With the same consideration, Yi et al. [37] proposed a gaze-only menu selection approach, utilizing a view-fixed peripheral menu layout. This design minimizes false triggers and enhances interaction speed by triggering appearance and selection based on the user's gaze proximity to the menu zone. However, their approach requires calibrations for each user to establish personalized boundaries for the layout. Enlightened by this work, we explore an implementation that leverages the same concept but without the need for pre-calibrations while preserving a preferred false triggering rate.

2.3 Head Orientation Featured Interactions

As head movement is more controlled and precise, head orientation-based methods have been suggested as an alternative to address the limitations of gaze-based methods, offering another avenue for hands-free interactions. Similar to gaze-based interactions, to provide a discrete and explicit confirmation mechanism for initiating a selection with head orientation and to mitigate the Midas Touch problem, integration of both dwelling [38] and pursuit [9] has been implemented. Other studies consider additional head factors, such as gestures and position. For instance, the work by Yu et al., [39], an approach called DepthMove, enables interactions based on head motions along the depth dimension. Specifically, users can shift their heads forward or backward to switch between distinct interfaces and engage with head rotations. Nevertheless, interaction involving head movement tends to be slower than gaze interactions [3, 6, 22], and it may introduce additional stress on the neck. Consequently, we believe that the optimal solution lies in combining the strengths of both gaze and head motions.

In the realm of hybrid techniques that incorporate head orientation for additional adjustment to gaze-based methods, the state-of-the-art method involves navigating the pointer by gaze and then smoothly reaching the target by the head for refined selection [18, 29, 31]. However, the procedure of refined selection by the head is likely to decelerate the overall activity. Beyond this method, innovative approaches such as Eye&Head Convergence [30] have been proposed. When the target is reached by gaze, the cursor expands to create a convergence area, and a head pointer emerges, facilitating confirmation by moving it into the convergence area. In line with this idea, Sidemark et al. [32] implemented an upgraded technique that utilizes gaze for pre-selection and head-crossing as a trigger. Both methods require the user to navigate gaze and head orientations precisely and hold one still while the other is moving, which may result in additional cognitive load, especially for fast selection scenarios in daily usage. Inspired by the prior works, we explore an approach that utilizes gaze for selection aligned with these works but incorporates a straightforward head movement as the triggering mechanism.

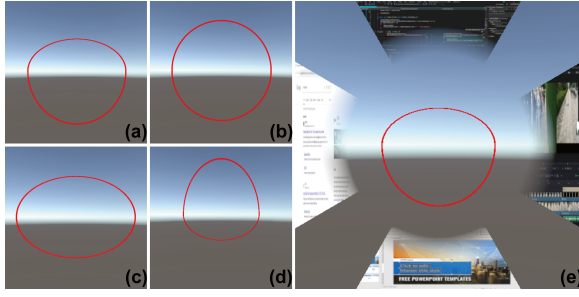


Figure 2: Images showing (a) the base shape of the threshold range, including halves of two ellipses joining together, (b-d) variations in threshold values allowing for different shapes of the threshold range, and (e) the layout of the peripheral interface with four thumbnails. Note that the threshold range is marked in red only for demonstration purposes; it is hidden in our implementation.

3 DESIGN AND IMPLEMENTATION OF GLANXR

The primary objectives of GlanXR are to 1) establish a fast and easy hands-free switching system for virtual screens, 2) reduce the required amplitude of eye and head movements during interactions to alleviate fatigue, and 3) offer an easily understandable and less distracting interface fixed at the periphery. To achieve these, we first implemented an algorithm to dynamically regulate an adaptive eye-head position threshold range by continuously collecting real-time eye-gaze data. If the position exceeds the range, it is detected as an unnatural eye-gaze movement utilized for screen selection. Secondly, a head-triggering mechanism was designed to confirm switching rapidly but robustly. Finally, a peripheral interface was developed to be less distracting and helpful for users to gain an overview of the contents and layouts on the screens.

3.1 Adaptive Eye-Head Position Threshold Range

The physical limitations of eye movements form a diamond-like shape, with shorter extents in the upward direction, while the downward direction is relatively longer [23]. Additionally, the range is nearly symmetric horizontally and has been considered symmetric for simplification in prior work [37]. Therefore, we design the base shape of the threshold range as halves of two ellipses joining together, merging at the top and bottom (Figure 2-(a-d)). To establish the horizontal, upward, and downward extents of the threshold range and dynamically resize it in real-time, the construction of datasets containing real-time axes data and the extraction of threshold values are essential.

To achieve these, we begin by gathering real-time combined eye-gaze data from the integrated eye tracker of the HMD, which is solely driven by eye movements. We calculate the angle values between the eye-gaze and the center to extract the eye-head position on the xy-axis in degrees. Additionally, we split the y-values into positive or negative for the top and bottom halves of ellipses. Then, according to the standard equation of an ellipse that is centered at the origin:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \quad (1)$$

where the a and b are the axes of the ellipse, respectively, we attempt to calculate the lengths of the axes of the ellipses where each gathered (x, y) is located. However, these conditions alone are insufficient to determine the lengths of the two axes. To address this, we applied ratios of approximately 28:44 for the axes when y is positive and 44:47 when y is negative, according to the physical limitations of eye movements [23]. This allowed us to establish the equations for the top-half and bottom-half ellipses:

$$\frac{x^2}{(44k_p)^2} + \frac{y^2}{(28k_p)^2} = 1 \quad (2)$$

and

$$\frac{x^2}{(44k_n)^2} + \frac{y^2}{(47k_n)^2} = 1 \quad (3)$$

in which the temporal multipliers k_p and k_n are applied separately for positive and negative conditions of y . Subsequently, we can determine the axes of the two halves of the ellipses through:

$$V_{top} = \frac{\sqrt{28^2x^2 + 44^2y_p^2}}{44}, H_{top} = \frac{\sqrt{28^2x^2 + 44^2y_p^2}}{28} \quad (4)$$

and

$$V_{bot} = \frac{\sqrt{47^2x^2 + 44^2y_n^2}}{44}, H_{bot} = \frac{\sqrt{47^2x^2 + 44^2y_n^2}}{47} \quad (5)$$

where the V_{top} and H_{top} are the vertical and horizontal axis of the top-half ellipse, the V_{bot} and H_{bot} are the vertical and horizontal axis of the bottom-half ellipse, the y_p and y_n are the positive and negative y -value of the eye-head position, and the x is the x -value of the eye-head position. These computed axes data are collected into datasets for threshold extraction. In addition, we combine the H_{top} and H_{bot} data into the same dataset and process it to calculate an averaged single threshold value, serving as the horizontal axis for both ellipses. This is intended to shape the threshold range as a closed contour, enabling the other side to automatically resize based on the ratios when the eye-head position is located on the top or bottom side (i.e., y is positive or negative).

