Analysis of Direct Selection in Head-Mounted Display Environments

Paul Lubos,* Gerd Bruder,† Frank Steinicke‡
Immersive Media Group
Department of Computer Science
University of Würzburg

Abstract

The design of 3D user interfaces (3DUIs) for immersive head-mounted display (HMD) environments is an inherently difficult task. The fact that usually haptic feedback is absent and that visual body feedback is missing, hinders an efficient direct interaction with virtual objects. Moreover, the perceptual conflicts, such as double vision and space misperception, as well as the well-known vergence-accommodation mismatch further complicate the interaction, in particular with virtual objects floating in the virtual environment (VE). However, the potential benefits of direct and natural interaction offered by immersive virtual environments (IVEs) encourage the research in the field to create more efficient selection methods.

Utilizing a Fitts' Law experiment, we analyzed the 3D *direct selection* of objects in the virtual 3D space as they might occur for 3D menus or floating objects in space. We examined the direct interaction space in front of the user and divided it into a set of interaction regions for which we compared different levels of selection difficulty. Our results indicate that the selection errors are highest along the view axis, less along the motion axis and marginal along the orthogonal plane. Based on these results we suggest some guidelines for the design of direct selection techniques in IVEs.

Index Terms: H.5.2 [Information Interfaces and Presentation]: User Interfaces—Input Devices and Strategies, Evaluation / Methodology; I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality

1 Introduction

The advent of affordable head-mounted displays (HMDs) renewed the interest in immersive virtual environments (IVEs). Such IVEs have the potential to offer natural and direct interaction with objects displayed in the virtual world. In particular, the workspace within the user's arm reach provides a region, in which the user can grab virtual objects similar to grabbing in the real world. However, some of the often observed drawbacks of virtual reality (VR) technologies, for instance, distance underestimation, complicate the design of 3D user interfaces (3DUIs) in IVEs [37]. Although, we have recently observed significant improvements in 3D input devices and motion tracking systems, using tracked human gestures and postures in "mid-air" introduces challenges to the design of high-performance interaction techniques in the 3D space [8, 44, 43].

Indeed, interacting with natural gestures in 3D space opens up new possibilities for exploiting the richness and expressiveness of the interaction. Users can control multiple degrees-of-freedom (DoFs) simultaneously, and exploit well-known real-world actions. However, as a matter of fact, interaction in the 3D mid-air is physically demanding and, therefore, often hinders user satisfaction and performance [8]. The increase in the DoFs that have to be controlled simultaneously as well as the absence of passive haptic feedback and resulting interpenetration and occlusion issues when "touching the void" [6, 7, 8] are often responsible for reduced performance. 3D object selection is one of the fundamental tasks in 3DUIs. It is usually the initial task followed by more complex user interactions such as 3D manipulations [3, 38].

In this context, *virtual hand techniques* are often considered to be the most natural way of directly selecting virtual objects as they map identically virtual tasks with real tasks, which stands in contrast to indirect selection [3, 4]. However, direct selection of a virtual object in a fully-immersive setup significantly differs from selecting a physical object in the real world [2]. For instance, users perceive the virtual environment (VE) stereoscopically with vergence-accommodation conflicts, and usually also cannot see their real body, but at most a virtual representation in form of a virtual hand, marker or 3D point in space.

Furthermore, even small imprecisions, inaccuracies and latency of the used tracking system may cause slight mismatches between visual appearance of the virtual hand and the user's proprioceptive and kinesthetic feedback [6, 7, 8]. Such a decoupling of motor and visual space during natural hand interaction may degrade performance due to the kinematics of point and grasp gestures in 3D space and the underlying cognitive functions [27, 46]. An increased vergence-accommodation conflict has been found to affect size and distance judgments, as well as judgments of interrelations between displayed virtual objects [7, 26]. So far, while the perceptual problems in fully immersive VR are known, it is assumed that the main error for selection tasks in HMD environments is in the movement direction [7]. We hypothesize that the main reason for errors lies in the distance estimation and is thus along the viewing axis.

Previous research focused on the selection of objects at arbitrary positions in the virtual space around the user and different guidelines have been proposed [10, 18]. However, due to the perceptual difficulties in HMD environments, it is necessary to revisit if the known guidelines from real workspaces and traditional interfaces apply for 3DUIs. One of the first steps towards that goal is to investigate whether there are any differences between the regions in the user's interaction space or not.

In this paper we compare direct 3D selection task performance in a HMD environment using a Fitts' Law experiment. We tested different regions as well as different levels of difficulties, using the metrics movement time, error rates, error distances, and resulting effective throughput as overall performance indicator. The results give interesting implications for the design of 3DUIs for direct selection of objects displayed in mid-air space. In summary, our contributions include the

- comparison of direct selection performance in different regions in the space around the user, and
- guidelines for designing 3D user interfaces for direct selection in the 3D space.

*e-mail: paul.lubos@uni-wuerzburg.de
†e-mail: gerd.bruder@uni-wuerzburg.de
‡e-mail: frank.steinicke@uni-wuerzburg.de

The remainder of this paper is structured as follows: Section 2 provides an overview about related work in the field of 3D selection in VEs. Section 3 describes the Fitts' Law experiment that we conducted. Results are described in Section 4 and discussed in Section 5.

2 Related Work

3D interaction in IVEs has been in the focus of many research groups over the last decades [2]. Although direct interaction provides the most natural type of interaction with virtual objects, it is often not possible to use direct interaction for objects that are not located within arm's reach.

Different indirect interaction techniques have been proposed, such as the Go-Go [34] and HOMER technique [5], which allows users to interact with virtual objects in vista space by nonlinear scaling of hand positions within arm's reach. In particular, these techniques make use of the entire reachable space of a user's arms during interaction with distant objects, which may become exhausting when interacting at a distance, and may thus result in degraded performance over time [2]. Moreover, it is often observed that indirect interaction techniques result in reduced performance during interaction with virtual objects located within arm's reach compared to direct interaction [32].

