This is the authors' preprint. The definitive version is available at https://dl.acm.org/doi/10.1145/3385956.3422102.

Ray-Casting-based 3D pointing and Dragging Interface for Naked-Eye Stereoscopic Displays

Raku Egawa
Shibaura Institute of Technology
Tokyo, Japan

Yucheng Qiu
Shibaura Institute of Technology
Tokyo, Japan

Takashi Ijiri
Shibaura Institute of Technology
Tokyo, Japan

ABSTRACT

In this work, a ray-casting-based three-dimensional (3D) pointing and dragging interface for naked-eye stereoscopic displays is proposed. When a user holds a stylus and points it to a display, the proposed system displays a ray, which extends from the stylus to the virtual space in the display. This ray can be used to interact with objects in the virtual space. By conducting a user study, we found that the proposed method allows users to perform 3D pointing with smaller hand movements compared to a hand-capture-based interface. A model, which extends Fitts's law, is also proposed. This model is capable of well predicting the time required for a 3D pointing task.

ACM Reference Format:
Raku Egawa, Yucheng Qiu, and Takashi Ijiri. 2020. Ray-Casting-based 3D pointing and Dragging Interface for Naked-Eye Stereoscopic Displays. In *26th ACM Symposium on Virtual Reality Software and Technology (VRST '20), November 1–4, 2020, Virtual Event, Canada.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3385956.3422102

1 INTRODUCTION

A naked-eye stereoscopic display is a device, which allows users to perceive depth by providing parallax images. A light-field display capable of providing dozens of images in different viewing directions is an example of this type of display. Mouse-based and hand-capture-based methods have been commonly used for interactions in three-dimensional (3D) scenes presented by a light-field display. However, the mouse-based method does not allow users to define depth. The hand-capture-based method allows users to specify 3D position and grab/release states using hand gestures. However, accurate 3D pointing is difficult with this method, since unintended hand movements are often caused by hand gestures.

Ray-casting-based methods have been widely used for 3D pointing in virtual reality environments [3, 5]. Ray-casting-based pointing is also considered useful for volumetric displays [1]. However, there is a limited number of research studies on ray-casting-based interactions in light-field display environments.

In this work, a ray-casting-based method for 3D pointing and dragging in a light-field display environment is presented. The proposed prototype is shown in Figure 1. When a user holds a stylus and points it to a display, a ray extending from the stylus to the virtual space in the display is presented. Using this ray the user can point, select, and drag objects. This method provides users with the feeling of direct manipulation as it seamlessly connects the stylus in the real space with the ray in the virtual space.

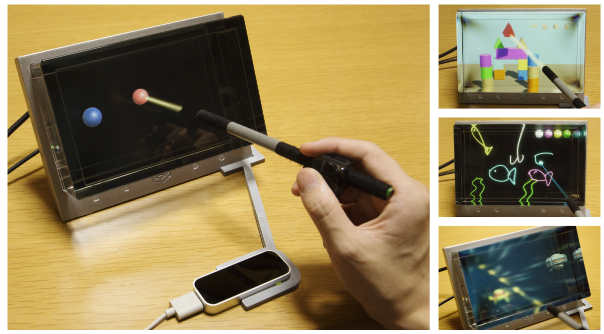


Figure 1: Ray-casting-based 3D pointing interface. A user can point and drag objects with a virtual ray extending from the stylus.

2 PROTOTYPE SYSTEM

The proposed prototype system consists of a stylus (which is a stick with a wireless button for clicking), Looking Glass (which is an 8.9-inch light-field display), and Leap Motion (which is a dual-lens infrared camera for tracking the stylus). We attach a reflective marker on the stylus and capture its 3D position using infrared stereo images, which are provided by the Leap Motion. Then, a ray is generated in the virtual space of the Looking Glass.

We implemented an application, which supports 3D pointing and dragging, for the user study presented below (Figure 1 left). A user can select an object by crossing the ray over it and pressing the button on the stylus. Then, the user can drag the object while keeping the button pressed. During dragging, the length of the ray is fixed. We also implemented other applications such as building-blocks manipulation, 3D drawing, and a shooting game (see Figure 1 right and supporting video for details).

3 EXPERIMENTS AND RESULTS

3.1 User Study

To evaluate the usability of the proposed method, we conducted a user study. We compared the proposed method with hand-capture-based methods using two different scaling parameters such as Hand_50 and Hand_100. For Hand_50 and Hand_100, the scale of the virtual space is 50% and 100% of the real space, respectively.

Three undergraduate students participated in the experiments. After a three-minute tutorial, each participant performed the following task; the system presents a target object (red sphere) and a

VRST '20, November 1–4, 2020, Virtual Event, Canada

ACM ISBN 978-1-4503-7619-8/20/11.
https://doi.org/10.1145/3385956.3422102

goal object (translucent-blue sphere) at random positions in the virtual space, and the user grabs the target, drags it, and releases it at the goal. The diameter of the target object was fixed to 15 mm. The diameter of the goal object varied randomly within the range 10–50 mm. Each participant performed the dragging operation for 100 times using the proposed ray-casting method (Raycast), Hand_50, and Hand_100. The order of methods was balanced among the participants.

3.2 Results and Discussion

We recorded the time and hand movement distance from the users grab and start dragging target objects until they release them. Due to equipment problems and measurement errors, it was not possible to record multiple tasks. As a result, 157 records were collected using our method, 283 records were collected using Hand_50, and 288 records were collected using Hand_100.

