Evaluation on User Perception Effect Based on Interaction Techniques in the Stereoscopic Environment

Dino Caesaron[1*], Rio Prasetyo Lukodono[2], and Yunita Nugrahaini Safrudin[3]

[1,3]Industrial Engineering Department, Telkom University
Bandung 40257, Indonesia
[2]Industrial Engineering Department, Universitas Brawijaya
Malang 65145, Indonesia
Email: [1]dino.caesaron@gmail.com, [2]rio_pl@ub.ac.id, [3]yunitanugrahainis@telkomuniversity.ac.id

Abstract—The interaction of user performance with three-dimensional (3D) objects has become an important issue in the recent development of virtual reality applications. Additionally, the basic conviction of current Virtual Reality (VR) supports the development of the viable interface between humans and machines. The research focuses on the user's interaction technique by considering two approaches (direct and indirect interaction techniques) for the users while interacting with three-dimensional objects. Numerous possible uses can benefit from virtual reality by considering a few fundamental visual and cognitive activities in the Virtual Environment (VE), such as the interpretation of space that users of clear and indirect perception are not well established. The experiment is performed in a stereoscopic environment using a reciprocal tapping task. Participants are expected to use direct pointing as well as indirect cursor techniques to select a stereoscopic spherical target. The results show that, in the sense of a direct interaction technique, user recognition of an object appears to converge in the center of a simulated area. Unfortunately, this convergence is not demonstrated in the indirect cursor situation. The pointing estimation from the users is more accurate when using the indirect interaction approach. The findings provide an understanding of the interaction characteristics done by the users in the stereoscopic environment. Importantly, developers of a virtual environment may use the result when developing effective user interface perception in specific interaction techniques.

Index Terms—User Perception, Interaction Techniques, Stereoscopic Environment

I. INTRODUCTION

VIRTUAL Reality (VR) has been launched and selected for different uses in a wide range of disciplines. For illustration, VR is used as an employment training program in military and surgical simulations, the entertainment field, and educational

research instruments. Interaction between the users and three-dimensional (3D) objects has been built by recent VR technology developments. Newer technologies and exhibits are dominated and used in two kinds of environment. It consists of stereoscopic and Head-Mounted Display (HMD). These technologies provide a more natural and visual impression of living in a Virtual Environment (VE), i.e., separating from the real physical reality environment of the users. Using HMD to study in the virtual system will help to understand the user perception works and the factor which impacts the depth perception in the environment [1]. However, devices like HMD typically generate inconvenience. For example, it is motion sickness, in which users generally perceive signal conflict with the gestures perceived by the vestibular system [2]. As an option, stereoscopic displays have recently risen in popularity as the market for hardware, applications, and imagery has been relatively limited. In addition, stereoscopic display environment has found less motion sickness than HMD for VEs, as viewed on a laptop screen and projection monitor systems [3].

The stereoscopic system provides each eye of the observer with slightly different images. These separate images will produce a slightly different view of each eye's scene via suitable technology called parallax. Then, it involves displaying on the monitor for two different images. Objects can be seen with zero, negative, and positive values. Parallax, referring to the points seen on, in front of, or behind the screen. Attached to the screen monitor, objects with zero parallaxes are ideally suitable for touch screen interaction [4].

On the other hand, contact with objects in front or behind the screen is more complicated. It represents tasks in reality that involve contact or touch objects. For example, in the real world, some tasks involve

putting a bare hand or a device in a deep plane (front or behind the screen). For this intent, the researchers explore the difficulty of interacting with objects as a simulated subject projected in front of the viewers at a specific depth within close and far-fields in front of the display screen.

