Evaluation of 3D Pointing Accuracy in the Fovea and Periphery in Immersive Head-Mounted Display Environments

Nikolaos Katzakis, Lihan Chen, Oscar Ariza, Robert J. Teather, and Frank Steinicke,

Abstract—The coupling between perception and action has seldom been explored in sophisticated motor behaviour such as 3D pointing. In this study, we investigated how 3D pointing accuracy, measured by a depth estimation task, could be affected by the target appearing in different visual eccentricities. Specifically, we manipulated the visual eccentricity of the target and its depth in virtual reality. Participants wore a head-mounted-display with an integrated eye-tracker and docked a cursor into a target. We adopted a within-participants factorial design with three variables. The first variable is *Eccentricity*: the location of the target on one of five horizontal eccentricities (left far periphery, left near periphery, foveal, right near periphery and right far periphery). The second variable is *Depth* at three levels and the third variable is Feedback Loop with two levels: open/closed. *Eccentricity* is refactored into *Motion Correspondence* between the starting location of the cursor and the target location with four levels: periphery to fovea, fovea to periphery, periphery to periphery, fovea to fovea. The results showed that the pointing accuracy is modulated mainly by the target locations rather than the initial locations of the effector (hand). Visible feedback during pointing improved performance.

Index Terms—Virtual Reality, Vision, Periphery, 3D, Pointing, Visual Feedback.

◆

1 INTRODUCTION

Daily actions such as reaching and grasping require coordinated control of both hand and eye movements. In this context, humans need to process the dynamically changing visual inputs, and often integrate information from other sensory modalities such as audition, somatosensation (proprioception, touch) etc. Previous studies have extensively explored reaching behavior in both real and virtual worlds, in which human observers usually fixate their gaze on the target. However, it is often the case that the targets for pointing and reaching are located in the periphery and not in the focus of gaze. To name a few, typing characters on a keyboard, using the mouse while browsing the web, opening a door and even reaching for food, human operators can complete the task without explicitly focusing on the targets, or with only a brief glimpse of the targets in the periphery. This vein of typical behaviors has also been observed in virtual reality (VR). In a virtual environment, given a certain amount of familiarization/training, humans behave in an automatic manner when pressing buttons on a controller, selecting items from a floating virtual menu and even manipulating 3D objects in space dexterously. All of the above actions demand little, if any, direct fixation on the specific object we are interacting with. Nevertheless, during the coupling of perception-action, we exploit additional multisensory cues (such as haptics and proprioception) to streamline the interactions. The sensory cues from auditory or tactile modalities could be weighted and integrated with the visual information to facilitate the action. [1], [2], [3].

A large body of vision research has focused on investigating the perceptual detection and discrimination of targets with the dichotomy of foveal vs. peripheral visual acuity. However, less is known about performance of interaction in the different areas of the peripersonal space, i. e., the space within arms reach. The peripersonal visual space comprises of a combination of central fovea area, parafoveal area and near peripheral and far peripheral areas. The central fovea is the region with the highest visual acuity and highest resolution. It plays a crucial role for several tasks such as reading or visual search where we need to localize specific objects. Peripheral vision takes place outside the parafoveal range. Our visual acuity decreases dramatically in the far periphery (above 30 degrees). However, the peripheral vision is still very important for the identification and recognition of well-known shapes, forms and even movements during action. [4].

Reaching tasks in VR in which the user is allowed to make saccades and track the target object into the fovea have been extensively explored [5], [6], [7]. However, as described above, there is a lack of understanding about how visuo-motor coordination works when interactions are primarily performed in the periphery, i. e., when users are not focusing on the objects; in both real-world and VR scenarios. To the best of our knowledge, two pieces of empirical research have investigated the role of peripheral stimulation in VE. Jones et al. [8] studied the perception of distance and spatial scale in a VE. Siderov et al. [9] studied stereoscopic depth discrimination thresholds as a function of spatial frequencies (low vs high) and retinal eccentricity (less than or larger than 10 degrees). Results indicate that distance judgments in virtual environments might be considerably similar to those in real world. Above 10 degrees, the stereoscopic depth discrimination thresholds for the high spatial frequency stimuli increase with eccentricity. With that said, depth estimation during 3D pointing has not been rigorously tested in VR. Until very recently, eye-tracker devices could not be easily integrated and used in typical VR display conditions such as head-mounted display (HMDs) or with stereoscopic glasses. Advancements in this area allow us to combine both technologies to deal with the scientific question of depth estimation along different visual eccentricities action in VR [10].

In this study, we explore how pointing accuracy and depth estimation is affected by location of the target in the periphery (eccentricity) during interaction. We conducted an experiment in which users were equipped with a HMD with an embedded eye tracking device. The participant's task was to dock a cursor into a target in a 3D VR environment. Across the trials, we varied visual eccentricities, depths and cursor visibility (feedback loop). Participants were constrained to always fixate their gaze on a central cross during the interaction. We evaluated their pointing performance and potential perceptual biases (overestimation or underestimation of depth) during the experiment.

2 RELATED WORK

Vision research [4], [11], [12], [13], [14], [15], [16], [17] has identified how visual acuity and stereoscopy degrades as a function of retinal eccentricity. Among the large body of evidence in vision research, the basic task paradigm is visual selection of the target. It is yet not known whether these findings could be transferred to virtual reality.

Selection is also a fundamental task in VR systems, much like in other user interfaces (UIs, e.g., the desktop). Unlike desktop interfaces, where selection is accomplished by a mouse-controlled/tracking cursor, selection in VR mostly mobilizes the entire arm, either to reach and grasp objects (e.g., with virtual hands) or to remotely point at them (e.g., ray-casting) [18], [19]. Virtual hand (and ray-based) selection techniques are often referred to as pointing tasks, as they specify a unique point or object in the environment, often preceding subsequent operations like object manipulation or travel [20]. Numerous 3D pointing studies have been conducted (see e.g., [5], [6], [21], [22], [23], [24]), yet none has specifically looked at selection of targets in peripheral vision.