The root mean square method, employing a time window of 30 seconds, is separately applied to the datasets to extract threshold values for these real-time data. It is sensitive to the magnitude of values, making it effective for capturing subtle fluctuations and significant variations in the datasets. Additionally, it imparts a smoothing effect, enhancing the stability of the analysis and facilitating the identification of overarching trends within the datasets. The time window is defined with further logical testing meant to capture meaningful patterns in eye-gaze behavior, considering the cognitive dynamics of fixation sequences during browsing, thus providing a balance between detecting rapid gaze shifts and assessing sustained attention. Once the thresholds for the axes are computed by the root mean square method, the equations for the two halves of the ellipses in the threshold range can be defined using Equation (1). Moreover, leveraging this formula, variations in threshold values allow for the generation and combination of different shapes, such as circles or complete ellipses.

When the eye-head position exceeds the defined threshold range, it is identified as an unnatural eye-gaze behavior and is subsequently employed to select screens from a peripheral interface. Since this threshold range can dynamically adapt to the user's real-time eye-gaze data, the eyes are not required to rotate to the far positions from the center. This mitigates potential eye fatigue that may accumulate when using similar approaches. Furthermore, as the user's natural eye-head position range may fluctuate over time and be influenced by tasks, screen content layouts, and individual differences, our algorithm can automatically adjust the threshold range in varying situations, ensuring robust performance.

3.2 Head-Triggering Mechanism

The head-triggering mechanism in our system is implemented to confirm the switching following the selection by eye-gaze with head movements. This process involves considering a head rotation as performed when it exceeds a velocity threshold of $20^\circ/s$. Considering that the head and eyes naturally move in the same direction during gaze-shifts, which may lead to false triggering, the system explicitly detects head rotations in the opposite direction to the eye-head position. Moreover, compared to orthogonal directions, the opposite direction is easier to understand and perform. To

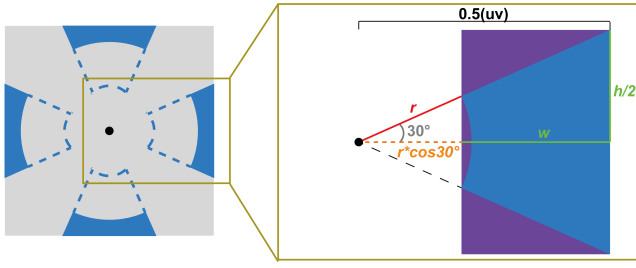


Figure 3: An illustration explains our calculation of the h and w of the zoomed small rectangles, referred to as purple, with the blue area representing the radial thumbnail clipped from the rectangle. All of these were calculated in uv coordinate of the screen space.

further prevent false triggering in the situation where the head is rotated before the eye-gaze could complete the movement to within the adaptive eye-head position threshold range, we applied a time window of 150ms following the head rotation. This duration is determined by leveraging the characteristic of synchronized eye-head movements, where for targets within view, head movements averagely start 150ms after the eye movement initiation [30]. If the eye-gaze remains fixated, with a velocity below 100°/s, beyond the threshold range for over 150ms after the head is rotated in the opposite direction, the switching will be successfully triggered. The velocity threshold values used are derived from demonstrations in previous studies [10, 30].

Simply put, the flow to trigger screen switching using our approach involves two steps: (1) Rotate the eyes toward the screen intended for switching, allowing the eye-head position to surpass the threshold range and maintain fixation. (2) Slightly rotate the head in the opposite direction. In addition, we add a restriction that screen switching can only be triggered when the head is facing the current screen in front. This restriction is essential for real-world usage scenarios where users may look around or interact with other devices, potentially leading to errors in the system.

This mechanism is designed to leverage the stable and controlled characteristics of head movements, serving as a robust confirmation following eye-gaze selections. Moreover, given that our approach necessitates only a slight head rotation without specific amplitude requirements, it minimizes stress on the neck and facilitates easy re-orientation to the front.

3.3 Peripheral Interface

GlanXR features a radial interface, primarily implemented by a shader programmed using High-Level Shading Language (HLSL), comprising thumbnails positioned along the circumference of the peripheral view (Figure 2-e). The thumbnails display the previews of each screen, aiming to assist users in conveniently and quickly understanding the current content on each screen. This is particularly helpful when multiple screens simultaneously show different windows of the same application, as thumbnails can prevent potential confusion when using icons. Moreover, the thumbnails on the interface can be re-positioned or added/reduced for further customization. In our implementation, we include four thumbnails, which are adequate for usage and provide the highest throughput [37]. To keep the resting thumbnails away from near-peripheral vision, thereby preventing visual disturbance while still allowing peripheral access, these thumbnails are initially positioned within a range of 35° away from the center to the edge (approximately 55°) of the HMD's field of view (FoV). Each has a field angle of 60° and is centered separately in the horizontal and vertical axes of the view. Furthermore, when a thumbnail is selected, it expands to the edge of the threshold range. This provides a clearer preview and enhances the user's awareness of the current threshold range, thereby helping prevent excessive eye movements.

To generate previews of the rectangular screens using radial thumbnails, we first define small rectangles that precisely cover the thumbnails in the shader (Figure 3). Given that the material coordinate system is scaled between 0-1 along the uv -axis, centered at (0.5, 0.5), and with half of the field angle (60°) of each thumbnail set to 30°, the height h and width w of the rectangles can be calculated as:

$$h = 0.5 \times \tan(30^\circ) \times 2 \quad (6)$$

$$w = 0.5 - r \times \cos(30^\circ) \quad (7)$$

where the r is computed by the threshold range value on the axes (or 35° converted to uv space if not selected), undergoing a proportional transformation to the material coordinate system. Also, the h remains a constant value as the field angle is predetermined, while the w can vary based on the dynamic threshold range or be fixed at the peripheral area when the thumbnail is not selected. Then, the center coordinates of the rectangles can be defined as follows:

$$center = (0.5, 0.5) - \frac{(w, h)}{2} \quad (8)$$

With this equation, the centers of the rectangles can remain centered in the middle of the thumbnails as the threshold range changes in real-time. Note that the definitions described above are for thumbnails located in the horizontal direction; for the vertical direction, the values of w and h should be swapped. Finally, we sample the pixels of the screen materials within the range of the small rectangles and clip them with the radial shape of the thumbnails.