The principle of direct and indirect interaction is used to create the characteristics of interaction between the users and the stereoscopic objects [5–7]. Users use their body (e.g., pointing or reaching by bare hand or tools), including gestures and eye gaze, to navigate through objects in stereoscopic environments that visual touch can be used. Indirect interaction also uses an intermediary instead of directly communicating with the object, such as using a controller to manipulate an icon (a virtual pointer, mouse, and others). It is to obtain items within stereoscopic environments. When users "touch or communicate" a stereoscopic object in Ves, considering the misjudgement in virtual space is one question posed in particular with direct interaction [8]. This analysis compares the various characteristics of direct and indirect interactions between stereoscopic objects in front of the display screen. To assign a stereoscopic object, the researchers examine the interaction with a direct pointing technique and evaluate it to the indirect method in which a user controls a virtual cursor to acquire a stereoscopic object.

It is essential to know the location of the items around them and how to interact or access them to perform most things, such as daily tasks that the operations shall provide a comprehensive spatial overview of the environment for proper navigation. The combination of cognitive abilities is needed for the case with several conditions like real-world orientation, object perception (what people see), position (where), and obstacle prevention (how) [9]. In general, two tasks of perception and vision can use to estimate perception of space. The origin of spatial perception and vision function has fascinated scientists with these two tasks physiologically and psychologically.

Locating an object in the Virtual World (VW) is different from the VE. The researchers report underestimating the location of virtual targets from the actual positions, although the participants can estimate a right metrics distance [10]. In the HMD case, it underestimates distance by almost 10% for VE [11]. One of the conditions causing the underestimate is the distance of more than 60 cm to estimate objects located in the near VE [10, 12]. On the other hand, some researchers also find that overestimation occurs in the perceiving visual sources from the sound localization position effect [13]. Factors that affect every estimation are the quality of VR condition, computer graphics,

sense of the present, and limited field of view [4, 14]. However, it depends on the purpose people use VR. For example, if people consider the high-risk case in VR, it can make them activate the mechanism to avoid familiarizing themselves with negative stimuli [15].

The research investigates the experience of frontal parameters with the interaction approaches of direct and indirect interactions within VEs. These two interaction concepts can be used to engage with stereoscopic objects in front of the display screen. Moreover, the interaction between the two methods and the reference objects is used to evaluate space interpretation at a frontal stage. The contribution of the research is investigating the role of direct or indirect interactions in frontal space perception. The results can provide VR developers with insights to select the right method of interaction for their application.

II. RESEARCH METHOD

The experiment is performed in a stereoscopic environment using the reciprocal tapping task. Participants are expected to use direct pointing as well as indirect cursor techniques to select a stereoscopic spherical target. Red spheres, arranged in a circle, represent targets in the experiment. In the order defined by ISO [16], eight targets are displayed one by one [7]. The respondents use the pointer stick in a direct position to put the tip of the stick in the middle of the targeted area. Equally, the virtual cursor has to be positioned over the target in VE in an indirect condition by adjusting the dual analog gamepad. When the target is defined, no visual input is given (e.g., color change) other than the next target.

The experiment uses a setup of 2×3 in the combination of level factors. The independent variable for the experiment considers interaction technique (direct or indirect interactions) and the target depth (90 cm, 120 cm, and 150 cm) of the participants. Then, the results collected from both experiments are structured to examine how close the position of a particular point along the x-axis and the y-axis is. It aims to analyze the general sense of frontal space.

The respondent's understanding of the depth role on space of frontal magnitude has not been studied in the research. It should be acknowledged that the inaccuracy of in-depth estimation may lead to the inaccuracy of the perception of distance. The researchers recruit participants by advertising through social media. In total, 14 healthy voluntary participants join the research. All the respondents are right-handed. None of the respondents uses glasses or contact lenses, and their Inter-Pupillary Distance (IPD) ranges from 6.5 cm to 7 cm (M = 6.57 cm). IPD is defined as the distance

of two centers of human eyes pupil [8]. Additionally, none of the participants have eye abnormalities, such as blind color, fatigue, amblyopia, or disruptions (stereopsis).

Both participants' stereoscopic vision capability is checked by viewing the closest reference distance (90 cm from the subject) before every experiment, and no participant fails the test. Most of the participants have no VR experience at all. Participants are allowed to take a break at any time between trials and encouraged to leave at any time if they feel VE-related discomfort to reduce the effects of fatigue.

