An Evaluation of 2D Target Expansion for Gaze-based Interactions

Xinyong Zhang
School of Information, Renmin University of China
Beijing, 100872, P.R. China
x.y.zhang@ruc.edu.cn

Hongbin Zha
Key Lab of Machine Perception, MOE, Peking University
Beijing, 100871, P.R. China
zha@cis.pku.edu.cn

Wei Ma
College of Computer Science and Technology, Beijing University of Technology
Beijing, 100124, P.R. China
mawei@bjut.edu.cn

CHI 2013 Workshop on "Gaze Interaction in the Post-WIMP World", April 27, 2013, Paris, France.
ACM 978-1-4503-1952-2/13/04.

Abstract
We present an experimental evaluation about target expansion for gaze input. Our experiments revealed that (1) target expansion can significantly improve the human performance in gaze-based applications; (2) that the visibility of expansion has notable impact on the stability of eye cursor, but overall it is the final sizes that determine the human performance; (3) and that the performance can be accurately predicted using the performance models we proposed [7, 8] even when targets are expanded in two dimensions.

Author Keywords
Target expansion; gaze input; 2D eye pointing.

ACM Classification Keywords
H.5.2. Information interfaces and presentation: User Interfaces – evaluation, theory and method.

Introduction
Interactive systems, which use the information of eye movements as augmented or direct control signals for human-computer interactions (HCI), have gained increasing acceptance. Related researches have attracted, and continue to attract, substantial attention in HCI community. We proposed a new performance model for eye pointing tasks [7], and we further extended the model to account for the tasks of pointing at general two-dimensional targets [8]. According to our work, increasing target size (e.g. expanding target) is more efficient than decreasing movement distance to improve the human performance.

McGuffin and Balakrishnan [2] as well as Zhai et al. [5] confirmed the advantages of target expansion for hand pointing. Miniotas et al.'s work [3] indicated that target expansion could also benefit eye pointing under one-dimensional task condition. Under two-dimensional task condition, it is reasonable to anticipate that eye pointing can also benefit from target expansion. However, there is lack of sufficient evidence to confirm this anticipation. Especially, some basic but crucial issues still have not been investigated in detail. For example, whether or not an animated expansion could seriously distract the user's attention so as to interrupt the continuity of dwell time (i.e. the process of command activation), and whether or not there are significant time differences between the acquisition of static targets and dynamic targets with equal final sizes.

In this paper, we present the results of two experiments to answer these concerns. One experiment investigated the effects of different factors related to target expansion, the other further verified the effectiveness of target expansion under general 2D target conditions.

Related Work

Used for traditional user interfaces

Fitts' law implies that decreasing movement distance and/or increasing target width can facilitate target acquisition. Therefore, target expansion has the potential to improve human performance in traditional user interfaces. McGuffin and Balakrishnan [2] also pointed out that target expansion was an effective strategy to accommodate the situation of increasing numbers but smaller size of commands and icons in the interfaces of modern software applications. They empirically explored the advantages of target expansion using a pointing experiment [2]. They reported that target expansion technique could improve human performance because overall performance, which still could be effectively modeled and predicted using Fitts' law, was not governed by the initial size W_i of the target but the final size W_f even when expansion took place as late as only 10% of the movement towards the target was uncompleted.

Zhai et al. [5] pointed out that the subjects were able to predict expansion and visualize the enlarged target from the beginning of the trials in McGuffin and Balakrishnan's experiment because the trials were explicitly grouped according to target states (static and dynamic). Therefore, they further investigated and validated the effectiveness of target expansion when subjects could not foresee the expansion as the target would randomly expand, shrink or remain the same.

Unfortunately, target expansion appears to be impracticable in the situation of tiled targets because it is very difficult to determine which target should be expanded owing to the relatively poor accuracy of endpoint prediction [1].

