Modeling Dwell-Based Eye Pointing Target Acquisition

Xinyong Zhang[◇ § †], **Xiangshi Ren**[‡], **Hongbin Zha**[§]
[◇]School of Information, Renmin University of China, Beijing, China
[§]Key Lab of Machine Perception, MOE, Peking University, Beijing, China
[†]MOM Lab, Beijing Institute of Technology, Beijing, China
[‡]Dept. of Information Systems Engineering, Kochi University of Technology, Kochi, Japan
x.y.zhang@ruc.edu.cn, ren.xiangshi@kochi-tech.ac.jp, zha@cis.pku.edu.cn

ABSTRACT

We propose a quantitative model for dwell-based eye pointing tasks. Using the concepts of information theory to analogize eye pointing, we define an index of difficulty (ID_{eye}) for the corresponding tasks in a similar manner to the definition that Fitts made for hand pointing. According to our validations in different situations, ID_{eye}, which takes account of the distinct characteristics of rapid saccades and involuntary eye jitters, can accurately and meaningfully describe eye pointing tasks. To the best of our knowledge, this work is the first successful attempt to model eye gaze interactions.

Author Keywords

Eye pointing, Fitts' law, information theory, modeling.

ACM Classification Keywords

H.5.2. [Information interfaces and presentation]: User interfaces - *theory and methods*.

General Terms

Experimentation, Human Factors, Measurement.

INTRODUCTION

With the advancement of eye tracking technologies, eye gaze interaction has gained increasing acceptance. Gaze input, with the advantages of being effortless, intuitive, speedy, and less risk of cumulative injuries [33], is a feasible solution to guarantee the accessibility of user interfaces for people with disabilities or those whose hands are being occupied with no chance to input. Besides eye typing [21], there have been a number of novel applications, such as reading assistant [12], coordination of multiple systems [5], game control [29], mobile computing [6], virtual environment [11] and information security [15]. Meanwhile, lots of studies reported attempts to improve the feasibility of gaze input [38]. However, there was still a lack of quantitative model specially developed for the dominant eye pointing tasks of gaze input.

Quantitative models are very important for the advancement of HCI. For example, the well-known Fitts' law model [7] is used as a theoretical tool for the innovation of interface designs [36]. It formulates movement times (MT) taken to manually acquire a target of width W at distance (amplitude) A as quickly and accurately as possible as follows [20]:

$$MT = a + b\log_2(A/W + 1) \quad (1)$$

where both a and b are regression coefficients. The logarithmic term is defined as the index of difficulty (ID) of manual pointing tasks. Unfortunately, Fitts' law appears to be unable to accurately and fully describe eye pointing tasks.

These situations motivated us to propose a new model. Our work is not based on Fitts' law, but we also use the notions of information theory [27], i.e. the theoretical basis of Fitts' law [18], to analogize eye pointing to a hypothetical signal transmission so as to develop our model. Since dwell time is used in general as a substitute for the common "click" due to the "Midas-Touch" problem [13], we confine our work to issues under the condition of dwell-based eye pointing.

RELATED WORK

As we aim to build a predictive model for the basic pointing tasks of gaze input, only the work of modeling different aimed pointing tasks in different situations was reviewed.

Models under One-Dimensional Target Conditions

The pointing task paradigm in Fitts' experiment was a one-dimensional task with no constraint on target height. ID originally defined by Fitts was in the form of $\log_2(2A/W)$. Before the widely-accepted Shannon formula [30] of ID in Equation 1, there already had been other theoretical modifications, such as the close resemblance of the Shannon formula, i.e. the form of $\log_2(A/W + 0.5)$, which could lead to plausible improvement for the correlation between MT and ID [26]. From the perspective of psychomotor behavior, the deterministic iterative-corrections model [4] can be used to take account of the logarithmic relationship between movement times and the term of A/W in Fitts' law, but Meyer et al. [22] argued its limitations and then proposed an optimized dual-submovement model, which assumes that a rapid aimed movement includes a primary submovement programmed to stop at the center of the target and an optional corrective submovement if the initial submovement misses the target. Derived from this model to approximately predict MT, ID then changes into the form of $\sqrt{A/W}$. Meyer et al. further generalized ID to account for multiple submovements so as to model MT as follows [23]:

$$MT = a + b\,(A/W)^{\frac{1}{2}} \Rightarrow a + b\,(A/W)^{\frac{1}{n}} \quad (2)$$

where n denotes the assumed number of submovements involved in rapid aimed movements. Actually, Kvålseth had early proposed a model with a substitution of an empirical constant for the order of $1/n$ as follows [17]:

$$MT = a + b\,(A/W)^c \quad (3)$$

Note that the two independent variables A and W in the formulas from Equation 1 to 3 theoretically have equal rate of contribution to the dependent variable MT (i.e. the property of scale independency [1]). That is to say, if both A and W decrease or increase at the same rate, MT will theoretically keep unchanged. Welford was aware that this potential relationship between A and W seemed to be unreasonable in some special situations [35]. He reformulated ID with two parts, as expressed by Equation 4, to capture the possible different contribution rates of A and W to MT. As expressed by Equation 5, another model proposed by Kvålseth also separates the contributions of A and W [17].

$$MT = a + b\log_2(A) - c\log_2(W) \quad (4)$$

$$MT = a \times W^b \times A^c \quad (5)$$

where a, b, and c are three regression coefficients.

Models under Multidimensional Target Conditions

Pointing to a concrete target is generally constrained by both its width and height. Thus, some researchers have sought for more appropriate models to take account of more factors i.e., factors other than W, to predict MT more accurately.

With respect to bivariate pointing, MacKenzie and Buxton [19] tested the effectiveness of four candidates to replace W in Equation 1. Those candidates were $min(W, H)$, W', $W+H$, and $W \times H$ (the symbol H denotes the target height, and W' denotes the intercept along the motion direction in the target). They reported that substituting $min(W, H)$ for W resulted in the best fit. Accot and Zhai [1] pointed out that the use of $min(W, H)$ made W and H mutually exclusive for MT prediction. It was unreasonable, they believed, that for a given combination of W and H, only one of the two factors was used to determine MT. Thus, Accot and Zhai systematically discussed the harmonic integration of W and H into ID in the form of $\log_2(\sqrt{(A/W)^2 + \eta(H/W)^2}+1)$.

For bivariate pointing, both movement angle and target shape should also be taken into consideration. Accot and Zhai's work as well as the proposal of MacKenzie and Buxton did not include these factors, but on the other hand, the integration of too many variables into the expression of ID could be troublesome. Grossman and Balakrishnan thus proposed a generalized ID, to which the probability of hitting target as the only parameter is mapped using a universal function, so as to take account not only of movement angle but also of the arbitrary shapes of targets [10]. They also extended Accot and Zhai's work for trivariate pointing [9] with the consideration of weights depending on movement angle.

Studies on Performance Modeling for Eye Pointing

Although no study had systematically explored the issue of modeling for eye pointing, some researchers had tested the validity of Fitts' law in this connection. For example, Ware and Mikaelian [34] used Fitts' law to describe their experimental results of eye-based selections, but they did not attempt to draw any formal conclusion. Regarding the narrow ID range in Ware and Mikaelian's experiment, Miniotas [24] carried out a one-dimensional task experiment with extended ID range, still reporting positively enough the fitness of Fitts' law model. Recently, Vertegaal's work seemed to confirm this point again [33]. But note that there was no small target used, weakening the representativeness of the result. In other studies [8, 25, 31, 37], however, it was indicated that the data from eye gaze interactions did not fit into the Fitts' law model as well as traditional pointing tasks did. Therefore, there is no general agreement in the literature about the validity of Fitts' law for eye pointing tasks.

