Affordance-Guided User Elicitation of Interaction Concepts for Unimodal Gaze Control of Potential Holographic 3D UIs in Automotive Applications

Maryia Kazhura*
Technische Universität Berlin
Volkswagen AG

Bea Vorhof†
Universität zu Lübeck

André Calero Valdez‡
Universität zu Lübeck

Figure 1: Figure 1: HoloLens 2 implementation of affordance guided gaze interaction concepts for eight exemplary interaction tasks, obtained from the user elicitation survey. The tasks were designed for use with a 240×135 mm large holographic 3D display, emulated with a HoloLens 2.

Abstract

Identifying usable and intuitive in¬teraction methods for novel display technologies, in settings where multiple interaction modes and devices can be used simultaneously (such as the car), remains a challenge for developers and user interface designers. The process can become even more complex when the target hardware is still in prototype stages of development and does not support usability tests in early design iterations. Using an affordance-guided user-centered elicitation survey with non-expert participants, we researched intuitive unimodal gaze interaction concepts to complete a series of interaction tasks with a 3D UI in a HoloLens 2 emulation of a 3D holographic passenger display inside a car.

Index Terms: Human-centered computing—Interaction design—Interaction design process and methods—User-centered design

1 Introduction

Current market reports expect a substantial increase in the so-called metaverse and Extended Reality (XR) or Cross Reality (CR) applications in the following years [47, 49]. Advancements in the field of autonomous driving technology also introduce new opportunities for immersive, non-driving-related infotainment and interaction experiences in future car interiors [54]. The demonstration of the head-worn CR device Microsoft HoloLens 2 in a moving vehicle shows a promising vision of how navigational content and HMI controls could be experienced in mixed reality while driving [3]. Augmented Reality HUDs (AR-HUDs), which have become state-of-the-art in multiple modern vehicles [44], could provide an additional platform for CR experiences, by allowing passengers to perceive and interact with driving-related information presented across multiple devices including interior displays and AR-HUD. In such single-user CR scenarios [51], the user interacts with multiple systems on different points on the Reality-Virtuality continuum. Furthermore, the development of immersive 3D display technologies, including holographic 3D displays [14], can further expand the design space for in-car interactions. In a holographic 3D display, virtual content is displayed in full natural 3D and placed in a large depth range in front of or behind a physical display screen, without loss of resolution or the need for extra head-worn hardware [43]. Hence, such novel display technologies could offer a more immersive and spatial display experience, with content not being restricted to a 2D representation on a flat display surface. Since no additional head-worn hardware is required, passengers could remain more aware of their surroundings while using the holographic 3D display, thus allowing them to interact with multiple devices simultaneously. Users could switch between multiple display devices to experience the same content in a potentially more enjoyable or informative way. For example, a

*e-mail: maryia.kazhura@volkswagen.de
†e-mail: bea.vorhof@student.uni-luebeck.de
‡e-mail: andre.calerovaldez@uni-luebeck.de

passenger could switch form viewing a navigational map on their smartphone, to viewing the same ma as a realistic 3D model with a route preview on a holographic display, or use a 3D representation of their car to configure the contents of exterior displays [5]. However, since touchscreen surfaces may not be suited for use with a holographic display, especially when the displayed content is placed in front of the touchscreen surface, research of novel interaction techniques and design ideas for holographic interfaces could improve future displays' usability. In addition, the interest in touchless interaction experiences increased in light of the recent COVID-19 pandemic [18, 38].

While mid-air gestures seem like the most suitable interaction modality for holographic interfaces, given that such techniques can be very similar to interactions with real-world objects and touchscreens and therefore easy to learn [23, 24], the realization of such techniques is associated with multiple challenges. Users may not know where and how to place their hands so they can be sensed best by the hand tracking cameras [11, 12] and the lack of haptic feedback on contact with a virtual object surface can introduce further uncertainty and cause the so-called undershooting and overshooting phenomena [4, 19]. Furthermore, continuous interactions may require users to keep their fingertip at a constant depth (e.g. during slider interactions [53], which can be particularly challenging in a moving vehicle. Unsupported mid-air interaction performed for an extended time period can also increase physical discomfort and muscle fatigue [16, 25].

