Disambiguating Ninja Cursors with Eye Gaze

Kari-Jouko Räihä and Oleg Špakov
Tampere Unit for Computer-Human Interaction (TAUCHI)
Department of Computer Sciences, FIN-33014 University of Tampere, Finland
{kjr, oleg}@cs.uta.fi

ABSTRACT

Ninja cursors aim to speed up target selection on large or multiple monitors. Several cursors are displayed on the screen with one of them selected as the active cursor. Eye tracking is used to choose the active cursor. An experiment with 13 participants showed that multiple cursors speed up the selection over long distances, but not over short distances. Participants felt the technique was fastest with 4 cursors per monitor, but still preferred to have only 1 cursor per monitor for their own use.

Author Keywords

Multiple cursors, Ninja cursors, selection, eye gaze.

ACM Classification Keywords

H.5.2. [Information interfaces and presentation (e.g., HCI)]: User Interfaces − Graphical user interfaces (GUI).

INTRODUCTION

Various approaches have been suggested for improving the standard point cursor in pointing and selecting tasks. Most techniques are based on either modifying target representation (e.g., [2]) or cursor representation (e.g., [5, 6]). A novel approach was suggested by Kobayashi and Igarashi [7]. Their Ninja cursors technique increases the number of cursors on the screen. The object representation is not changed, and the cursor representation is modified only to indicate which cursor on the screen is active.

Multiple cursors are attractive in that less mouse movement is needed, especially for distant targets. The drawback is that extra action is needed to indicate the desired cursor if the default cursor is not the one that the user intends to use. This slows down the interaction.

A different approach to facilitating fast selection is to use faster input modalities, eye gaze in particular. The MAGIC pointing technique [11] warps the mouse cursor to where the user is looking on the screen, in case there is a selectable object sufficiently close by. This can happen either liberally (always) or only after the user has touched the mouse. Warping is fast, but there are two drawbacks. In the liberal technique, the frequent jumping of the cursor may become distractive. In both cases, the inaccuracy of the eye tracker necessitates fine positioning by the mouse in the proximity of the target, which reduces the time saving brought by the quick warp.

To overcome the accuracy problem with selection by eye gaze, target representation techniques [1] and multimodal interaction [8, 9] have been suggested. Both can be effective with special user groups, but neither approach is able to speed up interaction in everyday conditions.

The original Ninja cursors were disambiguated by moving the cursor away from a target. This made the next cursor on the screen active. The process was repeated until the desired cursor became active. When eye tracking is added to the setup, the active cursor can be chosen simply by looking at one of the cursors. This is intuitive, since the cursor is normally looked at anyway when making a selection. It does not need any extra action by the keyboard or other modalities. Importantly, it also avoids the inherent accuracy problems of eye trackers: although the screen may be densely populated with selectable targets, the cursors are much more sparsely located. Even with 9 cursors on one screen, the accuracy to select one of them does not need to be better than 5 degrees. All commercial trackers achieve this with ease.

The use of an additional input channel to choose one cursor location from multiple candidates was previously introduced by Benko and Feiner [3]. Their M^3 (Multi-Monitor Mouse) technique uses head orientation as one option to indicate the active monitor. M^3 is different from out technique because it works with one cursor per monitor and does not easily generalize to multiple cursors per monitor. On the other hand, our technique does not easily generalize to four parallel monitors because of the limitations on viewing angle posed by the eye tracker.

We continue by presenting our design of gaze-enhanced Ninja cursors. We then describe an experiment where the efficiency of varying number of cursors was tested with different target sizes and different target densities. The paper concludes with a discussion of design implications.

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.
Copyright 2009 ACM 978-1-60558-246-7/09/04…$5.00.

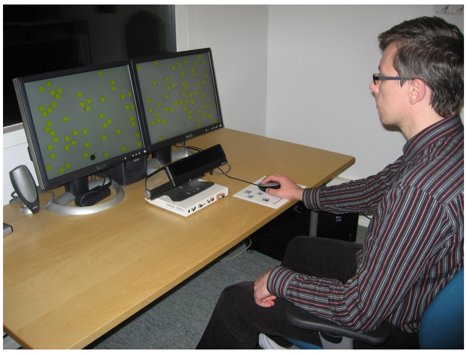

Figure 1. Monitors and eye tracker used in the experiment.

GAZE-ENHANCED NINJA CURSORS

In our implementation, the cursors are visualized as pink circles with diameter of 35 pixels. We support 1, 2, 4 and 9 cursors on each monitor. The screen area on each monitor is divided into a regular grid (2 × 2 or 3 × 3 for the 4 and 9 cursor conditions) so that each cell of the grid hosts one of the cursors. The grid is not shown to the user. One of the cursors is the active cursor, visualized using the normal cursor arrow (the upper right cursor in Figure 2). The cursor closest to the gaze coordinates returned by the eye tracker is selected as the active cursor.