Participants are asked to complete the pre-experiment consent questionnaire and outline the aims and protocols of the procedure. Participants are split into two groups. One group starts the experiment on the direct condition and the other on the indirect condition. Target depths are automatically allocated as soon as the procedure is chosen. The experiment is split into two phases of the trial (condition) to reduce subjects' learning impact and anxiety. For each interaction, 24 assessments are performed by each person (3 target depths $\times$ 8 targets). Overall, the procedure is performed within 60 minutes by each respondent.

III. RESULTS AND DISCUSSION

A. Experimental Setup

The VE, stereoscopic targets, and experimental activities have been established in the Unity 3D application. The applications modify the depth of the target. It depends on the depth of information of each person and the IPD. The VE, stereoscopic targets, and experimental activities have been established in the Unity 3D application. The depth of the target is modified by the applications, depending on the depth information of each person and IPD.

The experiment area is 4 m $\times$ 3 m $\times$ 2.5 m. It is surrounded by black curtains to avoid excessive light and generate an outstanding stereoscopic environment. The experimental space is equipped with tables and adjustable chairs. The researchers use the projection screen display for the trial, which is 130 cm wide $\times$ 100 cm high. At a set position, the participants are located 210 cm away relative to the screen. In ViewSonic PJD6251 DLP, XGA resolution (1024 $\times$ 768) in DPI is used for a stereoscopic projector located under the table and participants' eyes. The NVIDIA 3D glasses with an emitter-receiver have been used to capture the stereoscopic images projected by the projector. In addition, a chinrest is installed to hold the participants' heads so that the perception of the goal depth will not change.
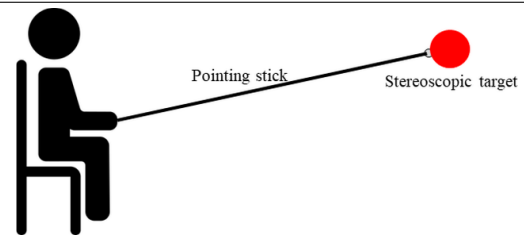

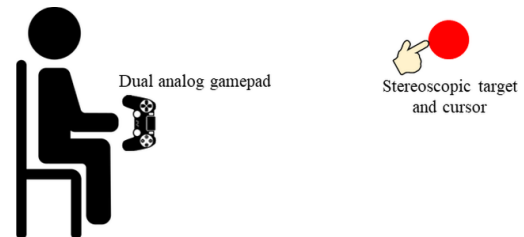Fig. 1. Illustration of direct pointing technique condition.

Fig. 2. Illustration of indirect cursor technique.

Figure 1 is an example of the process of precise pointing of the experiment. In direct pointing conditions, the researchers use three separate lengths of the stick: 80 cm, 110 cm, and 140 cm for interaction. For monitoring purposes, the researchers attach reflective markers to the tips of the pointing sticks. Then, Flex 13 OptiTrack motion system at 120 Hz frame rate is used with submillimeter precision and accuracy for monitoring the marker. In addition, to validate their actions (pointing), participants have to push the lower end of the handle with an attached wireless remote. The starting points of the stick tips are set at three different locations on the table to create a comfortable posture before each trial.

Figure 2 depicts the indirect cursor technique condition. The Sun-Yes R011 dual analog gamepad controller is used for interaction in the indirect configuration to guide the virtual cursor in the VE along the x-, y-, and z-axes. The virtual cursor's velocity is set to a comfortable and fast movement state that the researchers apply a value of 2 m/s for subsequent experiments. The data streams are obtained and processed by the PC from the marker locations, virtual cursors, and targets. For these reasons, all direct and indirect procedures are balanced and defined within global coordinates to conform with the same positional viewpoint. Figure 3 shows the experimental condition for direct and indirect cursor techniques.

