J. Eng. Technol. Sci., Vol. 54, No. 1, 2022, 220106

Evaluating Direct Pointing and Indirect Cursor Interactions with Fitts' Law in Stereoscopic Environments

Dino Caesaron[1], Chiuhsiang Joe Lin[2], Ilma Mufidah[1] & Anastasia Febiyani[3]

[1]Department of Industrial Engineering, School of Industrial Engineering, Telkom University, Bandung, 40257, Indonesia
[2]Department of Industrial Management, National Taiwan University of Science and Technology, No. 43, Sec. 4, Keelung Rd., Da'an Dist., Taipei City 10607, Taiwan
[3]Department of Industrial Engineering, Institut Teknologi Telkom Purwokerto, 53147, Indonesia
*E-mail: dinocaesaron@telkomuniversity.ac.id

Highlights:

- Movement time and throughput were analyzed for two modes of interaction (direct pointing and indirect cursor).
- Distance of target and difficulty level were considered.
- Direct pointing may be the best option for interacting with a stereoscopic target.
- The results contribute to the understanding and designing of effective user interactions in stereoscopic environments.

Abstract. The development of virtual environment research has reached the stage of human interaction with three-dimensional (3D) objects. In this study, Fitts' method was used for such interaction in a virtual environment, and the applicability of Fitts' law in 3D virtual environments was also considered. An experiment was using two modes of interaction (direct interaction and indirect interaction) that utilize different techniques depending on how users interact with a 3D object. Both interaction techniques were applied with three indexes of difficulty and three egocentric target distances (distance from the user to the target). Movement time and throughput were measured for each interaction technique. The results showed that the direct pointing technique is more efficient for interaction with targets close to the user, while the indirect cursor technique may be a viable option for targets further away from the user. The throughput was found to be significantly higher for the direct pointing technique compared to the indirect cursor technique. The results of mean movement time were highly correlated with the targets' index of difficulty for all interaction techniques, which is supporting evidence that Fitts' law can be applied to interactions in 3D virtual environments. Based on the results, developers of VE applications may relate to these findings in designing proper user interactions.

Keywords: *direct pointing; Fitts' law; human interaction; indirect cursor; movement time; throughput; virtual environment.*

Received April 27th, 2020, 1st Revision December 15th, 2020, 2nd Revision August 17th, 2021, Accepted for publication September 30th, 2021.

DOI: 10.5614/j.eng.technol.sci.2022.54.1.6

1 Introduction

Research on virtual environments (VEs) has been conducted since the early 1990s and has been applied with different scopes and objectives. At the beginning of the 2000s, Stanney, *et al.*[1] identified the usability of virtual environments and more recently Satter & Butler [2] adopted this research focusing on immersive virtual reality. In a medical scope, Rooij, *et al.* [3] did a deep analysis of virtual reality for children's therapy. Meanwhile, Cai, *et al*. [4] focused on the application of virtual reality for cyber-physical manufacturing. This shows that the virtual reality research scope has developed considerably over time.

Samsung, Microsoft, HTC and other mobile OS business players are trying to develop virtual reality applications in a wider scope. The application of virtual reality has been widened to encompass human interaction with three-dimensional (3D) objects rather than merely watching images of 3D models. Research on the human interaction with 3D objects frequently uses HMDs (head-mounted displays) combined with IVEs (immersive virtual environments). Unfortunately, Sharples, *et al*. [5] found some drawbacks regarding the use of such systems, such as nausea, vertigo, and musculoskeletal system discomfort. In addition, HMD does not facilitate more than one user at the same time. Thus, researchers recently have tended to use stereoscopic 3D displays (S3D) to replace HMDs. The use of S3D applications is increasing because it has relatively low cost. Moreover, it facilitates a parallax that allows separation of two images and thus provides more information. The additional information provided through the parallax allows the user to have depth perception of the object [6]. A 3D projection display is used that utilizes S3D and allows virtual image interaction using a real object.

The increasing popularity of user-virtual object interaction motivates researchers to assess user performance. Until now, studies have typically dealt with one or more 3D functions, such as manipulation, navigation, and selection [7]. This results in complex tests with users, 3D activities, various methods for interaction, etc. Focusing only on a 3D pointing function instead will simplify the procedure of the experiment. The pointing task consists of shifting a tip/cursor to a specific position by changing the tip/cursor coordinates in the translation axes (x, y, z) [8]. Nevertheless, even under these simplified conditions, attention toward advanced 3D interaction needs supplementary considerations, such as an evaluation of various application contexts along with quantitative aspects, i.e. performance. Bowman, *et al.* [9] has classified the metric of interaction performance into task performance and technique performance. While technique performance measures qualitative variables, including user experience, usability interface, and comfort, task performance focuses on quantitative variables such as movement/completion time of task and accuracy.