As an alternative to virtual hand approaches, pointing techniques have been proposed, which do not utilize direct interaction with virtual objects but instead cast rays [31], each approach offering solutions to different problems, such as occlusion selection by manipulation of the ray's curvature with the off hand [33]. Other pointing techniques utilize heuristics to bend rays towards targets [39] or cast volumes instead of shapes to increase the selection performance [24]. While pointing techniques offer an efficient solution to many VR selection problems, they require some level of abstraction and might be less natural as virtual hand metaphors.

Direct Mid-Air Selection

According to Mine et al. [32], direct interaction leads to significantly higher performance than manipulation of objects at a distance from the user's hand. Most results from similar studies agree on the point that optimal performance may be achieved when visual and motor spaces are superimposed or coupled closely [9, 22, 46].

However, it is still an open research question, how the position of virtual target objects located within arm's reach may affect interaction performance [16]. Direct interaction is subject to perceptual limitations, e. g., the vergence-accommodation mismatch, ghosting or double vision, which can result in strong misperception effects [6, 7, 8]. Depending on the location of virtual objects, users may be unable to discriminate interrelations or perceive distances to objects to be smaller or larger than they are displayed [16, 26]. Furthermore, Hofmann et al. [16] show that some regions within the user's arm reach are considered to be more comfortable than others by users without finding any differences in terms of performance. Such distortions do not appear in the real world and may be related to limitations of current technology to correctly reproduce natural cues from the real world in a perfect way [48].

Moreover, internal representations of the 3D space are influenced and updated by both visual as well as motor input, which may affect interaction performance [41, 49]. Due to varying energy expenditure between users based on differences in strength and endurance of arm muscles, interaction performance in mid-air within arm's reach in IVEs may be affected by different factors related to the ergonomics of direct interaction. In particular, contributing factors may include interaction duration, hand and arm postures, frequency of movements, and comfort [2].

Fitts' Law and Selection

Fitts' Law describes the tradeoff between speed and accuracy in selection tasks [13]. Selections by touching or grasping objects with a user's hands can be split up into two phases, the *ballistic* and the *correction* phase [25]. The ballistic phase consists of focusing on the target object and bringing the hand in the proximity of the goal by using proprioceptive motor control. After that, visual feedback is used in the correction phase in order to incrementally reduce the distance from the hand to the goal.

Fitts' Law predicts the movement time MT for a given target distance D and size W [13]. They are brought together in a log term which describes the difficulty of the task overall with $MT = a + b \cdot log_2(D/W+1)$. The values a and b can be empirically derived for different setups. The *index of difficulty* (ID) is given by the log term and indicates overall task difficulty; smaller or/and farther targets result in increased difficulty. The formula has been extended in order to get effective measures. The error rate is adjusted to 4% by resizing targets to their effective width W_e. This is supported by an international standard [17]. By calculating the average of the measured movement distances, D_e can be determined. With that, the effective throughput can be computed as a useful combination of speed and accuracy: $TP = log_2(D_e/W_e+1)/MT$.

The validity of Fitts' Law for 3D interaction has been researched in the last years. Results from studies of several research groups imply that Fitts' Law is indeed valid for the kinematics of arm movements in a 3D interaction space [11, 28, 29].

In addition, Wingrave and Bowman [51] showed that Fitts' Law still holds when pointing in VEs. They observed that D was related to the amplitude of the movement, and W to the visual size of the target. Poupyrev et al. [35] defined the size of an object W according to the vertical and horizontal angles, that the object occupies in the user's field of view (FOV).

Similarly, Kopper et al. [20] propose a modification of Fitts' Law as a model for human performance in distal pointing tasks. Their model is based on angular amplitude and angular width as they argue that, contrary to classic 2D Fitts' Law tasks, the objects are floating in 3D space and the sizes and distances depend on the user's position, which can be solved by using angular measurements.

Ha and Woo adopt Fitts' Law for 3D tangible augmented reality environments with virtual hand metaphors [15] by using the model established by Grossman et al. [14], which however was based on a 3D objects arranged on a 2D plane.

Murata and Iwase [1] proposed an extension of Fitts' Law for 3D pointing tasks introducing a directional parameter into the model. Further studies investigating whether object selection in 3D mid-air can be modeled by Fitts' Law are described by Ware and Balakrishnan [47]. Finally, Teather and Stuerzlinger [40] found that using the simple Euclidean distance is viable enough for their system using a fish-tank VR system.

3 Experiment

In this section we describe the Fitts' Law experiment in which we analyzed direct selection in the user's arm reach in an immersive HMD environment. Based on previous findings discussed in Section 2, we evaluate the following two hypotheses:

H1: Larger selection errors occur along the view direction than along the movement direction.

H2: Fewer selection errors and a higher effective throughput are present for selection targets located in lower interaction regions.

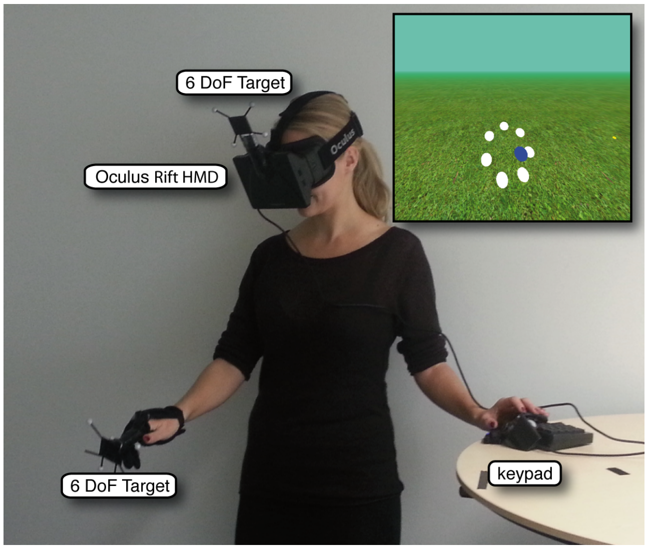


Figure 1: User during the experiment. The inset shows the user's view to the VE with the white and blue target spheres as well as the yellow sphere that visually represents the virtual finger tip.