Gaze, on the other hand, is an inclusive modality that enables a broader range of potential users to interact with computer systems and CR applications, by allowing users to use their hands to perform a different activity [42] simultaneously, such as using a touchscreen device or performing a mid-air gesture for additional input. Furthermore, users could interact with 3D content without having to learn a new set of 3D gestures when switching from a 2D touchscreen to a 3D hologram view. Since attention and eye movements go hand-in-hand [6, 17], explicit gaze interactions can benefit applications that require continuous attention [27], which would be particularly useful in autonomous driving scenarios or for non-driving passengers. As opposed to mid-air gestural interfaces, gaze interfaces could benefit users by allowing them to focus their attention on the displayed content without having to visually match their hand position with a target interaction zone in mid-air. Gaze patterns can further reveal specific areas of interest, which can enable individualized UIs inside the vehicle [17]. Implicit gaze controls can be used to predict which object the user intends to interact with, thus allowing the system to adjust the UI accordingly [15]. Some studies associated eye gaze interaction with faster task completion times compared to multimodal gaze-voice input [36], gaze input combined with a physical button or controller confirmation [40], hand pointing [48] or mouse selection [45]. Past studies also reported a lower cognitive workload of eye gaze compared to multimodal gaze-voice control [36] and a higher perceived ease compared to multimodal gaze-controller or gaze-button interaction [40], albeit with higher reported instances of eye fatigue and faster rising eye fatigue levels compared to head gaze [41]. Some researchers argue that the benefit of gaze control depends on the task at hand, reducing cognitive load in some tasks while increasing the cognitive load in others [13]. A general recurring challenge in applications that involve gaze controls is the so-called Midas Touch problem [21], which refers to users accidentally selecting items simply by looking at them [7].

In this contribution, we present unimodal gaze interactions for eight exemplary tasks with a potential holographic 3D display interface and discuss how these interactions could be applied to CR applications in automotive settings. This research project consisted of three phases. In Phase 1, we reviewed published literature to select suitable gaze interaction methods. In Phase 2, we conducted an elicitation survey with 64 non-expert participants to collect gaze interaction proposals for eight 3D interaction tasks (see Fig. 1 "priming"). In Phase 3, we analyzed the results from the elicitation survey. In Phase 4, we deduced interaction concepts based on the binned proposals from the elicitation survey. Each concept was implemented using the Microsoft MRTK [30, 31] version 2.7.2 and integrated with a HoloLens 2 emulation of a 240×135 mm large holographic 3D display (see Fig. 1 "results").

2 Method

2.1 Selection of gaze interaction methods (Phase 1)

After reviewing published literature on human-computer gaze interaction methods in 2D and 3D interfaces, we considered nine different gaze interaction methods suitable for use with our 3D display interface: (1) Eye pointing to trigger selections or display object-related information [1, 21, 28, 45]; (2) Smooth gaze pursuits of moving targets [9, 50]; (3) Gaze and blinking [29]; (4) Gaze and nodding [2]; (5) Combined head- and gaze-cursor movements [46]; (6) Gaze gestures [22, 33, 35]; (7) Exploration and selection of menu items via sidebars for gaze selection [7]; (8) Dwell-time techniques [8, 35, 37, 39, 52], for example to lock the degrees of freedom during 3D object manipulation [26], with dwell visualization presented on the object directly or around the cursor [10] and (9) locking of the gaze cursor to a target area [34, 55].

2.2 Interaction tasks and priming

We selected eight interaction tasks from the list of suitable use cases and interaction tasks with in-car holographic 3D displays explored in a previous work by Kazhura [24]. The selected tasks were presented in the following order (1-8): (1) *Search in menu*: turn a menu wheel until a desired item is at the center; (2) *Select item*: select a desired item from a selection of multiple items; (3) *Pull object closer*: move a 3D object in depth from an egocentric perspective towards oneself, until it reaches a desired position; (4) *Rotate object*: rotate a 3D object around its vertical axis until it reaches a desired rotation; (5) *Move slider*: move a temperature slider horizontally until it reaches a desired temperature; (6) *Move map*: move a 3D map until the map is centered around a target location; (7) *Zoom map*: zoom a 3D map until the desired view is reached; (8) *Select a target on a map*: select a desired target marker on the 3D map.