In the present study, we evaluated two different modes of interaction, i.e. direct pointing and indirect cursor interaction, through a pointing task in a 3D VE. We distinguish direct and indirect interaction as proposed by Mine [10] to specify how the user interacts with the 3D target. Direct interaction occurs when users utilize their body (e.g. hand pointing or reaching), including gestures, gaze direction, etc., to move the target. Indirect interaction means an icon is used to acquire or specify the location of the target.

Direct and indirect interaction techniques have both been applied in previous studies. For instance, Lin & Woldegiorgis [11] applied a direct pointing technique with a pointing stick to point to a virtual target within a stereoscopic display environment. Bruder, *et al.* [12] compared two direct interaction techniques, direct 3D mid-air and 2D direct touch screen, in target selection tasks within a stereoscopic tabletop setting. Another direct interaction technique was used in Swan, *et al*. [13], where direct matching and blind reaching were used to estimate the position of virtual and real targets on a stereoscopic screen. In the present study, these techniques (direct reaching, pointing, and matching) were considered to be direct interaction methods.

Indirect interaction occurs when the user moves a specific virtual icon as the pointer to acquire a virtual target [10]. The user controls a device or an intermediary, for example a mouse, joystick, etc., to control the icon. Researchers have conducted several experiments with indirect interaction techniques [14-16]. In Werkhoven & Groen [15], a virtual cursor was used to evaluate manipulation performance (positioning and grasping) in an immersive virtual environment (IVE). In their study, the participants were asked to position a virtual cursor (i.e. an arrow cursor) over a target by controlling a SpaceMouse 3D controller. Meanwhile, in Poupyrev & Ichikawa [14] a comparison between interaction metaphors of a virtual pointer and a virtual hand were used for an object selection task in IVE. The research by Deng, *et al.* [16] asked the participants to place a cursor with a ball shape into a spherical target area in a VE using a handheld controller. Although the mentioned previous studies analyzed the performance of interaction techniques, they reviewed the techniques differently, either with direct interaction or with indirect interaction. Therefore, this study attempted to compare both direct and indirect interaction in terms of their performance based on movement time and throughput.

Fitts' law has been successfully used for investigating user performance in 2D environments [17,18]. Moreover, recent researches have investigated user performance using a modified version of Fitts' law to model user control in 3D VEs [16,19] related to indirect interaction as defined above. However, few comparative studies have been done that focused on direct and indirect interaction. In addition, most of those researches investigated the performance of target acquisition in VEs using direct or indirect interaction separately. Moreover, to the best of our knowledge few studies covered an implementation of Fitts' law on user interaction with respect to direct and indirect interaction within VEs. Thus, this research tried to investigate the implementation of Fitts' Law on user interaction, in particular with respect to direct pointing and indirect cursor interaction within VEs. More specifically, this research identified the difference between direct pointing and indirect cursor interaction using a Fitts' law experiment.

The contributions of this research are: (1) the findings provide more reference information for VR developers to select an appropriate interaction technique in a 3D stereoscopic environment for their application; (2) motivated by the promise of direct interaction between users and 3D objects, our work provides additional insight to the existing knowledge with respect to the application of direct and indirect interaction techniques within VEs.

1.1 Fitts' Law and its Extensions

Fitts' law (Eq. (1)) is known as a predictive model for user performance in target selection tasks to predict the time required to rapidly move to a target (MT) for a given target width (W) and a required distance (A).

$$MT = a + b\, log_2 \left(\left| \frac{A}{W} + 1 \right| \right) \qquad (1)$$

where *a* and *b* are regression coefficients, while the logarithmic expression is the index of difficulty (ID): a smaller or larger distance from the target leads to smaller or greater difficulty. An extension of Fitts' law supported by the International Organization for Standardization (ISO) is the use of the term 'effective' [20]. By utilizing this method, the rate of error is adjusted to a specific value of 4% by resizing the target width to its effective width (We) and effective distance (*Ae*). The effective width and distance account for the task the users actually performed as opposed to the task that they were required to do [21]. Using effective width and distance allows effective throughput (*eTP*), which can be determined using the following equation:

$$eTP = log_2 \left(\left| \frac{Ae}{We} + 1 \right| \right) /\mathrm{MT} \qquad (2)$$

2 Methods and Materials

In the present study, we compared the performance of the direct pointing and indirect cursor techniques. We used a Fitts' law task with 3D targets presented at different egocentric distances with different negative parallax levels (target projected in front of the screen).