3.1 Participants

We recruited 27 subjects for our experiment. Nine of them were male and 18 were female (ages 19 - 25, M = 21.78, heights 1.58m - 1.86m, M = 1.72m). All of the subjects were students of human-computer interaction or media communication from the University of Würzburg, who received class credit points for the participation in the experiment. One of the subjects was left handed, the remaining 26 subjects were right handed. All but one of the subjects had normal or corrected vision.

Using the technique proposed by Willemsen et al. [50] we measured the interpupillary distance (IPD) of each subject before the experiment started (M = 6.26 cm, SD = 0.31 cm). Additionally, we used the Porta and Dolman tests to determine the sighting dominant eye of subjects [30]. The test revealed that seven subjects were left-eye dominant (3 male, 4 female), whereas 20 subjects were right-eye dominant (5 male, 14 female). We measured the arm length and the height of the subject's celiac plexus and calibrated the VE for each subject accordingly.

Only three subjects reported no experience with stereoscopic 3D, such as cinemas or TV. All other subjects reported at least some experience (rating scale 0 = yes, 4 = no, M = 2.63, SD = 1.28). Nine subjects reported that they have participated in HMD studies before, and eleven subjects reported experience with HMDs (rating scale 0 = no experience, 4 = a lot of experience, M = 0.96, SD = 1.32). The remaining seven subjects had no experience with HMDs. All subjects were naïve to the experimental conditions.

The mean of the total time per subject, including questionnaires and instructions was about 55 minutes. The mean time for performing the actual experiment with the HMD on the head was about 36 minutes. Subjects were allowed to take breaks at any time.

3.2 Materials

As illustrated in Figure 1, subjects were instructed to stand in an upright position facing away from a wall. A Razer Nostromo keypad was adjusted to a comfortable height for the non-dominant hand of the user. Subjects were instructed to keep their hand at that position during the experiment to confirm their selections.

The experiment was conducted with the user wearing an Oculus Rift (Developer Edition) HMD with an attached 6 DoF passive infrared (IR) target. The target was tracked by an Optitrack V120:Trio, a passive, factory-calibrated IR tracking system, allowing to track the position and orientation of the HMD and thus the subject's head. The Oculus Rift offers a horizontal FOV of approximately 90° and a vertical FOV of 110° at a resolution of 1280 × 800 pixels (640 × 800 for each eye).

Additionally, we attached another 6 DoF target to the index finger of the subject's dominant hand (see Figure 1). The tracking system uses three cameras with a resolution of 640×480 pixels, a FOV of 47° and a latency of 8.33ms. The virtual stimulus (see Figure 1(inset)) used in the experiment displayed a 3D scene, which was rendered with OpenGL on an Intel computer with an Core i7 3.8GHz CPU, 8GB of main memory and Nvidia GeForce GTX580 graphics card. For the Oculus Rift, we rendered the virtual scene side-by-side and applied a barrel distortion afterwards.

The targets in the experiment were represented by spheres. All but the current target sphere for one trial were always visible and colored white. The current target was colored blue when the marker was outside, and green when the marker was inside to give the subjects a visual clue. The spheres were lighted in the order specified by the ISO 9241-9 standard [17]. As illustrated in Figure 2 each trial consisted of an arrangement of 8 spheres, 7 target spheres which were forming a circle with each sphere at the same distance to the 8th start sphere which was located at a comfortable position in front of the subject's celiac plexus [45].

3.3 Methods

At first, subjects completed between 2 and 6 supervised training trials for the experimental phase to ensure that they understood the task correctly. Those training trials were excluded from the analysis.

We used a 2 × 9 × 3 × 2 design with the method of constant stimuli for the experiment trials. The two sphere radii, three distances and nine ring positions, as well as the two repetitions were uniformly and randomly distributed between all 108 trials for each participant.

Each trial consisted of sequential selections of all 7 targets in a full ring with recurring direct selections of the 8th target in the center position, resulting in a total of 16 selections per trial and 8 visible target spheres. Using both oral and written instructions, the subjects were guided to select the targets as quickly and accurately

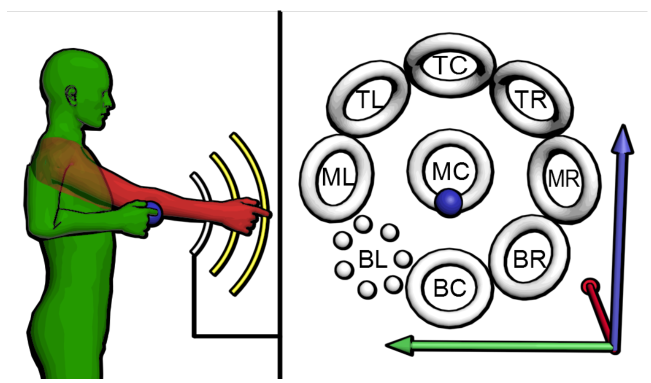


Figure 2: Schematic illustration depicting the layout of the spheres. The blue sphere depicts the starting position for all the trials. The circle segments show the three evaluated distances and the red arm depicts the subject's arm reaching for the furthest distance while the green person is in the starting position. The right side shows all the positions of the sphere rings on one depth level. The bottom-left (BL) ring shows an example configuration of seven target spheres during one trial. Each of the nine rings represents a setup of seven spheres used in another trial. Each trial consisted of the seven target spheres of one ring and the blue starting position.

as possible, as it is common in most Fitts' Law experiments [11, 28, 29].

While selecting a target correctly, subjects received visual feedback by changing the color of that target sphere from blue to green. Like previous studies on mid-air-selection [6, 7], the subjects had to confirm each selection by pressing a button on the keypad with their non-dominant hand to avoid any jitter caused by pressing buttons with their dominant hand. Alternatively, a pen might be used to point-and-click with the dominant hand, but Li et al. have shown that using the non-dominant hand might even improve performance for pen-based interfaces [23].

We computed the distance of the position of the index finger to the sphere center, which indicated a selection if the distance was less than the radius of the target sphere. If subjects performed a selection while the target sphere was not highlighted green, we recorded this as a selection error and advanced the trial state. The dependent variables were movement time, error distance (deviation from target center), error rate (percentage of targets missed), and effective throughput (see Section 2).