Used for eye-controlled user interfaces

The effectiveness of target expansion for one-dimensional pointing task motivated Miniotas et al. [3] to explore its effect for gaze-based interactions. They argued that the features of eye movements, such as ballistic jumps of the eyes and microsaccades during fixations for visual perception, probably make target expansion gain few benefits, even be a potential distraction. Therefore, the process of expansion was invisible for the subjects in their experiment. In other words, the appearance of the target always kept stationary but with a larger effective width. They reported that target expansion resulted in significant

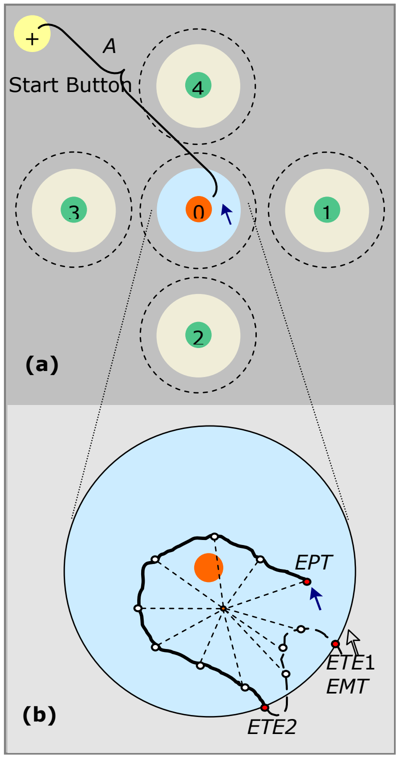

Figure 1. (a) Experimental interface, the dashed circles just show the final size when expanding; **(b)** the data recorded in eye pointing tasks.

decrease in error rate and time of target selection. Špakov and Miniotas [4] further investigated the feasibility of target expansion for eye-based pointing at menus. For the menu they designed, when the user fixates the desired item, its above or below adjacent item, within which the eye tracker reports a gaze point location due to the presence of calibration error, will be highlighted as the possible candidate. After a duration, the candidate expands in vertical directions, pushing the desired item down or up. Accordingly, the new gaze point location reported by the eye tracker also moves down or up because the user continuously fixates the desired one. According to the difference between the gaze locations before and after expansion, the real target item can be determined and expanded for final selection. This technique achieved an acceptable selection accuracy (totally 91%) even when there were significant calibration errors, but with an observable decrease in selection speed.

Note that the actual effect of target expansion in the experiments of Miniotas et al. correctly mapped the user's gaze point to his/her desired target, without obviously improving the efficiency of target selection. The reason of no speedup probably was that the protocol employed to process the gaze signals or the eye tracker's performance limitations partially hid the effectiveness of target expansion.

Experimental Evaluation

We carried out two experiments to investigate the issues of target expansion for gaze-based interactions.

Apparatus

Our experiments were conducted on a 2.7 GHz Pentium™ IV PC running Windows 2000 with a 19-inch CRT display at 1024 × 768 resolution. The gaze input device was a head mounted eye tracking system, EyeLink II, and was installed in a 700 Hz Pentium™ III PC. During the experiments, the gaze data were sampled at 250 Hz in pupil-only mode. The cursor was controlled using the improved speed reduction (iSR) method to alleviate the problem of eye jitters [6].

Experimental Interface, task and procedure

As Figure 1a shows, at the beginning of each trial, the trial start button, which was rendered as a 32-pixel-diameter solid circle but actually with an effective diameter of 120 pixels for the facilitation of trial start, randomly appeared at one of the predefined positions. The subject was instructed to fixate on the start button at first. After the eye cursor dwelt on the button continuously for 450 ms, a target group would appear at a given distance in diagonal direction. At the same time, the start button would disappear. The subject was asked to look at the desired target in the center of the group as soon as possible to select it by dwelling the cursor inside the target for 800 ms.

If the distance from the cursor to the center of any button was less than a predetermined threshold (the button's radius × the specified expansion factor), the button's effective area would be immediately expanded to the final size. There were two kinds of visual feedback for target expansion. One was that the appearance of the button dynamically expanded to the final size at a uniform speed (160 pixels/second). When the cursor left the button, it would shrink to its initial size. The other was that the button kept stationary in appearance. The policy of expanding target could eliminate the time cost of expansion, and the different visual feedback styles (*VF*) were used to reveal the possible visual distraction of dynamic expansion.

Measures used in the experiments

We analyzed the human performance using two main measures: ***Eye movement time (EMT)*** is defined as the time that the eye cursor moves toward a target and reaches it. *EMT* was approximately measured from the moment the target group appeared to the moment the eye cursor entered the target for the first time as Figure 1b shows. ***Eye pointing time (EPT)*** denotes the duration from the moment the trial started to the moment the target was "clicked".