Although Fitts' law can properly account for eye-hand coordination, it is reasonable to expect that Fitts' law would likely be unsuitable for eye pointing tasks, even if it might occasionally fit the data "well enough" [24], because the features of eye movements, namely rapid saccades and jittery fixations, and the feedback-control processes of eye pointing are fundamentally different from those of hand pointing. Unfortunately, no researcher has proposed a new model or any modification of Fitts' law or other unpopular models to appropriately reflect the unique characteristics of eye pointing. Therefore, we address this issue firmly anticipating the development of a theoretical tool for eye-based interactions.

ANALOGY OF EYE POINTING

Fitts' law is a cornerstone for modeling human performance in manual interactions because it can evolve to new forms or be used to derive new models for different interaction paradigms, such as 3D pointing [9], peephole pointing [3] and even an error model [36]. However, we gave up using Fitts' law to derive a new model for eye pointing tasks for the aforementioned reasons. On the other hand, we certainly derived inspiration from the "theoretical basis" of Fitts' law [7, 18, 26, 30] as we briefly explain in the following.

The Origin of Fitts' Law

The widely applied Fitts' law was not based on the knowledge of psychomotor behavior but on Shannon's theorem 17 about signal transmission [27]. In general, a communication system comprises five components: information source, transmitter, channel, receiver, and destination. The transmitter encodes messages produced in the information source into signals. The receiver inversely reproduces the messages from the signals transmitted over the channel. During this process, the signals are vulnerable to noise disturbance. In this kind of situation, the most crucial thing for the receiver is to select correct messages for the destination from a set of possible candidates. Shannon's theorem 17 describes the channel capacity (C) as follows:

$$C = B\log_2((S+N)/N) \quad (6)$$

where B denotes the bandwidth of the channel, and S and N express the average signal power of the transmitter and the average noise power, respectively. Fitts treated aimed hand pointing as a signal transmission process of the nervous system from the eyes to the hand. As can be seen, the definition of ID is analogous to that of channel capacity C by looking upon movement distance A as signal power and target width W as noise power. However, no hand is involved in eye pointing. Naturally, the corresponding information processing activity is different from that of hand pointing. From this

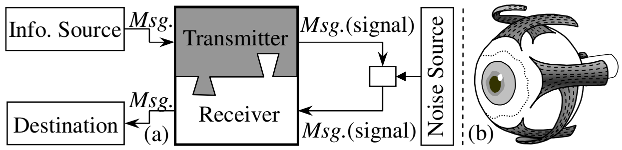

Figure 1. (a) A pseudo communication system. It has a special property in that the transmitter and the receiver are "integrated", without the need to encode messages. (b) Muscles driving the eye.

point of view, it is logical to conclude that Fitts' law could not properly model eye pointing tasks, and should not be expected to model them without significant modification.

Assumptions on Communication System

If eye pointing is also analogized to a kind of particular signal transmission processes, we anticipate that it still could be modeled in a similar manner to Fitts' law for hand pointing, but with additional assumptions which are demanded by the essentially different characteristics involved in eye pointing.

Signal Enhancement

Considering a special scenario where the transmitter is working at the given power of S, but the actual signal power is equivalent to S+S_0 due to the presence of an effective signal enhancement technique (S_0 is the gain of the technique), the channel capacity C will thus be given by

$$C = B\log_2((S + S_0 + N)/N) \tag{7}$$

A Pseudo Communication System

Imagine that there exists a system as Figure 1a illustrates, in which the transmitter is also, at the same time, the receiver, then it is unnecessary to encode and reconstruct the message when it is sent to the destination. That is to say, the message itself is a kind of special signal directly transmitted as an inherent property within the "composite unit" of the transmitter and receiver. Regardless of the mathematical derivation of Shannon's theorem 17, *we hypothesize that the logarithmic function can be considered to be the "operator" of the encoding and decoding process from the transmitter to the receiver*. Then the "channel capacity" (C') of the pseudo communication system will be devoid of this operator in its expression. Thus, C' can be given by

$$C' = B(S + N)/N \tag{8}$$

Furthermore, if the signal is enhanced, C' can be expressed using the following equation:

$$C' = B(S + S_0 + N)/N \tag{9}$$

As can be seen, C' is greater than the corresponding C in a real communication system. This result seems to be reasonable since the signal transmission in the pseudo system is expected to be more efficient than that in the real system.

Definition of Eye Pointing Tasks' Index of Difficulty

As Figure 1b illustrates, there are three pairs of antagonistic muscles that make the eye produce different motions. These muscles are controlled by the third, fourth, or sixth of twelve pairs of cranial nerves (i.e. the oculomotor, trochlear, or abducens nerves). The optic nerve that transmits visual information from the retina to the brain belongs to the second pair of cranial nerves. In other words, the nervous signals of the oculomotor system are produced and transmitted only in the cranial nerve system (brain), without involving the spinal nerve system, through which hands are controlled. Therefore, if we see the cranial nerve system as the transmitter and the spinal nerve system as the receiver for hand pointing, the former is both of them for eye pointing. This is a distinctly different signal transmission path to that of hand pointing. Furthermore, eye gaze can travel extremely quickly and easily from one point to another. This can probably be interpreted as a certain signal enhancement that contributes to saccadic eye movements. Therefore, we can analogize eye pointing to a kind of signal transmission in the pseudo communication system. The "channel capacity" C' in eye pointing can be described using Equation 9.

Thus, we initially define the index of difficulty in eye pointing tasks in a similar manner to the definition of Fitts ID:

$$ID_{eye} = (A + A_0 + W)/W \tag{10}$$

where A_0 is a constant being analogous to the component of S_0. Researchers repeatedly reported that movement time (or selection time) was not in proportion to A in eye pointing tasks [8, 28, 38]. In other words, the contribution rate of A was very low. This means that $A \ll A_0$. At the same time, $W < A$ in general. Hence, $\frac{W}{A_0} < \frac{A}{A_0} \ll 1$, and

$$ID_{eye} = A_0 \frac{1 + \frac{A}{A_0} + \frac{W}{A_0}}{W} \approx A_0 \frac{e^{\frac{A}{A_0}}}{W} \tag{11}$$

Moreover, it is almost impossible to stabilize the eye cursor on a single point because of the factor of involuntary eye jitter. This means that if the target decreases in size close to a certain threshold (μ), it will be increasingly difficult and finally impossible to finish the task. Thus, ID_{eye} is redefined as the form of $e^{\lambda A}/(W - \mu)$, leading to our model:

$$MT = a + b \times ID_{eye} = a + b \times (e^{\lambda A}/(W - \mu)) \tag{12}$$

where the symbols λ and μ are two empirical constants. We also modified the Fitts' law model as well as some other existing models that we reviewed (see Equations 2–5) to reflect this distinct feature of eye pointing. The formulations of these modifications are expressed as follows:

$$MT = a + b\log_2(A/(W - \mu) + 1) \tag{13}$$
$$MT = a + b(A/(W - \mu))^{0.5} \tag{14}$$
$$MT = a + b(A/(W - \mu))^{c} \tag{15}$$
$$MT = a + b\log_2(A) - c\log_2(W - \mu) \tag{16}$$
$$MT = a \times (W - \mu)^{b} \times A^{c} \tag{17}$$

We carried out two experiments to observe human capabilities when pointing with the eyes. Meanwhile, we used the results to verify the effectiveness of ID_{eye}, and to indicate whether or not it is more appropriate for eye pointing than the existing models and even their modifications.