With the calibrated celiac plexus height and the targets position we defined a starting position of 35cm in front of the user's celiac plexus. Furthermore, we defined 9 *interaction regions* (rings) around an axis straight away from the user, as illustrated in Figure 2. Those regions were chosen since we wanted to evaluate whether moving in a certain direction (despite the same ID) had any influence on the performance. The regions cover the user's interaction space. To ensure equal distances between the starting position, represented in the schematic by a blue sphere, we rotated the ring of spheres by 45 degrees from the center axis.

Each of the rings in Figure 2 represents a ring of spheres like the one denoted as BL (bottom left). In this Figure only the bottom left ring is shown as a ring of spheres to ensure better visibility. The visible sphere configuration in Figure 2 corresponds to the configuration visible in the inset of Figure 1. Each trial consisted of 8 spheres, the start sphere in the center (blue) and the 7 target spheres. In the following, we denote these interaction regions as top-left (TL), top-center (TC), top-right (TR), middle-left (ML), middle-center (MC), middle-right (MR), bottom-left (BL), bottom-center (BC), and bottom-right (BR). Each sphere on different distance levels has the same distance to the starting position. Using the measured length L from forearm to index finger, we calculated the distances of the spheres in the ring to the start position as follows: $D_1 = L \cdot 0.5$, $D_2 = L \cdot 0.65$, and $D_3 = L \cdot 0.8$.

To be able to compare the different target distances, we scaled the spheres to the two IDs 3.0 and 4.0. To achieve this, we computed two sphere sizes W for each distance D by rearranging Fitts' formula for the ID: $W = D/(2^{ID} - 1)$.

According to Fitts' Law, adapting the target size with respect to the distance between selections results in larger targets for longer selection distances, whereas the targets are smaller for shorter distances, thus resulting in the same task difficulty between the different interaction regions.

Questionnaires Additionally to the main experiment trials, we asked subjects to complete subjective questionnaires. Before and after the experiment subjects were asked to complete a Kennedy-Lane Simulator Sickness Questionnaire (SSQ) [19]. After the experimental phase subjects were asked to complete a Slater-Usoh-Steed (SUS) presence questionnaire [42].

4 Results

In the following section we summarize the results from the experiment. Due to severe simulator sickness two subjects had to be excluded. A third subject did not have corrected vision and was excluded, as well. Results from the remaining subjects were normally distributed according to a Shapiro-Wilk test at the 5% level. We analyzed the results with a repeated measure ANOVA and Tukey multiple comparisons at the 5% significance level. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity when Mauchly's test indicated that the assumption of sphericity had been violated.

4.1 Selection Performance

In this section we describe the analysis of the Fitts' Law selection performance of our experiment. In particular, we analyze the movement time between subsequent selections, rate of missed targets, error distances to target centers, and effective throughput. The effective throughput metric incorporates both the results for speed and accuracy to form an estimate of the bottom-line performance.

Movement Time The results for the movement time are shown in Figure 3(a). We found a significant main effect of position ($F(4.37, 100.61) = 9.48, p < .001, \eta_p^2 = .29$) on movement time. Post-hoc tests revealed that the movement time was significantly increased when objects were displayed at positions bottom positions BL, BC, and BR (see Figure 2) in comparison to the other positions TL, TC, TR, ML, MC, and MR ($p < .001$). Furthermore, the movement time for the three distances differs significantly ($F(2, 46) = 46.01, p < .001, \eta_p^2 = .67$). As expected, we found a significant effect of the ID on movement time ($F(1, 23) = 294.06, p < .001, \eta_p^2 = .93$). We found a significant interaction effect between position and distance ($F(7.04, 161.91) = 3.63, p = .001, \eta_p^2 = .14$). The average movement time during the experiment was M=1313ms (SD=371ms).

Error Rate The results for error rate are shown in Figure 3(b). The results show that the error rate for the three distances differs significantly ($F(1.59, 36.63) = 20.56, p < .001, \eta_p^2 = .47$). We found a significant effect of the ID on error rate ($F(1, 23) = 43.14, p < .001, \eta_p^2 = .65$). However, we found no significant main effect of the position ($F(8, 184) = .86, p = .552, \eta_p^2 = .04$) on the error rate. The average error rate during the experiment was M=8.8% (SD=11.3%).

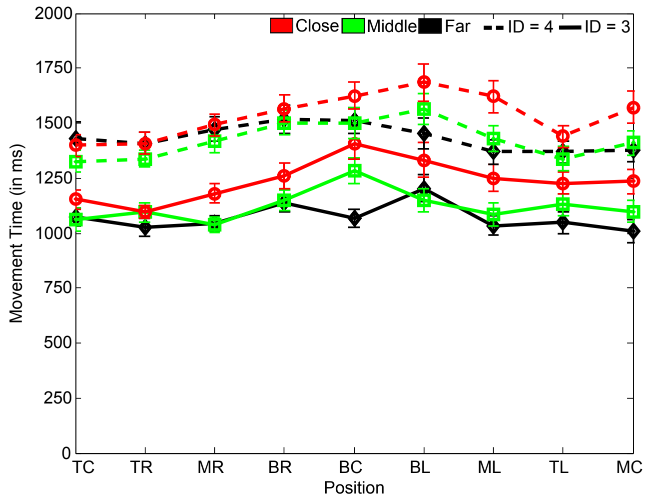
Error Distance The results for error distances between the center of each sphere and the finger position during selections are shown in Figure 3(c). The results show that the error distance for the three distances differs significantly ($F(1.34, 30.82) = 139.55, p < .001, \eta_p^2 = .86$). We found no significant main effect of position ($F(1.56, 35.88) = .75, p = .447, \eta_p^2 = .03$) on error distance. As expected, we found a significant effect of the ID on error distances ($F(1, 23) = 204.97, p < .001, \eta_p^2 = .90$). The average error distance during the experiment was M= 1.9cm (SD= 1.3cm).

