Modeling Error Rates in Spatiotemporal Moving Target Selection

Jin Huang
ISCAS, China
huangjin@iscas.ac.cn

Byungjoo Lee
KAIST, Republic of Korea
byungjoo.lee@kaist.ac.kr

ABSTRACT
When we try to acquire a moving target such as hitting a virtual tennis in a computer game, we must hit the target instantly when it flies over our hitting range. In other words, we have to acquire the target in spatial and temporal domains simultaneously. We call this type of task spatiotemporal moving target selection, which we find is common yet less studied in HCI. This paper presents a tentative model for predicting the error rates in spatiotemporal moving target selection. Our model integrates two latest models, the Ternary-Gaussian model and the Temporal Pointing model, to explain the influence of spatial and temporal constraints on pointing errors. In a 12-subject pointing experiment with a computer mouse, our model shows high fitting results with 0.904 $R^2$. We discuss future research directions on this topic and how it could potentially help the design in dynamical user interfaces.

KEYWORDS
Error Rates; Moving Target Selection; Spatial Pointing; Temporal Pointing.
INTRODUCTION
Moving target acquisition is one of the most fundamental interaction tasks in modern user interface. Understanding the selection precision in such tasks can assist the design and improve user experience of interactive systems with dynamic contents, such as video games, augmented reality (AR) and virtual reality (VR) applications [3–8]. Models have been proposed to predict the error rates in moving target acquisition, e.g., error models for cases requiring relatively large input movements [6, 11] and an error model for cases requiring negligible movements [7–9].

In practice, moving targets usually involve both spatial and temporal constraints. Take a simple tennis game as an example. In the game, player has to hit the tennis ball that flies to him. On one hand, the player has to swing his racket to hit the ball spatially; On the other hand, because the ball is flying from far to near and the player usually has only one chance to hit the ball, the player must estimate the right time to swing the racket. Either hitting the ball too early or too late will lead to a miss. The general form of this task, which we call spatiotemporal moving target selection, is widely found in many other types of games, such as first-person shooting games, or other types of interactive systems, like video surveillance systems and music-based applications. As far as we know, in HCI literature, there is no previous model predicates error rates as a function of the properties (i.e., constraints) in such tasks.

Considering the potential benefits of modeling user performance in spatiotemporal moving target selection, in this study, we first formulate a practicable task that reflects key elements in spatiotemporal moving target selection. And then, we built an integrated model to predict error rates in the task. We evaluated the performance of the proposed model in a 12-subject experiment, and finally summarized further research directions on this topic.

RELATED WORK
In static target selection, it is widely accepted that the rule of speed-accuracy trade-off first revealed by Fitts’ law [2] governs types of user performance such as speed, error and uncertainty [10]. For studies in error rates, Wobbrock et al. [12] proposed a model based on Fitts’ law parameters that predicted the error rates of pointing for static 1D and 2D targets. Zhou et al. [13], on the other hand, investigated user error in trajectory-based tasks [1] with temporal constraints, they found that users made more errors under the conditions with stricter time limits. These works provided references for investigating the error rates of moving targets in this paper.

In moving target selection, errors occur in both spatial and temporal domains. In the spatial domain, errors are usually defined as cursor clicks (i.e., spatial endpoints) that occur outside the target. In a recent study, a Ternary-Gaussian [6] was proposed that promised to predict spatial endpoint distribution by combining the uncertainty caused by size and speed of the target. In the temporal
domain, errors are represented as selection events (i.e., temporal endpoints) that triggered outside of a certain time window. To access the error rates in the temporal domain, the study of [7–9] introduced a task called Temporal Pointing and, proposed a model to predict the error rates in this task. Although these models had got precise predictions, they were limited by considering only the cases requiring negligible input motions. One recent study [11] has succeeded in predicting the pointing error rate even when the user’s motion is large based on the temporal pointing model. However, the model has a limitation that the error rate cannot be calculated only by the task condition given to the user. As a result, this paper would consider both spatial and temporal constraints and try to predict user performance in moving target selection under such constraints.