2.1 Interaction Techniques and Devices

For the first interaction mode, the direct pointing technique, the participants were asked to move a pointing stick with a reflective marker fixed to the tip and position the marker to the center of the target surface (Figure 1(A)). After that they had to press a wireless remote fastened to the bottom end of the pointing stick to validate the pointing action. A 6D motion capture system (Optitrack) captured and tracked the data of the reflective marker positions (x, y, and z positions) at 120 frame rates per second. For the second interaction mode, the indirect cursor technique, the participants were asked to control an analog gamepad to guide a virtual cursor (i.e. a hand cursor) to reach a target within a stereoscopic environment. The participants had to position the virtual cursor in the center of the target surface (Figure 1(B)). The analog gamepad allowed moving the virtual cursor along the x, y, and z-axes. The movement speed/sensitivity of the virtual cursor within a VE is related to the force exerted on the analog gamepad. By increasing the force on the analog gamepad, the speed of the virtual cursor increases. Before the experiment with the indirect cursor technique, a pre-test was carried out to get the optimum sensitivity value of the cursor for making rapid and precise movements. The pre-test resulted in a sensitivity value of 2 m/s. Data recording the position of the marker with the direct pointing technique, the virtual cursor with the indirect cursor technique, and the virtual target reference were collected and processed on a PC. The location of the virtual target and the motion capture system were calibrated and fixed within a global coordinate system for both the direct and indirect interaction techniques.

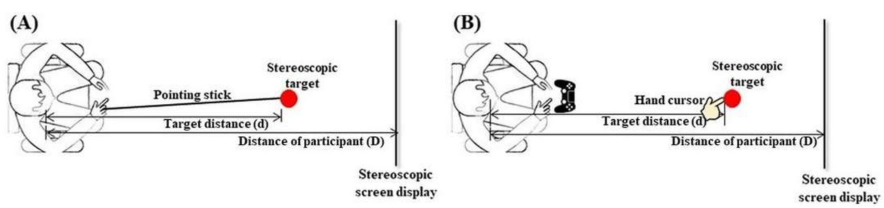

Figure 1 Illustration of the interaction techniques. The stereoscopic target appeared at a target distance (d) from the participant. (A) Direct pointing technique: the participant points at the target's surface with a pointing stick. (B) Indirect cursor technique: the participant places a virtual cursor on the target's surface by controlling a gamepad.

2.2 Experimental Design and Settings

Figure 2 provides an example of the arrangement of the target. A selection task setup according to the ISO 9241-9 standard was arranged in a circle with eight spherical targets rendered in red and displayed one by one in an order specified by the ISO standard [20].

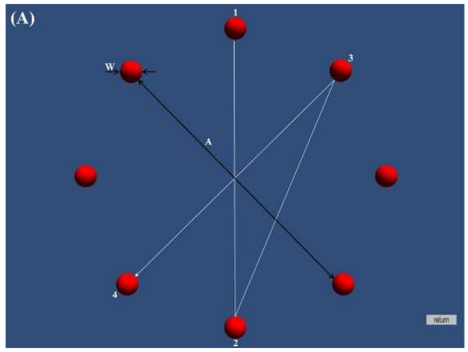

Figure 2 Illustration of the arrangement of the targets. Eight stereoscopic targets appeared one at a time in sequence. The figure also shows the target width (W) and the inter-target distance (A).

The perceived depth of the targets was varied by three parallax levels. The parallax determines the target position relative to a fixed projection screen location, with the participant seated at a distance of 210 cm (D) from the screen. Thus, the targets were shown at 90, 120, or 150 cm distance (d) from the participant (as shown in Figure 1).

The experiment was designed as within-subject, with a configuration of two (interaction technique) x three (egocentric distance) x six (index of difficulty) within-subject designs. The independent variables were: interaction technique (direct pointing vs. indirect cursor), egocentric distance (target displayed at a distance of 90, 120, and 150 cm), and index of difficulty (2.7, 3.5, 4.1, 4.5, 6.05, 6.15 bits). In addition, the six levels of the index of difficulty were graded into 'low', 'medium', and 'high' difficulty. The index of difficulty represents the difficulty of the task based on the target width (W) and the inter-target distance (A) [20,22]. As the distance to the inter-target gets longer, the difficulty increases. The dependent variables were movement time and effective throughput. The experimental room had a size of about 4 m x 3 m x 2.5 m and was covered with dark curtains to enable an excellent stereoscopic view without unnecessary light. The experimental room had a desk and an adjustable chair. The stereoscopic environment was projected on a 130 cm wide x 100 cm high projection screen, positioned from the participant at a fixed distance of 210 cm. The 3D projector was kept fixed during the experiment and was placed underneath the table outside of the participant's view.

2.3 Apparatus

The Unity 3D platform (version 4.3.4f1) was used to develop the stereoscopic environment, the targets and the experimental task. A 3D projector (ViewSonic PJD6251) projected the VE and ran for stereoscopic vision on a laptop that supported NVIDIA graphics. The participants were wearing a pair of 3D glasses integrated with an emitter to view the stereoscopic targets. The laptop adjusted the parallax based on the three levels of egocentric distance required and the measured value of interpupillary distance (IPD) from each participant.