EXPERIMENTAL PLATFORM

The gaze input device was the eye tracker of $EyeLink$ II. The pupil-only tracking mode at the sampling rate of 250 Hz was used. Its running platform as well as that of the following two experiments was the same as in our previous work [38]. According to our previous findings, we employed the improved speed reduction method (iSR) [38] as the main cursor control method to redress the jittery eye cursor. We also considered the situation that there was NO additional method used to alleviate the instability problem, i.e. the cursor was directly located at the subject's current gaze point.

EXPERIMENT 1: POINTING UNDER ONE DWELL TIME AND EYE CURSOR CONTROL CONDITION

In this study, both dwell time and cursor control method were not treated as the independent factors of experimental design but two conditions for the experiments. In this experiment, we used the dwell time of 800 ms to highlight and select targets, and the method iSR to control the eye cursor.

Task and Procedure

After seating the subject about 70 cm in front of the screen, the experimenter asked him/her to adjust the chair to an appropriate height and position so that the subject could look horizontally and directly at the center of the screen. The experimenter explained the task to the subject by demonstrating it on the screen. Then, he setup the eye tracker and calibrated it. As the task was very simple, no practice block was provided for the subject before the formal experiment.

At the beginning of each trial, the trial start button randomly appeared at one of the predefined positions on the screen's diagonals. This button was rendered as a 32-pixel-diameter solid circle but actually with an effective diameter of 120 pixels to facilitate the start of each trial. The subject was instructed to fixate on the start button at first. After the eye cursor had successfully dwelt on the button for 450 ms, the button group, including five circular buttons arranged as a "+" shape, immediately appeared at a given distance (i.e. amplitude A) in the diagonal direction with the simultaneous disappearance of the start button. Then, the subject needed to look at the center area of the desired target, which was arranged at the center of the group, as quickly as possible.

As Figure 2a shows, the given amplitude A was the distance between the centers of the start button and the desired target. The centers of all buttons were highlighted as small focus areas. There was no solid curve but two soft colors used to render the button edges, and even the start button's edge was invisible as illustrated using the dashed curve. Therefore, the effects of visual search and button edge could be largely eliminated, making the actual eye movement distance as accurate as possible. The gap between the desired target and each of the distractor buttons was fixed at the level of 40 pixels because this parameter had no significant main effect on eye pointing performance as previously reported [38]. Furthermore, the cursor was visible during the experiment processes. A specialized experiment [38] indicated that given an explicit instruction of staring at the highlighted center, the subject was able to avoid chasing the cursor. A visible cursor was helpful to monitor the calibration state and determine when a recalibration should be done.

For each trial, when the cursor entered a button, the dwell timer started/restarted. If the subject selected one of the four distractor buttons, a wrong-selection event was recorded. If no button had been selected five seconds after the trial was started (i.e. the button group was presented), a no-selection

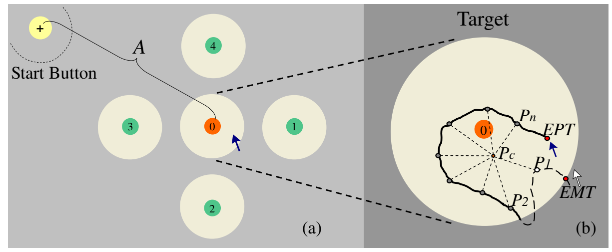

Figure 2. (a) Experimental interface. The desired target is placed at the center. (b) A trajectory of the cursor when the subject is fixating on the center area of the target. The time when the cursor enters the button for the first time is recorded as the end of eye movement time, and the time when the continuous dwell time is long enough so as to activate target selection is recorded as the end of eye pointing time. A series of points, $P_1, P_2, ..., P_n$, in the trajectory are sampled, with the aim of calculating the average distance from these points to their centroid (P_c). The dashed part of the curve means the dwell time is interrupted.

event was recorded. Each trial could be repeated for 5 times at most before being aborted with a recalibration performed if the subject did not successfully select the desired target.

Design

A repeated measures within-subject factorial design was applied. The independent factors (variables) were target diameter W (40, 56, 70, 86, 100 pixels), amplitude A (200, 380, and 850 pixels). A fully-crossed design of $3A \times 5W$ resulted in 15 combinations. Each combination had two trials performed respectively in any two of the diagonal directions on the screen, resulting in 30 trials per block. These trials were presented in a random order, but we avoided displaying the trial start button at the location where the target group had just been shown in the previous trial. Otherwise, it would seem that the target group had suddenly become the start button. There were 9 consecutive blocks for each subject. The subjects were able to finish this experiment within one session of about 30 minutes without a break in general.

Measures

We analyzed eye pointing performance from three aspects (i.e. three dependent variables): eye movement time, eye pointing time, and error rate. Eye movement time (*EMT*) is defined as the duration that the eyes are rapidly moving from the trial start button toward the target until arriving at it. During the experiment, as the cursor was slaved almost in real time to gaze points, *EMT* was approximately measured from the start of the trial to the moment the cursor entered the target for the first time (it was possible for the jittery cursor to enter the target more than one time) as Figure 2b shows. Eye pointing time (*EPT*) denotes the duration from the start of the trial to the moment the continuous dwell time (*DT*) satisfies the predetermined threshold (800 ms) to successfully select the target. Thus, $EPT \geq EMT + DT$ in general. In addition, we measured another dependent variable to represent the cursor's stability. As Figure 2b shows, when the subject is staring at the target, the cursor will not remain stationary on a single point but keeps jittering in a small undefined area, possibly leaving and reentering the target area several times. During this process, the cursor's positions were sampled at the frequency of 25 *Hz*. The average distance from those sampled points to their center was calculated to represent the average radius of the area (denoted by AR).

Participants

Eighteen able-bodied participants (8 females and 10 males), recruited in the university campus through advertisements, successfully completed this experiment. Their average age was 26.5 years in the range from 20 to 39 years. All of the participants had normal or correct-to-normal vision. Some of them had previously experience in a similar experiment.

Results

The error trials and the outliers, in which *EMT* is more than 4 standard deviations from the overall mean, totalling about 3.0% of the data, were excluded from our repeated measures ANOVA. No learning effect was found in this experiment.

Eye Movement Time and Eye Pointing Time (EMT & EPT)

We found that there were significant main effects on *EMT* for both A ($F_{2,34} = 46.96, p < .001$) and W ($F_{4,68} = 113.31, p < .001$), but the interaction effect was not statistically significant ($F_{8,136} = 1.80, p = .082$). The means of *EMT* were 604.9, 518.3, 430.4, 363.3 and 358.9 ms, decreasing with the increase of W from 40 to 100 pixels. Post hoc pairwise comparison tests indicated that there was significant difference between each pair of diameters except for the pair 86 and 100 pixels. For different A levels of 200, 380 and 850 pixels, *EMT* got increasing averages of 402.9, 430.4 and 532.2 ms, respectively. In all pairs of these A conditions, *EMT* had significant differences, but they indeed were not proportional to the changes in A, as reported before.

EPT was also significantly affected by both A ($F_{2,34} = 26.87, p < .001$) and W ($F_{4,68} = 103.67, p < .001$), with a significant interaction effect $A \times W$ ($F_{8,136} = 2.64, p < .05$). The averages of *EPT* were 1806.3, 1527.5, 1366.6, 1287.5 and 1260.2 ms respectively at the five diameter size levels, with similar pairwise comparison results to those of *EMT*. Grouped by amplitude conditions, *EPT* also increased with the changes of A from 200 to 850 pixels, getting the means of 1348.5, 1466.0 and 1534.4 ms, respectively. There were significant differences for *EPT* between each pair of amplitudes. Figure 3a, showing the means of both *EMT* and *EPT*, clearly indicates these results.

