2026 IEEE/SICE International Symposium on System Integration (SII)
January 11-14, 2026. Cancún, México

Using Eye Gaze and Body Location to Organize and Retrieve Information in Mixed Reality

Yuu Takahashi[1] and Takeshi Sasaki[1]

Abstract— **This study proposes a new method for organizing and accessing information in MR environments by anchoring data to parts of the human body (such as the head or hands). The system utilizes Microsoft HoloLens 2 and eye-tracking technology to enable users to intuitively access virtual information through gaze-based operations. This approach introduces a relative spatial coordinate system based on the user's body. User studies were conducted to verify the effectiveness of body-based tagging and information management in MR spaces. The results confirmed that this information access method is usable without issues, and it was revealed that rules exist governing the relationship between the access time and location of body-tagged information. These results suggest that body-based information tagging has the potential to enhance the usability of MR for everyday tasks.**

I. Introduction

As information becomes increasingly digitized, it has become common to organize and store large amounts of data in digital form, both in professional settings and in daily life. In particular, individuals who engage in manual tasks—such as engineers, healthcare professionals, and learners—often face challenges in accessing needed information without interrupting their work. In addition, quickly accessing specific information from a vast set of data remains a challenge [1]. Common solutions include pinning files to the desktop or organizing them into folders, but these methods can become inefficient as the number of items grows [2]. Even with keyword search capabilities, users still face cognitive and operational burdens in tagging, recalling, and efficiently retrieving data [3].

While desktop environments are designed around familiar 2D interfaces such as windows, icons, and folders, these organization strategies do not directly translate to more immersive computing environments. In particular, Mixed Reality (MR)—a technology that blends the physical and digital worlds—fundamentally changes how users interact with information. Unlike traditional desktops, MR lacks conventional visual metaphors and fixed screens, instead presenting data as spatial objects embedded in the user's environment [4] [5]. As MR devices such as the Microsoft HoloLens 2 and Apple Vision Pro become more widespread, it is expected that more digital information will be accessed and managed within MR spaces [4].

However, MR environments introduce new challenges for organizing and retrieving information effectively. The absence of standard desktop UI elements requires novel approaches to help users locate and interact with virtual content intuitively. Simply applying desktop-style organization (e.g., folders or search boxes) may not align with the embodied, spatial nature of MR interaction. Therefore, new methods are needed to support efficient information access in MR. To address these issues, we propose a system that allows users to spatially tag and organize information by associating it with parts of their body (such as their head or hands) through an MR interface. The system is implemented using the Microsoft HoloLens 2, a head-mounted display (HMD) equipped with transparent lenses that overlay digital content onto the physical world. By leveraging the device's built-in eye tracking capabilities, users can access virtual information simply by looking.

For example, placing a virtual object displaying weather forecasts diagonally above and in front of the user's head allows for quick access to the information—similar to how one might glance up at the sky to check the weather. Likewise, by placing virtual objects that show procedural instructions or reminders on the user's hand, the user can view them simply by glancing at their wrist or palm during a task, much like checking a wristwatch. We refer to the locations where virtual objects are placed and linked with information as "access points".

Associating information with specific parts of the body can enhance recall and intuitiveness, as the body itself provides a familiar and structured spatial reference. For example, since wristwatches are commonly worn on the wrist to check the time, anchoring time-related information near the wrist can evoke a strong cognitive association, making it easier for users to remember and access that information [6]. In this way, body-anchored access points leverage existing habits and embodied experiences to facilitate more natural and efficient interactions in spatial computing environments.