PROBLEM FORMULATION

Inspired by the previous studies [6–8], we formulate a time-constrained 1D moving target selection task to study this problem (see Figure 1 (a)). In the task, a user controls a 1D cursor (red vertical line) to acquire a moving target (gray vertical ribbon). The target has a specified width and moves from left to right at a specified speed. A selection region (marked yellow) is fixed on the vertical center of the target, and a horizontal time-line (marked blue) moves from top to bottom on the target at a specified speed. To accomplish the task, the user must control the cursor to acquire the target (i.e. spatial target) when the time-line is in the selection region (i.e. temporal target), otherwise the target acquisition will fail (see Figure 1 (b)).

An spatiotemporal target in the above task contains two types of constraints. The first type is the spatial constraints, which include 1) spatial distance \(D_s\): initial distance between the cursor and the vertical target; 2) spatial width \(W_s\): width of the target; and 3) spatial velocity \(V_s\): the moving speed of the target. The second type is the temporal constraints, including 1) temporal width \(W_t\): the time required for the horizontal time-line to pass through the selection region; 2) cue viewing time \(t_c\): the duration from the appearance of the time-line to its arrival at the selection region; and 3) period of input repetition \(P\): the duration between repeated selection. Finally, the problem we try to solve in this study is formulated as finding and evaluating a model that predicts the error rate for a spatiotemporal target with specified task contains of \(D_s, W_s, V_s, W_t, t_c\) and \(P\).

TENTATIVE MODEL

The model predicts the error rate \(E_{s,t}\) for a specified spatiotemporal target. The error rate is defined as the ratio of the number of failed selections to the total number of selections \(E_{s,t} = \frac{\text{number of failed selections}}{\text{total number of selections}}\). As it has been defined, to acquire the spatiotemporal target, the user must successfully acquire both spatial and temporal targets that constitute it. Using \(S_s\) and \(S_t\) to represent the spatial and temporal success rates, then the joint success rate can be written as \(S_s \times S_t\) by assuming
independence of the two events. Finally, the error rate for the spatiotemporal target can be obtained by subtracting the joint success rate from 1 as:

$$E_{s,t} = 1 - S_s \times S_t$$  \hspace{1cm} (1)

We use the **Ternary-Gaussian** model [6] to obtain $S_s$:

$$S_s = \frac{1}{2} \left[ \text{erf} \left( \frac{x_1 - \mu}{\sigma \sqrt{2}} \right) - \text{erf} \left( \frac{x_0 - \mu}{\sigma \sqrt{2}} \right) \right]$$  \hspace{1cm} (2)

where, $\mu = a + bV_s$, $\sigma = \sqrt{d + eV_s^2 + fW_s^2 + g \frac{V_s}{W_s}}$, $x_0$ and $x_1$ are the left and right boundaries of the spatial target, and $a$, $b$, $c$, $d$, $e$, $f$ and $g$ are free parameters.

We use the **Temporal Pointing** model [7] to obtain $S_t$:

$$S_t = \frac{1}{2} \left[ \text{erf} \left( \frac{1 - c\mu}{c_\sigma \sqrt{D_t}} \cdot \frac{W_t}{D_t} \right) + \text{erf} \left( \frac{c\mu}{c_\sigma \sqrt{D_t}} \cdot \frac{W_t}{D_t} \right) \right]$$  \hspace{1cm} (3)

where, $D_t = P/\sqrt{1 + (P/(1/e^{\nu t_c} - 1) + \delta)^2}$, and $c_\sigma$, $c\mu$, $\nu$, and $\delta$ are free parameters. In this experiment, the period of input repetition $P$ was not explicitly given to the participants, so it was empirically determined by averaging the intervals between clicks of participants.

**EXPERIMENT**

We conducted an experiment to evaluate the overall fit of the model with empirical data in a controlled spatiotemporal moving target selection task. The experiment followed a within-subject design with five fully crossed variables include 2 spatial distances ($D_s$: 384 and 768 px), 3 spatial widths ($W_s$: 24, 48, 144 px), 3 spatial speeds ($V_s$: 96, 192, 384 px/sec), 3 cue viewing time ($t_c$: 700, 1100, 1500 ms) and 4 temporal widths ($W_t$: 80, 160, 240, 320 ms). Twelve participants (6 males, 24.6 years old on average) were recruited. All the participants were right-handed, and they were familiar with computer mouse. The experiment was conducted on a Dell OptiPlex 9020 desktop computer, with an Intel Core i7 4 Quad core CPU at 3.6 GHz and a 23-inch (533.2×312mm) LED display at 1,920×1,080 resolution. The pointing device was a Dell MS111 mouse with 1000 dpi (see, Figure 2).