Technically, the cursors are implemented as windows of the Windows environment. Looking at a cursor actually means changing the focus from one cursor window to another, much like in the EyeWindows by Fono and Vertegaal [4].

Although the GUI part of the implementation is simple, a problem was caused by the fact that the coordinate space of the eye tracker is not the same as the coordinate space of the display. The bezel between the monitors has zero width in display space but a width of up to 3 cm in the eye tracked space. This required a new algorithm for mapping the gaze coordinates to display space.

To test the efficiency of the design we carried out a controlled experiment.

METHOD

Participants

We recruited 13 volunteers (7 female, 6 male) from an introductory HCI course at the University of Tampere. They received credit for a weekly assignment for participation. The mean age of the participants was 25.8 years, ranging from 20 to 44 (*SD* = 7.6). They had been using personal computers on average for 10.3 years (*SD* = 3.9).

Apparatus

The experiment was conducted using a two-monitor setup (Figure 1). The resolution of the monitors was 1280 × 1024 pixels. The participants were seated at a distance of 60 cm from the eye tracker. The X120 eye tracker by Tobii Technology was placed in front of the monitors to track the point of gaze in the two screens, resulting in a distance of 85–90 cm between the monitors and a participant's eyes. The width of the viewing field (covering the two monitors) was 857 mm and the height was 307 mm. One degree of viewing angle (the typical accuracy of eye trackers) corresponds to about 50 pixels on screen.

The mouse was a Logitech MX310 with two buttons and a scroll wheel. Mouse pointer speed was set higher than medium, at 80% of the maximum. The software was a Windows XP application that presented the task and saved the mouse coordinates, gaze coordinates, and timestamps at the start and end of each trial.

Task

Figure 2 shows the task (only one screen is shown). The objects are yellow circles and the target object is green. Three object densities (10, 100, and 400 objects) and three object sizes (circles with diameter of 32, 48, and 64 pixels) were used. The number of cursors was 1, 2, or 8 (that is, 1 altogether, 1 per monitor, and 4 per monitor).

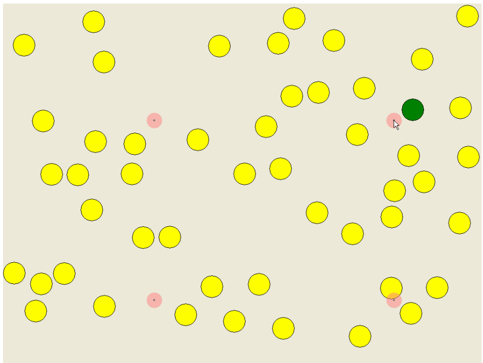

Figure 2. Task presented to participants. Four cursors and 50 objects with 48 pixel diameter on one screen are shown.

For each block the layout of the objects was randomized. Objects were placed on the screens in random positions, but not overlapping. For each trial the target object was randomly selected.

Participants were instructed to work normally, without emphasizing speed or accuracy. Errors were allowed.

Procedure

The eye tracker was first calibrated and the participants carried out two warm-up blocks, 15 trials in each. The first block had 2 cursors and 40 medium size targets, while the second had 8 cursors and 30 large targets. This showed the type of variations that participants could expect during the experiment, but the conditions were still different from those in the actual trials; so, that no trial condition benefited directly from a warm-up trial.

Each block had 10 trials. To start a block the participant clicked anywhere on the screen. This made one of the objects turn green, indicating the first target to select. Each subsequent click changed the target. After the tenth selection the layout changed. Participants were told to rest if desired before starting the next block by clicking anywhere on the screen.

After completing all trials participants filled in a short questionnaire. They were asked questions about the perceived speed of the cursor conditions and their opinions on usefulness and intuitiveness of multiple cursors.

Design

In summary, the experiment was a 3 × 3 × 3 within-subjects design with the following factors and levels:

Number of objects	10, 100, 400
Object size (diameter in pixels)	32, 48, 64
Number of cursors	1, 2, 8

The dependent variables were speed and error rate. The experiment involved 13 participants × 3 object densities × 3 object sizes × 3 cursor conditions × 10 trials = 3,510 trials.

RESULTS

Outliers

Some selection times were exceptionally long. They were caused by the participant losing sight of the cursor or not being able to find the target. However, only the two clearly slowest selections (with selection time over 3.5 seconds) were excluded from the analysis.