To achieve these goals, it is necessary to better support users' natural abilities and resources through actions in MR space [7].Therefore, as gaze-based access to information has been shown to enhance usability by enabling intuitive and efficient interaction [8], we propose gaze-based access to body-anchored information instead of relying on hand-based interaction. Hand-based operations can be impractical when information is attached to one's own body, such as the back of a hand or near the head, where reaching or touching may be difficult or physically awkward. Moreover, when multiple access points are densely distributed around the user's body in the limited surrounding space, using hands could lead to obstruction or unintended interactions. Gaze-based access,

[1]Yuu Takahashi and Takeshi Sasaki are with Dept. of Electrical Engineering and Computer Science, Shibaura Institute of Technology, 3-7-5 Toyosu, Koto-ku, Tokyo, Japan cy20236@shibaura-it.ac.jp, sasaki-t@ieee.org

on the other hand, enables users to quickly and selectively interact with virtual content without requiring large body movements or hand repositioning [9] [10] [11].

Although gaze interaction may face challenges such as depth ambiguity when multiple objects are aligned, it offers significant advantages for efficient and unobtrusive information retrieval in body-centered MR environments [12] [8]. This study focuses on verifying how information processing occurs in MR spaces in everyday life. Specifically, we will examine which locations are optimal for accessing information in terms of body-based tagging and information management in MR spaces, and what types of information are likely to be placed at tagged locations, based on information that is frequently checked in everyday life.

II. Related Work

A. Spatial Memory Based on Embodied Interaction

Previous research has demonstrated the cognitive benefits of placing digital information in specific spatial locations within a room. One such study introduced the concept of Spatial Knowledge Tags (SKTs), where physical objects such as home appliances serve as spatial anchors or "tags" for storing knowledge and information [13]. These objects act as memory triggers, allowing users to associate specific data with particular locations in the environment. The use of direct interaction—such as reaching out toward a tag—was shown to improve accessibility and recall by leveraging the user's embodied experience.

B. Hierarchical Spatial Memory Systems

Shinoki et al. [14] conducted a study in which a head-mounted display was used in conjunction with a wearable camera to visualize spatial memory, allowing users to switch between visual displays and physical actions via hand gestures recognized by the LeapMotion controller. In their system, the center of the measurement area was defined as the origin using LeapMotion, and various household items and devices were tagged within the environment. Both global and local spatial memories were displayed and confirmed using an AR interface. The study demonstrated that SKTs are more effectively represented as areas rather than points, and proposed a hierarchical structure where global SKTs serve as access points to retrieve and present local SKTs. This hierarchical tagging system was presented as a novel approach to supporting the formation and utilization of spatial memory in augmented reality environments.

The study demonstrated that Spatial Knowledge SKTs are more effectively represented as areas rather than points, and proposed a hierarchical structure where global SKTs serve as access points to retrieve and present local SKTs. This hierarchical tagging system was presented as a novel approach to supporting the formation and utilization of spatial memory in augmented reality environments.

C. Eye-body Coordination during Daily Activities for Gaze Prediction from Full-body Poses

Hu et al. [15] propose a model called Pose2Gaze, which enhances gaze prediction by incorporating both eye and body movements. While previous approaches primarily focused on eye or head motion, this study emphasizes the role of full-body posture in influencing gaze direction during everyday activities. The Pose2Gaze model leverages deep learning trained on datasets containing synchronized eye movements and body postures collected during actions such as sitting, walking, and object interaction, enabling gaze estimation from full-body pose data. The results demonstrate that this model outperforms traditional gaze prediction techniques, underlining the significance of bodily context in gaze estimation.

This work is particularly relevant to applications in virtual and augmented reality (VR/AR), where accurate understanding of user attention is essential. It offers promising directions for developing more intuitive and context-aware systems in human-computer interaction, as well as in assistive technologies.

D. Our Approach

Based on these previous studies, the approach of this study differs from conventional methods in that it uses the user's own body, specifically the head and hands, as the main reference points (tags) for spatial memory to link information for access. Conventional methods relied on static objects within a room, but this system achieves a relative spatial coordinate system by tracking the user's movements.

Furthermore, to enable relative positioning tagging based on the body, this study leverages MR and real-time body tracking to propose a method for optimizing information organization and access within the MR space.