In each trial, a participant clicked the “start” button on the left size of the window to start. After a 1000 ms interval, a moving target displayed, and participants were asked to control the 1D cursor to acquire the moving target. Participants could only acquire the target once per trial. If a target was successfully acquired, it flashed green to feedback. Participants repeated each condition 5 times for a total of $5 \times 2 \times 3 \times 3 \times 3 = 1080$ trials. With 12 participants, this resulted in 12960 total selections. The order of the trials were randomized.
RESULTS AND DISCUSSION

by using the 60 samples from 12 participants to calculate the error rate for each condition, we obtained 216 error rates in total. To determine the relationships between error rate and the task constraints, we tested statistical correlations between $E_{s,t}$, $S_s$ and $S_t$ on each of the 5 task constraints. Both Pearson and Spearman correlation coefficients were calculated and are displayed in Table 1.

The results showed significant correlations between $W_s$ and $S_s$ and between $V_s$ and $S_s$, and no significant correlation was found between $D_s$ and $S_s$. It was consistent with previous study [6], in which the target size and speed significantly affected the uncertainty in spatial moving target selection, but the initial distance did not. We also found significant correlation between $t_c$ and $S_s$. This could due to the effect of time urgency on spatial target selection as mentioned in [12]. Significant correlation between $W_t$ and $S_t$ was found, which was consistent with [7]. However, the correlation between $t_c$ and $S_t$ was weak. This might because we set up a relatively large range of $t_c$ compared with [7]. There was a significant correlation between $W_t$ and $S_t$, which indicated that increasing spatial width could increase temporal success rate. Finally, the joint error rate $E_{s,t}$ were significantly correlated to all the spatial and temporal constants expect the spatial distance. To test whether the majority of these effects had been reflected by our model, we evaluated the overall fit of the model with the empirical data. We reported the fitting results of the Ternary-Gaussian model on the data of spatial error rates, the Temporal Pointing model on the data of temporal error rates, and, the Ternary-Gaussian model, Temporal Pointing model and our proposed model on the data of joint error rates.

The $R^2$ value of the Ternary-Gaussian fit for the spatial error rates was 0.942, indicated that the model well explained the variability of data. The $R^2$ value of the Temporal Pointing fit for the temporal error rates was 0.764. The relative low fit might caused by the unexplained variability of data which introduced by spatial constraints. For the joint error rate, the fitting results were low for both Ternary-Gaussian ($R^2=0.475$, Figure 3 (a)) and Temporal Pointing ($R^2=0.337$, Figure 3 (b)) as we expected. Because both of them only consider one type of constraints in the domains they belong. On the contrast, as showed in Figure 3 (c), when combining the two models with Equation 1, the proposed model achieved high fit on the data with 0.904 $R^2$.

FUTURE WORK

As the first attempt to explain the error rates in spatiotemporal moving target selection, we chose to formulated the task in 1D space. However, by flexibly controlling the task properties, this form could formalize many common scenarios such as piano playing interfaces in practices. Further more, as the spatial and temporal constraints were considered separately in our model, it could be extended to 2D or 3D spaces by modifying only the spatial part of the model. Thus we considered this extension was predictable and would further expand the applicability of the model.
There are still limitations in this pilot study. First, the model was based on an assumption that the error rates in spatial pointing and temporal pointing are independent, which may not be hold in some extreme conditions. Second, the study did not investigate many other types of time constraints (e.g. no temporal distance, or infinite temporal width) that are also common in practices.

A more in-depth and systematic study of this topic could potentially contribute to aspects that are of interest to HCI. For example, it can help us to explore the potential correlations between user performances on spatial and temporal pointing. It can also be used to analyze user response strategies under the two types of constraints, e.g. figuring out which type of constraints will dominate user behavior under a certain circumstance.

ACKNOWLEDGMENTS

This work was supported by National Key R&D Program of China (Grant No. 2016YFB1001405), the National Natural Science Foundation of China (Grant No. 61802379) and Key Research Program of Frontier Sciences, CAS (Grant No. QYZDY-SSW-JSC041).

REFERENCES