Since VR systems typically co-locate the input and display spaces, selection performance has been influenced by several visual cues and sensory feedback mechanisms [25]. For example, early work investigated the influence of factors like stereo and head-tracking [21], [26]. Results in this line of work revealed that movement in depth is slower and less precise than movements in a plane parallel to the screen.

Other work [27], [28], [29] explored near field egocentric distance estimation with results indicating underestimation. In contrast, depth estimation within reaching distances in early literature [30], [31] found overestimation, while more recent work found a sort of *central tendency* effect in AR [32]. There is, it seems, a lack of consensus between published works and therefore more experiments are needed. In addition, none of the previous work has specifically distinguished between depth perception in the central fovea and the periphery, nor the possible interaction when pointing/motor actions happen across these areas.

Other work has compared pointing performance in real world and in VR. This was based on the observation that performance differences are likely attributable to perceptual differences between physical reality and VR displays [22], [23]. Notably, selection in VR systems suffers from various stereoscopic perceptual issues, including the well-known stereo convergence-accommodation mismatch, and double-vision due to targets and cursors residing in different depth planes [5]. These results, however, reinforce the importance of vision in selection tasks. Despite most evidence indicating that real-world selection performance is indeed superior to the one in VR [22], such studies cannot account for the findings to vision alone, since several other technical factors also contribute to the degradation in performance in VR. These factors include input latency, tracker noise [33], tracker registration [34], and tactile feedback during performance [35]. Nevertheless, visual information remains especially important, since the largest differences between selections in real and virtual worlds occurred during the correction phase of the motion, where visual feedback was involved in a tight feedback loop and hence improved the accuracy of the performance [22], [23], [24]. Overall, the cited studies indicate that participants experience difficulty when selecting targets outside of their foveal vision or when crossing from one region to another, however, performance remained consistent and was even enhanced when they acted in a closed loop with sufficient sensory feedback and within a small range of retinal eccentricities (below 10 degrees).

3 EXPERIMENT

The goal of our experiment was to understand the interplay between visual eccentricity of the targets and pointing accuracy. Participants were equipped with a HMD with integrated eye-tracker and had to dock a cursor into a target. The study was approved by the Ethics committee of Hamburg University.

3.1 Participants

Eighteen volunteers (age 22-38 years old, M=28.5) participated in the experiment. Most of the participants were students or staff members from the local department. All participants had normal or corrected to normal vision. Three of them wore glasses during the experiments. None of the participants suffered from a disorder of equilibrium. One of our participants reported strong eye dominance. No other vision disorders have been reported. Interpupillary distance was not individually adjusted. The experience of the participants with 3D stereoscopic displays (cinema, games etc.) was screened prior to the experiment. The mean rating score was 3.1, with a range of 1 (no experience) to 5 (much experience). Moreover, all participants had certain experience using head-mounted displays (HMDs) before.

3.2 Apparatus

Participants stood, wearing an HTC Vive head mounted display and held the wand with their dominant hand (Figure 1a). Their non-dominant arm was resting on a Razer Orbweaver keypad. The thumb button on the keypad was used to confirm a response and initiate the next trials. Inside the HMD, a Pupil Labs HTC Vive Binocular Add-on was mounted for eye tracking[1]. In the closed loop trials

---

[1] The pupil labs addon is equipped with a camera per eye, attachment rings with IR illuminators and tracking software stack. This results in 30 Hz gaze tracking.

participants could see a spherical cursor of 4 cm diameter, always co-located at the trigger position of the HTC Vive wand (Figure 1b). In the open loop trials participants could not see the cursor.

The cursor was positioned at the trigger rather than the controller's origin so that participants were essentially pointing with their fingertip and as such have a stronger sense of agency. Since participants used their off-hand for button presses their grip was constant.

3.3 Stimuli and Task

Participants were immersed in a VR environment with only a grid floor visible (Figure 1b). They were presented with a spherical semitransparent target. The target was hinged to the participant's head's frame of reference and thus remained in position even if participants moved their head (Figure 1h). This was implemented to ensure that the view and line of sight of the participants did not change during the task (i.e. no parallax cues). Standing allows them to have a comfortable pose in contrast to other alternatives, such as resting their head on a chin rest for a long time. Before the experiment, individual arm length was obtained and the target depth was setup with three levels: near, middle and far with reference to the arm length [36] (see 3.4 Variables).

Participants were instructed to move and match their cursor to a target sphere of equal diameter (4 cm) and then press the button on their off-hand. They were asked to maintain a balance between speed and accuracy. A top down view of the task is shown in Figure 1. As soon as they pressed the button, the behavioral data was recorded and the target jumped to the next position. An auditory stimulus was given to signal the button press. Target sizes were in consensus with previous pointing experiments [5], [6]. Each trial was given according to the condition by a combination of two positions, the start position of acting hand and an end position of the target.

In every trial there was a start area and an end area, one in the fovea and one in the periphery, both in the fovea or both in the periphery cf. Figure 1. All trials were randomized. Therefore, participants were tested for the four areas of interest: fovea-to-periphery, fovea-to-fovea, periphery-to-periphery, periphery-to-fovea.

The fixation point was located in the center of the field of view (FOV) and at the Z axis midpoint between all given target depths (Black X in Figure 1g). If the participant moved his or her gaze away from the fixation point during a trial (limited by 2 degree threshold), this trial was skipped.