B. Results

The researchers outline and interpret the findings of the experiment. The effects of interaction methods (direct and indirect techniques) on the perception of
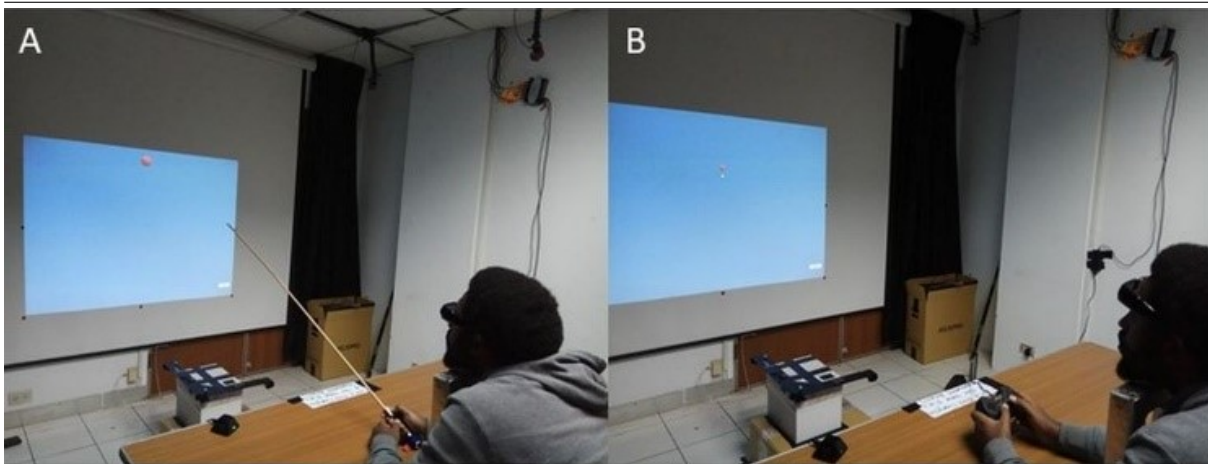
Fig. 3. Experimental conditions: (A) Direct interaction and (B) Indirect interaction technique.
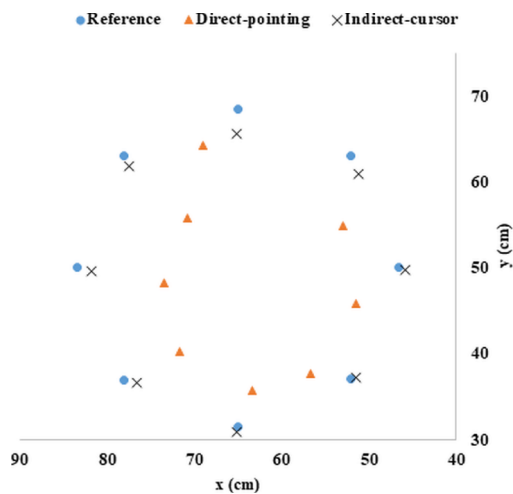

Fig. 4. The overall result of respondents' mean pointing estimations to target references.

frontal expansion are presented in stereoscopic environments. From Fig. 4, it can be observed that the overall judgment of the indirect cursor techniques is comparatively higher than the direct pointing technique to the stereoscopic target because the researchers do not calculate the accuracy score. However, to corroborate the analysis, the researchers review and find evidence for different characteristics of direct pointing and indirect cursor techniques. At least, there are two reasons that how depth cues, such as presence and occlusion, can affect the perception of the front view.

One possible reason for the accuracy differences is the higher sense of presence (i.e., virtual cursor), especially when the indirect cursor is used. The experiment from [17] finds that the user's performance in VEs can be improved due to the involvement of the users in the virtual cues (i.e., avatars). In the present analysis, while the simulated world is presented in a sparse environment (i.e., spheres in front of a dark blue background) for both conditions of interaction strategies, only the indirect cursor condition contains virtual cues (i.e., a virtual cursor). Its display can enhance the sense of presence when choosing a target related to a direct pointing condition. The target distance and position in the VE also influence the accuracy while using HMD and Screen Wide Display (SWD) [18]. The result is in line with previous research finding that a third-person view enables better spatial awareness, which leads to accurate interactions [19]. In the research, the indirect cursor condition is similar to the third-person view (perspective).