We also defined three measures to describe the features of eye-controlled cursor. As Figure 1b illustrates, during the process of each trial, the eye cursor's positions were sampled at the frequency of 25 Hz, and the number of ***entering target event (ETE)*** was recorded. The ***average distance*** (***AD***) from those sampled points to their center (the mean point) and the ***standard deviation*** of those corresponding distances (***SDD***) were calculated. These three new measures can reflect the eye cursor's stability. The possible visual distraction of target expansion can probably cause noticeable variation in these measures.

When one of the distractors was selected, a wrong-selection event was recorded. If none of the buttons was selected after the trial had begun for five seconds, a no-selection event would be logged. The current trial would be repeated 5 times at most if the subject could not successfully select the desired button.

When the experiments began, the subjects were seated at about 70 cm in front of the screen. Before setting up the eye tracker and calibrated it, the experimenter explained the task and the feature of target expansion by demonstrating it. There was no practice session.

Experiment 1: Expanding circular targets[1]

In order to simplify the experiment design and highlight the main effects of the crucial factors in target expansion, we used circular targets in this experiment.

Design

As illustrated in the bottom of Figure 3, the targets were manipulated to generate 17 diameter conditions including static diameter D_s and expanding diameter D_e with visible or invisible feedback. For D_s, expansion factor (*EF*) can be viewed as 1. There were 26 pairs of targets with equal final sizes. A mixed design of 3*A* × (2 *VF* × 6 D_e + 5 D_s) resulted in 51 combinations. All of them, with two trials for each, composed a block. The subjects were needed to perform 9 task blocks, within which the trials were presented in the same order. But it was randomly generated in advance. A group of subjects (8 females and 10 males with the average age of 26.5 years) successfully finished this experiment. After the experiment, the subjects were asked to complete a questionnaire to indicate their preferences about the different visual feedback styles.

1 The results of this experiment had been partially presented in a local conference in non-English.

Results

The error trials and the outliers that were more than 4 standard deviations from the overall mean *EMT*, totally 4.8% of the data, were excluded from our analysis. The data was divided into two parts according to the components of the mixed design: 3 *A* × 5 D_s and 3*A* × 2 *VF* × 6 D_e. The former was mainly used as a baseline for performance comparisons. The latter was used to reveal the effect of *VF*. No learning effect was found.

AD & *SDD*: *VF* had a significant main effect on *AD* ($F_{1,17} = 6.16$, $p < .05$) but not on *SDD*, and no interaction effect was found. The overall means of *AD* were 6.27 and 6.45 pixels, and *SDD* obtained the overall mean values of 4.03 and 4.00 pixels, respectively for the invisible and the visible expansions.

ETE: We found that there was a significant main effect for the factor *VF* on *ETE* ($F_{1,17} = 11.13$, $p < .005$) as well as a significant interaction effect *VF* × D_e ($F_{5,85} = 3.86$, $p < .005$). The overall mean values of *ETE* were 1.31 and 1.34 under the invisible and visible expansion conditions, respectively.

EMT & *EPT*: There were significant interaction effects between *VF* and D_e on both *EMT* ($F_{5,85} = 10.09$, $p < .001$) and *EPT* ($F_{5,85} = 3.13$, $p < .05$), but only the main effect of *VF* on *EPT* ($F_{1,17} = 7.67$, $p < .05$) was significant. The overall mean *EMT* values were 401.9 and 408.4 ms and those of *EPT* were 1329.7 and 1354.7 ms under the two expansion styles, respectively. Paired-samples *T*-tests indicated that there were 19 and 9 pairs (i.e. 24.4% and 11.5%) of targets that had equal final diameters but had significant differences in *EMT* and *EPT*, respectively, under all three *A* conditions.

Error Rate: No significant main effect for *VF* but an interaction effect *VF* × D_e ($F_{5,85} = 5.09$, $p < .001$) was

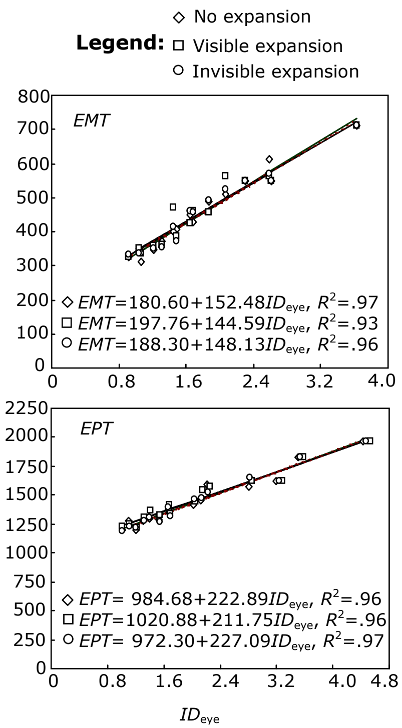

Figure 2. Regression lines of *EMT* and *EPT* in different situations of target expansion. Comparing the regression lines, we can find that the differences among them are very slight. This implies that visible expansion could result in visual distraction, but it almost does not decrease the overall human performance. The questionnaire in Experiment 1 indicated that almost all the subjects preferred to the visible expansion. Therefore, providing a visual feedback for expansion is still a good UI design.