In pilot tests, when participants moved their gaze away from the fixation point we simply made the target disappear and re-started the trial, forcing participants to move their cursor back to the start position of the trial and repeat the trial. However, we found that even if we forced participants to go back to the starting point, since they had saccaded and fixated on the target a few milliseconds ago, the target location was fresh in their memory. This made the trial easier. To avoid that, "failed" trials due to saccading were skipped and re-shuffled back into the remaining trials queue, to be performed later.

In early pilots of the experiment, the fixation point would move along the Z axis to always match the depth of the current target. i.e. the target would always be on the horopter[2]. This was attempted to avoid the problem of diplopia (i.e. seeing double cursors) when the fixation point is at a different depth than the manipulated cursor/target. Nevertheless, when the fixation point jumped to a different depth at each trial, despite staying at the center of the FOV, it triggered an 'uncontrolled' attentional process driven by a visual search to find the fixation point again. This mode of visual search was both frustrating and exhausting for participants, while at the same time it triggered the eye-tracker to mark the trial as invalid and shuffle it back in for later. Therefore we opted to lock the fixation point's depth.

Participants were encouraged to take a break whenever they wanted. Overall, each participant completed 132 trials (black arrows in Figure 1g x 3 repetitions) with a total of 2376 trials among all participants.

The process for each participant included pre-questionnaires, instruction, ten training trials, experiment and post-online-questionnaires. The participants needed to wear the HMD for approximately 20 minutes. The total duration for each participant was approximately 30 minutes. All participants reported, in post questionnaires, a high level of attentional engagement throughout the experiment.

3.4 Variables

We manipulated three variables in our experiment: (i) *Eccentricity*, (ii) *Depth* of the target, and (iii) *Feedback loop*, i.e. visibility of the user's cursor (open/closed loop).

Eccentricity refers to the angle, around a vertical axis centered between the eyes, of the target with respect to the center of the fovea. Eccentricity was manipulated with 10 degree increments at -20,-10,0,10,20 degrees. The fovea accounts for approximately $4 \sim 5^\circ$ in our central field of view (FOV) [13], [37]. Therefore we tested 0 degree, where the fovea is at its highest viusal acuity, 10 degrees i.e. the boundary of perifovea, and 20 degrees i.e. well within the periphery.

Depth refers to the depth of the target as a percentage of arm length. If α denotes the arm length of the participant, previous research [36] suggests that mid-air pointing performance works best at depths $0.618 \cdot \alpha$ followed by $0.854 \cdot \alpha$, we extrapolated the two depths in the direction of the eyes to obtain $0.382 \cdot \alpha$ as the shallowest target depth (i.e. $0.618 - (0.854 - 0.618) \cdot \alpha$).

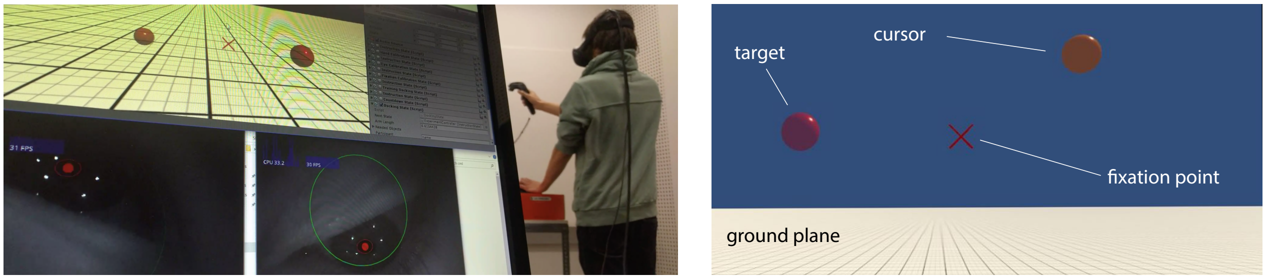
Feedback Loop refers to open/closed loop with regards to visual feedback. Open loop means participants did not see their cursor whereas closed loop means normal cursor display. Loop only affects the participant's cursor, not target visibility. The target was always visible.

Motion Correspondence is a refactoring based on the eccentricity of the start and end points. There were four levels: periphery to periphery, fovea to fovea, fovea to periphery and periphery to fovea.

We measured *total error* as the euclidean distance between the cursor and the target at the time the participant pressed the button $e_{total} = \sqrt{p_{target}^2 - p_{cursor}^2}$. The lower *total error* is, the closer the cursor was to the target when

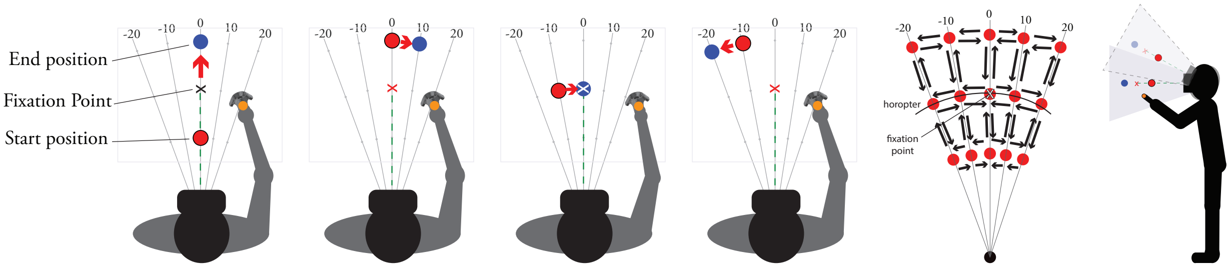
---

[2] a line or surface containing all those points in a space of which images fall on corresponding points of the retinas of the two eyes.



(a) Picture from the experiment showing a user with an HTC Vive head-mounted display (HMD) and motion controller, and the user's view of the virtual environment (VE) as well as the segmented camera images of the integrated eye tracking device.