Eye Cursor's Stability: Average Radius (AR) of Dwelling Area

There was a significant main effect on AR for diameter W ($F_{4,68} = 94.88, p < .001$) as well as amplitude A ($F_{2,34} = 11.32, p < .001$). Their interaction effect $A \times W$ was also significant ($F_{8,136} = 2.74, p < .01$). The means of AR significantly increased with the increase of W from 40 to 100 pixels, resulting in 4.5, 5.3, 6.0, 7.1, and 7.9 pixels, respectively. This trend implied that subjects probably focused on larger targets with less attention, and vice versa, as they might feel that it was easy to make the cursor dwell on a large target long enough. Furthermore, the averages under different A conditions were 5.7, 6.3, and 6.5 pixels, respectively. Post hoc pairwise comparison tests revealed significant differences between the A condition of 200 pixels and the other two. The increases of AR were probably due to the increasing residual stress in the eyes' muscles with the increase of the eyes' rotation angle. Figure 3b depicts the means of AR.

Error Rate

No matter how many times a trial was repeated due to failures, it was counted as only one error. Wrong-selection events seldom occurred (0.3%), thus only no-selection events were analyzed. The main effects of A ($F_{2,34} = 4.92, p < .05$) and W ($F_{4,68} = 69.19, p < .001$) and their interaction effect $A \times W$ ($F_{8,136} = 4.36, p < .001$) were significant in error rate. The mean error rate radically decreased from 15.4% to 1.3% with the increase in W from 40 to 100 pixels.

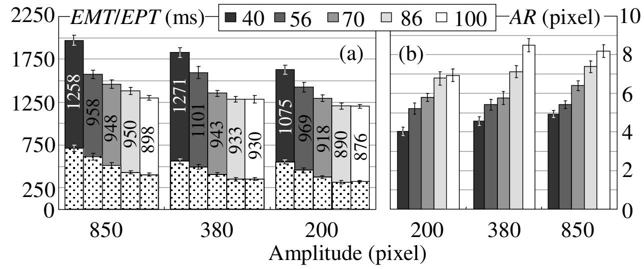

Figure 3. (a) *EMT* and *EPT* by amplitude A for different W levels. *EMT* is represented using the bottom section of each column, where the number denotes the corresponding difference between *EMT* and *EPT*. The difference was decreasing and gradually close to the given dwell time of 800 ms with the increase of W. (b) AR by amplitude A for different W levels. AR apparently kept increasing with the increase of W under each of the A conditions. The total mean of AR is 6.2 pixels.

Model Fitting

As Figure 3a shows, both *EMT* and *EPT* decreased in general with the increase of W and/or the decrease of A. This tendency is similar to that in manual pointing. Unfortunately, neither of them could be properly modeled using the Fitts' law model due to the relatively lower R^2 values of .704 and .542 for *EMT* and *EPT*, respectively. Thus, we fit the data of *EMT* and *EPT* to the six new formulations from Equation 12 to 17 as well as some of the reviewed models.

The results of model fitting were summarized in the left part of Table 1 (after the next page). As it shows, the ID_{eye} model was very suitable for both *EMT* and *EPT*. The other five new formulations (from Equation 13 to 17) also had a good fit to the data. The three models of Equations 1-3 especially achieved obvious improvements by introducing the term $-\mu$ into them. The other two original models of Equations 4 and 5 already had a good fit before being modified. However, the ID_{eye} model still can be viewed as the best one because (1) its R^2 was greater than that of the others in general for both *EMT* and *EPT*; and (2) the intercepts, i.e. the estimates of the regression coefficient a, were unreasonable in Equations 4, 5, and 15-17, e.g. the small intercepts in Equation 15, which could cause a negative prediction of *EMT* or a smaller prediction of *EPT* than the given dwell time under a specified task condition of little difficulty, and contrarily the big intercepts in Equation 4, which could lead to large predictions when the tasks are very easy to perform; and furthermore (3) the estimates of μ in the other five models were somewhat big, especially those for *EPT* which were about 35 pixels. If these estimates were right, it would be very difficult to finish the tasks when W is close to the level of 35 pixels. However, this was not a true finding because this diameter size had been successfully used in another experiment which is unreported in this paper. The estimates of μ in the ID_{eye} model were reasonable as interpreted in detail below. In brief, other models with good fitness in fact did not properly capture the essential of eye pointing.

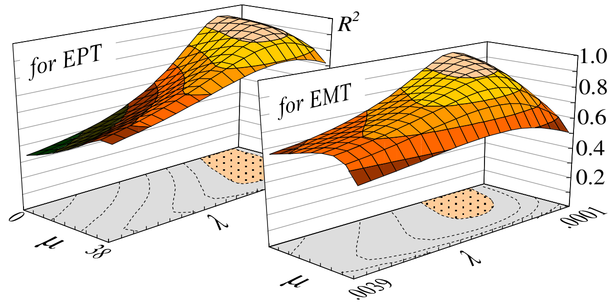

Figure 4. ID_{eye} **model's fitness depending on** μ **and** λ**. With the Considerations of rationality and sufficiency, the range from 0 to 38 pixels for** μ **and the range from .0001 to .0039 for** λ**, due to the tiny variations among** ID_{eye} **values when** $\lambda < .0001$ **and contrarily the huge differences when** $\lambda \geq .0040$**, are predefined as the "feasible region". The dotted floor areas represent the "feasible solutions".**

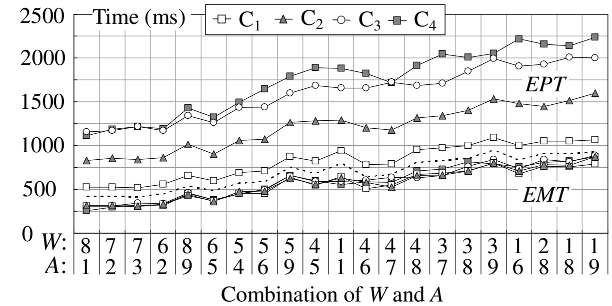

Figure 5. ***EMT*** **and** ***EPT*** **by combination of** W **and** A **under different conditions of dwell time and cursor control method (denoted by** C_n**). The dashed line is a separation line. The numbers from 1 to 8 for** W **denote the diameters of 40, 42, 46, 52, 68, 84, 100, and 116 pixels, respectively; and those from 1 to 9 for** A **denote the amplitudes of 200, 350, 500, 700, 850, 1000, 1100, 1200, and 1300 pixels, respectively.**

Although we could use least-square regression to simultaneously estimate the parameters of ID_{eye} model: a, b, μ and λ, the last two empirical constants logically should be predetermined and be universally compatible across different eye pointing experiments, at most with variations in a very limited range. With respect to μ, when it is used for *EMT*, it seems that μ should be 0 because eye jitter does not affect the speed of saccadic eye movements. When μ is used for *EPT*, we can directly refer to the observation of AR depending on the meanings of μ. The total mean of AR (symbolized using $\overline{AR}$) was about 6.2 pixels, which means that the eye cursor in general was jittering in small areas with the average radius of 6.2 pixels when dwelling on targets (see Figure 3b for the concrete radius levels of jittering under different conditions). In other words, the average diameter size of approximately 12.4 pixels ($2\overline{AR}$) was the likely threshold μ. Fortunately as shown in Table 1, the estimates of μ (0 pixel for *EMT*, and 11.3 pixels for *EPT*) confirmed these two points. At the same time, the ID_{eye} model could get almost the same best fit for EPT at both μ levels of 12.4 and 11.3 pixels, with R^2 approximate to .958. Regarding λ, it is the inverse of A_0 as expressed in Equation 11. Thus, it is reasonable to assume that $\lambda < \frac{1}{1000}$. Its estimates, .00071 and .00052 for *EMT* and *EPT* respectively, also confirmed this assumption.