Effective Throughput The results for the effective throughput are shown in Figure 3(d). We found a significant main effect of position ($F(8, 184) = 12.07, p < .001, \eta_p^2 = .34$) on throughput. Post-hoc tests revealed a significant decrease of the effective throughput when objects were displayed at bottom positions BL, BC, and BR (see Figure 2) in comparison to the other positions TL, TC, TR, ML, MC, and MR ($p < .001$). The results show that the throughput for the three distances differs significantly ($F(1.46, 33.56) = 9.61, p < .001, \eta_p^2 = .30$). Furthermore, we found a significant interaction effect between position and distance ($F(9.48, 218.11) = 2.26, p < .02, \eta_p^2 = .09$). Matching the expectations from Fitts' Law, we found a significant effect of the ID on throughput ($F(1, 23) = 35.97, p < .001, \eta_p^2 = .61$). The average throughput during the experiment was M=1.98bps (SD=.44bps).

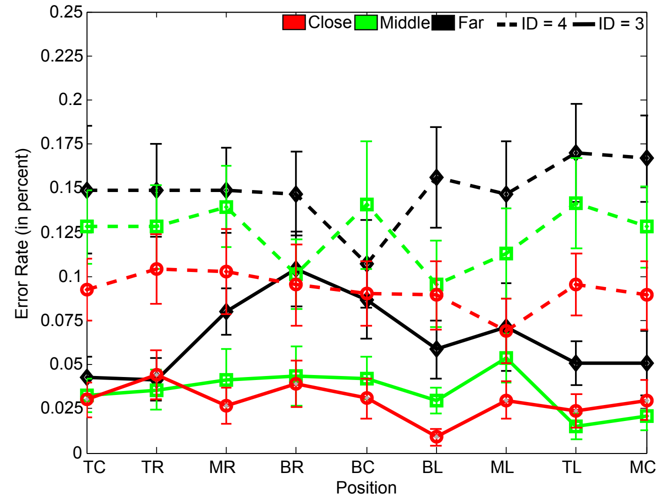
4.2 Error Distribution

In this section we analyze the distribution of selection errors along

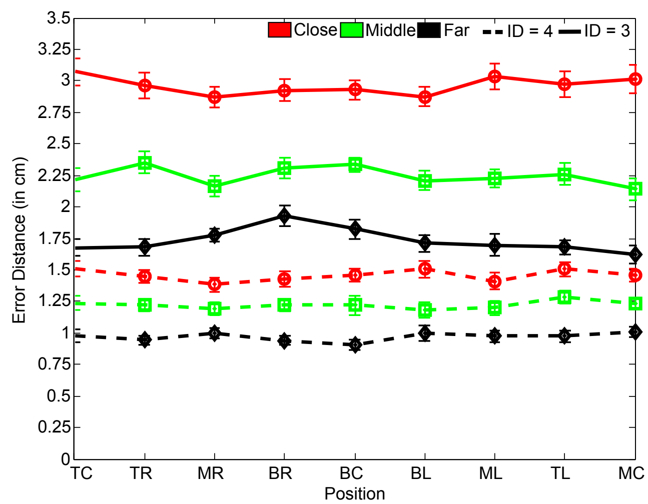
- the finger's movement vector between selections,
- the view axis from the dominant eye to the target, and



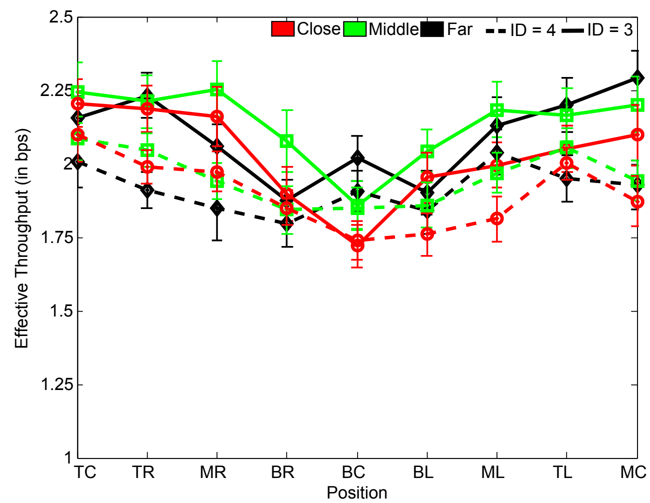
(a) Movement time



(b) Error rate



(c) Error distance



(d) Effective throughput

Figure 3: Plots of the experiment's results. The x-axis show the positions and the y-axis the a) movement time in ms, b) error rate in percent, c) euclidean error distance in cm and d) the effective throughput in bits per second (bps). The black lines show the closest distance between starting position and the target spheres. The dark gray lines stand for the middle distance and the light gray lines for the longest distance. Solid lines represent the low ID and dotted lines the high ID.

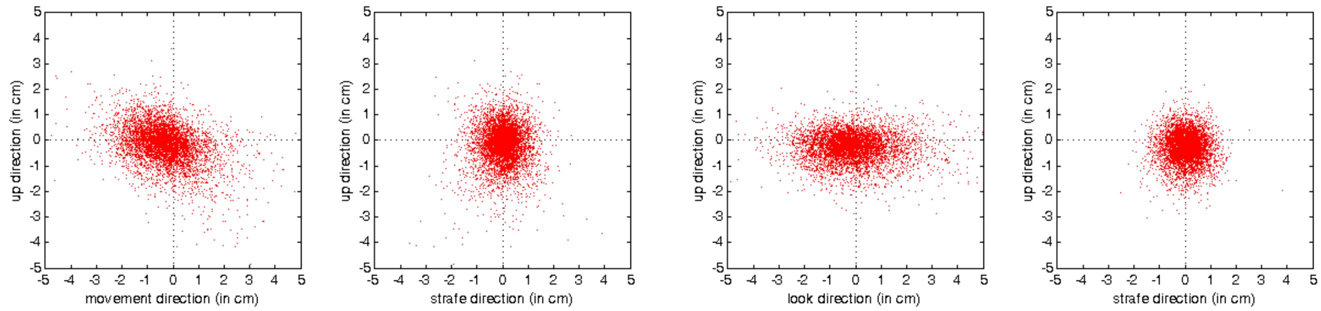
- the view axis from the non-dominant eye to the target.