(b) Screenshot of the experiment. On the left, in red, the fixed size target that needed to be matched. On the right side, in orange, the participant's cursor located at his or her fingertip.



(c) Fovea to Fovea: Participants started the trial by docking into the *start position* and then moving to dock into the *end position*. The eye tracker ensured the participant's gaze remained on the fixation point.

(d) Fovea to Periphery: The start position was in the fovea region but the end position was in the periphery.

(e) Periphery to Fovea: Starting in the periphery (10 or 20 degrees eccentricity) and ending in the fovea region.

(f) Periphery to Periphery: Starting position was in the periphery and ending position was also in the periphery.

(g) Start and end position were limited to neighboring areas. i.e. There was never a start in the left periphery that ended in the right periphery etc. as seen in the illustration.

(h) The targets were always anchored to the HMD and moved with it.

Fig. 1: Setup of the Experiment.

the button was pressed. We also measured *depth error* as the depth component of the aforementioned euclidean distance (i.e. only Z-axis error $e_{depth} = p_{cursor}.z - p_{target}.z$.). A positive depth error means participants over-estimated the depth while a negative one means underestimation.

4 RESULTS

We attempted to fit [38] the *eccentricity* on Total Error with a standard quadratic model $\varepsilon = \varphi^2$ (ε stands for error and φ stands for angle). Fitting this standard quadratic formula resulted in a good fit ($F_{2,2445} = 46.45, p < 0.001, r^2 = 0.88$) cf. Figure 2 middle and table 1.

Angle	Error	SD	SE
-20	3.11 cm	0.0246	0.001300
-10	2.38 cm	0.0186	0.000777
0	2.31 cm	0.0192	0.000801
10	2.43 cm	0.0185	0.000773
20	3.42 cm	0.0264	0.001392

TABLE 1: Results for eccentricity effect on error.

We had five levels of visual *eccentricity*. We separated the repeated measures analysis of variance (ANOVA) into "start eccentricity" and "end eccentricity". For the *depth error*, the main effect of visual eccentricity at the start location was not significant, $F_{4,68} = 1.161$, *p*=0.336. However, the main effect was borderline significant at the end location, $F_{4,68} = 2.337$, *p*=0.064. Bonferroni corrected comparison showed the error was larger in the center (1.56 cm) than the one in left 10 degree of periphery (1.14 cm), *p*=0.038. This means that visual eccentricity of the end position for the targets is critical for pointing performance in this experiment.

There was a significant effect of target *depth* (near, middle and far) on *total error* ($F_{1,17} = 6.47, p = 0.02, \eta^2 = 0.27$). Results can be seen in Figure 3. Throughout the experiments, participants generally overestimated the depth of the target while the least overestimation was observed near the fixation point. The main effect of depth on *depth error* was not significant ($F_{2,34} = 2.373, p = 0.109, \eta^2 = 0.122$). Therefore, in the current task, although the depth of the target affected *total error*, it did not affect Z-axis over/under estimation.

Feedback Loop had a significant effect on depth error ($F_{1,17} = 9.416, p = 0.007, \eta^2 = 0.356$) . Closed loop manipulation increased 3D pointing performance and reduced both types of error. An aggregate plot can be seen in Figure 2 right.

We additionally explored the effect of motion direction (i.e. forward or sideways) on our dependent variables, *total error* and *depth error*. We grouped the differences of angles between starting points and end points, there were three conditions: "startless10 " (start angle was 10 degrees smaller than the one for end angle), "equal" and "startlarger10" (start angle is 10 degree larger than the one for end angle). The main effect of motion correspondence was not significant, $F_{2,34} = 1.729$, *p*=0.193. Therefore, the moving