We used a rendered model of an automotive dashboard with an integrated holographic 3D display, followed by an explanation of the basic principles of holographic 3D visualization to prime participants to the affordances of an in-car holographic 3D display. Furthermore, we prepared extensive video explanations of the nine selected unimodal gaze interaction methods, as well as a step-by-step visual representation of each task execution. The task visualization included rendered images of the target 3D UI embedded in a hypothetical car interior displaying the following steps: the default state of the UI prior to interaction; a visual representation of the active target; a visual representation of the selected target; a visualization of the target as it is being moved/rotated/zoomed; a visualization of the released but still active target when it is no longer being manipulated; and the end of the interaction with the target being no longer activated.

2.3 Sample

We split the elicitation survey into two parts with four tasks each, to reduce the time demand and workload for each participant. Hence, we recruited two groups of volunteering participants with similar age distributions, with most participants born between 1996 and 1964 (73,53% in Group A and 76,33% in Group B). Each participant group completed one part of the survey: Group A (N = 34, F =15, M = 19) completed tasks one to four and Group B (N = 30, F =14, M = 16) completed tasks five to eight. Most participants (Group A: 88%; Group B: 73%) used gaze controls less than once per year.

2.4 Elicitation survey (Phase 2)

To collect the proposals, we conducted the survey remotely. Participants received a virtual presentation that contained information about the aims and background of the study, holographic 3D displays and how they could be used in a car, videos presenting the gaze interaction methods described above, and an introduction to each task's goal and its step-by-step visualization. Data was collected using online questionnaires linked in the presentation. The first questionnaire gathered information about participants' age, gender, driving habits, technology usage and media consumption. After filling out the demographic questionnaire, participants would proceed to view the material about the first of four tasks. A link to the task questionnaire was placed at the end of each task description. In the task questionnaire, participants were asked to describe their gaze interaction idea using the following instruction: "How would you complete this task using only your eye-gaze to control the user interface?". Participants were encouraged to refer to the gaze interaction methods described in the presentation. In addition, participants were asked to rate the ease and self-descriptiveness of their proposed interaction method on a 10-point scale. Ease referred to how easy the interaction would be to execute, while self-descriptiveness was the degree to which the interaction is self-explanatory, as described in the standard ISO 9241-110:2020 [20]. At the end of each task questionnaire, participants rated the usefulness of unimodal gaze interaction for the completion of the given task.

3 Results

3.1 Analysis of elicitation survey results (Phase 3)

In a first step, the 179 proposed interaction ideas were reviewed and filtered to remove incomplete responses that failed to describe a gaze interaction. The resulting number of analyzed proposals for each task was as follows: *Search in menu* (22), *Select item* (17), *Pull object closer* (14), *Rotate object* (15), *Move slider* (20), *Move map* (15), *Zoom map* (13) and *Select a target on a map* (13). The remaining responses were in part extensive and detailed, allowing us to use the set of responses to deduce meaningful interaction concepts. The number of gaze usefulness ratings obtained from each task was: *Search in menu* (23), *Select item* (21), *Pull object closer* (17), *Rotate object* (19), *Move slider* (24), *Move map* (17), *Zoom map* (17) and *Select a target on a map* (17). This shows how only a portion of the recruited participants completed all online questionnaires and provided usable data.

We binned the filtered proposals based on the similarity of the described methods during the single stages of interaction. For example: we grouped proposals that described looking at various additional UI objects or buttons to move or rotate an element (3D arrows, a 3D scale, buttons with icons, etc.) into the same category. We then ranked the binned groups according to their average ratings of ease and self-descriptiveness. If multiple categories had an equal number of occurrences, we prioritized the category with a higher ease and self-descriptiveness rating.

With respect to the obtained gaze usefulness ratings (see Fig. 2), we found that gaze control was considered most useful for selection tasks, while being rated less useful for the tasks *Zoom map*, *Search in menu*, *Rotate object* and *Pull object closer*.