We could not find any significant main effect for errors in the three directions ($F(1.03, 23.76) = .61, p = .449, \eta_p^2 = .03$). However, we found a significant interaction effect between the three distances and the three considered directions ($F(2.04, 46.96) = .86.90, p < .001, \eta_p^2 = .79$). Post-hoc tests revealed that for the shortest distance errors significantly differed between movement direction and dominant eye ($t(23) = -5.26, p < .001$), as well as between movement direction and non-dominant eye ($t(23) = -5.52, p < .001$), but not between dominant and non-dominant eye ($t(23) = .80, p = .432$). However, we could not find any significant differences for the larger distances. The error distributions are shown in Figure 4.

The results for the short distances show that errors were distributed mainly along the view direction and less along the movement direction. The lack of significant difference for larger distances can be explained by the fact that for these targets the view directions and movement directions converged.

Movement Direction We found a significant difference between errors in movement direction and along the orthogonal axes ($t(23) = 2.81, p < .02$). The error distribution along the finger's movement vector for each selection is shown in Figure 4(a). The average error in movement direction was M=1.21cm (SD=.69cm) and the average error along the orthogonal axes to the movement direction was M=1.09cm (SD=.40cm). For the shortest distance, the average error in movement direction was M=.83cm (SD=.13cm) and along the orthogonal axes M=.93cm (SD=.14cm).

Dominant Eye The error distribution along the vector from the subject's dominant eye to the target position for each selection is shown in Figure 4(b). We found a significant difference between errors in the dominant eye's view direction and along the orthogonal axes ($t(23) = 2.59, p < .02$). The average error in the dominant eye's view direction was M=1.20cm (SD=.60cm). The average error along the orthogonal axes to the dominant eye's view direction was M=1.10cm (SD=.50cm). For the shortest distance, the average error in the dominant eye's view direction was M=.93cm



(a) Movement direction

(b) Dominant eye direction

Figure 4: Scatter plots showing the error distribution of selection points for the closest tested distance in the experiment over all three IDs: Errors in (a) movement direction from the starting position to the target position as well as the axes orthogonal to the movement direction, and (b) the viewing direction from the dominant eye towards the target as well as the axes orthogonal to the viewing direction during the experiment.

(SD=.15cm) and along the orthogonal axes M=.83cm (SD=.14cm).

Non-dominant Eye The error distribution along the vector from the subject's non-dominant eye to the target position for each selection is shown in Figure 4(c). We found a significant difference between errors in the non-dominant eye's view direction and along the orthogonal axes ($t(23) = 2.53, p < .02$). The average error in the non-dominant eye's view direction was M=1.19cm (SD=.59cm). The average error along the orthogonal axes to the non-dominant eye's view direction was M=1.11cm (SD=.51cm). For the shortest distance, the average error in the non-dominant eye's view direction was M=.92cm (SD=.14cm) and along the orthogonal axes M=.83cm (SD=.13cm).

4.3 Questionnaires

Before and after the experiment, we asked subjects to judge their level of simulator sickness as well as their subjective sense of presence. While we measured an average pre-SSQ score of M = 10.29 (SD = 9.80), the post-SSQ score was M = 34.15 (SD = 21.99). The increase in simulator sickness over the time of the experiment was significant ($t(23) = -5.83, p < .001$). The mean SUS-score for the reported sense of feeling present in the virtual scene was M = 3.33 (SD = 1.52), which indicates a reasonably high level of presence.

5 Discussion

Fitts' Law was developed with a focus on two-dimensional interfaces, offering valuable and precise estimations on the potential time necessary to fulfill a task and hence the performance [12]. Afterwards, MacKenzie et al. have shown that Fitts' Law also applies to 3D spaces [27], and Teather and Stuerzlinger expanded the application of the law to VEs [40]. All of those works focused on the errors along the movement direction, since the very definition of Fitts' Law includes the width along the axis of motion. However, our results imply that in HMD environments the errors along the visual axis are dominant. Our results suggest that the errors are larger in perception space than in motor space.

These results supports our hypothesis H1, i. e., the movement direction is less a cause of errors than the view direction, which is in line with results found for touch interaction in stereoscopic tabletop environments [6, 7]. An increase of the interaction distance causes a convergence of the motion axis and the visual axis in touch selection tasks. This could explain why at increased distances we could not find a significant difference. However, the errors do not disappear, but rather accumulate at a distance, which might be the reason for the misconception that the errors are mainly due to the movement direction.

While analyzing the distribution of errors along the viewing direction, we calculated a 95% confidence ellipsoid [36], which is visualized in Figure 4.3. On the axes strafe, up, and view the center is located at position $(0.16, -0.47, -0.59)$ with radii $(1.83, 2.11, 3.91)$. Considering the errors were made mostly along the viewing axis, further investigations might show that increased tolerance along the viewing axis would allow faster and more precise selections in HMD environments.

Moreover, we found differences regarding the performance depending on the location of the interaction spaces as illustrated in Figure 3(d). Contrary to our hypothesis H2, our results show that the apparently more comfortable lower regions offer significantly reduced performance compared to the top positions, despite the same distances to the starting position and the same index of difficulty.

Since all constellations had the same two indexes of difficulty in motor space, this behavior could be caused by the increased dis-

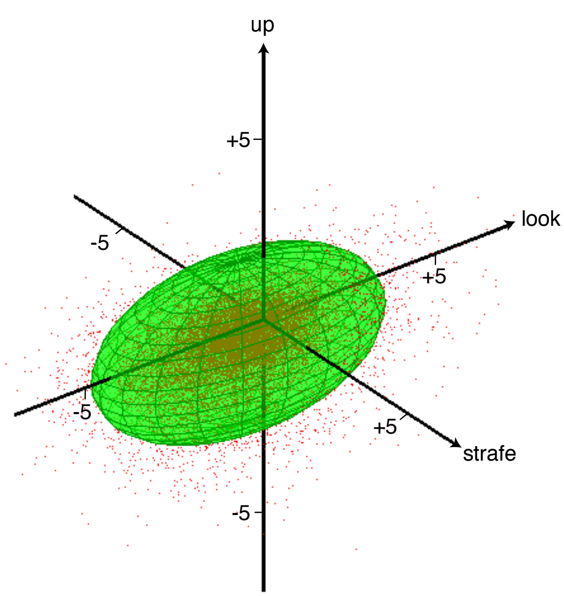


Figure 5: 95% confidence ellipsoid encompassing the hit points of the experiment around the center of the targets.

tance along the view axis and thus the increased difficulty in the perceptual space as discussed above. Despite the increased radii of the spheres at distance, which was done to ensure that the ID remained the same at different distances, the low resolution of the Oculus Rift HMD may also be considered a contributing factor to the effect for more distant targets due to the reduced visual acuity. Future work may focus on evaluating higher resolution displays, which will help to improve the understanding of the reasons for the observed effects. Furthermore, future work should consider if microlens array HMDs, which address the accommodation-convergence conflict [21], could provide better distance estimations.