Considering the potential distraction posed by the peripheral interface, further measures are implemented to enhance user experience. Specifically, the thumbnails are designed to be semi-transparent with soft edges. Moreover, to address instances where the user's eye-head position occasionally and rapidly exceeds the threshold range, triggering a selection, we apply a SmoothDamp technique. This enables a gradual fade-in effect that prevents flicker and provides awareness of the threshold range.

4 USER STUDY

Our primary design considerations for GlanXR were to (1) ensure faster switching between multiple virtual screens, (2) reduce the amplitude of head rotations while maintaining a relatively stable head position, and (3) mitigate distractions to the viewing of screen contents, as compared to other conditions.

In the experiment, we compared GlanXR to a traditional hands-on method; a taskbar with icons, and two other hands-free methods; an expansive multi-screen layout (head-based) and a gaze-selecting icons interface (gaze-based), as shown in Figure 1-(e-h). Additionally, because clicking icons on the taskbar is a common method for screen/window switching in both physical and virtual workspaces, and it does not involve any additional techniques, the taskbar icons condition in this experiment served as a baseline (control) condition. Participants were asked to perform assigned tasks in fast-switching and working simulation scenarios. We examined their completion time, head and eye behavior, and subjective preference with the four conditions.

4.1 Equipment

We used an HTC Vive Pro Eye with a FoV of 110° and a framerate of 90 fps, a Lenovo Legion Y740 PC, and a Logi M550 mouse. The HMD was integrated with a Tobii eye tracker operating at 120Hz. The experimental environment was developed and driven by Unity 2023.1.12f1. Additionally, for the working simulation scenario, an additional Intel Realsense D435i camera was employed.

4.2 Participants

Twenty-four naive individuals (9 female, 15 male, mean age 25.21, SD 4.85, range 19-35) were recruited to participate in the study.

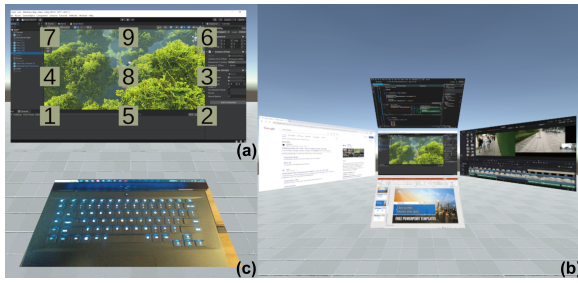


Figure 4: Images showing (a) an example of randomly placed buttons on a screen, (b) an overview of the virtual screens implemented in the **Working Simulation Scenario**, and (c) the stream view of the keyboard in the virtual environment.

All participants had normal or corrected-to-normal vision and little to no experience with hands-free interactions in VR. They were informed that they were allowed to quit the study at any time. Furthermore, they were asked about any medical history related to their eyes, neck, or cranial nerves. The experiment was conducted under the approval of a university Institutional Review Board (IRB).

4.3 Experimental Design

We used a repeated measures within-subjects design in the experiment to test GlanXR against three other conditions in two different scenarios. For all four conditions in both scenarios, the screens were centered to the participant's natural head position and positioned one meter ahead of the head in the virtual environment during initialization. Additionally, the screens were sized as 1.6 meters wide and 0.9 meters high in a 16:9 aspect ratio, ensuring a clear view of the contents without being too large in the field of view. Participants were allowed to adjust their position freely to achieve a comfortable view. The order of conditions was assigned based on a balanced Latin square to minimize order effects, while the order of scenarios was alternated.

In the **baseline condition**, we simulated a taskbar with icons representing screens. The size and locations of these icons were aligned with the actual taskbar in the Windows operating system, and it was placed at the bottom of the screens. Participants had to click on the icons using a controller or mouse to switch between screens. In the **expansive multi-screen condition**, all screens were simultaneously displayed. One screen was positioned at the center, with the others surrounding it in the up, down, left, and right directions. The screens were rotated to ensure that their center points were equally one meter from the initial head position. In the **gaze-selecting condition**, a gaze-based interface was implemented, enabling hands-free screen switching by directing the gaze toward icons positioned above the screen. Similar icons to those used in the baseline condition were employed to represent the screens. To help improve the ease of selection, the icons were intentionally designed to be relatively large and spaced apart. Specifically, the size of these icons was set to 0.3 by 0.3 meters, and the distance between the icons and from the icons to the screen was set to 0.1 meters.

Two virtual environments were built to evaluate the conditions in both fast-switching and working simulation scenarios. The fast-switching scenario focused on assessing performances during the screen-switching procedure. The conditions require participants to engage in continuous fast switching, irrespective of the contents displayed on the screens. Meanwhile, the working simulation scenario aimed to assess participant performances in real-world usage of multi-screen virtual workspaces, involving relatively complex tasks such as reading comprehension and text input.

4.3.1 Fast-Switching Scenario

In the virtual environment for the fast-switching scenario, we logically picked five appropriate applications based on their common

usage in daily work settings and used their screenshots as backgrounds for the virtual screens to help distinguish between them. Then, interactive interfaces with buttons were attached to each screen. The buttons on these interfaces were numbered 1-9, and their layout was designed to cover the entire screen (Figure 4-a). In each trial, a panel displays randomly selected target buttons and screens at the beginning. The panel reappears to show the next target when clicking the correct button. Furthermore, the buttons were programmed to rearrange randomly after the correct button was clicked. Therefore, participants always had to navigate their gaze across the entire screen to find the target button. After correctly clicking 50 target buttons, the trial automatically ends.

Participants were asked to switch to the target screen and click on the target button as fast as possible using a controller. Moreover, the controller was also employed to click the taskbar icons to trigger screen switching in the baseline condition. This prevents potential bias caused by switching between the mouse and controller, ensuring a consistent focus on rapid screen switching.