3.2 Final interaction concepts (Phase 4)

We deduced the final concepts presented in Figure 1 based on the top three categories from the binning procedure, but considering additional factor such as technical feasibility, ergonomics, and affordances of holographic 3D-displays. We then implemented the interaction concepts according to the description below, using the Microsoft MRTK. Interaction with the system was enabled, when the user's gaze point was inside the UI-area within the bounds of the virtual display. Looking away from the virtual display resulted in a deactivation of all selected objects. Based on a pilot test with

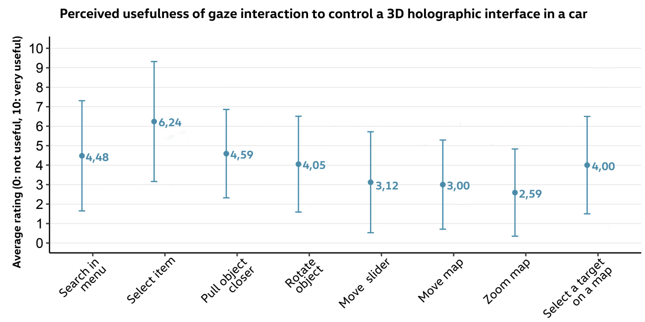

Figure 2: Average gaze usefulness ratings across eight tasks. Error bars depict standard deviations

10 participants we selected a threshold of 300 ms for gaze fixation that triggered visual highlighting of the fixated object and the start of a dwell timer. We selected 1000 ms for the dwell timer duration. Gaze selections therefore required a total fixation duration of 1300 ms. Gaze selections were confirmed with audio feedback. Visual feedback for the gaze interactions was designed according to insights from literature (e.g. [10]) and recommendations from the MRTK [32]. Our final gaze interaction concepts for each task are:

- *Search in menu*: User selects an arrow above the menu using dwell. During gaze fixation of the selected arrow, the menu wheel rotates in that direction.
- *Select item*: User selects an item using dwell.
- *Pull object closer*: User selects the 3D object using dwell. Upon selection, a 3D scale appears below the object, with a nearer arrow at the front and a farther arrow behind the object. User can use dwell to select an arrow and then continue fixating the selected arrow to move the object in that direction for the duration of arrow fixation.
- *Rotate object*: Upon gaze contact with the bounding box around the 3D object, a 3D rotation icon is displayed in the lower left corner of the box. User enters rotation mode by selecting the rotation icon using dwell. In rotation mode, two arrows appear below the box and user can look at an arrow to rotate the object in that direction for the duration of arrow fixation.
- *Move slider*: User selects the box on the slider via dwell. User can then fixate a point on the slider bar to trigger another dwell timer. When the dwell is completed, the slider gradually moves to the selected position.
- *Move map*: Upon gaze fixation on the map, four bars appear along the map edges. User can select an edge bar via dwell and move the map by looking at the selected edge. Once the target location is in view, the user can select it via dwell and the map automatically centers around the selected location.
- *Zoom map*: User can select one of the zoom buttons near the right edge of the map using dwell. User can zoom in or out (depending on the selected button) while fixating the selected button.
- *Select a target on a map*: User selects a target using dwell.

4 Discussion

Our interaction concepts were deduced from the proposed ideas of non-expert participants, based on affordance guided priming material. Our step-by-step visualization of each task left it open to interpretation, as to whether the task could be completed in a continuous or discrete manner. Some participants proposed discrete methods for tasks that others would resolve continuously (e.g., moving an object in steps vs. moving it continuously). We might have gotten different results had we used other visuals to convey the task goals. While we expect our interaction concepts to be intuitive, self-descriptive, and easy to use based on the ratings of the elicited proposals, we suggest viewing the results from such surveys as inspiration and guidance rather than explicit design instructions. The presented concepts may allow in-car passengers to interact with a 3D view of their currently used application on a holographic 3D display while simultaneously using their hands to gesture, or to control a touchscreen device to switch between different applications displayed on a 2D display. Since the proposed methods are limited in their complexity and can only be performed sequentially, some applications may even require the use of additional modalities (e.g. a physical controller device, smartphone or voice command). In addition, the presented gaze interactions could also be applied to future AR-HUDs during autonomous driving, allowing passengers to actively interact with the virtual content in the driving scene, while using other interaction modes to interact with multiple displays and linked devices in the interior (e.g. Smartphones or Smartwatches) in single or multi-user CR scenarios.