The second reason is an occlusion. A visual conflict happens when selecting a virtual target by using the direct pointing technique. It can be another reason for less accuracy in the direct pointing technique [20, 21]. When the participant tries to reach or point a virtual target with a pointing stick, the virtual target or the pointing stick appears blurred. On the other side, in the indirect cursor condition, both virtual target and virtual cursor are displayed stereoscopically, which may reduce the visual conflicts [22, 23].

The horizontal and vertical inaccuracies of the direct method in the stereoscopic environments draw all the targets towards the middle, rendering the total space compactly and smaller than expected. Therefore, it can infer that the perception of the frontal extent in stereoscopic environments which are lower than expected, especially in using direct technology.

The result is substantially in line with previous research, showing that the view of the observer's space in a stereoscopic environment is smaller than expected [24, 25]. However, for the indirect cursor
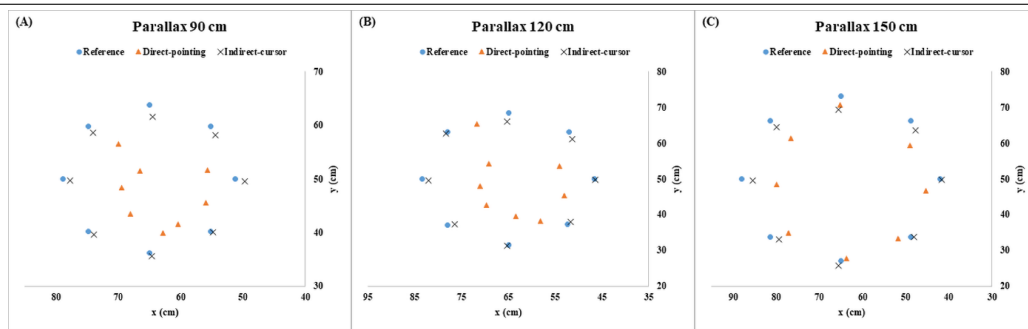
Fig. 5. The result of the estimation for pointing target references displayed at (A) 90 cm, (B) 120 cm, and (C) 150 cm in front of the participants.

technique, the compressed perception is not so obvious. These differences can be explained by the fact that both virtual target and virtual cursor are displayed stereoscopically within VE in this experiment. Therefore, a similar stereoscopic environment can reduce the misperception of space compression [22].

The various depths of the estimated target in this analysis are also compared for the pointing estimates. As seen in Fig. 5A-C, it gives the position judgment and appears more accurate to all the target depths measured than the direct technique. Furthermore, for the target shown in front of the participants at 90 and 120 cm, it points judgment in a direct situation. It shows the same results as the average mean, indicating the compressed frontal dimension. However, at 150 cm from the participants, the corresponding target depth finds that the two techniques are comparably accurate. For this context, it can be generalized that, although the target location is perceived to be compressed at depths of 90 and 120 cm, primarily when direct pointing conditions are used, the perceived 150 cm space is more precise for the two techniques used. Therefore, the depth at which a target is projected within the stereoscopic display's negative parallax seems to affect the perception of frontal space. However, further statistical analysis is needed to support these current findings.

The reasons for this problem are still debatable, but the researchers agree in part that the longer the target is displayed, the stronger the impression will be obtained. The term accommodation-converge mismatch can explain this phenomenon [26]. If the stereoscopic object is displayed closer to the eye, a significant mismatch (conflict) 'accommodation-converge' appears compared to if the object is displayed at a further distance. Also, it is ambiguous if the respective figures rely on the depth of information provided by different perceptual modalities [27]. Another consideration is the lighting supporting in the VE, there are no different results for the perceptual accuracy between real and VE while the luminous is controllable [28].