4.3.2 Working Simulation Scenario

In the virtual environment for the working simulation scenario, five virtual screens were implemented, each showing a questionnaire page, a Wikipedia page, a weather application, a spreadsheet, and an email application (Figure 4-b). In addition, an interface that streams the view of the keyboard by a camera was applied and calibrated to match the actual position of the keyboard (Figure 4-c).

Concerning input devices, we utilized a standard mouse and keyboard configuration, aligning with typical real-world computing environments during work. Also, in the baseline condition, the participants had to use the mouse to click on the taskbar icons to switch between screens. As the primary task, participants were required to answer questions on the questionnaire as fast as possible by gathering information from the other screens. Furthermore, the questionnaire consists of seven entries: four single-choice entries, two numeric input entries, and a text input entry. This design encourages common computer interactions, such as cursor navigation, clicking, typing, and content retrieval. Considering that participants might remember the content, we administered four questionnaires, one after each trial, with the order counterbalanced.

4.4 Measurements

During the experiment, we gathered the data necessary to calculate a comprehensive set of dependent measures as follows:

Completion time: The duration from the start of each trial until task completion, where task completion criteria varied between scenarios: 50 correct button clicks in the fast-switching scenario and completion of the questionnaire in the working simulation scenario. This measure was employed to assess the time users spent performing tasks in both scenarios using different methods.

Switching frequency: The number of switches participants make to complete the task. This provides insights into how precisely participants could locate and switch to a target screen.

Average angular speed of head rotations in $^{\circ}/s$: The averaged differences in head orientations between frames, measured in degrees per second. This measure was applied to determine how much a participant's head averagely rotated in each trial.

Average angle difference between head orientation and face-forward (i.e., head deviation from face-forward) in degrees: We first gathered the head orientation data from the HMD and then calculated the average angle between the orientation data and a forward vector (0,0,1). This measure represents whether participants tended to maintain their heads in a natural position facing forward.

Average eye-head position in degrees: The averaged eye-head position data between frames, which reveals the average amplitude participants' eyes moved from the center.

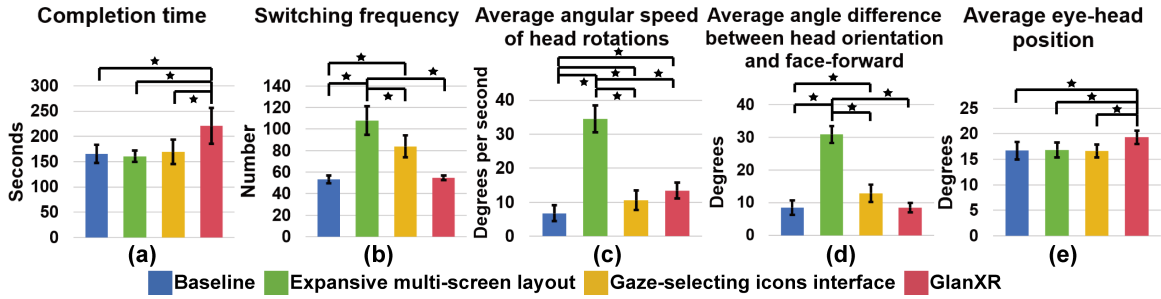


Figure 5: Analysis of the data collected from the **Fast-Switching Scenario**, including (a) completion time and (b) switching frequency, (c) average angular speed of head rotations, (d) average angle difference between head orientation and face-forward, and (e) average eye-head position.

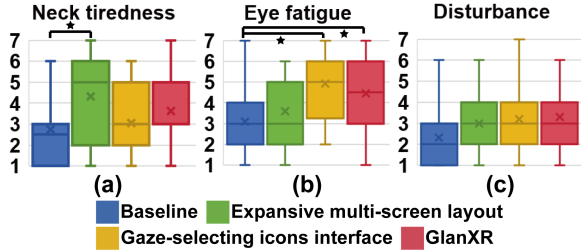


Figure 6: Analysis of the subjective ratings for (a) neck tiredness, (b) eye fatigue, and (c) disturbance, in the **Fast-Switching Scenario**.

Post-trial subjective ratings: Participants were asked to rate the following subjective questions with seven-scale ratings from 1 (“Strongly disagree”) to 7 (“Strongly agree”), including 1) My neck feels tired. 2) I feel eye fatigue. 3) I felt distracted by this method.

The quantitative measures listed above were recorded with the experimental program and the HMD’s built-in sensors. To estimate the switching frequency in the expansive multi-screen condition, we define a switch as occurring when the participant’s head faces a new screen for more than two seconds.

4.5 Procedure

As each participant joined our experiment, we briefly introduced the experiment and informed them about the tasks and requirements. Then, the participants adjust their chair position and put on the HMD with assistance from an experimenter. The experiment included four trials in two scenarios (eight trials in total), each representing a single condition. Before the first trial began, participants were required to complete eye calibration and a practice trial for the proposed approach. The practice trial had a three-minute limit, but participants could terminate it once they felt sufficiently familiar with the approach. After each trial, participants were asked to fill out a questionnaire and were allowed to take a two-minute break. Additionally, before the first trial of the working simulation scenario started, we required participants to check if the streaming view of the keyboard was appropriately placed in the environment. The experiment took around 60 minutes to complete on average.

5 RESULTS

In this section, we describe the results of our experiment concerning the two scenarios and the four conditions. We first applied Shapiro-Wilk tests and QQ plots to the collected data, revealing non-normal distributions. This suggests that the effects of participants’ abilities and inherent limitations may played a significant role. In particular, participants were likely to show varying capabilities in tasks such as inputting and reading speed, familiarity with different methods, and hand-head-eye coordination. Therefore, non-parametric tests were used for statistical analysis. Additionally, Friedman tests and pairwise Wilcoxon post-hoc tests with Bonferroni correction were conducted, along with the computation of Kendall’s W and r values

to indicate effect sizes. In the figures, the heights of the bars are the means, with error bars indicating standard deviation. Significance was determined at a level of 0.05, denoted by ‘*’.