Error Rates

There were two types of errors. The first occurred when a non-active cursor was on top of the target. In other words, in the 2 cursor condition the active cursor was on a different monitor from the target in 16 trials out of 1,170, and in the 8 cursor condition the active cursor was one of the 7 other cursors in 29 trials. These trials were excluded from the analysis below since they were caused by system malfunction. This will be explained in the discussion.

After removing these trials, 1.37% erroneous selections (target being missed) remained in the 1 cursor case, 0.43% in the 2 cursor case, and 0.62% in the 8 cursor case.

Speed

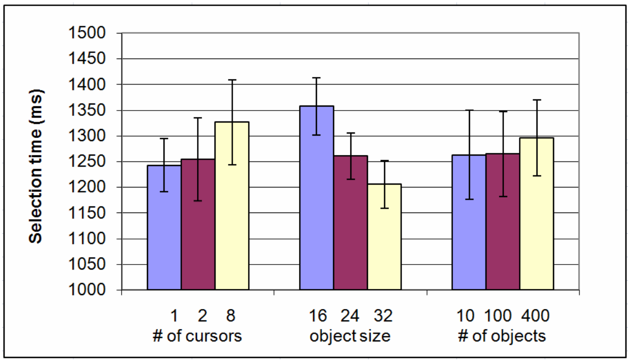

Figure 3. Selection time by independent variable.

Average selection times, grouped by the levels of the three independent variables, are shown in Figure 3. A three-way ANOVA showed a highly significant main effect of object size on selection time ($F_{2,24} = 31.07$, $p < .001$). No other significant main effects were found. Independent samples *t*-tests showed that 1 cursor was significantly faster than 8 cursors ($t(12) = -1.81, p < .05$) and that 2 cursors also were significantly faster than 8 cursors ($t(12) = -1.93, p < .01$).

Figure 4 plots the dependence of selection time on target distance for the three cursor conditions.

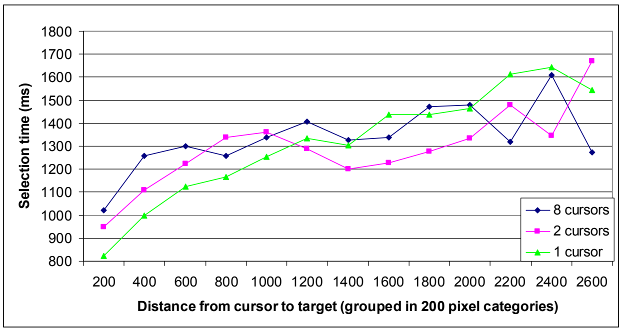

Figure 4. Selection time for each cursor condition as a function of target distance.

The selection times are affected by both the number of cursors and the distance the cursor needs to move. In the 1 cursor condition, the distance is the same as the distance of two consecutive targets. With 4 cursors per monitor the optimal distance is never bigger than the diagonal of a quarter of the screen. As seen in Figure 4, selection times grow steadily in the 1 cursor condition, but when the distance to move is larger than the width of a monitor, multiple cursors start to outperform a single cursor.

A two-way ANOVA showed that the number of cursors × target distance interaction effect is statistically significant ($F_{24,288} = 4.174, p < .0001$).

Subjective preferences

In the post-test questionnaire, seven participants felt they were fastest in the 8 cursor condition and six in the 2 cursor condition. None perceived the single cursor condition as fastest. Nevertheless, only two would have opted for 8 cursors when given the choice. Eight would have chosen 2 cursors (1 per monitor) and three would have preferred a single cursor.

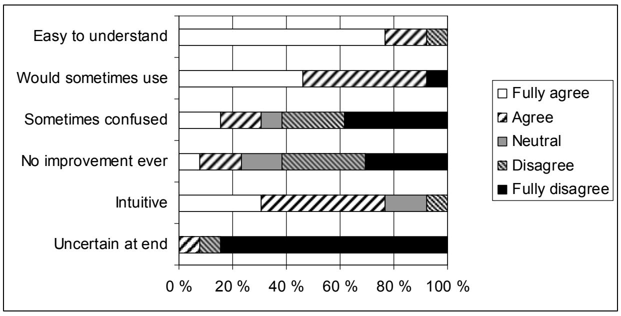

Figure 5. Subjective opinions of the participants.

Participants were asked to indicate on a five-point scale (from fully agree to fully disagree) their agreement with six claims about multiple cursors. Three claims were positive to multiple cursors and three were unfavorable. The distributions of answers are shown in Figure 5.