In the direct pointing technique, three different lengths of wooden light sticks (80, 110, and 140 cm) attached with 0.6 cm reflective markers at the tip were used to point to the stereoscopic target at distances of 90, 120, and 150 cm, respectively. The different lengths of the stick were used in order to maintain an ergonomic posture while pointing at the targets. The weights of the sticks did not significantly differ so that these would not influence the performance of the participants. A PS2 dual analog gamepad was used for the indirect cursor technique to control the virtual cursor inside the stereoscopic environment. The left button of the analog gamepad was used to control the virtual cursor, including diagonal movement, in two-degree-of-freedom (DoF) translation (left-right, and up-down). The analog gamepad's right button controlled the virtual cursor depth (forward-backward). Of the other four buttons on the gamepad grip only the 'X' button was activated (to confirm the pointing actions). The initial/starting points of the virtual cursor as well as the stick tips were held fixed in three separate positions to create a relaxed pose before the participant conducted the direct pointing task. If we consider the 3D projector as a reference of measurement, the relative positions of the tips were (82, 0, 82), (82, 0, 90), and (82, 0, 120), respectively for the targets displayed at 90, 120, and 150 cm from the participant. Both the stick tip and the virtual cursor position were resumed to their initial location after each trial. During the experiment, the participants sat on an adjustable chair to preserve the sitting height and their chin was placed on a chinrest on the tabletop to eliminate the likelihood of variance in depth perception due to movements of the participant's head.

2.4 Participants

Eighteen participants were recruited by promotion on social media. All participants were ensured to be naïve to the purposes of the experiment. The recruited participants were twelve males and six females, aged 23 to 30 (mean = 25.22, SD = 2.26). All participants reported they had normal vision and were right-handed. Their IPDs were among 6.5 to 7 cm (mean = 6.56, SD = 0.16). Most of the participants had no experience at all with the use of VR systems. Before participating in the experiment, all participants were asked to fill out a consent form. Prior to the experiment, all participants were tested to observe the closest target distance at 90 cm from them. None of them failed this test.

2.5 Procedure

Before the experiment, the participants read and completed a consent form explaining the objectives of the study, the activity guidelines and the research procedures. The participants obtained the same verbal summary of the study after reading the instructions while their IPDs were being tested. The IPD measurement value determined the parallax setting for showing the stereoscopic target.

The experiment was split into two parts: one for the direct pointing technique and the other for the indirect cursor technique. Once the interaction mode (either direct pointing or indirect cursor) was chosen, the egocentric distance and task difficulty were randomly assigned. Once the interaction mode and the distance were selected, the participants had to complete all six IDs in completely randomized trials. We separated each interaction technique trial by at least two days in order to minimize the participants' fatigue and learning effect [23-25]. Before the experiment began, each participant performed a training session (around 3 min) for each interaction technique to let them familiarize themselves with the experimental setting and apparatus. The training session was followed by the first experimental session on the selected interaction technique.

The participants sat on a chair 210 cm away from the projection screen during the training session and wore 3D glasses. They were told to look at the stereoscopic environment to get a clear impression of the stereoscopic targets. They were also advised to quit when they felt discomfort associated with the stereoscopic environment at any time. At the beginning of each trial, the virtual targets were continuously displayed and then disappeared when the participant reached the target. The participants were instructed to point/select as fast as possible at the target surface while still maintaining accuracy. There was no feedback (e.g. sound, color change, etc.) when the target was selected or dismissed other than the next target showing up. For each interaction condition each participant completed 18 trials (3 egocentric distances * 6 IDs) for a total of 144 targets (3 egocentric distances * 6 IDs * 8 targets). The total time per participant was approximately 90 min, including completion of the consent form, the training session and the experiment.

3 Results

A repeated-measures (rm) ANOVA was performed on each dependent variable, i.e. movement time and effective throughput, with as independent variables technique, egocentric distance, and ID. If the results were significant, a post-hoc Tukey HSD test was conducted ($\alpha = 0.05$). The Greenhouse-Geisser correction was used to correct the degrees of freedom (df) when Mauchly's test indicated that the sphericity assumption was violated.

3.1 Movement Time

Figure 3A shows the results for movement time. The effect of interaction technique on movement time was significant (F [1,17] = 22.86, $p < 0.001$). The participants needed less time when using direct pointing interaction (mean = 2.22s, SD = 0.11s) compared to indirect cursor interaction (mean = 3.30s, SD = 0.24s). When using both the direct pointing and the indirect cursor technique, the movement time increased with ID. The rmANOVA confirmed that the movement time was significantly influenced by ID (F [1.64, 27.79] = 90.64, $p < 0.001$). The post-hoc Tukey analysis showed that the movement time was increased significantly when the targets were displayed at high ID ($p < 0.001$) in comparison to medium and low IDs. The rmANOVA results also showed that the movement time was significantly different for different egocentric distances (F [1.99, 33.93] = 8.58, $p < 0.05$). During the experiment, the average movement time was lower (mean = 2.60s, SD = 0.14s) when the targets were displayed at a distance of 90 cm compared to the targets being displayed at distances of 120 cm (mean = 2.67s, SD = 0.16s) and 150 cm (mean = 3.00s, SD = 0.17s). Significant two-way interaction was found between interaction technique and ID (F [2, 34] = 79.65, $p < 0.001$).