III. System Design: Information Organization and Access

This chapter explains the system for organizing information linked to the body in MR space and the methods for accessing that information. To create this system, we use MRTK (Mixed Reality Toolkit), the HoloLens development SDK released by Microsoft for developing HoloLens 2 using Unity .

A. Information Organization Using Body-anchored Tags

1) Acquiring Coordinates of Body Parts and Access Points: To bind virtual information to specific body parts in MR space, it is necessary to obtain the spatial coordinates of both the user's body and the displayed access points.

In our system, we use the HoloLens 2 hand tracking feature to acquire the coordinates of the user's both hands. The head position is derived from the device's own position. Information access points in the MR environment are initially placed relative to the HoloLens device's origin point.

2) Linking Information to Body Parts: To anchor information reliably to body parts, we implemented a custom system instead of using MRTK's default Orbital Solver. The Orbital Solver cannot correctly inherit wrist rotations and therefore keeps objects aligned to world coordinates, causing drift when the hand rotates.

To overcome this limitation, our method explicitly computes the access point's world pose using body-part transformation matrices:

$$T_{\text{anchor}}^{\text{world}} = T_{\text{body}}^{\text{world}} \cdot T_{\text{anchor}}^{\text{local}} \quad (1)$$

Here, $T_{\text{body}}^{\text{world}}$ is obtained from HoloLens hand tracking, and $T_{\text{anchor}}^{\text{local}}$ is the stored offset at the moment of attachment. This approach preserves both translation and rotation relative to the user's body, enabling access points to maintain stable posture and behave as if physically attached even during rapid or complex movements.

B. Accessing Body-anchored Information

1) Eye Tracking-based Access: The proposed system utilizes the eye-tracking feature of HoloLens 2 for hands-free interaction. Using the MRTK EyeTracking API, we continuously track the user's gaze to determine what virtual objects they are looking at. To ensure reliable performance, users are required to complete the eye calibration setup before use.

By using gaze for input, users can access virtual information without using their hands, allowing for intuitive interaction and improving overall usability in MR environments.

2) Access Points in Eye Tracking: Access points are activated when the user maintains gaze for 1.5 seconds, in order to prevent accidental activation while using the system. This mechanism ensures that a brief or unintended glance does not trigger actions.

Additionally, access points are normally invisible (transparent) and only become visible when they are being looked at. This design minimizes visual clutter in the user's field of view.

C. Assigning Access Points to Specific Body Parts

In this system, access points can be anchored to specific body parts—namely the head, right hand, or left hand. To make the process clear and consistent, anchoring is performed in the following three steps:

1) Selecting an Access Point: The user first selects one of the generated access points in the MR environment. Each access point can be repositioned freely in space before anchoring, allowing the user to choose a suitable location based on visibility and intended usage.

2) Choosing the Body Part to Attach To: Next, the user chooses a body part using one of the three UI buttons arranged vertically: "AttachHead," "AttachHandR," and "AttachHandL." When the user presses one of these buttons, the system identifies the corresponding body-part transform—head pose, right-hand pose, or left-hand pose—obtained from the HoloLens tracking system. At this moment, the system records the relative transformation between the access point and the selected body part. This stored offset defines the access point's local position and orientation in the user's body-centered coordinate system.

Fig. 1: Example of access points anchored to the head

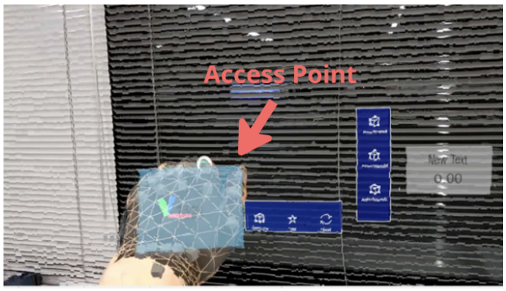

Fig. 2: Example of access points anchored to the hand

3) Anchoring and Following the Body Part: Once anchored, the access point becomes transparent and begins to follow the movement and rotation of the assigned body part. Because the system uses the transformation matrices of the tracked body part, the access point maintains a stable relative position and orientation, even when the user moves or rotates their hand or head.