In general, an enumeration process can be used to search for the proper combination of (μ, λ), which makes the model achieve satisfying fitness. Figure 4 plots R^2 values of the model in terms of (μ, λ). As it can be seen, the "feasible solutions" of (μ, λ), of which the criteria we set are $R^2 \geq .90$, locate in a single continuous area (being dotted). Within each of the dotted areas, R^2 can converge to its peak at a single "optimal solution", e.g. the estimate of (0, .00071) for *EMT* or (11.3, .00052) for *EPT*, due to there seeming convexity in the curved surface. Meanwhile, the two small dotted areas are similar with a certain overlap if represented together. These situations ensure the reliability of the estimated parameters in the new model, without the possibility of leading to conflicting interpretations of μ and/or λ.

EXPERIMENT 2: POINTING UNDER DIFFERENT DWELL TIME AND EYE CURSOR CONTROL CONDITIONS

ID_{eye} defined for eye pointing indeed worked very well for the data in Experiment 1. However, there were still some inescapable doubts. First, it was unclear whether or not the new model was still the best and most reliable one within a wider ID_{eye} range under different conditions. Second, since λ is not a measurable parameter unlike μ, it was essential to reveal whether or not the changes of λ are negligible under different conditions especially with respect to a large A range, i.e. whether or not λ can remain unchanged in different situations. Third, we had to confirm whether it was a fixed principle for eye pointing (or just an occasional phenomenon) that ID_{eye} model can sufficiently represent *EMT* and *EPT* when $\mu = 0$ and $2\overline{AR}$, respectively. To dispel these uncertainties, we carried out this further experiment.

Experiment Setup

These doubts, especially the third one, necessitated different observations of $\overline{AR}$. We anticipated that $\overline{AR}$ would likely vary with the change of dwell time because AR is measured when the cursor is dwelling on targets. Three dwell time lengths, including 800 ms, 500 ms, and a very short duration of 200 ms, regardless of the "Midas Touch" problem, were used. Furthermore, we previously had provided a hard evidence indicating that an effective cursor control method can cause a significant difference in $\overline{AR}$ [38]. Thus, we also expected a certain variation of $\overline{AR}$ with the change of cursor control method. In addition, in order to extend ID_{eye}'s range, we replaced the 19-inch CRT display (1024×768 pixels) used in Experiment 1 with a LCD (1440×900 pixels), maintaining almost the same pixel size as in the former. The purpose of this experiment was not to systematically explore the effects of different factors on eye pointing, so we did not use a fully-crossed design. The manipulated levels of W and A are indicated in the bottom of Figure 5, where there are 21 combinations that can produce a relatively equal distribution of ID_{eye} values in the corresponding range. The combinations of dwell time and cursor control method were (200 ms, iSR), (500 ms, iSR), (800 ms, iSR), and (800 ms, NO). They were denoted using C_1, C_2, C_3, and C_4, respectively.

The experiment task and procedure were the same as in Experiment 1. The combinations of W and A with 2 trials for each composed a block. There were 8 blocks under each condition of C_{1-4}. These conditions were divided into two sessions held over consecutive days with each lasting about 40 minutes. Their order was counterbalanced across all the subjects. Twenty subjects (9 females and 11 males with the average age of 23 years) successfully finished their tasks.

Model	Experiment 1 (results indicated above for *EMT* and below for *EPT*)									Experiment 2											
	a		*b*		*c*, λ		μ		R^2	μ (Estimate)				*a* (Estimate)				R^2			
	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.	Estimate	Std.Err.		C_1	C_2	C_3	C_4	C_1	C_2	C_3	C_4	C_1	C_2	C_3	C_4
Eq.12	181.57 975.91	17.43 107.40	12098 18475	1089 9011	.00071 .00052	.00008 .00013	0.0 11.3	11.8 9.2	.969 .958	0.0 0.0	6.8 3.8	2.3 0.0	0.0 0.4	153.4 320.0	187.3 620.0	184.3 873.3	151.0 811.0	.918 .931	.939 .950	.963 .970	.970 .957
Eq.13	142.57 937.28	34.64 61.16	79.00 116.20	11.16 17.48			33.2 37.8	2.7 1.1	.922 .940	36.3 36.8	36.9 37.0	35.3 36.2	33.8 36.5	86.20 274.9	82.90 514.3	48.10 764.5	-11.28 681.5	.881 .899	.909 .933	.924 .948	.945 .957
Eq.14	235.65 1126.19	33.58 58.74	57.99 68.88	12.71 18.36			28.9 35.3	3.7 1.9	.900 .911	29.1 30.1	30.8 30.9	28.5 29.9	25.9 29.8	216.6 431.2	222.7 716.4	195.4 1005	144.7 991.2	.852 .856	.891 .905	.927 .927	.945 .902
Eq.15	-219.51 421.76	725.92 864.04	443.69 646.40	678.95 799.02	0.1563 0.1524	0.1690 0.1295	32.4 37.3	3.1 1.4	.923 .944	35.9 35.3	35.8 36.3	32.4 35.1	30.8 35.2	-1104 0.00	-349.9 -527.3	-10.4 -14.5	-136.6 0.00	.877 .888	.908 .931	.932 .948	.951 .945
Eq.16	776.12 1502.00	550.99 302.74	63.73 88.87	9.32 16.27	153.02 169.92	77.50 44.64	13.7 33.4	22.4 4.2	.943 .956	33.5 33.0	35.8 35.4	35.3 33.5	34.4 33.2	426.1 795.4	290.7 868.9	149.6 1319	35.9 1477	.883 .908	.905 .931	.916 .948	.938 **.963**
Eq.17	1665.32 1671.68	2933.53 407.05	-0.6103 -0.1980	0.3434 0.0529	0.2066 0.0907	0.0240 0.0138	0.0 29.5	32.7 4.8	.961 **.967**	22.7 25.0	29.1 31.3	27.9 28.8	24.0 25.6	805.9 1179	482.0 1042	344.9 1527	324.3 2082	.888 .911	.914 .938	.935 .958	.953 **.962**
Eq. 1	134.81 893.02	59.75 147.17	112.17 195.10	20.16 49.66					.704 .542					30.92 211.5	18.64 422.6	-15.82 664.9	-73.12 552.7	.706 .695	.720 .735	.784 .766	.837 .766
Eq. 2 (n=2)	178.45 971.07	48.82 124.12	106.32 184.10	17.82 45.29					.733 .560					118.6 313.3	109.0 554.5	85.17 824.1	41.54 765.8	.766 .753	.785 .793	.850 .824	.891 .805
Eq. 3	299.36 1172.90	98.19 266.29	26.39 49.65	46.49 130.84	0.89 0.86	0.53 0.78			.742 .564					339.1 562.1	352.0 874.3	335.0 1189	273.9 1112	.815 .796	.846 .836	.900 .862	.917 .817
Eq. 4	1105.96 3233.29	127.37 289.98	63.73 88.87	9.03 20.57	198.35 421.00	16.54 37.64			.942 .923					1557 2151	1575 2699	1388 3307	1253 4018	.854 .879	.861 .891	.878 .918	.909 .937
Eq. 5	1665.32 4869.96	365.30 760.35	-0.6103 -0.4200	0.0411 0.0293	0.2066 0.0905	0.0230 0.0163			.961 .950					3759 4648	3821 4424	2258 4777	1648 7029	.882 .903	.901 .919	.924 .944	.948 .954