The Tukey post-hoc test showed that the participants took significantly longer when using the indirect cursor technique for selecting the targets with the highest ID ($p < 0.001$) in comparison to the medium and lowest IDs. There was no significant difference in movement time for different IDs when the direct cursor technique was used. We also found significant two-way interaction between egocentric distance and interaction technique (F [1.64, 27.92] = 8.18, $p < 0.005$). The post-hoc test showed that the participants took significantly longer to select the targets at the farthest distance of 150 cm ($p < 0.001$) when using the direct pointing technique compared to the targets being at 90 and 120 cm. When using the indirect cursor technique, no significant difference was found in movement time for different egocentric distances. Also, there was no significant interaction difference observed between egocentric distance and ID.

3.2 Effective Throughput

Figure 3B shows the results of effective throughput. The throughput encompassed both speed and accuracy, with higher scores being equivalent to better performance. The effect of interaction technique on throughput was significant (F [1, 17] = 9.08, $p < 0.01$). The direct pointing throughput was higher (mean = 2.36 bps, SD = 0.12 bps) than that for the indirect cursor technique (mean = 1.96bps, SD = 0.14bps). The egocentric distance also had a significant effect (F [1.90, 32.30] = 31.35, p < 0.001) on throughput. The post-hoc tests revealed that the throughput was substantially decreased when the targets were seen at 150 cm (p < 0.005) compared to the other distances. The effect of ID on throughput was also significant (F [1.72, 29.26] = 123.88, p < 0.001). In addition, the Tukey post-hoc tests showed that the effective throughputs were substantially different (p < 0.001) for all IDs. A significant interaction difference was observed between technique and egocentric distance (F [1.38, 23.49] = 4.05, p < 0.05) based on throughput. The post-hoc tests revealed that the throughput was significantly lower (p < 0.005) when selecting targets displayed at 150 cm compared to 120 cm and 90 cm with the direct pointing technique. For the indirect cursor technique, we found that the throughput for the target displayed at 150 cm was significantly lower (p < 0.05) than for targets at 90 and 120 cm.

A significant two-way interaction was found between technique and ID (F [1.16, 19.79] = 7.15, p < 0.05) based on throughput. The Tukey post-hoc test showed that the throughput was significantly different (p < 0.001) for all IDs, with a high ID yielding the highest throughput, a low ID yielding the lowest throughput value, and a medium ID falling between the two, for both techniques. Furthermore, significant interaction was found between egocentric distance and ID. The result from the post-hoc analysis showed a difference in throughput for all three egocentric distances and IDs, i.e. higher throughput for 90 cm than for 120 cm and 150 cm, and a low ID giving a higher throughput compared to the other IDs.

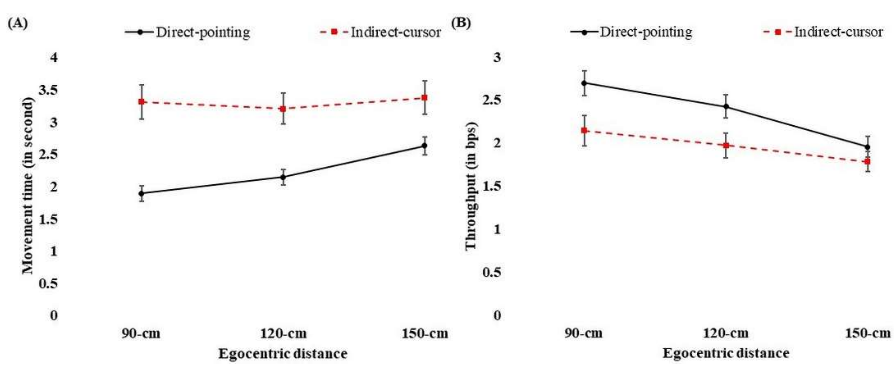

Figure 3 Results for two interaction techniques. The x-axis shows the target egocentric distances (displayed at 90,120, and 150 cm), while the y-axis shows (A) movement time, and B (throughput). The error bars show the standard error.

3.3 Fitts' Law Modeling

Fitts' law can be implemented as a predictive model. Linear regression is used to predict the time for a participant to acquire the target and click (judge) on it [19, 26]. We conducted this analysis for both techniques at the three egocentric target distances and IDs. The analysis results confirmed that movement time could be predicted (as expressed by R^2) very well from the index of difficulty for both the direct pointing technique (for distances 90, 120, and 150 cm respectively, R^2 = 0.998, 0.942, 0.978) and the indirect cursor technique (for distances 90, 120, and 150 cm respectively, R^2 = 0.827, 0.907, 0.931). This suggested that Fitts' model can be applied to the direct and indirect interaction techniques and used to describe 3D interactions in VEs. The regression lines for the movement time are showed in Figure 4.