5.1 Fast-Switching Scenario

5.1.1 Completion Time and Switch Frequency

For the fast-switching scenario, regarding completion time, Friedman test revealed significance between the four conditions ($\chi^2 = 26.67$, $p < 0.001$, $W = 0.65$) (Figure 5-a). Further differences were found in GlanXR against the baseline ($p < 0.001$, $r = 0.78$), the expansive multi-screen layout ($p < 0.001$, $r = 0.87$), and the gaze-selecting icons interface ($p < 0.01$, $r = 0.79$). These findings show that participants performed the task more slowly when using GlanXR than the other methods. Regarding switch frequency, significance was also laid between the conditions ($\chi^2 = 57.96$, $p < 0.001$, $W = 0.35$) (Figure 5-b). Post-hoc tests indicated significant effects in the expansive multi-screen layout against the baseline ($p < 0.001$, $r = 0.87$), the gaze-selecting icons interface ($p < 0.05$, $r = 0.75$), and GlanXR ($p < 0.001$, $r = 0.87$). Additionally, significance was also observed in the gaze-selecting icons interface against the baseline ($p < 0.001$, $r = 0.88$) and GlanXR ($p < 0.001$, $r = 0.84$). These results revealed that the participants could locate and switch to the target screen more precisely with both the baseline method and GlanXR.

5.1.2 Head and Eye Behaviors

To determine how much the hands-free screen switching methods affected participants’ head rotation and orientation in the fast-switching scenario, Friedman tests were applied to reveal that there were significant differences in both of the head’s average angular speed data ($\chi^2 = 48.74$, $p < 0.001$, $W = 0.40$) (Figure 5-c) and average angle difference between head orientation and face-forward ($\chi^2 = 49.63$, $p < 0.001$, $W = 0.34$) (Figure 5-d). Further pairwise Wilcoxon tests showed that for the head’s average angular speed, significant effects lay in the baseline against the other methods, including the expansive multi-screen layout ($p < 0.001$, $r = 0.87$), the gaze-selecting icons interface ($p < 0.05$, $r = 0.57$), and GlanXR ($p < 0.001$, $r = 0.74$). Additionally, significant effects were found in the expansive multi-screen layout vs. the gaze-selecting icons interface ($p < 0.001$, $r = 0.88$) and GlanXR ($p < 0.001$, $r = 0.87$). For the average angle difference between head orientation and face-forward, significance was found in the expansive multi-screen layout vs. the baseline ($p < 0.001$, $r = 0.87$), the gaze-selecting icons interface ($p < 0.001$, $r = 0.88$), and GlanXR ($p < 0.001$, $r = 0.88$). Additionally, significant effects were also observed in the gaze-selecting icons interface compared to the baseline ($p < 0.01$, $r = 0.56$) and GlanXR ($p < 0.01$, $r = 0.75$).

These findings suggest that hands-free methods required additional head rotations compared to the baseline, with the expansive multi-screen layout requiring the most. Moreover, on average, participants maintained a more natural forward-facing head position

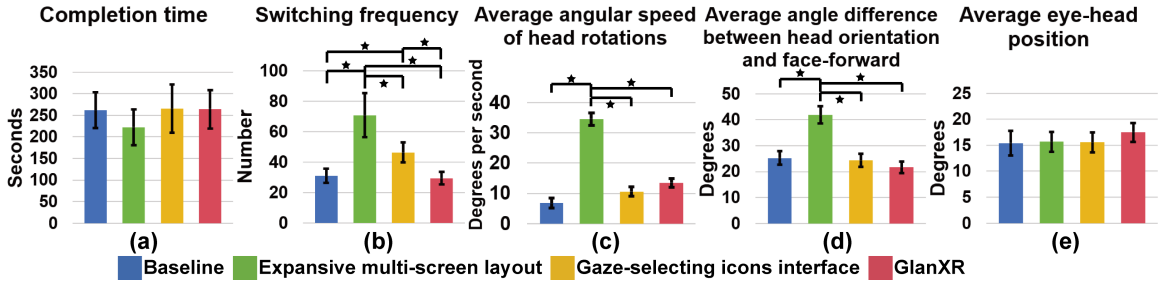


Figure 7: Analysis of the data collected from the **Working Simulation Scenario**: (a) completion time and (b) switching frequency, (c) average angular speed of head rotations, (d) average angle difference between head orientation and face-forward, and (e) average eye-head position.

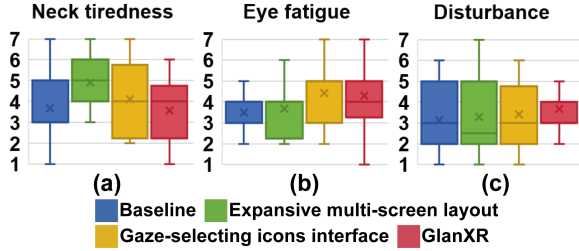


Figure 8: Subjective ratings analysis for (a) neck tiredness, (b) eye fatigue, and (c) disturbance, in the **Working Simulation Scenario**.

when using GlanXR, among the hands-free methods. Besides head behaviors, the analysis of average eye-head position data also revealed significance ($\chi^2 = 16.45$, $p < 0.001$, $W = 0.61$) (Figure 5-e). Further differences were found in GlanXR against the baseline ($p < 0.05$, $r = 0.62$), the expansive multi-screen layout ($p < 0.01$, $r = 0.65$), and the gaze-selecting icons interface ($p < 0.01$, $r = 0.82$). This indicates that, on average, participants moved their eyes in a larger amplitude with GlanXR.

5.1.3 Subjective Preference

Through the analysis of the data collected from post-trial questionnaires, significant effects were identified among the four conditions concerning “neck tiredness” ($\chi^2 = 13.89$, $p < 0.01$, $W = 0.58$) (Figure 6-a) and “eye fatigue” ($\chi^2 = 22.07$, $p < 0.001$, $W = 0.59$) (Figure 6-b). No significance was observed in relation to “disturbance” ($\chi^2 = 7.32$, $p = 0.06$) (Figure 6-c). For “neck tiredness”, significant differences were found in the expansive multi-screen layout vs. the baseline ($p < 0.05$, $r = 0.69$), indicating that participants felt more stress in their necks when using the expansive multi-screen layout than the taskbar with icons. For “eye fatigue”, significant differences were found in the baseline vs. the gaze-selecting icons interface ($p < 0.01$, $r = 0.84$) and the baseline vs. GlanXR ($p < 0.05$, $r = 0.60$), showing that both methods using gaze input induced additional eye fatigue compared to the baseline.