4.1 Future work

Since gaze can be efficiently combined with gesture or voice input to solve specific challenges in AR/VR applications [42, 54], our survey further collected interaction proposals for potential multimodal controls of the suggested tasks, combining gaze input with mid-air gestures, voice control, or other modalities that participants could imagine using in combination with gaze. Using the same approach, we deduced multimodal gaze-supported interaction concepts with mid-air gestures and voice commands. While these results are out of scope for this contribution, we plan to evaluate the task performance, user experience, and usability of the presented unimodal gaze interactions and compare them with a multimodal gaze-hand tracking interface. Furthermore, we will compare mid-air hand interactions with multimodal gaze-hand and gaze-voice interactions to further investigate the potential benefits of gaze input for interaction with immersive 3D UIs.

4.2 Limitations

One major limitation of our approach is the remote elicitation approach, which required participants to elicit ideas for a technology they had little to no experience with, based on low-fidelity priming materials (descriptions, videos, and images as opposed to an immersive CR experience of each task's affordances). This low-fidelity approach could have limited participants' creativity and the perception of gaze controls' usefulness for 3D UIs. Furthermore, the generalizability of our results is limited by the specific priming material used to elicit the proposals, as the proposals were heavily influenced by the affordances of the visualized UI. In addition, the study was anonymous, thus limiting communication between us and participants. Participants could not make quick inquiries about the presented methods or openly discuss a specific topic or idea. We can see how this impacted our results based on the number of complete and usable responses, in contrast to the total sample size of each group. It is also important to note that while we surveyed non-experts and split the tasks into two groups, experienced AR/VR users might propose different and more consistent approaches with a greater focus on established UI/UX guidelines and technical feasibility. The cultural diversity of our sample was also limited since we recruited participants from a single region in Germany. People with other demographic backgrounds may suggest different interaction ideas or dismiss certain ideas entirely. The gaze method 9 (locking the gaze cursor to a target area) was presented in less detail compared to the other methods, which may have made it more challenging for participants to understand and include the method in their proposals. Finally, the presented UIs do not include much visual clutter. It is therefore debatable whether the concepts are transferable to more complex holographic interfaces.

5 Conclusion

Non-experts who have had little experience with XR or CR technologies may not see the benefits of gaze control for certain tasks yet and may find it challenging to propose suitable interaction ideas. We were able to use the results from the elicitation study to develop user-centered gaze control-based 3D UIs for future evaluation. Based on our experience, we encourage the inclusion of participants' ideas in early design iterations of novel UIs. However, we recommend considering a more direct and interactive CR survey approach, to help participants whose imagination and creativity may be limited. For example, by letting users experience the priming UI in XR or by using a visual editor on a tablet combined with a mixed or virtual reality visualization of the edited concept. However, we believe that priming non-experts to the affordances of the intended system can help obtain more meaningful results.

Acknowledgments

This study was funded by the Volkswagen Aktiengesellschaft. The results, opinions and conclusions expressed in this publication are not necessarily those of Volkswagen Aktiengesellschaft.

References

[1] F. Alt, S. Schneegass, J. Auda, R. Rzayev, and N. Broy. Using eye-tracking to support interaction with layered 3D interfaces on stereoscopic displays. In *Proc. IUI*, pp. 267–272. Association for Computing Machinery, New York, 2014. doi: 10.1145/2557500.2557518

[2] A. Ataya, W. Kim, A. Elsharkawy, and S. Kim. Gaze-head input: Examining potential interaction with immediate experience sampling in an autonomous vehicle. *Applied Sciences*, 10(24):9011, Dec. 2020. doi: 10.3390/app10249011

[3] D. Bach. With their hololens 2 project, microsoft and volkswagen collaborate to put augmented reality glasses in motion, May 2022.

