Binocular Cursor: Enabling Selection on Transparent Displays Troubled by Binocular Parallax

Joon Hyub Lee, Seok-Hyung Bae
I^{2}DEA Lab, Department of Industrial Design, KAIST
291 Daehak-ro, Yuseong-gu, Daejeon 305-701, Republic of Korea
joonhyub.lee | seokhyung.bae @kaist.ac.kr

ABSTRACT
Binocular parallax is a problem for any interaction system that has a transparent display and objects behind it, as users will see duplicated and overlapped images. In this note, we propose a quantitative measure called Binocular Selectability Discriminant (BSD) to predict the ability of the user to perform selection task in such a setup. In addition, we propose a technique called Binocular Cursor (BC) which takes advantage of this duplicating and overlapping phenomenon, rather than being hampered by it, to resolve binocular selection ambiguity by visualizing the correct selection point. An experiment shows that selection with BC is not slower than monocular selection, and that it can be significantly more precise, depending on the design of BC.

Author Keywords: Transparent display; binocular parallax.
ACM Classification: H.5.2 [Information Interfaces and Presentation]: User Interfaces - interaction styles.
General Terms: Design; Human Factors.

INTRODUCTION
Transparent displays are on the verge of commercialization. Users can look at physical and virtual objects through a transparent display, and combined with touch capability and augmentation techniques, they will be able to interact with it, even without additional gears.

However, unlike opaque displays, transparent displays are affected by binocular parallax, which occurs because a person's left and right eyes are horizontally offset and therefore see two different images, with 'convergence' determining how the two images are combined. When a person converges on the near object, the images are combined to produce a focused image for the near object, creating a duplicated and overlapped image for the distant object, and vice versa. When the user looks concurrently at objects at different distances through a transparent display and interacts with them, the parallax can potentially degrade the usability.

We performed a simple demonstration (Figure 1) with two cameras separated by 7 cm (simulating eyes) placed 45 cm from the fingertip (simulating comfortable arm length interaction on a transparent tablet), and a 4 cm diameter Ping-Pong ball (simulating the bounding sphere of a behind object). In this configuration object distance as small as 30 cm caused significant parallax, rendering even basic interaction such as pointing difficult.

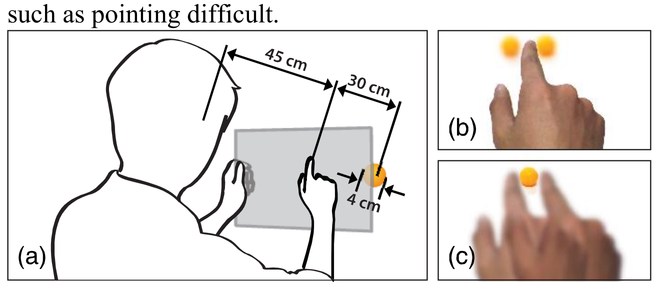

Figure 1. Binocular parallax can render interaction through a transparent tablet difficult (a). Images from left and right eyes are combined to create focused images for the fingertip (b) and the behind object (c).

In this note, we explain an approach of quantifying this problem, 'Binocular Selectability Discriminant (BSD)', a concept first introduced in an earlier work-in-progress publication [7], and an interaction technique, 'Binocular Cursor (BC)', that eliminates the ambiguity caused by the parallax, through appropriate visualization of the selection point.

RELATED WORK
A transparent display can serve as an augmentation window and display useful overlay information in a stationary setup, such as machining parameters over a CNC machining tool [11] and 3D annotation over a holographic 3D model [1]. In addition, when it is mobile and touch-enabled, a user can use the 'transparent tablet' to look at out-of-reach visual content such as a large-scale 3D model of a city, and manipulate it comfortably [9], through image plane interaction [12]. However, such interactions assume monocular vision, and binocular vision can cause ambiguity (Figure 1).

Some setups that use transparent panels as reflectors [3, 4] are configured such that when the overlay image from the source is reflected, it is at the same depth as the behind objects. These setups are unaffected by binocular parallax. We note that only the setups that use a transparent display as is [6], not as a reflector, are affected.

Stereoscopic 3D displays with touch input are also troubled by binocular ambiguity [14]. When selectable objects were displayed behind the screen stereoscopically, and the users were asked to select them by touching on the screen surface, they tended to select with the finger seen by the dominant eye, and also with the middle point in between the two duplicated finger images. It was found that selection performance could be enhanced by interpreting a selection as the touching with a specific point in between the two points.