IV. CONCLUSION

The research decides to investigate frontal experiences in the stereoscopic context using two direct-point and indirect-cursor interaction techniques. Three different target depths have been used to view the target in the negative parallax. The results reveal that the perception of space is found to be compressed in the frontal stereoscopic environment. It is indicated by the pointing estimations, which are concentrated to the center concerning the target references. It is also found that choosing a simulated target using an indirect cursor is more precise (in terms of how close the target-pointing calculation) than the technique in the direct pointing.

The research has limitations for the under three distinct depths. First, the experiment only considers the frontal extent. Therefore, the other planes, such as sagittal and transverse, should be further investigated. Second, a statistical analysis is needed to support the findings. Finally, a specific design is used for the equipment, display, and mission. Hence, more analysis across various experimental designs and configurations will be required to underpin the generality of the outcome.

REFERENCES

[1] F. El Jamiy and R. Marsh, "Distance estimation in virtual reality and augmented reality: A survey," in *2019 IEEE International Conference on Electro Information Technology (EIT)*. Brookings, SD, USA: IEEE, May 20–22, 2019, pp. 063–068.

[2] J. Mittelstaedt, J. Wacker, and D. Stelling, "Effects of display type and motion control on cybersickness in a virtual bike simulator," *Displays*, vol. 51, pp. 43–50, 2018.

[3] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson, "Virtual Reality Induced Symptoms and Effects (VRISE): Comparison of Head Mounted Display (HMD), desktop and projection display systems," *Displays*, vol. 29, no. 2, pp. 58–69, 2008.

[4] J. Chen and C. Or, "Assessing the use of immersive virtual reality, mouse and touchscreen in pointing and dragging-and-dropping tasks among young, middle-aged and older adults," *Applied Ergonomics*, vol. 65, pp. 437–448, 2017.

[5] M. R. Mine, "Virtual environment interaction techniques," University of North Carolina at Chapel Hill, Tech. Rep., 1995.

[6] C. J. Lin, B. T. Abreham, and B. H. Woldegiorgis, "Effects of displays on a direct reaching task: A comparative study of head mounted display and stereoscopic widescreen display," *International Journal of Industrial Ergonomics*, vol. 72, pp. 372–379, 2019.

[7] C. J. Lin, D. Caesaron, and B. H. Woldegiorgis, "The accuracy of the frontal extent in stereoscopic environments: A comparison of direct selection and virtual cursor techniques," *PLOS ONE*, vol. 14, no. 9, pp. 1–20, 2019.

[8] J. M. Loomis and J. M. Knapp, "Visual perception of egocentric distance in real and virtual environments," in *Virtual and adaptive environments: Applications, implications, and human performance issues*, L. J. Hettinger and M. W. Haas, Eds. CRC Press, 2003, pp. 21–46.

[9] E. A. Maguire, N. Burgess, and J. O'Keefe, "Human spatial navigation: Cognitive maps, sexual dimorphism, and neural substrates," *Current Opinion in Neurobiology*, vol. 9, no. 2, pp. 171–177, 1999.

[10] C. Armbrüster, M. Wolter, T. Kuhlen, W. Spijkers, and B. Fimm, "Depth perception in virtual reality: Distance estimations in peri-and extrapersonal space," *Cyberpsychology & Behavior*, vol. 11, no. 1, pp. 9–15, 2008.

[11] M. Gonzalez-Franco, P. Abtahi, and A. Steed, "Individual differences in embodied distance estimation in virtual reality," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. Osaka, Japan: IEEE, March 23–27, 2019, pp. 941–943.

[12] C. J. Lin, L. Y. Cheng, and M. C. Wang, "Performance of estimating depth in projection based stereoscopic virtual display," *Journal of the Society for Information Display*, vol. 23, no. 2, pp. 76–83, 2015.

[13] C. Valzolgher, M. Alzhaler, E. Gessa, M. Todeschini, P. Nieto, G. Verdelet, R. Salemme, V. Gaveau, M. Marx, E. Truy *et al.*, "The impact of a visual spatial frame on real sound-source localization in virtual reality," *Current Research in Behavioral Sciences*, vol. 1, pp. 1–10, 2020.