[4] L.-W. Chan, H.-S. Kao, M. Y. Chen, M.-S. Lee, J. Hsu, and Y.-P. Hung. Touching the void: Direct-touch interaction for intangible displays. In *Proc. CHI*, pp. 2625–2634. Association for Computing Machinery, New York, 2010. doi: 10.1145/1753326.1753725

[5] A. Colley, J. Häkkilä, B. Pfleging, and F. Alt. A design space for external displays on cars. In *Proc. AutomotiveUI*, p. 146–151. Association for Computing Machinery, New York, 2017. doi: 10.1145/3131726.3131760

[6] A. T. Duchowski. *Eye tracking methodology: Theory and practice*. Springer, Cham, third ed., 2017. doi: 10.1007/978-3-319-57883-5

[7] C. Elmadjian and C. H. Morimoto. Gazebar: Exploiting the midas touch in gaze interaction. In *Extended Abstracts CHI*. Association for Computing Machinery, New York, 2021. doi: 10.1145/3411763.3451703

[8] A. Esteves, Y. Shin, and I. Oakley. Comparing selection mechanisms for gaze input techniques in head-mounted displays. *International Journal of Human-Computer Studies*, 139:102414, Feb. 2020. doi: 10.1016/j.ijhcs.2020.102414

[9] A. Esteves, D. Verweij, L. Suraiya, R. Islam, Y. Lee, and I. Oakley. Smoothmoves: Smooth pursuits head movements for augmented reality. In *Proc. UIST*, pp. 167–178. Association for Computing Machinery, New York, 2017. doi: 10.1145/3126594.3126616

[10] M. Fernandez, F. Mathis, and M. Khamis. Gazewheels: Comparing dwell-time feedback and methods for gaze input. In *Proc. NordiCHI*. Association for Computing Machinery, New York, 2020. doi: 10.1145/3419249.3420122

[11] E. Freeman. *Interaction techniques with novel multimodal feedback for addressing gesture-sensing systems*. PhD thesis, University of Glasgow, Glasgow, UK, 2016.

[12] E. Freeman. Ultrasound haptic feedback for touchless user interfaces: Design patterns. In O. Georgiou, W. Frier, E. Freeman, C. Pacchierotti, and T. Hoshi, eds., *Ultrasound Mid-Air Haptics for Touchless Interfaces*, pp. 71–98. Springer International Publishing, Cham, CH, 2022. doi: 10.1007/978-3-031-04043-6

[13] P. Gauselmann, C. Frings, and T. Tempel. Cognitive offloading benefits eye gaze interaction. *Applied Cognitive Psychology*, pp. 1–5, June 2023. doi: 10.1002/acp.4098

[14] R. Häussler, Y. Gritsai, E. Zschau, R. Missbach, H. Sahm, M. Stock, and H. Stolle. Large real-time holographic 3D displays: enabling components and results. *Applied optics*, 56(13):F45–F52, May 2017. doi: 10.1364/AO.56.000F45

[15] J. M. Henderson. Gaze control as prediction. *Trends in cognitive sciences*, 21(1):15–23, Dec. 2017. doi: 10.1016/j.tics.2016.11.003

[16] J. D. Hincapié-Ramos, X. Guo, P. Moghadasian, and P. Irani. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proc. CHI*, pp. 1063–1072. Association for Computing Machinery, New York, 2014. doi: 10.1145/2556288.2557130

[17] W. J. Horrey, C. D. Wickens, and K. P. Consalus. Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. *Applied journal of experimental psychology*, 12(2):67–78, June 2006. doi: 10.1037/1076-898X.12.2.67

[18] S. Huang, S. P. B. Ranganathan, and I. Parsons. To touch or not to touch? comparing touch, mid-air gesture, mid-air haptics for public display in post covid-19 society. In *SA Posters*. Association for Computing Machinery, New York, 2020. doi: 10.1145/3415264.3425438

[19] J. H. R. Isaac, M. K. V., and M. M. Effect of visual awareness of the real hand on user performance in partially immersive virtual environments: Presence of virtual kinesthetic conflict. *International Journal of Human–Computer Interaction*, 36(16):1540–1550, June 2020. doi: 10.1080/10447318.2020.1768667

[20] ISO/TC 159/SC 4. Ergonomics of human-system interaction — Part 110: Interaction principles. Standard ISO 9241-110:2020, International Organization for Standardization, Geneva, CH, May 2020.