found. The average error rates were 1.7% and 2.1%, under the invisible and visible expansion conditions.

Performance modeling

Although *T*-tests revealed that there were 28 pairs of equal-final-diameters that leaded to significant differences in *EMT* or *EPT*, we only found 5 pairs where the visible or invisible expansion resulted in significant differences compared with the baseline (static targets). This means that both *EMT* and *EPT* were governed by the final target diameter (D_f) in general.

Previously, we proposed a new performance model for eye pointing tasks as follows:

$$T = a + b \times ID_{eye} = a + b \times \frac{e^{\lambda A}}{D - \mu} \qquad (1)$$

where the symbols λ and μ are two empirical constants, and *a* and *b* are two regression coefficients. The variables *A* and *D* denote movement distance and target size (diameter), respectively. ID_{eye} is defined as the index of difficulty of eye pointing task. We found that the grand mean of *AD* can be directly used as the constant μ. As Figure 2 shows, when the final target sizes were used, the performance could be accurately modeled ($R^2 \geq .93$). For *EMT*, ID_{eye} was calculated using $80 \times e^{.0007A}/D_f$; for *EPT*, ID_{eye} was calculated using $80 \times e^{.0005A}/(D_f - \mu)$, where μ =12.4, 13.0 and 12.7, respectively, under the conditions of no expansion, visible expansion and invisible expansion.

Experiment 2: Expanding rectangular targets

The first experiment clarified that the acquisition of expanding targets is governed by the final sizes and it can be accurately predicted using Equation 1. The targets used in Experiment 1 were circular, however, those in real user interfaces are rectangular in general. Recently, We developed an extension of Equation 1

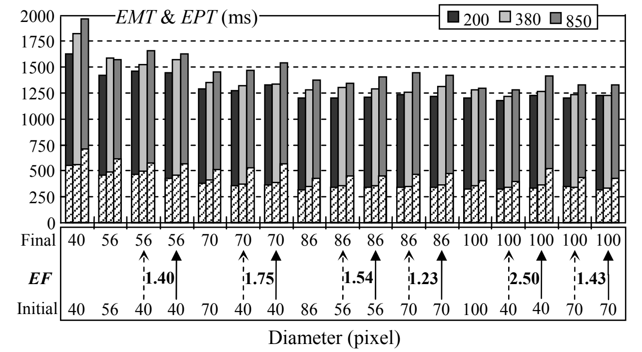

Figure 3. *EMT* and *EPT* by diameter condition for different *A* levels. The dashed arrow denotes invisible expansion; the solid one denotes visible expansion. *EF* denotes expansion factor.

to account for the acquisition of rectangular targets [8]. The new model is as follows:

$$T = a + b \times 80\, e^{\lambda A} \sqrt{\frac{\omega}{(W - \mu)^2} + \frac{1 - \omega}{(H - \mu)^2}} \qquad (2)$$

where ω is a weighted parameter in the range of [0, 1], *W* and *H* denote target width and height, respectively. We found that the parameter ω was able to reveal the asymmetrical impacts of *W* and *H* on eye pointing. The purpose of this experiment was to verify whether Equation 2 could be used to model the acquisition of expanding rectangular targets.

Design

As Table 1 shows, there were 15 combinations of target width, height and expansion factor (*EF*), resulting in 15 different final target sizes. These combinations fully crossed with 2 movement distances (400 and 900 pixels). A task block consisted of 120 trials (15 final sizes × 2 distances × 4 diagonal directions × 1 trial). Twenty one subjects took part in the experiment and they successfully finished 7 task blocks.