6 Conclusion and Future Work

In this paper we investigated interaction spaces for 3DUIs in HMD environments. Our Fitts' Law experiment shows that the main factor causing errors in 3D selection is the visual perception and not, as presumed, the motor movement direction during selections. The error in the movement direction, which was presumed to be paramount, induces only part of the error.

We found significant more errors along the viewing axis than along the orthogonal axes. With these results, we suggest increased tolerance along the viewing axis, and proposed an ellipsoid cursor. Further research could investigate the effects of an increased error tolerance along the viewing axis, which could potentially improve selection task accuracy in IVEs.

Moreover, our results indicate that the bottom regions, which are farther away along the viewing axis, are worse for precise tasks, despite the fact that they have similar IDs compared to visually close regions.

We propose the following guidelines for immersive HMD environments:

G1: 3D selection tasks, which require high precision, should be restricted to a level close to the eyes despite more comfortable interaction in the lower regions.

G2: For accurate selections of spherical targets, such as points, we suggest inflating the selection space using an ellipsoid cursor as illustrated in Figure 4.3, which covers the 95% confidence region for touch selections.

Our results suggest that the ellipsoid cursor shown in Figure 4.3 can significantly improve selection performance over the tested interaction regions. However, future work should focus on developing a model of the size and form of the ellipsoid in relation to position and viewing distance within arm's reach and beyond.

Future experiments should evaluate such perceptually-inspired 3D selection regions of virtual objects, which have the potential to result in improved task performance. Moreover, further studies are required to elucidate the contributing factors to such distributions of touch points, which may include the effective visual acuity provided by the display as well as the well-known spatial misperception in stereoscopic VEs.

Acknowledgements

This work was partly funded by the German Research Foundation (DFG).

References

[1] M. A and H. Iwase. Extending Fitts' Law to a three-dimensional pointing task. *Human Movement Science*, 20:791–805, 2001.

[2] F. Argelaguet and C. Andujar. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics*, 37(3):121–136, 2013.

[3] D. Bowman. *Interaction Techniques for Common Tasks in Immersive Virtual Environments: Design, Evaluation, and Application*. PhD thesis, Georgia Institute of Technology, 1999.

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

[5] D. A. Bowman and L. F. Hodges. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In *In Symposium on Interactive 3D Graphics (I3D '97)*, pages 35–38. ACM Press, 1997.

[6] G. Bruder, F. Steinicke, and W. Stuerzlinger. Effects of visual conflicts on 3D selection task performance in stereoscopic display environments. In *Proc. of ACM 3DUI*, pages 115–118, 2013.

[7] G. Bruder, F. Steinicke, and W. Stuerzlinger. Touching the void revisited: Analyses of touch behavior on and above tabletop surfaces. *Lecture Notes in Computer Science: Human-Computer Interaction - INTERACT 2013*, 8117:278–296, 2013.

[8] 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. of ACM CHI*, pages 2625–2634, 2010.

[9] J. P. Djajadiningrat. Cubby: what you see is where you act. interlacing the display and manipulation spaces. 1998.

[10] S. A. Douglas, A. E. Kirkpatrick, and I. S. MacKenzie. Testing pointing device performance and user assessment with the iso 9241, part 9 standard. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 215–222. ACM, 1999.

[11] A. Y. Dvorkin, R. V. Kenyon, and E. A. Keshner. Reaching within a dynamic virtual environment. *J. NeuroEng. Rehabil.*, 23(4), 2007.

[12] P. Fitts. The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. *Journal of Experimental Psychology*, 47(6):381–391, 1954.

[13] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psych.*, 47:381–391, 1954.

[14] T. Grossman and R. Balakrishnan. Pointing at trivariate targets in 3d environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 447–454. ACM, 2004.

[15] T. Ha and W. Woo. An empirical evaluation of virtual hand techniques for 3d object manipulation in a tangible augmented reality environment. In *3D User Interfaces (3DUI), 2010 IEEE Symposium on*, pages 91–98. IEEE, 2010.

[16] M. Hofmann, R. Brger, N. Frost, J. Karremann, J. Keller-Bacher, S. Kraft, G. Bruder, and F. Steinicke. Comparing 3d interaction performance in comfortable and uncomfortable regions. In *Proceedings of the GI-Workshop VR/AR, 3-14*, 2013.

[17] International Organization for Standardization. *ISO/DIS 9241-9 Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices*, 2000.

[18] B. ISO. 9241 1: 1997. *Ergonomic requirements for office work with visual display terminals*.

[19] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psychol.*, 3(3):203–220, 1993.

[20] R. Kopper, D. A. Bowman, M. G. Silva, and R. P. McMahan. A human motor behavior model for distal pointing tasks. *International journal of human-computer studies*, 68(10):603–615, 2010.

[21] D. Lanman and D. Luebke. Near-eye light field displays. *ACM Transactions on Graphics (TOG)*, 32(6):220, 2013.

[22] D. K. Lemmerman and J. J. LaViola. Effects of interaction-display offset on user performance in surround screen virtual environments. In *Virtual Reality Conference, 2007. VR'07. IEEE*, pages 303–304. IEEE, 2007.

[23] Y. Li, K. Hinckley, Z. Guan, and J. A. Landay. Experimental analysis of mode switching techniques in pen-based user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 461–470. ACM, 2005.