Model	Expt.1	C_1	C_2	C_3	C_4
Eq.12	.969 .958	.918 .929	.938 .948	.963 .965	.970 .944
Eq.13	.704 .643	.706 .727	.720 .781	.784 .816	.827 .828
Eq.14	.733 .666	.766 .781	.785 .835	.850 .869	.891 .855
Eq.15	.742 .667	.815 .805	.846 .854	.900 .883	.917 .856
Eq.16	.942 .934	.854 .885	.861 .901	.878 .928	.909 **.947**
Eq.17	.961 .957	.882 .906	.901 .925	.924 .951	.948 **.959**

Table 1. Model fitting results. For Experiment 2, only R^2 and the estimates of *a* and μ are listed due to the page limitation. The embedded sub-table lists the results of R^2 when presetting μ=0 for *EMT* and $2\overline{AR}$ for *EPT*.

Results and Model Fitting

The data were processed in the same manner as before. Figure 5 plots the averages of *EMT* and *EPT*. It shows that there was no significant difference in *EMT* ($F_{3,57} = 1.56, p = .208$) under the different conditions (C_{1-4}), indirectly supporting the aforementioned point that eye jitters almost have no effect on *EMT*. The corresponding observations of $\overline{AR}$ were 4.07, 5.77, 6.16, and 7.16 pixels, respectively. There was a significant main effect for the factors of dwell time ($F_{2,38} = 120.80, p < .001$) and eye cursor control method ($F_{1,19} = 162.51, p < .001$) on *AR*, so the experiment really captured the variations of $\overline{AR}$ as we anticipated.

The right part of Table 1 shows the results of model fitting. It indicates that our ID_{eye} model was still the best under all the four conditions. Furthermore, it was the most stable model according to the changes of R^2 under the different conditions. Regarding the estimates of μ, they did not closely match their corresponding $2\overline{AR}$ values in the ID_{eye} model, but, in the other five modified models, they were still as large as in Experiment 1. When we preset $\mu = 0$ for *EMT* and $\mu = 2\overline{AR}$ for *EPT* [1] to fit the data of this experiment, there was almost no deterioration for the fitness of ID_{eye} models, but it did result in obvious decreases in R^2 for the other models as illustrated in the embedded sub-table. That is to say, although the peak fitness of ID_{eye} models did not always occur under the conditions when $\mu = 0$ for *EMT* or $\mu = 2\overline{AR}$ for *EPT*, they were enough to lead to good and also the best fitness in general. Thus, $2\overline{AR}$ and 0 are two appropriate options of μ for *EPT* and *EMT*, respectively.

[1]Considering the interaction effect of $A \times W$ on AR as revealed in Experiment 1, we also checked the fitting results when using the averages of AR of different $A \times W$ combinations in place of $2\overline{AR}$, but we did not find improvements of fitness in general.

With respect to λ, its estimates were .00047, .00053, .00066, and .00068 for *EMT*; or .00036, .00043, .00041, and .00034 for *EPT* under the conditions from C_1 to C_4, respectively, when μ was also estimated in the same manner. However, when μ was preset, the estimates of λ were .00047, .00049, .00064, and .00068 for *EMT*; or .00040, .00048, .00050, and .00039 for *EPT*, respectively. We can see that the estimates for *EMT* overall are relatively bigger than those for *EPT*. Note the fact that the empirical constant λ (i.e. $\frac{1}{A_0}$) is only related to ballistic saccades, which, we think, mainly depend on the instinctive physiological response of the antagonistic muscles (see Figure 1b). Therefore, λ essentially should be unique in different situations without the need to vary for the absolute maximum of R^2. We also explored the functional relation between R^2 and both of λ and μ, and got similar results as Figure 4 illustrates. This implies that it is feasible to use a common approximation of these estimates to obtain satisfying fitness (R^2). Here, their average, about .0005, is a proper option because when we simultaneously preset $\lambda = .0005$ and $\mu = 0$ for *EMT* and $2\overline{AR}$ for *EPT*, our ID_{eye} models were still the best (see Figure 6 in the next page).

DISCUSSION

So far, we have confirmed ID_{eye} model's accuracy, reliability, and stability for eye pointing tasks, and detailedly explained the decisions about the two empirical constants (μ and λ) used in the definition of ID_{eye}. However, the form of Equation 12 probably results in the incorrect understanding about ID_{eye} that it is not independent of measurement units (in this paper, 5 pixels were about 1.65 mm or .135 degrees of visual angle). Actually, according to Equation 11, the factor of units has been eliminated. From this point of view, we can use the following form of Equation 18 to clearly indicate the independence of measurement units. Based on this equation, Figure 6 replots the data of *EMT* and *EPT* as well as their corresponding regression lines in different situations.

$$MT = a + b \times A_0 \frac{e^{A/A_0}}{W - \mu} \approx a + b \times 2000 \frac{e^{A/2000}}{W - 2\overline{AR}} \quad (18)$$

Using the notions of information theory, we built the ID_{eye} model for eye pointing like Fitts analogized hand pointing to signal transmission. It is essential to point out that Equations 8 and 9 are not intended to deny the classical Shannon formula but to hypothesize a pseudo communication system that would break through the limitation of information channels and more properly represent the oculomotor system. Proving the hypotheses is beyond the scope of this paper, but the effectiveness of ID_{eye} derived from Equation 9 is strongly supported by our experimental results. Our hypotheses do indeed appear to be suitable to eye pointing.

We did not construct ID_{eye} from the perspective of psychomotor behavior, but we can get some useful insights from Meyer et al.'s model (Equation 2). They pointed out that when the number of submovements grows toward infinity, it leads to a logarithmic relationship between MT and the term of A/W similar to Fitts' law in form [23]. In the situation of eye pointing, however, it seems to be true that only one primary movement (saccade) in general is needed to make the eyes rapidly fixate on the desired target, which is specifically marked to prevent subjects from visual search. That means we can let $n = 1$ when using Equation 2 to account for eye pointing, causing eye movement time to be linearly related to the term of A/W. Fitting the data in Experiment 1 and the data under the four conditions (C_{1-4}) of Experiment 2, we found that this relationship was not precise enough as the corresponding R^2 values were .741, .806, .831, .892, and .917 for *EMT*, and .564, .790, .830, .856, and .816 for *EPT*, respectively. According to our experimental results, this relationship properly expresses the contribution of W to eye movement time, but the factor A needs to be taken into account more exactly on the base of Meyer et al.'s model. With respect to the features of eye movements, i.e. rapid ballistic saccades and involuntary eye jitters, it is natural to think of introducing an empirical constant (A'_0) into the term of A/W to weaken A's contribution and another one (μ') to compensate for eye jitters from W's contribution. Therefore, we can develop a formula which is expressed as follows:

$$MT = a + b \times (A'_0 + A)/(W - \mu') \quad (19)$$

Equations 18 and 19 essentially approximate to each other when $A \ll A'_0$ and A_0. Our regression analyses confirmed this point. That is to say, ID_{eye} can also be suitably interpreted using Meyer et al.'s theory. However, we prefer the perspective of the pseudo communication system. This does not mean that we view information theory as the real theoretical basis for eye pointing, but it does emphasize the distinct neural control of eye pointing from that of hand pointing and provides an analogy similar to that of Fitts' law.