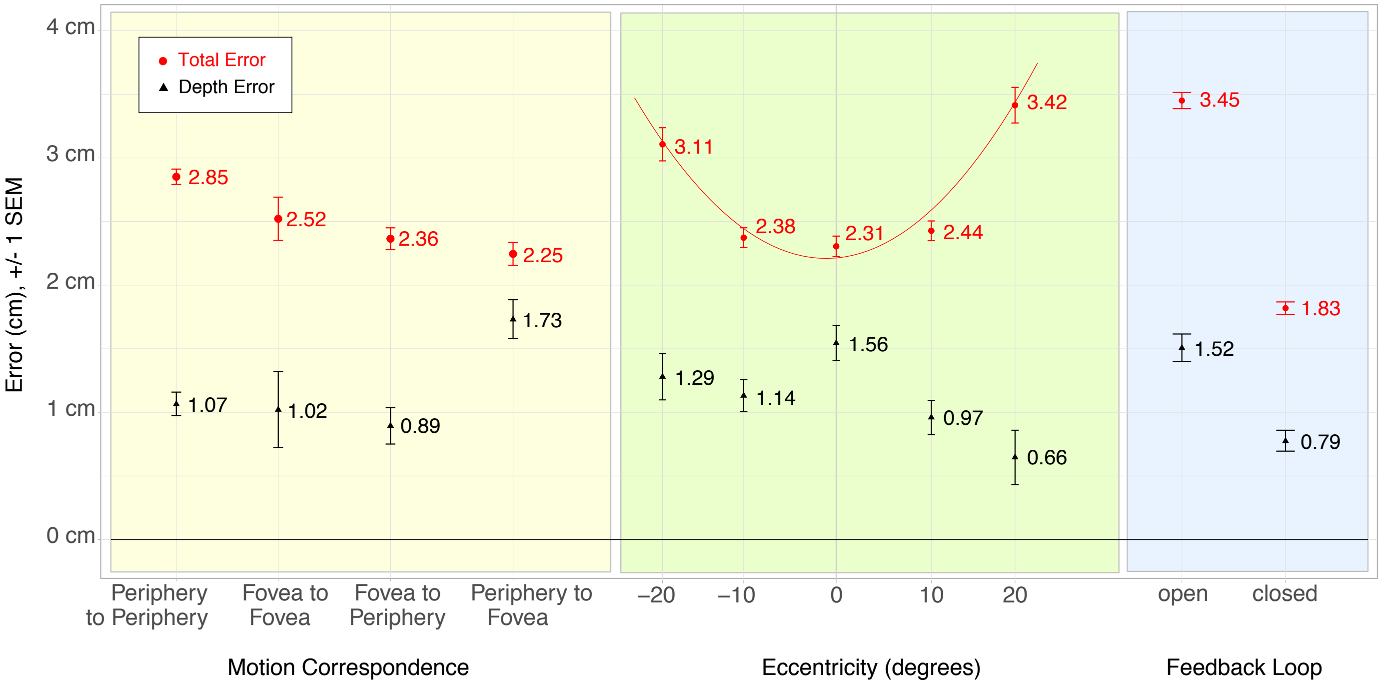


Fig. 2: Depth error and Total Error with regards to *motion correspondence*, *eccentricity*, *feedback loop*. In *Total Error*, sign is unimportant but with *Depth Error*, positive values mean over-estimation (cursor was placed *behind* the target. The vast majority of trials were over-estimated.

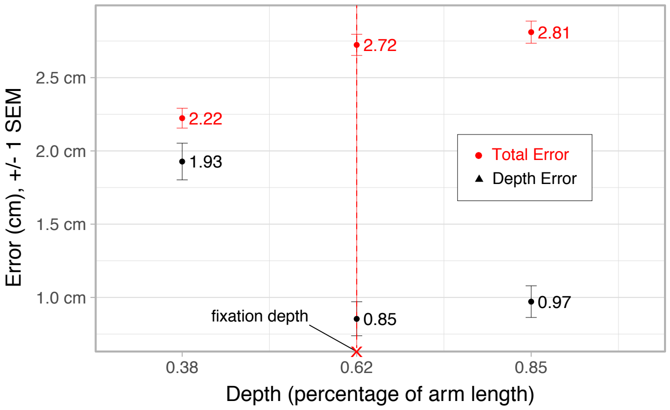


Fig. 3: *Depth* effect on Total Error and Depth Error.

trajectory made no difference, this indicates a good control for the current testing conditions.

5 DISCUSSION

This study contributes three main findings: First, for a 3D pointing task within arm's reach in VR, human observers overestimated the perceived depth of the target. This overestimation was robust across all experimental conditions. In contrast, some previous studies have shown that egocentric depth perception tends to be underestimated in VR, especially over walkable distances; objects are perceived smaller and closer than they should [28], [39]. Contrary to these studies, depth (distance) in this work was always within arm's reach. Depth overestimation within reaching distances, has also been reported in early literature [30], [31]. However, none of these previous works has specifically distinguished between depth perception in the central fovea and the periphery, and the interaction between both.

Imperceptible depth cues [40] could be responsible for the discrepancies observed in this experiment. Since targets were anchored to the HMD there were no cues from motion parallax and shadows on the ground were outside the field of view of the observers. The specular highlight from the virtual light on the cursor and target shifted minimally within this short distance. The question remains, however, why did participants have the largest depth overestimation in the periphery to fovea condition? We postulate that the nature of the task could be partially responsible for this. Consider the following:

When starting and finishing a trial in the foveal region, participants are moving the cursor along the depth axis. They slide their cursor up and down along this depth axis until they are satisfied that their cursor is inside the target and is not occluding it. When a movement begins in the periphery, however, participants are moving their cursor in an arc-like trajectory tangentially to the horopter. Upon finishing that movement, even if their cursor is slightly behind the target, they don't take the extra effort to switch and correct along the depth axis. It is likely that they leave their cursor where it landed, and that turned out to be on average 1.73 cm behind the target. This begs the question, however, why did was the same not occur on the fovea to periphery trials? A possible explanation is that when starting in the fovea participants have a better estimation of the target depth and are able to perform a more precise arc-like motion to reach the target. When starting in the periphery, the start position of the trial already presents a challenge and therefore the arc-like motion to come to the fovea results in further offset [30], [41].

The second finding is that participants performed better when they pointed to a target appearing in the fovea area than they did in the periphery area (Figure 2 left). Vision research has addressed well a linear degradation of acuity with distance [42], [43]. The degradation we are seeing in this experiment as a function of target depth suggests that the fixation point depth has a great influence in matching the target depth, but not for overall pointing accuracy (total error). Interestingly, we found that *depth error* (the perceptual bias) was smaller near the fixation point (such as left 10 degree of eccentricity) than the one on the sharp central fovea (0 degree). This is possibly due to visual overlapping between the 'target' and 'fixation mark', while the left visual bias of viewing and the reduced overlapping render the performance being better in the periphery near the central fovea [44], [45], [46]. VR interface designers could use this information to layout elements on the horopter for best depth accuracy when depth performance is important.

The third main finding is the role of sensory feedback during the pointing. Cursor visibility (*Feedback loop*) manipulation also shows that the importance of visual feedback, that has been shown with 3D pointing tasks [25], [47], [48] also holds in the periphery. Surprisingly, lack of visual feedback only caused a *total error* of 1,62 cm with a *depth error* of 0,73 cm. i.e. Participants were able to match the targets quite well despite not seeing their own cursor. In addition to this, cursor invisibility accounts for the increased depth overestimation in the fovea (Figure 2 Eccentricity, 0 degree column). i.e. Motions that ended up in the fovea were much less accurate when the cursor was invisible.