When the user looks directly at the access point, it transitions from transparent to semi-transparent, allowing interaction through gaze. Fig. 1, Fig. 2 and Fig. 3 illustrate examples of access points anchored to the head and to the front or back of the hand. Because objects placed on the palm side are not visible from the back of the hand (and vice versa), the anchoring orientation affects usability and must be chosen carefully.

IV. Evaluation Methodology and Experimental Design

A. Evaluation Approach

To evaluate the effectiveness of the proposed MR system in tagging and searching information based on the body, we developed an experimental protocol to measure the speed at which users access various types of virtual information linked to the body. The experiments were conducted with participants aged between 19 and 23, all of whom were male and familiar with using digital devices such as personal computers. However, except for Participant B, all participants were new to MR-based interaction, and only Participant C was left-handed.

During the measurement process, the system randomly displays the name of the target access point and starts a timer. Participants fix their gaze on the specified access point, and when their gaze is maintained for 1.5 seconds,

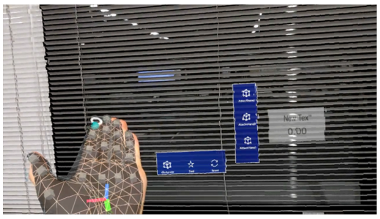

Fig. 3: View of access objects anchored to the back of the hand from the palm

the object lights up red to indicate successful access, and the timer stops. This procedure allows us to quantify access time, which serves as the primary metric for evaluating the efficiency of body-based tagging.

For this experiment, we prepared access points using information commonly encountered in daily life, such as "time," "day of the week," "date," "weather," "blood pressure," and "weight." Participants were asked to measure the time taken to access each access point and record which body position they associated with each access point. This evaluation approach investigates the cognitive and ergonomic aspects of body-based information tagging and provides insights into the optimal mapping between information types and body part locations.

B. Experimental Design

1) Experiment 1: Verifying Access Feasibility: In this preliminary experiment, we measured the access time to access points linked to the head and right hand to confirm that users could correctly and consistently access information anchored to their bodies.

As an experimental method, when the user pressed a button labeled "Test" in the MR space, the system randomly instructed the user to access a point anchored to either the head or wrist. The user was instructed to maintain their gaze on the correct object for 1.5 seconds, and when it lit up red, the visual timer stopped, indicating successful access. The time from pressing the button to successful access was recorded.

2) Experiment 2: Measurement of Access Time Linked to Location of Access Point: The main purpose of this experiment was to verify the mechanisms of information processing in MR spaces in everyday life. Specifically, we investigated how frequently checked information is managed in MR spaces using body-based tagging. In particular, we examined the optimal location for accessing information and the tendency for tagged information to be placed in specific locations.

This experiment consisted of two main steps.

a) Step 1: Anchoring Access Points: Participants used the "Generate" button to create six different access points, each representing one of the target information types. Each point was physically moved to a suitable location and then anchored to a selected body part (head, left hand, or right hand) using the appropriate "Attach" buttons.

The participants were advised not to move their body parts during the anchoring moment, as coordinates were captured at the time the attach button was pressed.

b) Step 2: Access Time Test: After anchoring all six access points, participants began the access time test by pressing the "Test" button. This action displayed the name of one of the access points on a panel, along with a timer. Participants then had to gaze at the specified access point until it turned red, indicating successful access.

V. Results and Discussion

A. Experiment 1: Feasibility of Accessing Body-anchored Information

1) Results: The results of this experiment are summarized in Fig. 4, which shows the average access time across participants. All participants were able to successfully access the intended access point in under one second, excluding the required 1.5 seconds for gaze confirmation.

This short access time suggests that participants were able to consistently recall the spatial location of the access points and precisely direct their gaze toward them.