CHI 2013, April 27–May 2, 2013, Paris, France.

However, even with such a remedy, selection is still ambiguous for the user, making it difficult to select from a crowded scene or perform more complex tasks. To overcome this problem, some selection techniques to disambiguate selection have been suggested [13], such as displaying a cursor selectively to the dominant eye only. Unfortunately, such a technique is specific to stereoscopic setups only and inapplicable for direct touch input. Thus, an unambiguous selection technique specific to touch-capable transparent displays remains to be studied.

BINOCULAR SELECTABILITY DISCRIMINANT

We define Binocular Selectability Discriminant (BSD), which quantifies the extent of binocular parallax and tests whether it will cause a problem when a user performs a selection task on a transparent tablet. We chose selection because it is often essential for higher level tasks.

In the simplified model (Figure 2), a user with an eye-to-eye distance of '*L*' (~7 cm) holds the transparent display at distance '*D*' (~45 cm), and attempts to select an object with width '*w*' and distance '*d*' from the display by placing his finger on the display. When the user converges on the object, the image of the finger is duplicated with distance '*p*' apart. From similar triangles, *p* is *Ld/D*.

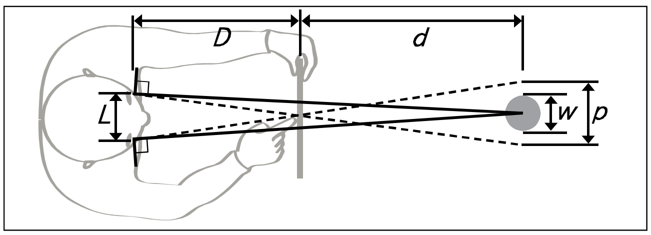

Figure 2. Diagram for deriving Binocular Selectability Discriminant (BSD).

For selection to be possible, *p* must be smaller than *w*, not considering the finger thickness. The discriminant writes:

$$\mathrm{BSD} \equiv \frac{w}{d} - \frac{L}{D}$$

Large BSD corresponds to easy selection (Figure 3a), BSD of 0 to barely possible selection (Figure 3b), and BSD of less than 0 to impossible selection (Figure 3c). This simple model allows us to predict the user's ability to make an unambiguous selection depending on the above parameters.

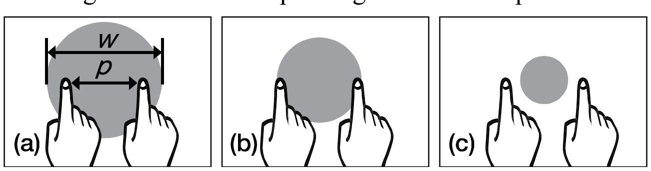

Figure 3. Easy to select (BSD > 0) (a), barely able to select (BSD = 0) (b), unable to select unambiguously (BSD < 0) (c).

BINOCULAR CURSOR

Difficulty of Visualizing Selection Point

When binocular ambiguity is beyond a certain threshold (BSD < 0), explicit visualization of the selection point can enable unambiguous selection. However, such a visualization is difficult: when the user attempts to select an object, the user converges on the object [10], causing any imagery displayed on the nearer transparent tablet to become duplicated and overlapped. To avoid this, visualization can be placed at the object distance so that the user is able to converge on the object and visualization simultaneously, but this requires additional hardware such as a network of projectors configured to cover the entire selection space. Such a setup can be undesirable, especially for mobile use.

Binocular Cursor (BC) Concept

We present Binocular Cursor (BC), which visualizes the selection point directly on the transparent tablet, taking advantage of the duplicating/overlapping phenomenon, rather than being hampered by it. When a user attempts to select a behind object, two partial cursors appear on the left and right of the finger. These partial cursors are arranged such that when the user converges on the distant object, the partial cursors appropriately duplicate and overlap to complete the cursor (Figure 4). BC appears upon touch and is maintained while the user moves it around, and selection is made when the finger lifts, as in an offset cursor [15]. Even with completed BC, the partial cursors are still visible, but we assume that users will focus on the completed cursors and not pay attention to the peripheral artifacts.