[24] J. Liang and M. Green. JDCAD: A Highly Interactive 3D Modeling System. *Computers & Graphics*, 18(4):499–506, 1994.

[25] G. Liu, R. Chua, and J. T. Enns. Attention for perception and action: task interference for action planning, but not for online control. *Exp. Brain Res.*, 185:709–717, 2008.

[26] J. M. Loomis and J. M. Knapp. Visual perception of egocentric distance in real and virtual environments. In L. J. Hettinger and M. W. Haas, editors, *Virtual and adaptive environments*, volume Virtual and adaptive environments. Mahwah, 2003.

[27] C. L. MacKenzie, R. G. Marteniuka, C. Dugasa, D. Liskea, and B. Eickmeiera. Three-dimensional movement trajectories in fitts' task: Implications for control. *Q.J. Exp. Psychology-A*, 39(4):629–647, 1987.

[28] C. L. MacKenzie, R. G. Marteniuka, C. Dugasa, D. Liskea, and B. Eickmeiera. Three-dimensional movement trajectories in fitts task: Implications for control. *Q.J. Exp. Psychology-A*, 34(4):629–647, 1987.

[29] I. S. MacKenzie and P. Isokoski. Fitts throughput and the speed-accuracy tradeoff. In *ACM CHI*, pages 1633–1636, 2008.

[30] A. P. Mapp, H. Ono, and R. Barbeito. What does the dominant eye dominate? a brief and somewhat contentious review. *Percept. Psychophys.*, 65(2):310–317, 2003.

[31] M. Mine. Virtual Environments Interaction Technqiues. Technical Report TR95-018, UNC Chapel Hill Computer Science, 1995.

[32] M. R. Mine, F. P. Brooks Jr, and C. H. Sequin. Moving objects in space: exploiting proprioception in virtual-environment interaction. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 19–26. ACM Press/Addison-Wesley Publishing Co., 1997.

[33] A. Olwal and S. Feiner. The Flexible Pointer: An Interaction Technique for Selection in Augmented and Virtual Reality. In *ACM Symposium on User Interface Software and Technology (Conference Supplement)*, pages 81–82, 2003.

[34] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa. The Go-Go Interaction Technique: Non-Linear Mapping for Direct Manipulation in VR. In *ACM Symposium on User Interface Software and Technology (UIST)*, pages 79–80, 1996.

[35] I. Poupyrev, S. Weghorst, M. Billinghurst, and T. Ichikawa. A framework and testbed for studying manipulation techniques for immersive vr. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 21–28. ACM, 1997.

[36] R. Steeves and C. Fraser. Statistical post-analysis of least-squares adjustment results. In *CISM Adjustment and Analysis Seminars, EJ Krakiwsky (Ed.)*, 1987.

[37] F. Steinicke, G. Bruder, K. Hinrichs, M. Lappe, B. Ries, and V. Interrante. Transitional environments enhance distance perception in immersive virtual reality systems. In *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, APGV '09, pages 19–26, New York, NY, USA, 2009. ACM.

[38] F. Steinicke, T. Ropinski, and K. Hinrichs. Object Selection in Virtual Environments with an Improved Virtual Pointer Metaphor. In *Proceedings of International Conference on Computer Vision and Graphics (ICCVG)*, pages 320–326, 2004.

[39] F. Steinicke, T. Ropinski, and K. Hinrichs. Object selection in virtual environments using an improved virtual pointer metaphor. In *Computer Vision and Graphics*, pages 320–326. Springer, 2006.

[40] R. J. Teather and W. Stuerzlinger. Pointing at 3D targets in a stereo head-tracked virtual environment. In *Proc. of IEEE 3DUI*, pages 87–94, 2011.

[41] J. A. Thomson. Is continuous visual monitoring necessary in visually guided locomotion? *J. Exper. Psych. Hum. Percep. Perform.*, 9(3):427–443, 1983.

[42] M. Usoh, E. Catena, S. Arman, and M. Slater. Using presence questionaires in reality. *Presence-Teleop. Virt.*, 9(5):497–503, 1999.

[43] D. Valkov, F. Steinicke, G. Bruder, K. Hinrichs, J. Schöning, F. Daiber, and A. Krüger. Touching floating objects in projection-based virtual reality environments. In *EGVE/EuroVR/VEC*, pages 17–24, 2010.

[44] D. Valkov, F. Steinicke, G. Bruder, and K. H. Hinrichs. 2D touching of 3D stereoscopic objects. In *Proc. of ACM CHI*, pages 1353–1362, 2011.

[45] S. Vetter, J. Bützler, N. Jochems, and C. Schlick. Fitts' law in bivariate pointing on large touch screens: Age-differentiated analysis of motion angle effects on movement times and error rates. In C. Stephanidis, editor, *Universal Access in Human-Computer Interaction. Users Diversity*, volume 6766 of *Lecture Notes in Computer Science*, pages 620–628. Springer Berlin Heidelberg, 2011.

[46] Y. Wang and C. L. MacKenzie. Effects of orientation disparity between haptic and graphic displays of objects in virtual environments. In *INTERACT*, volume 99, pages 391–398, 1999.

[47] C. Ware and R. Balakrishnan. Reaching for objects in vr displays: lag and frame rate. *ACM Trans. Comput.-Hum. Interact.*, 1(4):331–356, Dec. 1994.

[48] P. Willemsen, M. B. Colton, S. H. Creem-Regehr, and W. B. Thompson. The effects of head-mounted display mechanical properties and field of view on distance judgements in virtual environments. *ACM Trans. Appl. Percpt.*, 6(2), 2009.

[49] P. Willemsen, A. A. Gooch, W. B. Thompson, and S. H. Creem-Regehr. Effects of stereo viewing conditions on distance perception in virtual environments. *Presence-Teleop. Virt.*, 17(1):91–101, 2008.

[50] P. Willemsen, A. A. Gooch, W. B. Thompson, and S. H. Creem-Regehr. Effects of stereo viewing conditions on distance perception in virtual environments. *Presence-Teleop. Virt.*, 17(1):91–101, 2008.

[51] C. Wingrave and D. Bowman. Baseline factors for raycasting selection. In *Proceedings of HCI International*, 2005.