2) Discussion: The data indicates that access to hand-anchored objects was slightly faster than access to head-anchored objects. This may explain why users were not yet accustomed to using eye movements as an input method. Participants tended to prefer moving their hands in the direction of their gaze rather than moving their gaze horizontally. As a result, hand-anchored objects (especially those placed in front of the face) were easier and faster to access.

To quantify this difference, we defined the average access time of participants to hand-anchored objects as M_1 and the average access time of participants to head-anchored objects as M_2. We calculated the effect size in access time using Cohen's d with the hand as the anchor, as shown in the equation

$$d = \frac{M_1 - M_2}{\alpha}. \tag{2}$$

The average difference in access time was 0.112 seconds, with a pooled standard deviation α of 0.156 seconds. This yielded an effect size of $d = 0.174$. While this value does

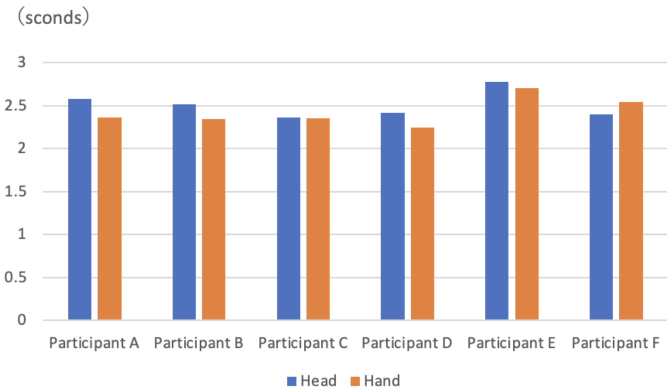

Fig. 4: Comparison of average access time for head-anchored and hand-anchored access points

TABLE I: Average Access Time by Body Part (Seconds)

Body Part	Average Access Time (seconds)
Head (21)	4.222
Left Hand (9)	4.033
Right Hand (6)	6.103

TABLE II: Number of Access Points Linked to the Back and Palm of the Hand

Body Part	Date	Blood Pressure	Time	Day of Week	Weight	Weather
Palm	0	6	0	1	1	0
Back	1	0	3	0	2	2

not indicate a statistically significant or strong effect, it does suggest a small but noticeable tendency in which hand-anchored access points may be accessed slightly faster than head-anchored ones.

Although the effect is modest, this trend may reflect the possibility that hand-based anchoring provides more accessible or visually salient reference points in gaze-based MR interaction. In particular, the hand's flexibility of placement and its natural alignment with a user's line of sight could contribute to this slight advantage. Further investigation with a larger sample size would be needed to more conclusively understand this relationship.

B. Experiment 2: Access Time by Information Type and Placement

1) Results: In this experiment, we measured the speed at which participants accessed access points and which parts of their bodies they associated with each access point, and organized the results as shown in Table I as the number of access points linked to the body parts in parentheses. The measurement results revealed a tendency for access points associated with the left hand to be accessed most quickly.

Fig. 5 shows the percentage of access points associated with each body part for each type of information. It was found that all participants associated "blood pressure" information with the hand, while information such as "day of the week" and "date" tended to be associated with the head.

2) Discussion: Access points anchored to the hands exhibited several notable characteristics. Access times for information anchored to the left hand were generally faster than those for the right hand. One possible explanation is that most participants (except Participant C) were right-handed and operated the system buttons with their dominant hand,

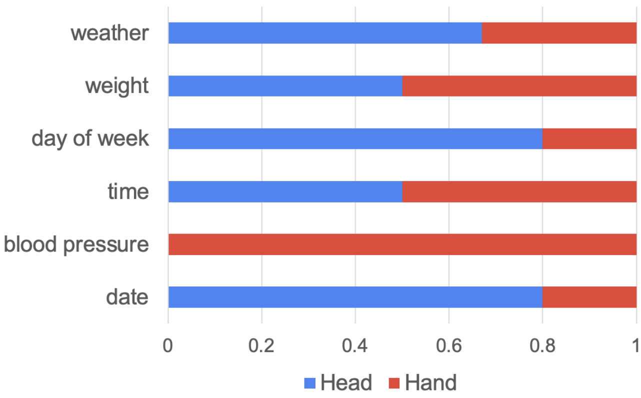

Fig. 5: Percentage of locations linked to the body by access point type

making the non-dominant (left) hand more stable and thus more suitable as a consistent spatial reference. Participants also noted that access points anchored to the non-dominant hand were easier to remember and access.