3.2.1 Hand movement. The hand-movement distance for the pointing tasks using three different methods is shown in Figure 2. The horizontal axis represents the distance between the target and goal objects when starting dragging; the data were divided into equally spaced bins of 20 mm size each and the average of hand-movement distance in each bin was plotted. It was found that the hand movement of the proposed method was the smallest compared with those of Hand_50 and Hand_100. Using the hand-capture method, users can drag objects by actually moving their hands. Additionally, the proposed Ray-casting method allows users to rotate their wrist to mainly drag objects. Based on this difference, the proposed method can achieve a hand-movement reduction for the dragging operation.

3.2.2 Extension of Fitts's Law. To explain a user's behavior regarding 3D dragging with the proposed method, we propose a predictive model, which extends Fitts's law. We first apply the original Fitts's law [2] as a baseline:

$$MT = a + b\,ID \tag{1}$$

$$ID = \log_2\left(\frac{A}{W} + 1\right), \tag{2}$$

where MT represents the movement time and ID is the index of difficulty of the pointing task. A is the distance between the target and goal objects and W is the diameter of the goal object. a and b are coefficients obtained from the experimental results. The regression analysis with this model Eq. (1-2) on the data collected using our method showed a low determination coefficient, $R^2 = 0.326$.

Using the proposed method, the user can specify the direction and depth by rotating their wrist and moving their hand, respectively. To take both into account, we extend Fitts's law as follows:

$$MT = a + b\,ID_{angle} + c\,ID_{depth} \tag{3}$$

$$ID_{angle} = \log_2\left(\frac{\alpha}{\omega} + 1\right) \tag{4}$$

$$ID_{depth} = \log_2\left(\frac{D}{W_t + W_g} + 1\right), \tag{5}$$

where MT represents the movement time. ID_{angle} represents the index of difficulty with respect to the direction specification by wrist rotation; it is designed based on [4]. α is the angle between the centers of the target and goal objects when starting dragging

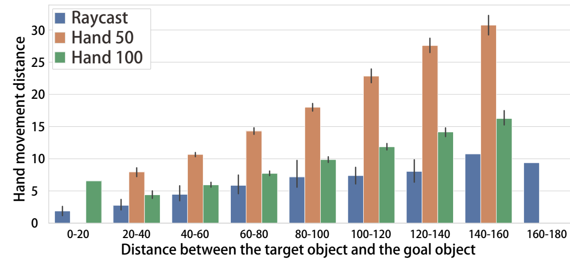
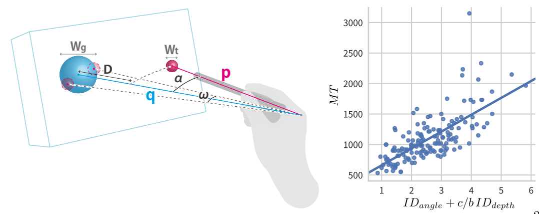


Figure 2: Hand-movement distance for dragging objects.



$a = 382.86, b = 276.37, c = 573.42 \quad R^2 = 0.500$

Figure 3: Illustration of a 3D pointing task (left) and regression analysis results using the proposed model Eq. (3-5). The horizontal axis of the chart represents $ID_{angle} + c/b\,ID_{depth}$.

and ω is the angle of the centers of two objects when they make contact (Figure 3 left). ID_{depth} represents the index of difficulty with respect to the movement in depth, where D is the absolute difference between the depths of the target and goal objects and W_t and W_g are the diameters of the target and goal objects. a, b, and c are coefficients obtained from the experimental results.

Figure 3 right shows the results of regression analysis with the proposed model Eq. (3-5) on the collected data using our method. This model provided a higher determination coefficient, $R^2 = 0.500$, than the original Fitts's model Eq. (1-2). This result suggests that the time required for 3D pointing using the ray-casting method can be explained well with the proposed model.

4 CONCLUSION

In this work, a ray-casting-based method for 3D pointing and dragging in light-field display environments was proposed. The proposed method was compared with commonly used hand-capture-based methods. We found that the proposed method reduces the hand movement for a 3D pointing task. A model, which extends Fitts's law, was also proposed. It was found that this model is capable of well predicting the time required for 3D dragging. Our current limitation is that, when the viewpoint moves up or down, the stylus and the ray are not aligned in a straight line. We would like to solve this issue by performing head tracking. Other future work includes larger scale user studies and evaluation of users' mental workload when using the proposed method.

ACKNOWLEDGMENTS

Part of this paper was presented at the 186th Human-Computer Interaction Conference in Japanese [6]. This work was funded by the Grant-in-Aid for Scientific Research of Japan (15H05924).

REFERENCES

[1] Ravin Balakrishnan, George W Fitzmaurice, and Gordon Kurtenbach. 2001. User interfaces for volumetric displays. *Computer* 34, 3 (2001), 37–45.

[2] Paul M Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology* 47, 6 (1954), 381–391.

[3] Jaemoon Jung, Hanjun Park, Dongwook Hwang, Minseok Son, Donghyun Beck, Juhee Park, and Woojin Park. 2014. A Review on Interaction Techniques in Virtual Environments. *Proc. IEOM 2014* (2014), 1582–1590.

[4] Regis Kopper, Doug A Bowman, Mara G Silva, and Ryan P McMahan. 2010. A human motor behavior model for distal pointing tasks. *International journal of human-computer studies* 68, 10 (2010), 603–615.

[5] Mark R. Mine. 1995. Virtual Environment Interaction Techniques. *Chapel Hill Computer Science Technical Report* TR95-018 (1995).

[6] Raku Egawa, Takashi Ijiri. 2020. Raycast-based Selection and Dragging Interface for Naked-Eye Stereoscopic Display (in Japanese). *IPSJ SIG Technical Report* 2020-HCI-186, 15 (2020).