Actions in the periphery are typically preceded by a saccade and a reach/grasping planning sequence in the brain. By depriving this saccade from participants they are left without the planning component and can only rely on the cues at hand that, at certain depths, are affected by diplopia. It remains unclear how the results would change if the fixation point changed its depth to match the depth of the target.

The cursors had 0.5 alpha opacity with one specular highlight from the scene lighting. Colour occlusion (orange sphere in front of red sphere) might have been used as a strategy by participants. i.e. participants could have been "trying" various depths until they see the orange colour of the cursor covering the red target. We postulate that this strategy was most probably not used because the majority of trials were overestimated.

A potential improvement to this study might include first measuring interpupillary distance and eccentricity limits of each eye (nasal occlusion). Given these limits it might make sense to assume that stereoscopy will start degrading rapidly past those limits because the brain can only depend on monocular cues. Finally, in this study we only investigated targets horizontally across the horopter and it remains to be seen how these results change when targets are laid out vertically.

6 CONCLUSION

We presented the first controlled study of 3D pointing in the periphery in virtual reality. The empirical evidence suggests that human observers still maintain the high accuracy of 3D pointing to the target within the arm reach in the VR environment, though with general biases of depth overestimation in the current task. The visibility of the visual feedback during the perception-action coupling loop is critical for precise interaction, especially when gaze-following behavior could be largely constrained in a number of daily scenarios. Movements starting in the periphery landing in the foveal region exhibited significantly more depth overestimation than all other motion correspondence cases.

In ongoing studies, we include additional forms of open/closed loop in other sensory modalities (such as no force feedback vs. force feedback), and implement 3D pointing in a desktop version with different yet controlled action trajectories, to test the generalization of these findings.

ACKNOWLEDGMENTS

This work was partially funded by the German Research Foundation (DFG) and the National Science Foundation of China (NSFC) in project Crossmodal Learning, TRR-169.

REFERENCES

[1] G. M. Goodwin, D. I. McCloskey, and P. B. Matthews, "Proprioceptive illusions induced by muscle vibration: contribution by muscle spindles to perception?" *Science*, vol. 175, no. 4028, pp. 1382–1384, 1972.

[2] M. Gonzalez-Franco, D. Perez-Marcos, B. Spanlang, and M. Slater, "The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment," in *Virtual Reality Conference (VR), 2010 IEEE*. IEEE, 2010, pp. 111–114.

[3] N. Katzakis, J. Tong, O. Ariza, L. Chen, G. Klinker, B. Röder, and F. Steinicke, "Stylo and handifact: Modulating haptic perception through visualizations for posture training in augmented reality," in *Proceedings of the 5th Symposium on Spatial User Interaction*, ser. SUI '17. Brighton, UK: ACM, 2017, pp. 58–67. [Online]. Available: http://doi.acm.org/10.1145/3131277.3132181

[4] H. Strasburger, I. Rentschler, and M. Jüttner, "Peripheral vision and pattern recognition: A review," *Journal of vision*, vol. 11, no. 5, pp. 13–13, 2011.

[5] R. J. Teather and W. Stuerzlinger, "Pointing at 3D Target Projections with One-Eyed and Stereo Cursors," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 159–168.

[6] R. J. Teather and W. Stuerzlinger, "Pointing at 3D Targets in a Stereo Head-tracked Virtual Environment," in *3D User Interfaces (3DUI), 2011 IEEE Symposium on*. IEEE, 2011, pp. 87–94.

[7] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa, "The go-go interaction technique: non-linear mapping for direct manipulation in vr," in *Proceedings of the 9th annual ACM symposium on User interface software and technology*. ACM, 1996, pp. 79–80.

[8] J. A. Jones, J. E. Swan II, and M. Bolas, "Peripheral stimulation and its effect on perceived spatial scale in virtual environments," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 4, pp. 701–710, April 2013.

[9] J. Siderov and R. S. Harwerth, "Stereopsis, spatial frequency and retinal eccentricity," *Vision Research*, vol. 35, no. 16, pp. 2329 – 2337, 1995. [Online]. Available: http://www.sciencedirect.com/science/article/pii/0042698994003078

[10] O. H.-M. Lutz, C. Burmeister, L. F. dos Santos, N. Morkisch, C. Dohle, and J. Krüger, "Application of head-mounted devices with eye-tracking in virtual reality therapy," *Current Directions in Biomedical Engineering*, vol. 3, no. 1, pp. 53–56, 2017.

[11] S. Gould, J. Arfvidsson, A. Kaehler, B. Sapp, M. Messner, G. R. Bradski, P. Baumstarck, S. Chung, A. Y. Ng *et al.*, "Peripheral-foveal vision for real-time object recognition and tracking in video." in *IJCAI*, vol. 7, 2007, pp. 2115–2121.

[12] K. Latham and D. Whitaker, "Relative roles of resolution and spatial interference in foveal and peripheral vision," *Ophthalmic and Physiological Optics*, vol. 16, no. 1, pp. 49–57, 1996.

[13] L. Loschky, G. McConkie, J. Yang, and M. Miller, "The limits of visual resolution in natural scene viewing," *Visual Cognition*, vol. 12, no. 6, pp. 1057–1092, 2005.

[14] M. B. McCamy, J. Otero-Millan, S. L. Macknik, Y. Yang, X. G. Troncoso, S. M. Baer, S. M. Crook, and S. Martinez-Conde, "Microsaccadic efficacy and contribution to foveal and peripheral vision," *Journal of Neuroscience*, vol. 32, no. 27, pp. 9194–9204, 2012.

[15] R. Rosenholtz, "Capabilities and limitations of peripheral vision," *Annual Review of Vision Science*, vol. 2, pp. 437–457, 2016.