Furthermore, We showed the distribution table of information types on the back and palm sides of the hand II as the number of access points linked within parentheses.All participants consistently anchored "blood pressure" information to the palm side of the hand (near the wrist), whereas "time" was associated only with the back side. These results appear to reflect intuitive spatial memory associations based on daily experiences—blood pressure is typically measured on the arm, while time is often accessed by glancing at a wristwatch worn on the back of the hand.

Access points anchored to the head exhibited several distinct characteristics. The spatial positions of these head-anchored access points were categorized into six regions: upper left, lower left, upper center, lower center, upper right, and lower right. The number of access points assigned to each region is summarized in Table III. The results revealed a tendency for head-anchored access points to be biased toward the right side. This right-side bias may be attributed to the influence of eye dominance, as access in this system was performed using eye tracking. Most participants were right-handed, and previous studies have shown that right-handed individuals tend to have right-eye dominance [16]. This may have contributed to the observed spatial preference for placing access points on the right side of the head. In addition, time-related information also showed a tendency to be positioned on the right side. This may be influenced by the common UI design convention in which the current time is displayed on the right side of computer screens, potentially reinforcing spatial memory associations through habitual visual exposure.

Information was categorized into two groups: body-related information (e.g., blood pressure and weight) and non-body-related information (e.g., date, day of the week, time, and weather). The average access time to access points for each category are summarized in Table IV. Based on these results and participant feedback, a tendency was observed for body-related information to be accessed more quickly. This finding suggests that information with strong semantic or habitual associations with specific body parts—such as when information is anchored to the hands or head—tends to be more memorable and easier to access. These spatial

TABLE III: Number of Access Points per Location Anchored in the Head

Top Left	Bottom Left	Top Center	Bottom Center	Top Right	Bottom Right
3	3	4	0	6	4

TABLE IV: Average Access Speed for Related and Unrelated Information

Category	Average Access Speed (seconds)
Related Information (12)	4.855
Unrelated Information (24)	5.048

placements appear to reflect intuitive spatial memory associations grounded in users' everyday experiences. For example, blood pressure is typically measured on the arm, and time is often checked by looking at a wristwatch, reinforcing familiar spatial mappings.

These results indicate that not only physical ease of access but also the cognitive ease of recall plays a significant role in determining access time. In body-anchored interfaces, information that is naturally associated with particular body parts may be more easily encoded into spatial memory, allowing for faster recall and selection through gaze-based interaction. Consequently, the findings suggest that body-based tagging systems can leverage embodied cognition to support intuitive and memorable information organization in MR environments.

VI. Conclusion

In this study, we proposed a body-anchored information organization system for MR environments, leveraging HoloLens 2 and its eye-tracking capabilities. The system enables users to intuitively organize and access virtual information by associating it with specific parts of the body, such as the head or hands.

Through user experiments, we confirmed that information related to the body—such as blood pressure or weight—is easier to associate with and recall when anchored to corresponding body parts. Additionally, access time tended to improve when information was placed on the non-dominant hand, possibly due to its relative stability during interaction. These results suggest that body-based spatial tagging can enhance both the memorability and accessibility of virtual information in MR spaces.

The proposed system enables quick and intuitive access to information in everyday scenarios such as personal task management, context-based notifications, and health monitoring, and is expected to enhance the effectiveness of information presentation in MR spaces. The experimental results suggest that information placement reflecting the regularity observed in the placement of access points may contribute to improving users' access time to information. By embedding information in a spatial context consistent with the human body, users can intuitively recall the location of information, enabling efficient information retrieval. As a result, MR interfaces are expected to function as more natural and cognitively efficient tools in both professional and personal settings.