Width	Height	*EF*
40	40	1.0
40	40	1.6
40	40	2.2
40	40	3.0
40	40	4.0
40	160	1.0
50	100	1.0
50	100	1.6
60	40	1.0
60	40	1.4
60	40	2.0
80	50	1.0
80	50	1.6
80	50	2.0
150	50	1.0

Table 1. The combinations of initial target width, height and *EF*.

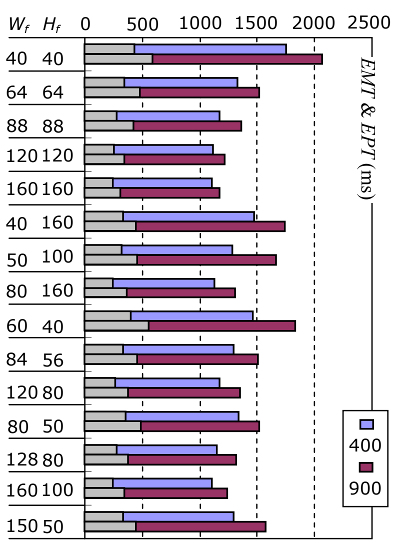

Figure 4. *EMT* and *EPT* by final target size.

Results

The error trials and the outliers, totally 3.0% of the data, were excluded from the analysis. Figure 4 shows the averages of *EMT* and *EPT* by final target size. Model fitting indicated that *EMT* and *EPT* could be accurately predicted using the following two regression equations. Their R^2 values were .948 and .975, respectively.

$$EMT = 175.1 + 89.7 \times 80e^{.001A}\sqrt{\frac{0.39}{W_f^{\,2}} + \frac{0.61}{H_f^{\,2}}} \quad (3)$$

$$EPT = 945.1 + 156.2 \times 80e^{.001A}\sqrt{\frac{0.54}{(W_f - 12.7)^2} + \frac{0.46}{(H_f - 12.7)^2}} \quad (4)$$

Discussion and Conclusions

The analyses of *EMT* and *EPT* confirmed that the performance in eye pointing was negatively affected by visible expansion. Checking those 28 diameter pairs that caused significant differences of *EMT* or *EPT* when the final sizes were equal, we further found that the majority of them occurred in the similar situations when the target visibly expanded to a large diameter, such as 86 and 100 pixels, especially from 40 to 100 pixels (4 pairs for *EPT*, and 8 pairs for *EMT*). According to the significant interaction effect $VF \times D_e$, this could plausibly be interpreted to mean that the visual distraction became increasingly serious when the target expanded to bigger size.

Although expanding target sizes has greater benefit than decreasing movement distance [7], this does not mean that the UI designer should expand targets as large as possible. According to our regression equations, we can figure out that when D_f is greater than 85 pixels, it is unnecessary to expand targets any more because of the negligible "marginal benefit" of expanding.

This study solved our concerns about target expansion when used for gaze input. We found that the human performance in eye pointing tasks was governed by the final target sizes and could be accurately predicted using the models we previously proposed.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC, Grant No. 61070144). It was also supported in part by the NSFC of Grant No. 61003105.

References

[1] Lank, E., Cheng, Y.C.N. and Ruiz, J. Endpoint prediction using motion kinematics. In *Proc. CHI 2007*, ACM Press (2007), 637-646.

[2] McGuffin, M. and Balakrishnan, R. Acquisition of expanding targets. In *Proc. CHI 2002*, ACM Press (2002), 57-64.

[3] Miniotas, D., Špakov, O. and MacKenzie, I.S. Eye gaze interaction with expanding targets. In *CHI 2004 Extended Abstracts*, ACM Press (2004), 1255-1258.

[4] Špakov, O. and Miniotas, D. Dynamic target expansion to facilitate eye-based pointing at menus. Information Technology And Control}, Vol.34, No.2 (2005), 135-139.

[5] Zhai, S., Conversy, S., Beaudouinlafon, M. and Guiard, Y. Human on-line response to target expansion. In *Proc. CHI 2003*, ACM Press (2003), 177-184.

[6] Zhang, X., Ren, X. and Zha, H. Improving eye cursor's stability for eye pointing tasks. In *Proc. CHI 2008*, ACM Press (2008), 525-534.

[7] Zhang, X., Ren, X. and Zha, H. Modeling dwell-based eye pointing target acquisition. In *Proc. CHI 2010*, ACM Press (2010), 2083-2092.

[8] Zhang, X. and Zha, H. Human performance modeling for two-dimensional dwell-based eye pointing. In *Proc. PACIS 2012*, Paper 96.