[16] D. Whitaker, K. Latham, P. Mäkelä, and J. Rovamo, "Detection and discrimination of curvature in foveal and peripheral vision," *Vision Research*, vol. 33, no. 16, pp. 2215–2224, 1993.

[17] J. Xing and D. J. Heeger, "Center-surround interactions in foveal and peripheral vision," *Vision research*, vol. 40, no. 22, pp. 3065–3072, 2000.

[18] F. Argelaguet and C. Andujar, "A Survey of 3D Object Selection Techniques for Virtual Environments," *Computers & Graphics*, vol. 37, no. 3, pp. 121–136, 2013.

[19] D. Bowman, E. Kruijff, J. J. LaViola Jr, and I. P. Poupyrev, *3D User Interfaces: Theory and Practice, CourseSmart eTextbook*. Addison-Wesley, 2004.

[20] D. A. Bowman, D. B. Johnson, and L. F. Hodges, "Testbed evaluation of virtual environment interaction techniques," *Presence: Teleoperators and Virtual Environments*, vol. 10, no. 1, pp. 75–95, 2001.

[21] J. Boritz and K. S. Booth, "A study of interactive 3d point location in a computer simulated virtual environment," in *Proceedings of the ACM symposium on Virtual reality software and technology*. ACM, 1997, pp. 181–187.

[22] L. Liu, R. van Liere, C. Nieuwenhuizen, and J.-B. Martens, "Comparing aimed movements in the real world and in virtual reality," in *Virtual Reality Conference, 2009. VR 2009. IEEE*. IEEE, 2009, pp. 219–222.

[23] K. Nieuwenhuizen, L. Liu, R. van Liere, and J.-B. Martens, "Insights from dividing 3d goal-directed movements into meaningful phases," *IEEE computer graphics and applications*, vol. 29, no. 6, pp. 44–53, 2009.

[24] O. Ariza, G. Bruder, N. Katzakis, and F. Steinicke, "Analysis of proximity-based multimodal feedback for 3d selection in immersive virtual environments," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, March 2018, pp. 327–334.

[25] R. Teather and W. Stuerzlinger, "Visual aids in 3d point selection experiments," in *Proceedings of the 2Nd ACM Symposium on Spatial User Interaction*, ser. SUI '14. New York, NY, USA: ACM, 2014, pp. 127–136. [Online]. Available: http://doi.acm.org/10.1145/2659766.2659770

[26] J. Boritz and K. S. Booth, "A study of interactive 6 dof docking in a computerised virtual environment," in *Virtual Reality Annual International Symposium, 1998. Proceedings., IEEE 1998*. IEEE, 1998, pp. 139–146.

[27] P. E. Napieralski, B. M. Altenhoff, J. W. Bertrand, L. O. Long, S. V. Babu, C. C. Pagano, J. Kern, and T. A. Davis, "Near-field distance perception in real and virtual environments using both verbal and action responses," *ACM Trans. Appl. Percept.*, vol. 8, no. 3, pp. 18:1–18:19, Aug. 2011. [Online]. Available: http://doi.acm.org/10.1145/2010325.2010328

[28] A. Jones, J. E. Swan, G. Singh, and E. Kolstad, "The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception," in *Virtual Reality Conference, 2008. VR '08. IEEE*, 2008, pp. 9–14.

[29] C. Walter, G. Bruder, B. Bolte, and F. Steinicke, "Evaluation of field of view calibration techniques for head-mounted displays and effects on distance estimation," in *Proceedings of the GI Workshop on Virtual and Augmented Reality (GI VR/AR)*, 2012, pp. 13–24. [Online]. Available: http://basilic.informatik.uni-hamburg.de/Publications/2012/WBBS12

[30] J. M. Foley, "Stereoscopic distance perception," *Spatial Displays and Spatial Instruments*, pp. 9–18, 1989.

[31] J. M. Foley, "Binocular distance perception." *Psychological review*, vol. 87, no. 5, p. 411, 1980.

[32] G. Singh, J. E. Swan, II, J. A. Jones, and S. R. Ellis, "Depth judgment measures and occluding surfaces in near-field augmented reality," in *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, ser. APGV '10. New York, NY, USA: ACM, 2010, pp. 149–156. [Online]. Available: http://doi.acm.org/10.1145/1836248.1836277

[33] R. Teather, A. Pavlovych, W. Stuerzlinger, and I. S. MacKenzie, "Effects of tracking technology, latency, and spatial jitter on object movement," in *3D User Interfaces, 2009. 3DUI 2009. IEEE Symposium on*. IEEE, 2009, pp. 43–50.

[34] D. W. Sprague, B. A. Po, and K. S. Booth, "The importance of accurate vr head registration on skilled motor performance," in *Proceedings of Graphics interface 2006*. Canadian Information Processing Society, 2006, pp. 131–137.

[35] K. Chun, B. Verplank, F. Barbagli, and K. Salisbury, "Evaluating haptics and 3d stereo displays using fitts' law," in *Haptic, Audio and Visual Environments and Their Applications, 2004. HAVE 2004. Proceedings. The 3rd IEEE International Workshop on*. IEEE, 2004, pp. 53–58.

[36] P. Lubos, G. Bruder, O. Ariza, and F. Steinicke, "Touching the sphere: Leveraging joint-centered kinespheres for spatial user interaction," in *Proceedings of the 2016 Symposium on Spatial User Interaction*. ACM, 2016, pp. 13–22.

[37] A. M. Larson and L. C. Loschky, "The contributions of central versus peripheral vision to scene gist recognition," *Journal of Vision*, vol. 9, no. 10, p. 6, 2009. [Online]. Available: +http://dx.doi.org/10.1167/9.10.6

[38] J. M. Chambers, *Statistical Models in S*. Taylor and Francis, 1992.