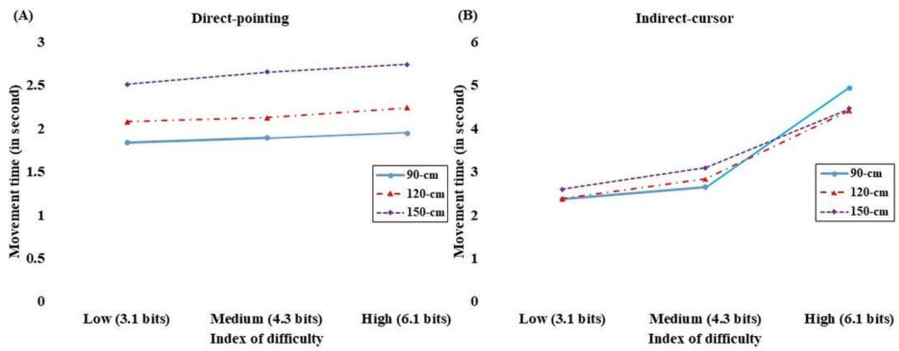

Figure 4 Regression lines of the measured movement time at three target distances (displayed at 90, 120, and 150 cm) for (A) the direct pointing technique and (B) the indirect cursor technique.

4 Discussion

The present study investigated the direct pointing technique and the indirect cursor technique on 3D target selection performance using a stereoscopic display system based on movement time, effective throughput, and slope of Fitts's model. In direct interaction, the user utilizes the human body (e.g. hand pointing, eye gaze), while in indirect interaction, the user utilizes a virtual control (e.g. a cursor) controlled with any physical or virtual controller (e.g. a joystick or a gamepad) to select 3D targets in a VE. Even though the two techniques are different (direct and indirect) with respect to how the user interacts with a 3D target, a typical pattern exists in the users' performance with both interaction techniques. The results showed that both the direct pointing technique and the indirect cursor technique followed Fitts' law model, with the movement time linearly increasing with the target's ID. This relationship, therefore, supports that Fitts' law can be generalized for interactions with 3D targets in stereoscopic environments with a similar trend as in 2D environments.

A clear effect of interaction technique on movement time was found in this study. The participants were faster when using the direct pointing technique than when they used the indirect cursor technique. This difference may be explained by the characteristics of interaction with direct pointing, which provides more ease of use and a more natural way of interaction with targets displayed in negative parallax compared to indirect cursor interaction. The results from this study also showed a significant effect between interaction technique and egocentric distance on movement time, with the direct pointing technique outperforming the indirect cursor technique at all three distances (90, 120, and 150 cm). The participants took longer to perform the target acquisition task at 150 cm than the other two distances. This implies that direct pointing needs more time, as the target moves away from the user.

In our experiment, the direct pointing technique had the highest effective throughput at all tested egocentric distances, indicating that this technique should be the first option for the general purpose of user interaction on a stereoscopic display. However, the performance results of this technique in terms of throughput decreased drastically (about 20%) for a target displayed at 150 cm. The performance with the indirect cursor technique decreased much more slowly, reducing only around 9% at a distance of 150 cm. This findings suggests that the performance of the direct pointing technique decreased more drastically for large negative parallax compared to the indirect cursor technique. Nonetheless, this recommendation is based on the effective throughput metric, with the basic assumption that selection errors within the VR application area have only a low intensity.

Soukoreff & MacKenzie [27], who surveyed nine experiments to evaluate input devices in accordance with ISO 9241-9 [20], have suggested that the throughput values of the input device fall within a narrow range. The range of throughput values was 1.6-2.55bps for a joystick, 0.99-2.9bps for a touchpad, and 3.7-4.9bps for a mouse. In the present study, a gamepad was employed for the indirect cursor technique. The mean throughput value for this technique was 1.96bps, which is still in the range of the throughput value for a joystick as input device. Although there is a slight difference between how a joystick and a gamepad is held, we consider both inputs as indirect interaction methods. As compared to the mouse throughput value mentioned in Soukoreff & MacKenzie [27], the direct pointing technique had lower throughput (2.55bps). This low throughput value may be accounted for by the different experimental environment. All of the studies reported in Soukoreff & MacKenzie [27] used a 2D computer-based environment while in the present study we used a 3D stereoscopic environment. The low throughput value obtained in our study also reveals that the target acquisition task in a 3D VE is more complicated than in a 2D environment. Target acquisition in a 3D VE requires more degrees of freedom in movement. For instance, Fitts' task in a 2D computer-based environment requires one- or two-axis movement (along the x- and y-axes), whereas in our experiment three-axis movement was required (along the x-, y-, and z-axes). Therefore, movement in a 3D environment is more reactive to indices of difficulty than movement in a 1D or 2D environment.

The current study was not designed to evaluate factors related to the error rate or the accuracy of the interaction technique. We used the assumption of the error rate staying low as the participants kept their accuracy while doing the pointing task, both with the direct pointing technique and the indirect cursor technique. In our experimental procedure, due to some technical aspects of the 3D software system, the participant could continue the task even if they made an error (mispointing), and the task was completed when the number of clicks reached a predefined number (i.e. for eight targets). While today's throughput usually includes an adjustment for accuracy, it is possible and common to calculate throughput taking into account only movement time and index of difficulty [18,21] for Fitts' law experimental tasks. However, for a comprehensive understanding of the characteristics of both interaction techniques, further experimental investigations are needed that include accuracy (error rate) to obtain complete performance metrics.