Furthermore, we did not derive ID_{eye} from the physics of eye movements as presented by Komogortsev and Khan [14], considering that the properties of performance models, such as accuracy, utility and simplicity, are very significant for HCI [30, 36]. In fact, Fitts' law and its evolutions [1, 3,

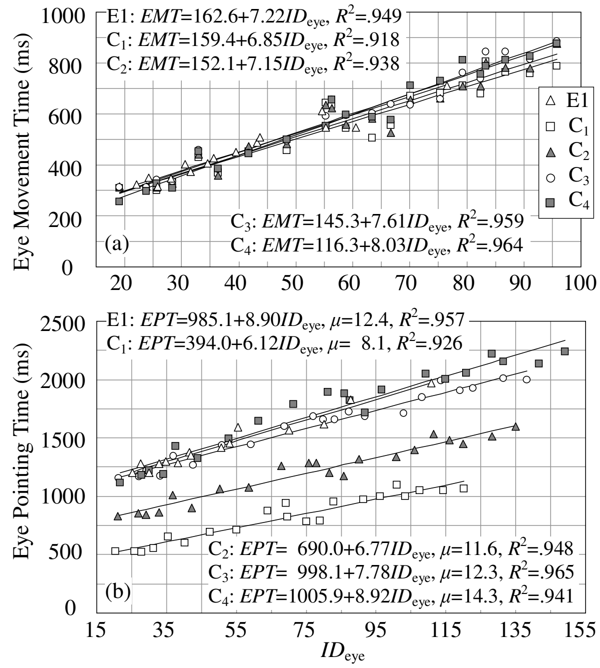

Figure 6. Regression lines of (a) *EMT* and (b) *EPT* in different situations (E1 denotes Experiment 1). ID_{eye} was calculated using Equation 18. We could divide 25 in the definition of ID_{eye} to make it comparable with the standard Fitts ID (in the range from 1.45 to 5.07). Regarding *EPT*, each estimate of the coefficient a satisfied the potential requirement that it should be, but not too, larger than the given dwell time.

9] are not based on the physics of hand movements, and there is also a lack of psychomotor theory to satisfactorily explain them [30]. Our model retains these properties and further exhibits some useful implications for the designs of eye-controlled user interface and the comparison of different eye tracking technologies as follows:

- Accot and Zhai argued that scale independence was one of the desirable model properties [1]. However, ID_{eye} reveals that ***eye-controlled interfaces do not have the property of scale independence as hand-controlled interfaces do***. As Table 1 shows, the models (Equations 1-3) maintaining this property had poor fitness in general, but their modifications (Equations 13-15) and those (Equations 4 and 5) originally without this property were able to fit the data well enough. This implied that scale independence is not a potential characteristic for eye-controlled interfaces.
- Eye-based interactions, as a kind of promising interaction paradigm for HCI, partly benefit from the speed of saccadic eye movements. The empirical constant λ in ID_{eye} appropriately reflects this feature by a very small decimal, which directly shows the low contribution rate of A to eye pointing time. Decreasing A and/or increasing W can effectively decrease eye pointing time as implied in Fitts ID for hand pointing, but ID_{eye} further indicates that ***the benefit of increasing W is greater than that of decreasing A***. Therefore, this particularly informs designers that target expansion is more appropriate for the improvement of human performance in eye-controlled interfaces, and that zooming interface (i.e. increasing both of W and A) also can decrease eye pointing time [2] although it is almost useless in general to facilitate hand pointing.

- For eye pointing, an unstable cursor is inevitable because of inherent eye jitters and the noise inherent in current eye tracking devices. ID_{eye} takes account of this issue by introducing a parameter μ, indicating that ***addressing the issue of eye cursor's stability is also an important way to improve human performance***. Experiment 2 revealed that $\overline{AR}$ could increase with the extension of dwell time. This implies that when dwell time becomes longer, the probability of interrupting its continuity will increase so that the tasks will be more difficult. Comparing the situations of C_{1-3} as plotted in Figure 6b, we can find that these situations properly represent the fact that shortening dwell time, practically with the concern of the "Midas-Touch" problem, not only can directly decrease *EPT*, but it also can decrease ID_{eye} to further reduce *EPT*. A comparison of the situations of $C_{3,4}$ indicates that the result that the method iSR can effectively reduce *EPT* by stabilizing eye cursor (i.e. decreasing $\overline{AR}$), as we previously reported [38], is repeatable. Using ID_{eye}, we can "theoretically" explain why this result appeared, and confirm the necessity of this kind of solution.
- As mentioned above, one of the reasons causing the cursor to be unstable is the noises of the eye tracker. The noise levels produced in different eye tracking technologies may be different. Thus, the parameters of the ID_{eye} model, especially μ, would likely vary for different technologies, reflecting the potential performance differences. That is to say, ***the new model can be used as an evaluation tool for the comparison of different gaze input devices.***

The effectiveness of ID_{eye} is clear and believable not only because of its good fitness but also, and more significantly, because its semantics match the characteristics of eye pointing. Comparing the regression lines in Figure 6b, we further found that the component of dwell time (*DT*) was potentially included in the coefficient a, and it was almost linearly related to *EPT*. Although we preferred to view *DT* as a preset "system variable", we still considered the situation when *DT*, as an independent variable, was linearly integrated into Equation 18. Using the *EPT* data under the iSR condition ($DT =$ 200, 500, and 800 ms) in Experiment 2 and treating *EMT* as a special case of *EPT* (i.e. $DT = 0$), we could get the following regression equation ($R^2 = .982$):

$$MT = 139.2 + 1.125DT + 6.953ID_{eye} \qquad (20)$$

This result further indicated the validity of the ID_{eye} model. Some of the models that we tested, such as Equation 4 and its modification, fit mathematically well enough to the data and sometimes even better. However, we still think they are dubious for eye pointing because the estimates of a or μ in different situations, as summarized in Table 1, were not in the reasonable range, implying that the model variables did not correctly capture the characteristics of eye pointing.

Regarding the experimental interface, it was not designed based on ISO 9241-9. According to the device scope, to which this ISO is applicable, gaze input device appears not to be covered. The multidirectional tapping task specified in this ISO is used to take account of the effect of direction [30]. However, this effect is negligible for gaze input as the eyes can effortlessly rotate unlike the hand. More importantly, this task needs the subject to continually perform the trials until the end of a round. We found that the subject often intended to blink to comfort his/her eyes, and that the experimenter occasionally needed to recalibrate the eye tracker. Thus, it was useful that the experiment processes could be temporarily suspended at the end of any trial.

We need to point out that the device we used is very precise to uncover the characteristics of eye movements as evaluated by van der Geest and Frens [32]. That is to say, ID_{eye} model is not specifically developed for the concrete eye tracker but for the general eye movements involved in eye pointing. Furthermore, the functional relationship in ID_{eye} models is also independent of cursor control methods. We had fitted this model to the data in our previous experiment [38], in which four different cursor control methods (including iSR) and the dwell time of 1000 ms were used. For those methods, its fitness was still satisfying with R^2 of .919, .954, .982, and .960 for *EMT*; and .971, .983, .988, and .978 for *EPT*, respectively. Therefore, we believe that ID_{eye} models are fully generalizable for eye pointing. Finally, our study differentiated eye movement time from the total pointing time so that it can provide useful insights for modeling non-dwell-based eye pointing, such as MAGIC pointing [37] and EyePoint [16] or pointing based on voluntary frowning actions [31].