References

[1] D. Vianna, A.-M. Yong, C. Xia, A. Marian, and T. Nguyen, "A tool for personal data extraction," in *the 2014 IEEE 30th International Conference on Data Engineering Workshops*, 2014, pp. 80–83.

[2] J. D. Dinneen and C. Julien, "The ubiquitous digital file: a review of file management research," *Journal of the Association for Information Science and Technology*, vol. 71, no. 1, 2020, pp. E1–E32.

[3] D. Vianna, V. Kalokyri, A. Borgida, A. Marian, and T. Nguyen, "Searching heterogeneous personal digital traces," in *the Association for Information Science and Technology*, vol. 56, no. 1, 2019, pp. 276–285.

[4] T. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," *IEEE Translation Journal on Magnetics in Japan*, vol. 2, no. 8, 1987, pp. 740–741.

[5] T. Papadopoulos, K. Evangelidis, T. H. Kaskalis, G. Evangelidis, and S. Sylaiou, "Interactions in augmented and mixed reality: an overview," *Applied Sciences*, vol. 11, no. 18, 2021, pp. 2–22.

[6] M. T. Raghunath and C. Narayanaswami, "User interfaces for applications on a wrist watch," *Personal and Ubiquitous Computing*, vol. 6, no. 1, 2002, pp. 17–30.

[7] S. Zhao, F. Tan, and K. Fennedy, "Heads-up computing moving beyond the device-centered paradigm," *Communications of the ACM*, vol. 66, no. 9, 2023, pp. 56–63.

[8] L. Sidenmark and H. Gellersen, "EyeHead: synergetic eye and head movement for gaze pointing and selection," in *the 32nd Annual ACM Symposium on User Interface Software and Technology*, 2019, pp. 1161–1174.

[9] N. A. Rappa, S. Ledger, T. Teo, K. W. Wong, B. Power, and B. Hilliard, "The use of eye tracking technology to explore learning and performance within virtual reality and mixed reality settings: a scoping review," *Interactive Learning Environments*, vol. 30, no. 7, 2022, pp. 1338–1350.

[10] J. Hadnett-Hunter, G. Nicolaou, E. O'Neill, and M. Proulx, "The effect of task on visual attention in interactive virtual environments," *ACM Transactions on Applied Perception*, vol. 16, no. 3, 2019, pp. 1–7.

[11] K. Ishibashi, Z. Li, and K. Yada, "Application of deep learning to eye tracking video for estimating sales area where consumer looked," in *the 34th Annual Conference of the Japanese Society for Artificial Intelligence*, 2020, pp. 1–2.

[12] I. B. Adhanom, P. MacNeilage, and E. Folmer, "Eye tracking in virtual reality: a broad review of applications and challenges," *Virtual Reality*, vol. 27, 2003, pp. 1481–1505.

[13] M. Niitsuma, H. Hashimoto, and H. Hashimoto, "Spatial memory: an aid system for human activity in intelligent space," in *the 2006 IEEE International Conference on Robotics and Automation*, 2006, pp. 4258–4263.

[14] S. Shinoki and M. Niituma, "Hierarchical spatial memory for densification," in *the JSME Conference on Robotics and Mechatronics*, vol. 2015, 2015, pp. 2A1-P02.1–2A1-P02.4, (in Japanese).

[15] Z. Hu, J. Xu, S. Schmitt, and A. Bulling, "Pose2Gaze: eye-body coordination during daily activities for gaze prediction from full-body poses," *IEEE Transactions on Visualization and Computer Graphics*, 2024, pp. 1–12.

[16] D. Bourassa, I. C. McManus, and M. P. Bryden, "Handedness and eye-dominance: a meta-analysis of their relationship." *Laterality*, vol. 1, no. 1, 1996, pp. 5–34.