5.2 Working Simulation Scenario

5.2.1 Completion Time and Switch Frequency

Like the fast-switching scenario, we applied a Friedman test to the collected data. However, no significance was found between the four conditions ($\chi^2 = 2.35$, $p = 0.50$) (Figure 7-a), revealing that there was no significant difference in task completion speed when using the methods in the working simulation scenario. Regarding switch frequency, significance was observed between the conditions ($\chi^2 = 43.76$, $p < 0.001$, $W = 0.37$) (Figure 7-b). Post-hoc tests indicated significant effects in the expansive multi-screen layout against the baseline ($p < 0.001$, $r = 0.81$), the gaze-selecting icons interface ($p < 0.05$, $r = 0.67$), and GlanXR ($p < 0.001$, $r = 0.85$). In addition, significance was also found in the gaze-selecting icons interface against the baseline ($p < 0.001$, $r = 0.79$) and GlanXR ($p < 0.001$,

$r = 0.81$). These results suggested that the participants could also locate and switch to the target screen more precisely with both the baseline method and GlanXR, as compared to the other two methods in this scenario.

5.2.2 Head and Eye Behaviors

Regarding head behaviors including rotation and orientation, Friedman tests revealed that there were significant differences in both of the head’s average angular speed data ($\chi^2 = 50.55$, $p < 0.001$, $W = 0.65$) (Figure 7-c) and average angle difference between head orientation and face-forward ($\chi^2 = 53.60$, $p < 0.001$, $W = 0.75$) (Figure 7-d). Additional pairwise Wilcoxon tests showed that for the head’s average angular speed, significant effects were located in the expansive multi-screen layout against the other methods, including the baseline ($p < 0.001$, $r = 0.88$), the gaze-selecting icons interface ($p < 0.001$, $r = 0.88$), and GlanXR ($p < 0.001$, $r = 0.88$). For the average angle difference between head orientation and face-forward, significance was found in the expansive multi-screen layout against the baseline ($p < 0.001$, $r = 0.88$), the gaze-selecting icons interface ($p < 0.001$, $r = 0.88$), and GlanXR ($p < 0.001$, $r = 0.88$). These findings indicate that when working in the virtual workspace with the multi-screen method, participants tended to rotate their heads more and maintain them in a position further from the forward, compared to the other methods. Furthermore, the analysis of average eye-head position data also revealed significance ($\chi^2 = 19.65$, $p < 0.001$, $W = 0.87$) (Figure 7-e). However, no pairwise significant effects emerged after the post-hoc Wilcoxon test with Bonferroni correction.

5.2.3 Subjective Preference

By analyzing the post-trial questionnaires data from this scenario, significance was found among the four conditions concerning “neck tiredness” ($\chi^2 = 13.36$, $p < 0.01$, $W = 0.55$) (Figure 8-a) and “eye fatigue” ($\chi^2 = 12.08$, $p < 0.01$, $W = 0.59$) (Figure 8-b). No significance was observed in relation to “disturbance” ($\chi^2 = 4.33$, $p = 0.23$) (Figure 8-c). For “neck tiredness”, significant differences were found in the expansive multi-screen layout vs. the baseline ($p < 0.05$, $r = 0.64$) and the expansive multi-screen layout vs. GlanXR ($p < 0.05$, $r = 0.66$), indicating that participants felt more stress in their necks when using the expansive multi-screen layout compared to the taskbar with icons and the proposed approach. Moreover, no further significant effects between the methods were revealed after applying the post-hoc test for “eye fatigue”.

5.3 False Triggering Rate

In line with the work by Yi et al. [37], participants were required to report errors (i.e., false triggering of switches) in the trials where GlanXR was applied in both scenarios. A total of 2025 screen switches were triggered by the 24 participants using GlanXR in both scenarios, with 24 occurrences (mean 0.5, SD 1.19) being false triggers. These data were integrated to compute a false triggering rate of 1.19%.

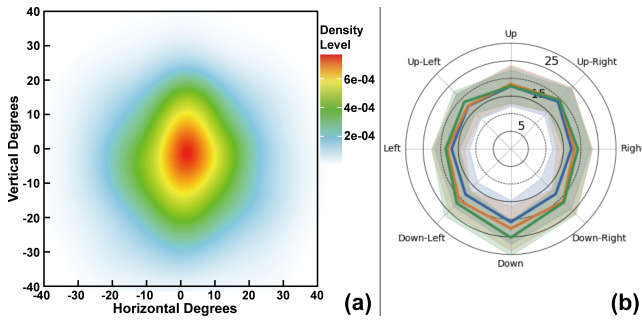


Figure 9: (a) Our density heat map illustrating the eye-head positions in degrees of all participants in each frame during the fast-switching scenario using the expansive multi-screen layout, and (b) The map by Sidemark and Gellersen [30].

5.4 Range of Eye-Head Position

To investigate the natural distribution range of eye-head positions, we extracted data on eye-head positions in angles per frame from the fast-switching scenario using the expansive multi-screen layout, where participants were most likely to look equally in all directions. We generated a density heat map based on the data, as shown in Figure 9. The range involves approximately 25° on the left, right, and top sides, and 35° on the bottom side.

6 DISCUSSION

According to the results of task completion time metrics in the fast-switching scenario, the participants performed slower using GlanXR than the other methods. We believe that this result was influenced by the acceptance and understanding levels of the participants of GlanXR. Although we provided a three-minute practice trial, it did not seem efficient enough for most participants to master the method completely. Therefore, in situations that require rapid and frequent switching, the robust design of GlanXR is likely to slow them down. On the other hand, for the working simulation scenario, there was no significant difference among the conditions, indicating that for daily use situations, in which screen switching is less frequent and has a longer interval between the switches, using GlanXR is as fast as other methods. Furthermore, the expansive multi-screen layout and the gaze-selecting icons interface were observed to have significantly higher levels than the taskbar (baseline) and GlanXR. This suggests that participants could not instantly locate the target screen and required more steps to navigate to it using the two methods.