[21] R. J. K. Jacob. What you look at is what you get: Eye movement-based interaction techniques. In *Proc. CHI*, pp. 11–18. Association for Computing Machinery, New York, 1990. doi: 10.1145/97243.97246

[22] M. Jenke, L. Huppenbauer, and T. Maier. Investigation of continual operation procedures via a user centered gaze control by means of flexible gaze gestures. *Zeitschrift für Arbeitswissenschaft*, 72(1):23–34, Mar. 2018. doi: 10.1007/s41449-017-0067-6

[23] H.-C. Jetter, H. Reiterer, and F. Geyer. Blended interaction: understanding natural human–computer interaction in post-wimp interactive spaces. *Personal and Ubiquitous Computing*, 18:1139–1158, Oct. 2014. doi: 10.1007/s00779-013-0725-4

[24] M. Kazhura. Exploring new depths: How could passengers interact with future in-car holographic 3D displays? In *Proc. MobiTAS (HCII)*, pp. 35–61. Springer-Verlag, Berlin, Heidelberg, 2022. doi: 10.1007/978-3-031-04987-3

[25] J. H. Kim, H. Ari, C. Madasu, and J. Hwang. Evaluation of hologram distances in reducing shoulder stress during augmented reality interactions. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 64(1):868–871, Feb. 2020. doi: 10.1177/1071181320641201

[26] C. Liu, A. Plopski, and J. Orlosky. Orthogaze: Gaze-based three-dimensional object manipulation using orthogonal planes. *Computers & Graphics*, 89:1–10, June 2020. doi: 10.1016/j.cag.2020.04.005

[27] F. Lu, S. Davari, L. Lisle, Y. Li, and D. A. Bowman. Glanceable ar: Evaluating information access methods for head-worn augmented reality. In *Proc. VR*, pp. 930–939. IEEE, Piscataway, NJ, 2020. doi: 10.1109/VR46266.2020.00113

[28] A. McNamara and C. Kabeerdoss. Mobile augmented reality: Placing labels based on gaze position. In E. Veas, T. Langlotz, J. Martinez-Carranza, R. Grasset, M. Sugimoto, and A. Martín, eds., *Proc. ISMAR*, pp. 36–37. IEEE, Piscataway, NJ, 2016. doi: 10.1109/ISMAR-Adjunct.2016.0033

[29] X. Meng, W. Xu, and H.-N. Liang. An exploration of hands-free text selection for virtual reality head-mounted displays. In H. Duh, ed., *Proc. ISMAR*, pp. 74–81. IEEE, Piscataway, NJ, 2022. doi: 10.1109/ISMAR55827.2022.00021.

[30] Microsoft. Mixed Reality Toolkit: Unity software package (v2.7.2), 2021.

[31] Microsoft. What is Mixed Reality Toolkit 2? online article, 2022.

[32] Microsoft. Eye-gaze-based interaction on hololens 2: Online article, 2023.

[33] E. Mllenbach, M. Lillholm, A. Gail, and J. P. Hansen. Single gaze gestures. In *Proc. ETRA*, pp. 177–180. Association for Computing Machinery, New York, 2010. doi: 10.1145/1743666.1743710

[34] A. Murata and W. Karwowski. Automatic lock of cursor movement: Implications for an efficient eye-gaze input method for drag and menu selection. *IEEE Transactions on Human-Machine Systems*, 49(3):259–267, June 2018. doi: 10.1109/THMS.2018.2884737

[35] O. Namnakani, Y. Abdrabou, J. Grizou, A. Esteves, and M. Khamis. Comparing dwell time, pursuits and gaze gestures for gaze interaction on handheld mobile devices. In *Proc. CHI*. Association for Computing Machinery, New York, 2023. doi: 10.1145/3544548.3580871

[36] M. Parisay, C. Poullis, and M. Kersten-Oertel. Eyetap: Introducing a multimodal gaze-based technique using voice inputs with a comparative analysis of selection techniques. *International Journal of Human-Computer Studies*, 154:102676, Oct. 2021. doi: 10.1016/j.ijhcs.2021.102676

[37] Y. T. Paulus and G. B. Remijn. Usability of various dwell times for eye-gaze-based object selection with eye tracking. *Displays*, 67:101997, Apr. 2021. doi: 10.1016/j.displa.2021.101997

[38] J. Pearson, G. Bailey, S. Robinson, M. Jones, T. Owen, C. Zhang, T. Reitmaier, C. Steer, A. Carter, D. R. Sahoo, and D. K. Raju. Can't touch this: Rethinking public technology in a covid-19 era. In *Proc. CHI*. Association for Computing Machinery, New York, 2022. doi: 10.1145/3491102.3501980