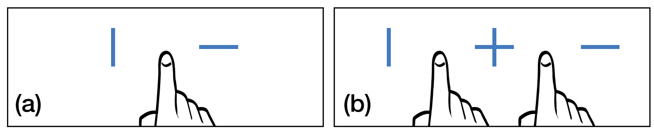

Figure 4. Partial cursors (a) adequately duplicate and overlap to create a complete Binocular Cursor (BC) when the user converges on the behind object (b).

Binocular Cursor Design

Partial cursors can be horizontally shifted to visualize different selection points. The two most intuitive selection points are at the dominant eye image of the finger (DE BC) (Figure 5) and the middle of the two finger images (ME BC, where ME stands for 'middle eye') (Figure 4) [14]. While horizontal and vertical partial cursors can switch sides for ME BC due to symmetry, they cannot for DE BC, because the horizontal line needs to be placed on the finger and the vertical line on the left (right DE) or right (left DE) to avoid finger occlusion. In addition, DE BC requires more space and selection near the display edges can be troublesome, as one of the partial BC can go outside the screen. But this may not be problematic as the users would normally point their mobile devices to the object of interest, and not deliberately use peripheral regions for interaction.

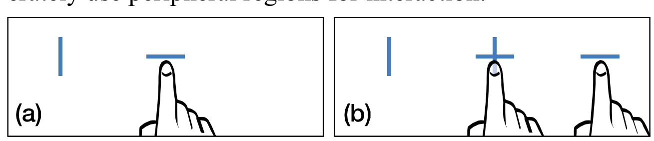

Figure 5. Partial cursors horizontally shifted (a) to place BC on the dominant eye image of the finger, converged (b).

Moreover, BC can be designed to create an area cursor (○BC) that can increase the effective target size (Figure 6), with which users can select smaller targets faster with less error [2], compared to a crosshair BC (+BC). The size of the area cursor can be adjusted depending on the object density in the user's view of the scene through the tablet.

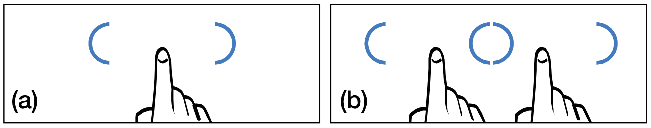

Figure 6. Partial cursors designed to create an area BC (a), converged (b).

Data Acquisition

To implement BC, the system requires the 3D positions of the two eyes, and the positions and sizes of the selectable objects behind the tablet. While the eye positions relative to the tablet can be obtained directly using commercial sensors, different strategies are needed to obtain sizes and distances of selectable objects of different types. When selectable objects are virtual, e.g. 2D and 3D objects displayed on a distant screen, or holograms [1, 9, 11], their sizes and relative positions are usually known. When selectable objects are physical, the surrounding environment needs to be scanned and segmented [5]. When selectable objects are environmental, the tablet's position and orientation relative to a pre-surveyed environment are needed.

EVALUATION

An experiment was conducted to verify 6 hypotheses:

H1. BCs will be as quick as ONE EYE.
H2. BCs will be as precise as ONE EYE.
H3. DE BCs will be quicker than ME BCs.
H4. DE BCs will be more precise than ME BCS.
H5. ○BCs will be quicker than +BCs.
H6. ○BCs will be more precise than +BCs.

Apparatus & Implementation

We created a transparent tablet using a 15″ pressure-sensitive transparent overlay panel, a commercially available component that adds basic single touch capability to an LCD screen. Since it is cheap and opaque enough (20% opaque), we used it as a rear-projection screen, upon which BCs were projected (Figure 7).

To focus on testing the feasibility of BC, the degree of freedom was minimized: the position of the tablet was fixed by mounting it on a custom-built profile structure, and the positions of the user's eyes were also fixed with a chin rest planted 45 cm in front of the tablet (Figure 7a).

The experiment was conducted in a large lecture room with low lighting condition. The targets simulating bounding spheres of selectable objects of variable sizes were projected onto the matt, white wall with another projector, with the target distance varied by moving the setup closer to or farther away from the wall (Figure 7a).

We used 4-point calibration to align the 4 corners of the target projection area on the distant wall with that of the transparent tablet in the perspective of each of the two eyes.

Participants

12 volunteers (3 female, 9 male), with age ranging from 22 to 32 participated in the experiment. All were right-handed. We used the Porta test and the Dolman test to determine the eyedness [8]. 11 were right and 1 was left eye dominant.