5 Conclusion

Our overall results revealed the effect of different interaction techniques, different parallaxes of the virtual target, and different IDs on movement time and effective throughput in performing pointing tasks within a stereoscopic environment. The direct pointing technique produced a shorter movement time for all three parallaxes in comparison to the indirect cursor technique. The target displayed (150 cm) farthest from the participants took a significantly longer time than the other two depths (90 cm and 120 cm). These results suggest that the time taken for direct pointing decreases more as the target is displayed farther away from the participant. In the present study, the direct pointing technique yielded the best performance (in terms of effective throughput) for all targets displayed in negative parallax. With this technique, the target selection times benefits from the naturalness and efficiency of the arm movement. In addition, the performance of direct pointing decreased more rapidly than that of the indirect cursor technique with increasing distance of the targets displayed. This result suggests that when the interaction distance to the target is quite far (above 1 m) the indirect cursor technique can be a viable option. Moreover, developers of VEs could apply the findings of this study to improve the user experience by choosing an appropriate interaction technique and identifying at what target location the interaction can be more effective and efficient. However, the performance described in this study was based on effective throughput measurement under the assumption that selection errors were relatively low for both interaction techniques.

In our experiment, the ISO 9241-9 standard was applied using a pointing task to evaluate two interaction techniques in terms of how users interact with a 3D target. The results of the present study demonstrated that Fitts' law can be generalized to interaction with 3D targets within stereoscopic environments in a similar fashion as with interactions in 2D environments (e.g. mouse, touchscreen). The present study provided essential information on how performance varies for different interaction techniques and stereoscopic parallaxes of virtual objects with various difficulty indexes. The findings of the present study may be used as a rational guideline for the choice of interaction technique in 3D stereoscopic display setups. The research focused on performance metrics of two different interaction techniques (direct and indirect) with a stereoscopic display. Future interaction studies in stereoscopic environments may be needed to investigate kinematic and user behavior, such as perceived space, accuracy, and movement when employing different interaction techniques.

There are some limitations to be mentioned here. First, this study was unable to analyze the accuracy of the interaction techniques in a Fitts' law task. The regression lines captured should represent all IDs, however, here we only show the average IDs. Further investigation and experimentation into these issues are strongly recommended. Second, the planes of motion in the Fitts' law task only addressed performance in the frontal plane (in front of the participant) under three egocentric distance conditions. Performance in other planes (lateral and transverse) may be important and should be further investigated. Lastly, it is essential to note that the present study only considered task performance with a stereoscopic projection display, utilizing a specific tracking system, an input device, within a specific range of egocentric distances. Further experimentation with replication of the methodology over a number of stereoscopic displays and spatial conditions (egocentric and exocentric spatial relationships) are needed to generalize the results.

Acknowledgement

This study was partially supported by *Direktorat Penelitian Pengembangan Masyarakat* (PPM), Telkom University and the Ministry of Science and Technology of Taiwan (MOST 104-2221-E-0111-063-MY3).

References

[1] Stanney, K.M., Mollaghasemi, M., Reeves, L., Breaux, R. & Graeber D.A., *Usability Engineering of Virtual Environments (VEs): Identifying Multiple Criteria that Drive Effective VE System Design*, International Journal of Human-Computer Studies, **58**(4), pp. 447-481, 2003.

[2] Satter, K. & Butler, A., *Competitive Usability Analysis of Immersive Virtual Environments in Engineering Design Review*, Journal of Computing and Information Science in Engineering, **15**(3), pp. 1-12, 2015.

[3] Rooij, M.V., Lobel, A., Harris, O., Smit, N. & Granic, I., *DEEP: A Biofeedback Virtual Reality Game for Children At-Rrisk for Anxiety*, Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, pp. 1989-1997, 2016.

[4] Cai, Y., Starly, B., Cohen, P. & Lee, Y.S., *Sensor Data and Information Fusion to Construct Digital-twins Virtual Machine Tools for Cyber-Physical Manufacturing*, Procedia Manufacturing, **10**, pp. 1031-1042, July 2017.

[5] Sharples, S., Cobb S., Moody A. & Wilson J.R., *Virtual Reality Induced Symptoms and Effects (VRISE): Comparison of Head Mounted Display (HMD), Desktop and Projection Display Systems,* Displays, **29**(2), pp. 58-69, March 2008.

[6] McIntire, J.P., Havig, P.R. & Geiselman, E.E., *Stereoscopic 3D displays and Human Performance: A Comprehensive Review*, Displays, **35**(1), pp. 18-26, Jan. 2014.