[39] K. Pfeuffer, Y. Abdrabou, A. Esteves, R. Rivu, Y. Abdelrahman, S. Meitner, A. Saadi, and F. Alt. Artention: A design space for gaze-adaptive user interfaces in augmented reality. *Computers & Graphics*, 95(4):1–12, Apr. 2021. doi: 10.1016/j.cag.2021.01.001

[40] K. Pfeuffer, L. Mecke, S. Delgado Rodriguez, M. Hassib, H. Maier, and F. Alt. Empirical evaluation of gaze-enhanced menus in virtual reality. In *Proc. VRST*. Association for Computing Machinery, New York, 2020. doi: 10.1145/3385956.3418962

[41] Y. Y. Qian and R. J. Teather. The eyes don't have it: An empirical comparison of head-based and eye-based selection in virtual reality. In *Proc. SUI*, pp. 91–98. Association for Computing Machinery, New York, 2017. doi: 10.1145/3131277.3132182

[42] I. Rakkolainen, A. Farooq, J. Kangas, J. Hakulinen, J. Rantala, M. Turunen, and R. Raisamo. Technologies for multimodal interaction in extended reality—a scoping review. *Multimodal Technologies and Interaction*, 5(12):81, Dec. 2021. doi: 10.3390/mti5120081

[43] S. Reichelt, R. Häussler, G. Fütterer, and N. Leister. Depth cues in human visual perception and their realization in 3D displays. In B. Javidi, S. Jung-Young, J. T. Thomas, and D. D. Desjardins, eds., *Proc. SPIE DEFENSE, SECURITY, AND SENSING*, vol. 7690. SPIE, Bellingham, WA, 2010. doi: 10.1117/12.850094

[44] Research in China. Automotive head-up display (HUD) industry report, 2022, Aug. 2022.

[45] L. E. Sibert and R. J. K. Jacob. Evaluation of eye gaze interaction. In *Proc. CHI*, pp. 281–288. Association for Computing Machinery, New York, 2000. doi: 10.1145/332040.332445

[46] L. Sidenmark, D. Potts, B. Bapisch, and H. Gellersen. Radi-eye: Hands-free radial interfaces for 3D interaction using gaze-activated head-crossing. In *Proc. CHI*. Association for Computing Machinery, New York, 2021. doi: 10.1145/3411764.3445697

[47] Statista, Inc. AR hardware - worldwide: Online article, 2022.

[48] V. Tanriverdi and R. J. K. Jacob. Interacting with eye movements in virtual environments. In *Proc. CHI*, pp. 265–272. Association for Computing Machinery, New York, 2000. doi: 10.1145/332040.332443

[49] J. Trenker. Metaverse market report, Feb. 2023.

[50] M. Vidal, A. Bulling, and H. Gellersen. Pursuits: Spontaneous interaction with displays based on smooth pursuit eye movement and moving targets. In *Proc. UbiComp*, pp. 439–448. Association for Computing

Machinery, New York, 2013. doi: 10.1145/2493432.2493477
[51] N. Wang and F. Maurer. A design space for single-user cross-reality applications. In *Proc. AVI*. Association for Computing Machinery, New York, 2022. doi: 10.1145/3531073.3531116
[52] C. Ware and H. H. Mikaelian. An evaluation of an eye tracker as a device for computer input2. *SIGCHI Bull.*, 17(SI):183–188, May 1986. doi: 10.1145/30851.275627
[53] K. Waugh, M. McGill, and E. Freeman. Push or pinch? exploring slider control gestures for touchless user interfaces. In *Proc. NordiCHI*. Association for Computing Machinery, New York, 2022. doi: 10.1145/3546155.3546702
[54] G. Wiegand, C. Mai, K. Holländer, and H. Hussmann. InCarAR: A design space towards 3D augmented reality applications in vehicles. In *Proc. AutomotiveUI*, pp. 1–13. Association for Computing Machinery, New York, 2019. doi: 10.1145/3342197.3344539
[55] S. Zhang, Y. Tian, C. Wang, and K. Wei. Target selection by gaze pointing and manual confirmation: performance improved by locking the gaze cursor. *Ergonomics*, 63(7):884–895, May 2020. doi: 10.1080/00140139.2020.1762934