[14] W. B. Thompson, P. Willemsen, A. A. Gooch, S. H. Creem-Regehr, J. M. Loomis, and A. C. Beall, "Does the quality of the computer graphics matter when judging distances in visually immersive environments?" *Presence: Teleoperators & Virtual Environments*, vol. 13, no. 5, pp. 560–571, 2004.

[15] M. Shin, S. Lee, S. W. Song, and D. Chung, "Enhancement of perceived body ownership in virtual reality-based teleoperation may backfire in the execution of high-risk tasks," *Computers in Human Behavior*, vol. 115, pp. 1–11, 2021.

[16] ISO, "ISO 9241-9:2000: Ergonomic requirements for office work with Visual Display Terminals (VDTs)–part 9: Requirements for non-keyboard input devices," 2000. [Online]. Available: https://www.iso.org/standard/30030.html

[17] H. M. Sun, S. P. Li, Y. Q. Zhu, and B. Hsiao, "The effect of user's perceived presence and promotion focus on usability for interacting in virtual environments," *Applied Ergonomics*, vol. 50, pp. 126–132, 2015.

[18] C. J. Lin, B. T. Abreham, D. Caesaron, and B. H. Woldegiorgis, "Exocentric distance judgment and accuracy of head-mounted and stereoscopic widescreen displays in frontal planes," *Applied Sciences*, vol. 10, no. 4, pp. 1–16, 2020.

[19] G. Gorisse, O. Christmann, E. A. Amato, and S. Richir, "First-and third-person perspectives in immersive virtual environments: Presence and performance analysis of embodied users," *Frontiers in Robotics and AI*, vol. 4, pp. 1–12, 2017.

[20] J. R. Berard, J. Fung, B. J. McFadyen, and A. Lamontagne, "Aging affects the ability to use optic flow in the control of heading during locomotion," *Experimental Brain Research*, vol. 194, no. 2, pp. 183–190, 2009.

[21] W. Thompson, R. Fleming, S. Creem-Regehr, and J. K. Stefanucci, *Visual perception from a computer graphics perspective*. CRC press, 2011.

[22] G. Bruder, F. Steinicke, and W. Sturzlinger, "To touch or not to touch? Comparing 2D touch and 3D mid-air interaction on stereoscopic tabletop surfaces," in *Proceedings of the 1st Symposium on Spatial User Interaction*, 2013, pp. 9–16.

[23] G. Bruder, F. Steinicke, and W. Stürzlinger, "Effects of visual conflicts on 3D selection task performance in stereoscopic display environments,"

in *2013 IEEE Symposium on 3D User Interfaces (3DUI)*. Orlando, FL, USA: IEEE, March 16–17, 2013, pp. 115–118.

[24] J. W. Kelly, W. Hammel, L. A. Sjolund, and Z. D. Siegel, "Frontal extents in virtual environments are not immune to underperception," *Attention, Perception, & Psychophysics*, vol. 77, no. 6, pp. 1848–1853, 2015.

[25] C. J. Lin and B. H. Woldegiorgis, "Egocentric distance perception and performance of direct pointing in stereoscopic displays," *Applied Ergonomics*, vol. 64, pp. 66–74, 2017.

[26] P. J. Werkhoven and J. Groen, "Manipulation performance in interactive virtual environments," *Human Factors*, vol. 40, no. 3, pp. 432–442, 1998.

[27] P. Maruhn, S. Schneider, and K. Bengler, "Measuring egocentric distance perception in virtual reality: Influence of methodologies, locomotion and translation gains," *PLOS ONE*, vol. 14, no. 10, pp. 1–24, 2019.

[28] K. Chamilothori, J. Wienold, and M. Andersen, "Adequacy of immersive virtual reality for the perception of daylit spaces: Comparison of real and virtual environments," *Leukos*, vol. 15, no. 2-3, pp. 203–226, 2019.