[7] Bowman, D.A., Kruijff, E., LaViola, J.J. & Poupyrev, I., *3D User Interfaces: Theory and Practice*, Addison Wesley Longman Publishing Co., Inc., 2004.

[8] Raynal, M., Dubois, E. & Schmitt, B., *Towards Unification for Pointing Task Evaluation in 3D Desktop Virtual Environment,* Springer Berlin Heidelberg, 2013.

[9] Bowman, D.A., Johnson, D.B. & Hodges, L.F., *Testbed Evaluation of Virtual Environment Interaction Techniques*, Presence Teleoperators & Virtual Environments, **10**(1), pp. 75-95, August 2001.

[10] Mine, M.R., *Virtual Environment Interaction Techniques*, University of North Carolina at Chapel Hill, 1995.

[11] Lin C.J., Woldegiorgis B.H., *Egocentric Distance Perception and Performance of Direct Pointing in Stereoscopic Displays*, Applied Ergonomics, **64**, pp. 66-74, May 2017.

[12] Bruder, G., Steinicke, F. & Sturzlinger, W., *To Touch or not to touch? Comparing 2D touch and 3D Mid-Air Interaction on Stereoscopic Tabletop surfaces*, Proceedings of the 1st symposium on Spatial user interaction, pp. 9-16, July 2013.

[13] Swan, J.E., Singh, G. & Ellis, S.R., *Matching and Reaching Depth Judgments with Real and Augmented Reality Targets*, IEEE Transactions on Visualization and Computer Graphics, **21**(11), pp. 1289-1298, July 2015.

[14] Poupyrev, I. & Ichikawa, T., *Manipulating Objects in Virtual Worlds: Categorization and Empirical Evaluation of Interaction Techniques*, Journal of Visual Languages & Computing, **10**(1), pp. 19-35, Feb. 1999.

[15] Werkhoven, P.J. & Groen J., *Manipulation Performance in Interactive Virtual Environments*, Human Factors, **40**(3), pp. 432-442, Sep. 1998.

[16] Deng, C.L., Geng, P., Hu, Y.F. & Kuai, S.G., *Beyond Fitts's Law: A Three-Phase Model Predicts Movement Time to Position an Object in an Immersive 3D Virtual Environment*, Human Factors, **61**(3), pp. 879-894, March 2019.

[17] Bi, X., Li, Y. & Zhai, S., *Fitts Law: Modeling Finger Touch with Fitts' Law*, the SIGCHI Conference on Human Factors in Computing Systems, ACM, pp. 1363-1372, April 2013.

[18] MacKenzie, I.S. & Teather, R.J., *FittsTilt: The Application of Fitts' law to Tilt-based Interaction*, 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design, ACM, pp. 568-577, Oct. 2012.

[19] Lin, C.J., & Widyaningrum, R., *Eye Pointing in Stereoscopic Displays.* Journal of Eye Movement Research, **9**(5), pp. 1-14, May 2016.

[20] ISO, *ISO 9241-9: Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)-Part 9-Requirements for Non-keyboard Input Devices2000*, Internasional Organization Standardization, pp 1-47, Feb. 2000.

[21] Lin, C.J., Chen, H.J., Cheng, P.Y. & Sun, T.L., *Effects of Displays on Visually Controlled Task Performance in Three-Dimensional Virtual Reality Environment*, Human Factors and Ergonomics in Manufacturing & Service Industries, **25**(5), pp. 523-533, Sept. 2014.

[22] Park, K.S., Hong, G.B. & Lee, S., *Fatigue Problems in Remote Pointing and the Use of an Upper-Arm Support,* International Journal of Industrial Ergonomics, **42**(3), pp. 293-303, May 2012.

[23] Yue, K., Wang, D., Hu, H. & Fang, S., *The Correlation Between Visual Fatigue and Duration of Viewing as Assessed by Brain Monitoring*, Journal of the Society for Information Display, **26**(7), pp. 427-437, June 2018.

[24] Lin, C.J., Ho, S.H. & Chen, Y.J., *An Investigation of Pointing Postures in a 3D Stereoscopic Environment*, Applied Ergonomics, **48**, pp. 154-163, May 2015.

[25] Chao, C.J., Wu, S.Y., Yau, Y.J., Feng, W.Y. & Tseng, F.Y., *Effects of Three-Dimensional Virtual Reality and Traditional Training Methods on Mental Workload and Training Performance*, Human Factors and Ergonomics in Manufacturing & Service Industries, **27**(4), pp. 187-196, May 2017.

[26] Vertegaal, R., *A Fitts Law Comparison of eye Tracking and Manual Input in the Selection of Visual Targets*, Proceedings of the 10th international conference on Multimodal interfaces, ACM, p. 241-248, Oct. 2008.

[27] Soukoreff, R.W. & MacKenzie, I.S., *Towards a Standard for Pointing Device Evaluation, Perspectives on 27 years of Fitts' Law Research in HCI*, International Journal of Human-Computer Studies, **61**(6), pp. 751-789, 2004.