[39] C. Armbrster, 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, pMID: 18275307. [Online]. Available: https://doi.org/10.1089/cpb.2007.9935

[40] N. Gerig, J. Mayo, K. Baur, F. Wittmann, R. Riener, and P. Wolf, "Missing depth cues in virtual reality decrease performance of three-dimensional reaching movements," in *Proceedings of the 11th International Symposium on Computer Science in Sport (IACSS 2017)*, M. Lames, D. Saupe, and J. Wiemeyer, Eds. Cham: Springer International Publishing, 2018, pp. 113–123.

[41] J. Foley, "Error in visually directed manual pointing," *Perception & Psychophysics*, vol. 17, no. 1, pp. 69–74, 1975.

[42] J. P. Gorham, B. B. Bruce, and A. K. Hutchinson, "Comparing distance visual acuity measurement with a novel eye chart and the landolt c chart in a population of children aged 6–18years," *International Ophthalmology*, vol. 37, no. 6, pp. 1345–1348, Dec 2017. [Online]. Available: https://doi.org/10.1007/s10792-016-0384-3

[43] J. Siderov and R. S. Harwerth, "Stereopsis, spatial frequency and retinal eccentricity," *Vision research*, vol. 35, no. 16, pp. 2329–2337, 1995.

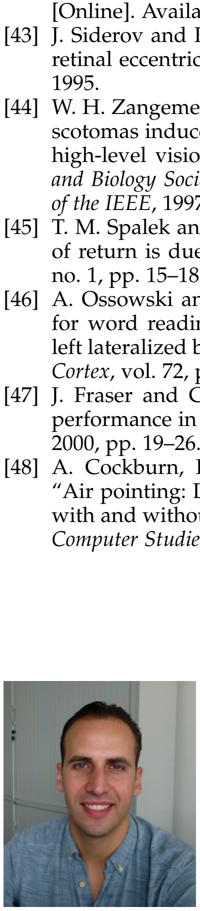
[44] W. H. Zangemeister and U. Oechsner, "Virtual reality hemianopic scotomas induce eccentric fixation scanpath strategies to optimize high-level vision in healthy subjects," in *Engineering in Medicine and Biology Society, 1997. Proceedings of the International Conference of the IEEE*, 1997, pp. 2807–2813 vol.6.

[45] T. M. Spalek and S. Hammad, "The left-to-right bias in inhibition of return is due to the direction of reading." *Psychol Sci*, vol. 16, no. 1, pp. 15–18, 2005.

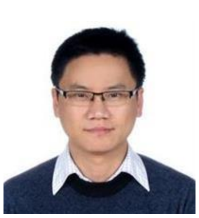
[46] A. Ossowski and M. Behrmann, "Left hemisphere specialization for word reading potentially causes, rather than results from, a left lateralized bias for high spatial frequency visual information." *Cortex*, vol. 72, pp. 27–39, 2015.

[47] J. Fraser and C. Gutwin, "The effects of feedback on targeting performance in visually stressed conditions," in *Graphics interface*, 2000, pp. 19–26.

[48] A. Cockburn, P. Quinn, C. Gutwin, G. Ramos, and J. Looser, "Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback," *International Journal of Human-Computer Studies*, vol. 69, no. 6, pp. 401–414, 2011.



Nikolaos Katzakis is a PostDoc research associate in the department of Informatics, University of Hamburg. He earned his Ph.D. in Information Systems Engineering from Osaka University in 2015. Following a brief stay in the Technical University of Munich in 2015 he has been working in the University of Hamburg. His research interests include computer graphics, augmented reality and deep learning.



Lihan Chen Lihan Chen is an Associate Professor at School of Psychological and Cognitive Sciences, Peking University. He obtained his Bachelor degree in Biomedical Engineering from Zhejiang University in 1999 and Master degree in General Psychology from Zhejiang University in 2005. He got his PhD degree in Experimental Psychology at Ludwig Maximilians University Munich in 2010.His major research interests include multisensory time perception, cross-modal correspondence, tactile perception and telepresence.

Oscar Ariza Nunez is a doctoral candidate at the Human-Computer Interaction Group of the Department of Informatics at the University of Hamburg. Oscar Ariza earned a Masters in Computer and Systems Engineering at Universidad de Los Andes. His research interests include wearable technology for 3D interaction, analysis of selection movements and meaningful haptic feedback for VR and AR environments.

Robert J. Teather Dr. Teather is an assistant professor in the School of Information Technology at Carleton University. He completed his PhD at York University in 2013, and a postdoctoral fellowship at McMaster University in 2015. His research specialization is in 3D user interfaces, VR, and user interfaces for computer games. His PhD work focused on developing standardized methods for the empirical comparison of input devices for 3D interaction primarily in order to compare mouse and 3D tracker-based input. To this end, Dr. Teather has established himself as an expert in comparing drastically different input devices in common fundamental interaction tasks (e.g., target selection), across varying system configurations (e.g., display properties such as stereo graphics, or system properties such as latency). He was awarded an NSERC Discovery Grant to study walk-up usable 3D interactive systems in 2016, and a Canada Foundation for Innovation John Evans Leaders Fund to build a lab studying interactions in digital media, in 2017.

Frank Steinicke is a professor for Human-Computer Interaction at the Department of Informatics at the University of Hamburg. His research is driven by understanding the human perceptual, cognitive and motor abilities and limitations in order to reform the interaction as well as the experience in computer-mediated realities. Frank Steinicke regularly serves as panelist and speaker at major events in the area of virtual reality and human-computer interaction and is on the IPC of various national and international conferences. He serves as the program chair for IEEE VR 2017/2018. Furthermore, he is a member of the Steering committee of the ACM SUI Symposium and the GI SIG VR/AR, and currently editor of the IEEE Computer Graphics and Applications Department on Spatial Interfaces.