Regarding head behaviors, the expansive multi-screen layout required participants to rotate their heads more and maintain their head orientation further from forward in both scenarios. It could potentially cause additional stress on the neck, supported by participants' ratings of "neck tiredness", which were higher than the baseline. Preventing false triggering is one of our primary goals in designing GlanXR, and the observed result of 1.19% is highly satisfactory. This is especially notable considering the large number of switches performed in the experiment, and it remained relatively low compared to other hands-free selecting approaches (12% [42], 12% [8], 10% [16], and 3.6% [37]). The ratings for "disturbance" have also partially confirmed the robustness of GlanXR from the participants' perspective, as it did not show many errors that would make them feel disturbed. Concerning "disturbance", the subjective ratings indicated that our adaptive threshold range algorithm, along with peripheral interface design, could effectively minimize potential disruptions for users while viewing screen content.

The range of eye-head position plays a crucial role in coordinating the movement of the eyes and head, and must be carefully considered in developing interfaces based on eye-gaze and head movements. From our experimental data, illustrated in the density heat

map of eye-head positions (Figure 9-a), the average range was observed to be nearly the same amplitude in the left, right, and top directions but extended further in the bottom direction. This finding aligns with the demonstration in the study by Sidemark and Gellersen [30] (Figure 9-b), further justifying our vertically asymmetrical design of the adaptive eye-head position threshold range, which employs two halves of ellipses.

Concerning our fundamental design considerations for GlanXR, despite the results indicating that it did not fulfill all of them, it still exhibited reasonable performance. Firstly, GlanXR did not significantly improve participants' performance concerning faster screen switching and task completion than the other conditions. However, it showed comparable speed to the other conditions in the working simulation scenario, indicating that it would not slow users if applied to virtual workspace applications in practical situations. Secondly, GlanXR and the gaze-selecting icons method strongly outperform the state-of-the-art expansive multi-screens method in reducing the amplitude of head rotations while maintaining a relatively natural head position. In addition, both have the potential to minimize users' neck stress at a level comparable to hands-on methods while offering hands-free interactions. Thirdly, GlanXR was not found to significantly disturb participants, even though it was the only method that involved an interface fixed in the view among the tested methods. This result suggests that our design, which focuses on mitigating visual distractions, has performed as expected. Generally speaking, our findings suggest that GlanXR has several advantages that could make it a preferable alternative or reduce fatigue for certain applications. Additionally, a passive advantage is that peripheral interfaces like this alleviate the virtual clutter problem due to the reduced number of windows.

7 CONCLUSION AND FUTURE WORK

In this paper, we proposed a novel hands-free screen switching approach - GlanXR, designed to assist users switch between virtual screens in a easy and robust manner. To achieve this, we utilized rapid and less energy-intensive eye-gaze as a selecting mechanism, coupled with less jittery and more controlled head movement to confirm the selections. Additionally, we integrated a real-time adaptive eye-head position threshold range and a periphery fixed interface featuring interactive thumbnails that deliver previews of the screens. We conducted a within-subjects user study involving two scenarios with 24 participants to evaluate the proposed approach compared to traditional hands-on, head-based, and gaze-based methods. Results from this study revealed that GlanXR is robust, ensures precise switches, and is unlikely to disturb natural gaze behaviors while viewing screen contents. Additionally, both a gaze-based method and our approach effectively mitigated neck stress compared to a head-based method without a significant impact on inducing eye fatigue in participants. One future direction might be to add an adaptive threshold related to the head rotation speed for the eye fixation time window instead of the fixed 150ms. The faster the head rotates, the less eye fixation time needed.

Although this work primarily focuses on the research background and limitations of current screen switching methods, it is evident that our approach can be adapted to other configurations and serve as a hands-free interactive protocol. For instance, the thumbnails can be easily replaced with buttons for operating commands. We hope that this work will inspire the development of other hands-free interfaces that leverage eye-gaze, head movements, and their interaction to assist fast and robust interactions in VR and AR workspaces. Additionally, we encourage the exploration of techniques that improve comfort during prolonged VR usage.

ACKNOWLEDGMENTS

This research was funded in part by JSPS grant 21H03482 and NSF grant 2223035.