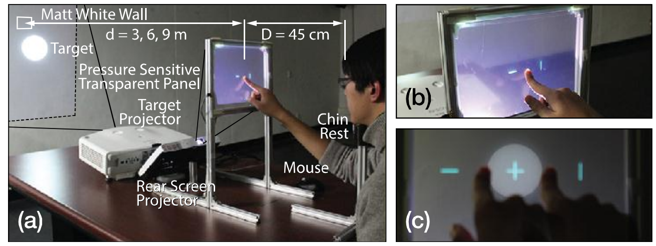

Figure 7. The experiment setup (a); rear-projected, touch-sensitive transparent display (b); a combined image to simulate the user's view of BC while converging on the target (c).

Procedure & Task

After the pretests, a warm-up session was held for each participant for about 10 minutes. During this session, the eye-to-eye distance (*L* in Figure 2), the distance between the centers of pupils, was measured and finely adjusted until the participant could see a correctly converging BC.

The participant selected targets appearing at random locations by touching on the tablet with the index finger of the dominant hand (Figure 7b). The participant was instructed to use a mouse at a specified position and to right-click with the dominant hand to initiate each selection task. The participant lifted the finger off the screen when they judged that a selection was made. No visual feedback was given for correct or incorrect selection to allow for selection error. Each experiment lasted about 40 minutes.

Design

A repeated measures within-participant design was used. The independent variables were: target distance (3, 6, 9 m, *d* in Figure 2 & Figure 7a); cursor type (ONE EYE, +ME, +DE, ○ME, ○DE); target size (apparent diameters 1/16, 1/8, 1/4, 1/2, 1, 2 times the finger images separation distance *p* in Figure 3a & Figure 7c). Each participant made 450 selections (3 target distances × 5 cursor types × 6 target sizes × 5 blocks).

We counterbalanced the presentation order of target distance with a balanced Latin square. At each target distance, the participant first selected with the non-dominant eye closed (ONE EYE) without BC, and then with different BCs in a counterbalanced presentation order. The sizes of BCs were set to span 1/4 of the apparent distance between the finger images (Figure 7c). For each cursor type, differently sized targets appeared in a random order, each appearing 5 times.

Results & Discussion

Selection time (ms), defined as the time taken from the right-click initiation to the last lifting of the finger, and selection error, defined as the on-screen distance (mm) between the nearest correct selection point to the touched point, were the dependent variables.

Repeated measures ANOVA showed significant main effects for cursor type ($F_{4,44} = 19$, $p < .01$) and target size ($F_{5,55} = 98$, $p < .01$) on selection time, but not for target distance ($F_{2,22} = .015$, $p = .86$) (Figure 8).

Post-hoc pairwise comparison test showed that selection time differences were not significant between ONE EYE and ○DE (p = 1.0), not between ONE EYE and ○ME (p = .052), significant between ONE EYE and +ME (p < .01), and between ONE EYE and +DE (p < .01), thus *partially confirming H1*. The differences were significant between ○ME and ○DE (p < .05), but not between +ME and +DE (p = 1.0), thus *rejecting H3*. In addition, the differences were significant between +ME and ○ME (p < .05), and between +DE and ○DE (p < .01), thus *confirming H5*.

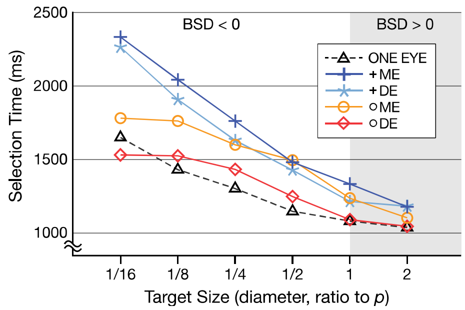

Figure 8. Selection time for cursor types and target sizes.

Repeated measures ANOVA showed significant main effects for cursor type ($F_{4,44}$ = 17, p < .01) target size ($F_{5,55}$ = 38, p < .01) on selection error, and target distance ($F_{2,22}$ = 3.7, p < .05) (Figure 9).