DISCUSSION

Kobayashi and Igarashi ran an experiment [7] that used mostly the same parameters as ours. With 8 cursors their selection time grew from about 1 s with sparsely positioned large targets to more than 2 s with densely positioned small targets. In our case the similar change is from 1.2 to 1.45 s. With fewer cursors the trend is similar but less pronounced. This indicates that using eye gaze disambiguation improves the technique for varying screen conditions, since it requires the same amount of interaction in all cases. In simple setups, where the default cursor often is the desired one in the original version, the cognitive load brought by the eye tracker makes our technique slightly slower.

There were two types of errors in the test. First, in 2% of the cases where multiple cursors were used, the selection would have been correct had the correct cursor been active. The reason for an incorrect cursor being used was that cursor selection is based on individual gaze samples. Using fixations would have been too sluggish. Individual samples have the problem that if the gaze point is momentarily lost by the tracker (which does happen), the previous known position is used. An improved disambiguation algorithm is needed to solve this problem.

Nevertheless, the results are encouraging. Multiple cursors outperform a single cursor when the target is far from the current location. They also yield fewer selection errors. Low error rate is usually an indication that accuracy was emphasized over speed, suggesting that further speed improvements could have been achieved had the participants been instructed differently. Moreover, our setup is favorable to the single cursor condition (because of the fast mouse gain setting). Adjusting the mouse gain depending on the number of cursors, and including more training could bring additional improvements in efficiency.

It was interesting to note that the participants perceived the 8 cursor condition as fastest, contrary to the measured times. A large majority found the technique intuitive and felt comfortable using it. Walk-up usability seems high and makes the technique a potential choice, if supported by the technology available.

On the other hand, based on the short use of the technique in the test, only 2 out of 13 participants would have chosen more than 2 cursors for their own use. The cognitive difficulties of crossing bezels [10] seem to play a role; most would have been happy with having just one cursor per monitor.

CONCLUSIONS

Multiple cursors disambiguated with eye gaze outperform a single cursor when the distance to target from the current cursor location is high, i.e., when the target is not on the same monitor as the cursor. Multiple cursors are affected by object size and object density as a single cursor.

Further study is needed to analyze eye and mouse movements in synchrony. Knowing how much before the mouse cursor the eyes hit the target, and whether the eyes and the mouse move at the same time, could lead to an improved disambiguation algorithm. We will also run further experiments that include more than 8 cursors. Another ongoing investigation is an experiment for comparing Ninja cursors with MAGIC pointing.

ACKNOWLEDGMENT

We thank Poika Isokoski and I. Scott MacKenzie for their comments on the manuscript. This work was partially supported by the EYE-TO-IT project of EU/FP6/IST/FET, by the COGAIN Network of Excellence, and by the Academy of Finland (project 111658).

REFERENCES

1. Bates, R. and Istance, H. O. Zooming interfaces! Enhancing the performance of eye controlled pointing devices. In *Proc. ASSETS 2002*, 119–126.
2. Baudisch, P., Cutrell, E., Robbins, D., Czerwinski, M., Tandler, P., Bederson, B., and Zierlinger, A. Drag-and-Pop and Drag-and-Pick: Techniques for accessing remote screen content on touch- and pen-operated systems. In *Proc. INTERACT 2003*, 57–64.
3. Benko, H. and Feiner, S. Multi-monitor mouse. In *CHI '05 Extended Abstracts*, 1208–1211.
4. Fono, D. and Vertegaal, R. EyeWindows: evaluation of eye-controlled zooming windows for focus selection. In *Proc. CHI '05*, 151–160.
5. Grossman, T. and Balakrishnan, R. The Bubble Cursor: Enhancing target acquisition by dynamic resizing of the cursor's activation area. In *Proc. CHI '05*, 281–290.
6. Guiard, Y., Blanch, R., and Beaudouin-Lafon, M. Object pointing: a complement to bitmap pointing in GUIs. In *Proc. Graphics Interface 2004*, 9–16.
7. Kobayashi, M. and Igarashi, T. Ninja Cursors: using multiple cursors to assist target acquisition on large screens. In *Proc. CHI '08*, 949–958.
8. Kumar, M., Paepcke, A., and Winograd, T. EyePoint: practical pointing and selection using gaze and keyboard. In *Proc. CHI '07*, 421–430.
9. Miniotas, D., Špakov, O., Tugoy, I., and MacKenzie, I. S. Speech-augmented eye gaze interaction with small closely spaced targets. In *Proc. ETRA 2006*, 67–72.
10. Robertson, G., Czerwinski, M., Baudisch, P., Meyers, B., Robbins, D., Smith, G., and Tan, D. Large display user experience. *IEEE Computer Graphics & Application*, July/August 2005, 44-51.
11. Zhai, S., Morimoto, C., and Ihde, S. Manual and gaze input cascaded (MAGIC) pointing. In *Proc. CHI '99*, 246–253.