REFERENCES

- [1] Virtual desktop. <https://www.vrdesktop.net/>, 2016.
- [2] Immersed. <http://immersed.com>, 2017.
- [3] R. Bates and H. O. Istance. Why are eye mice unpopular? a detailed comparison of head and eye controlled assistive technology pointing devices. *Universal Access in the Information Society*, 2:280–290, 2003.
- [4] E. Bizzi, R. E. Kalil, and P. Morasso. Two modes of active eye-head coordination in monkeys. *Brain Research*, 1972.
- [5] E. Bizzi, R. E. Kalil, and V. Tagliasco. Eye-head coordination in monkeys: evidence for centrally patterned organization. *Science*, 173(3995):452–454, 1971.
- [6] J. Blattgerste, P. Renner, and T. Pfeiffer. Advantages of eye-gaze over head-gaze-based selection in virtual and augmented reality under varying field of views. In *Proceedings of the Workshop on Communication by Gaze Interaction*, pp. 1–9, 2018.
- [7] D. B. Chaffin. Localized muscle fatigue—definition and measurement. *Journal of Occupational and Environmental Medicine*, 15(4):346–354, 1973.
- [8] A. Esteves, E. Velloso, A. Bulling, and H. Gellersen. Orbits: Gaze interaction for smart watches using smooth pursuit eye movements. In *Proceedings of the 28th annual ACM symposium on user interface software & technology*, pp. 457–466, 2015.
- [9] A. Esteves, D. Verweij, L. Suraiya, R. Islam, Y. Lee, and I. Oakley. Smoothmoves: Smooth pursuits head movements for augmented reality. In *Proceedings of the 30th annual acm symposium on user interface software and technology*, pp. 167–178, 2017.
- [10] Y. Fang, R. Nakashima, K. Matsumiya, I. Kuriki, and S. Shioiri. Eye-head coordination for visual cognitive processing. *PloS one*, 10(3):e0121035, 2015.
- [11] A. M. Feit, S. Williams, A. Toledo, A. Paradiso, H. Kulkarni, S. Kane, and M. R. Morris. Toward everyday gaze input: Accuracy and precision of eye tracking and implications for design. In *Proceedings of the 2017 Chi conference on human factors in computing systems*, pp. 1118–1130, 2017.
- [12] E. G. Freedman. Coordination of the eyes and head during visual orienting. *Experimental brain research*, 190:369–387, 2008.
- [13] J. H. Fuller. Head movement propensity. *Experimental Brain Research*, 92:152–164, 1992.
- [14] J. P. Hansen, A. S. Johansen, D. W. Hansen, K. Ito, and S. Mashino. Command without a click: Dwell time typing by mouse and gaze selections. In *Interact*, vol. 3, pp. 121–128, 2003.
- [15] B. Hu, I. Johnson-Bey, M. Sharma, and E. Niebur. Head movements during visual exploration of natural images in virtual reality. In *2017 51st annual conference on information sciences and systems (ciss)*, pp. 1–6. IEEE, 2017.
- [16] A. Huckauf and M. H. Urbina. Object selection in gaze controlled systems: What you don’t look at is what you get. *ACM Transactions on Applied Perception (TAP)*, 8(2):1–14, 2011.
- [17] R. J. Jacob. What you look at is what you get: eye movement-based interaction techniques. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 11–18, 1990.
- [18] S. Jalaliniya, D. Mardanbegi, and T. Pederson. Magic pointing for eyewear computers. In *Proceedings of the 2015 ACM International Symposium on Wearable Computers*, pp. 155–158, 2015.
- [19] D. Kern, A. Mahr, S. Castronovo, A. Schmidt, and C. Müller. Making use of drivers’ glances onto the screen for explicit gaze-based interaction. In *Proceedings of the 2nd international conference on automotive user interfaces and interactive vehicular applications*, pp. 110–116, 2010.
- [20] J. F. Knight and C. Baber. Neck muscle activity and perceived pain and discomfort due to variations of head load and posture. *Aviation, space, and environmental medicine*, 75(2):123–131, 2004.
- [21] J. F. Knight and C. Baber. Effect of head-mounted displays on posture. *Human factors*, 49(5):797–807, 2007.
- [22] M. Kytö, B. Ens, T. Piumsomboon, G. A. Lee, and M. Billingham. Pinpointing: Precise head-and eye-based target selection for augmented reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2018.
- [23] W. J. Lee, J. H. Kim, Y. U. Shin, S. Hwang, and H. W. Lim. Differences in eye movement range based on age and gaze direction. *Eye*, 33(7):1145–1151, 2019.
- [24] F. Lu, S. Davari, L. Lisle, Y. Li, and D. A. Bowman. Glanceable ar: Evaluating information access methods for head-worn augmented reality. In *2020 IEEE conference on virtual reality and 3D user interfaces (VR)*, pp. 930–939. IEEE, 2020.
- [25] P. Majaranta, U.-K. Ahola, and O. Špakov. Fast gaze typing with an adjustable dwell time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 357–360, 2009.
- [26] B. S. Oommen, R. M. Smith, and J. S. Stahl. The influence of future gaze orientation upon eye-head coupling during saccades. *Experimental Brain Research*, 155:9–18, 2004.
- [27] J. Orlosky, C. Liu, K. Sakamoto, L. Sidenmark, and A. Mansour. Eye-shadows: Peripheral virtual copies for rapid gaze selection and interaction. To appear in the *Proceedings of the IEEE conference on Virtual Reality*.
- [28] A. M. Penkar, C. Lutteroth, and G. Weber. Designing for the eye: design parameters for dwell in gaze interaction. In *Proc. of the 24th Australian Computer-Human Interaction Conference*, pp. 479–488, 2012.
- [29] L. Sidenmark, C. Clarke, X. Zhang, J. Phu, and H. Gellersen. Outline pursuits: Gaze-assisted selection of occluded objects in virtual reality. In *Proceedings of the 2020 chi conference on human factors in computing systems*, pp. 1–13, 2020.
- [30] L. Sidenmark and H. Gellersen. Eye, head and torso coordination during gaze shifts in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 27(1):1–40, 2019.
- [31] L. Sidenmark, M. Parent, C.-H. Wu, J. Chan, M. Glueck, D. Wigdor, T. Grossman, and M. Giordano. Weighted pointer: Error-aware gaze-based interaction through fallback modalities. *IEEE Transactions on Visualization and Computer Graphics*, 28(11):3585–3595, 2022.
- [32] L. Sidenmark, D. Potts, B. Bapich, and H. Gellersen. Radi-eye: Hands-free radial interfaces for 3d interaction using gaze-activated head-crossing. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–11, 2021.
- [33] C. J. Snijders, G. H. Van Dijke, and E. Roosch. A biomechanical model for the analysis of the cervical spine in static postures. *Journal of biomechanics*, 24(9):783–792, 1991.
- [34] Z. C. Thumser, B. S. Oommen, I. S. Kofman, and J. S. Stahl. Idiosyncratic variations in eye-head coupling observed in the laboratory also manifest during spontaneous behavior in a natural setting. *Experimental Brain Research*, 191:419–434, 2008.
- [35] Z. C. Thumser and J. S. Stahl. Eye-head coupling tendencies in stationary and moving subjects. *Experimental brain research*, 195:393–401, 2009.
- [36] M. Vidal, K. Pfeuffer, A. Bulling, and H. W. Gellersen. Pursuits: eye-based interaction with moving targets. In *CHI’13 Extended Abstracts on Human Factors in Computing Systems*, pp. 3147–3150, 2013.
- [37] X. Yi, Y. Lu, Z. Cai, Z. Wu, Y. Wang, and Y. Shi. Gazedock: Gaze-only menu selection in virtual reality using auto-triggering peripheral menu. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 832–842. IEEE, 2022.
- [38] C. Yu, Y. Gu, Z. Yang, X. Yi, H. Luo, and Y. Shi. Tap, dwell or gesture? exploring head-based text entry techniques for hmds. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 4479–4488, 2017.
- [39] D. Yu, H.-N. Liang, X. Lu, T. Zhang, and W. Xu. Depthmove: Leveraging head motions in the depth dimension to interact with virtual reality head-worn displays. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 103–114. IEEE, 2019.
- [40] W. H. Zangemeister and L. Stark. Types of gaze movement: variable interactions of eye and head movements. *Experimental neurology*, 77(3):563–577, 1982.
- [41] G. Zhao, J. Orlosky, J. Gabbard, and K. Kiyokawa. Hazardsnap: Gazed-based augmentation delivery for safe information access while cycling. *IEEE Transactions on Visualization and Computer Graphics*, 2023.
- [42] Z. Zhe, F. W. Siebert, A. C. Venjakob, and M. Roetting. Calibration-free gaze interfaces based on linear smooth pursuit. *Journal of Eye Movement Research*, 13(1), 2020.