Post-hoc pairwise comparison test showed that selection error differences were not significant between ONE EYE and +ME (p = .65), and between ONE EYE and +DE (p = 1.0). ○ME and ○DE were both significantly more precise than ONE EYE (p < .01), thus *confirming H2*. The differences were not significant between +ME and +DE (p = 1.0) and between ○ME and ○DE (p = 1.0), thus *rejecting H4*. Lastly, the differences were significant between +ME and ○ME (p < .05) and between +DE and ○DE (p < .01), thus *confirming H6*.

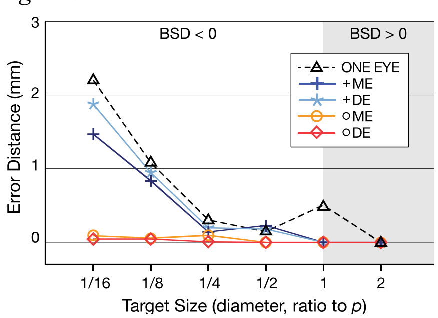

Figure 9. Error distance for cursor types and target sizes.

We compared BCs against monocular selection because it was the only comparable technique with which users could select unambiguously. However, many people cannot wink voluntarily, and even for those who can, it can be fatiguing if it lasts for more than a short period of time. In our experiment, some participants were allowed to block one eye with the non-selecting hand, but in practice, they would not be able to do so because he/she would be holding the mobile transparent tablet with it. Still, monocular image plane interaction by itself is effective and efficient [12], and the fact that selection with BC can be as quick and also more precise, without the inconvenience, shows BC's usefulness.

CONCLUSION AND FUTURE WORK

Unlike opaque displays, transparent displays are inherently susceptible to usability degradations caused by binocular parallax. We suggested a measure (BSD) that can quantify this problem and a technique (BC) that rather uses the parallax to enable unambiguous selection, with competent performances. In the near future, we will implement BC with higher DOF so that both the face and transparent tablet can move, and evaluate it in actual application.

REFERENCES

1. Bimber, O. (2006). Augmenting holograms. *IEEE CG&A*, 26(5), 12-17.
2. Grossman, T. & Balakrishnan, R. (2005). The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. *Proc. CHI'05*, 281-290.
3. Hachet, M., Bossavit, B., Cohé, A., & de la Rivière, J. (2011). Toucheo: multitouch and stereo combined in a seamless workspace. *Proc. UIST'11*, 587-592.
4. Hilliges, O., Kim, D., Izadi, S., Weiss, M., & Wilson, A. (2012). HoloDesk: direct 3D interactions with a situated see-through display. *Proc. CHI'12*, 2421-2430.
5. Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., & Fitzgibbon, A. (2011). KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. *Proc. UIST'11*, 559-568.
6. Lee, J. & Boulanger, C. (2012). Direct spatial interactions with see-through 3D desktop, *SIGGRAPH'12 Poster*.
7. Lee, J., Bae, S., Jung, J., & Choi, H. (2012). Transparent display interaction without binocular parallax. *Adj. Proc. UIST'12*, 97-98.
8. Mapp, A., Ono, H., & Barbeito, R. (2003). What does the dominant eye dominate? A brief and somewhat contentious review. *Perception & Psychophysics*, 65(2), 310-317.
9. Schmalstieg, D., Encarnação, L., & Szalavári, Z. (1999). Using transparent props for interaction with the virtual table. *Proc. I3D'99*, 147-153.
10. Sherstyuk, A., Dey, A., Sandor, C., & State, A. (2012). Dynamic eye convergence for head-mounted displays improves user performance in virtual environments. *Proc. I3D'12*, 23-30.
11. Olwal, A., Gustafsson, J., & Lindfors, C. (2008). Spatial augmented reality on industrial CNC machines. *Proc. SPIE'08 Electronic Imaging*.
12. Pierce, J., Forsberg, A., Conway, M., Hong, S., Zeleznik, R., & Mine, M. (1997). Image plane interaction techniques in 3D immersive environments. *Proc. I3D'97*, 39-44.
13. Teather, R. & Stuerzlinger, W. (2011). Pointing at 3D targets in a stereo head-tracked virtual environment. *Proc. 3DUI'11*, 87-94.
14. Valkov, D., Steinicke, F., Bruder, G., & Hinrichs, K. (2011). 2D touching of 3D stereoscopic objects. *Proc. CHI'11*, 1353-1362.
15. Vogel, D. & Baudisch, P. (2007). Shift: a technique for operating pen-based interfaces using touch. *Proc. CHI'07*, 657-666.