CONCLUSIONS

Using a pseudo communication system, we distinguished the signal transmission in the nervous system controlling eye movements from that of controlling hand movements, and developed a new model for eye pointing that parallels Fitts' law for hand pointing from the perspective of information theory. This model takes account of two kinds of eye movements involved in eye pointing: rapid saccade and jittery fixation. We presented an experimental study to carefully validate our model's effectiveness. We justified the new model not only based on its satisfying fitness but also because of its appropriate and meaningful interpretations of eye pointing. Our work can provide some useful insights for HCI researchers to understand the human capability of eye pointing so as to benefit the designs of eye-controlled interface.

ACKNOWLEDGMENTS

The first author thanks Prof. Guangming Hou and Prof. Jingping Wang for their great supports, especially the necessary working and living conditions they provided in BIT.

REFERENCES

1. Accot, J. and Zhai, S. Refining Fitts' law models for bivariate pointing. In *Proc. CHI 2003*, ACM Press (2003), 193–200.
2. Bates, R. and Istance, H.O. Zooming interfaces! enhancing the performance of eye controlled pointing devices. In *Proc. the fifth international ACM conference on assistive technologies*, ACM Press (2002), 119–126.
3. Cao, X., Li, J.J. and Balakrishnan, R. Peephole pointing: modeling acquisition of dynamically revealed targets. In *Proc. CHI 2008*, ACM Press (2008), 1699–1708.
4. Crossman, E. and Goodeve, P. Feedback control of hand-movement and Fitts' law. *Quarterly Journal of Experimental Psychology*, Vol.35A (1963), 251–278.

5. Dickie, C., Hart, J., Vertegaal, R. and Eiser, A. LookPoint: an evaluation of eye input for hands-free switching of input devices between multiple computers. In *Proc. OZCHI 2006*, ACM Press (2006), 119–126.
6. Drewes, H., Luca, A.D. and Schmidt, A. Eye-gaze interaction for mobile phones. In *Proc. Mobility Conference 2007*, ACM Press (2007), 364–371.
7. Fitts, P.M. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, Vol.47 (1954), 381–391.
8. Gajos, K.Z., Wobbrock, J.O. and Weld, D.S. Automatically generating user interfaces adapted to users' motor and vision capabilities. In *Proc. UIST07*, ACM Press (2007), 231–240.
9. Grossman, T. and Balakrishnan, R. Pointing at trivariate targets in 3D environments. In *Proc. CHI 2004*, ACM Press (2004), 447–454.
10. Grossman, T., Kong, N. and Balakrishnan, R. Modeling pointing at targets of arbitrary shapes. In *Proc. CHI 2007*, ACM Press (2007), 463–472.
11. Haffegee, A. and Barrow, R. Eye tracking and gaze based interaction within immersive virtual environments. In *Proc. ICCS 2009, Part II*, LNCS 5545 (2009), 729–736.
12. Hyrskykari, A., Majaranta, P., Aaltonen, A. and Räihä, K.-J. Design issues of idict: a gaze-assisted translation aid. In *Proc. ETRA 2000*, ACM Press (2000), 9–14.
13. Jacob, R. What you look at is what you get: eye movement-based interaction techniques. In *Proc. CHI'90*, ACM Press (1990), 11–18.
14. Komogortsev, O.V. and Khan, J.I. Eye movement prediction by Kalman filter with integrated linear horizontal oculomotor plant mechanical model. In *Proc. ETRA 2008*, ACM Press (2008), 229–236.
15. Kumar, M., Garfinkel, T., Boneh, D. and Winograd, T. Reducing shoulder-surfing by using gaze-based password entry. In *Proc. the 3rd symposium on usable privacy and security*, ACM Press (2007), 13–19.
16. Kumar, M., Paepcke, A. and Winograd, T. EyePoint: practical pointing and selection using gaze and keyboard. In *Proc. CHI 2007*, ACM Press (2007), 421–430.
17. Kvålseth, T. An alternative to Fitts' law. *Bulletin of the Psychonomic Society*, Vol.16, No.5 (1980), 371–373.
18. MacKenzie, I.S. A note on the information-theoretic basis for Fitts' law. *Journal of Motor Behavior*, Vol.21 (1989), 323–330.
19. MacKenzie, I.S. and Buxton, W. Extending Fitts' law to two-dimensional tasks. In *Proc. CHI'92*, ACM Press (1992), 219–226.
20. MacKenzie, I.S. Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction*, Vol.7 (1992), 91–139.
21. Majaranta, P. and Räihä, K.-J. Twenty years of eye typing: systems and design issues. In *Proc. ETRA 2002*, ACM Press (2002), 15–22.
22. Meyer, D.E., Smith, J.E.K., Kornblum, S., Abrams, R.A. and Wright, C.E. Optimality in human motor performance: ideal control of rapid aimed movements. *Psychological Review*, Vol.95 (1988), 340–370.
23. Meyer, D.E., Smith, J.E.K., Kornblum, S., Abrams, R.A. and Wright, C.E. Speed-accuracy tradeoffs in aimed movements: toward a theory of rapid voluntary action. In Jeannerod M., editor, *Attention and Performance XIII* (1990), 173–226.
24. Miniotas, D. Application of Fitts' law to eye gaze interaction. In *CHI 2000 Extended Abstracts*, ACM Press (2000), 339–340.
25. Miniotas, D., Špakov, O. and MacKenzie, I.S. Eye gaze interaction with expanding targets. In *CHI 2004 Extended Abstracts*, ACM Press (2004), 1255–1258.
26. Seow, S.C. Information theoretic models of HCI: a comparison of the Hick-Hyman law and Fitts' law. *Human-Computer Interaction*, Vol.20 (2005), 315–352.
27. Shannon, C.E. A mathematical theory of communication. *The Bell System Technical Journal*, Vol.27 (1948), 379–423, 623–656.
28. Sibert, L. and Jacob, J. Evaluation of eye gaze interaction. In *Proc. CHI 2000*, ACM Press (2000), 281–288.
29. Smith, J.D. and Graham, T.C. Use of eye movements for video game control. In *Proc. ACE 2006*, ACM Press (2006), Article No. 20.
30. Soukoreff, R.W. and MacKenzie, I.S. Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *Int. J. Human-Computer Studies*, Vol.61 (2004), 751–789.
31. Surakka, V., Illi, M. and Isokoski, P. Gazing and frowning as a new human-computer interaction technique. *ACM Transactions on Applied Perceptions*, Vol.1, No.1 (2004) 40–56.
32. van der Geest, J.N. and Frens, M.A. Recording eye movements with video-oculography and scleral search coils: a direct comparison of two methods. *Journal of Neuroscience Methods*, Vol.114 (2002), 185–195.
33. Vertegaal, R. A Fitts' law comparison of eye tracking and manual input in the selection of visual targets. In *Proc. ICMI'08*, ACM Press (2008), 241–248.
34. Ware, C. and Mikaelian, H.H. An evaluation of an eye tracker as a device for computer input. In *Proc. CHI'87*, ACM Press (1987), 183–188.
35. Welford, A.T., Norris, A.H. and Shock, N. W. Speed and accuracy of movement and their changes with age. *Acta psychologica*, Vol.30 (1969), 3–15.
36. Wobbrock, J.O., Cutrell, E., Harada, S. and MacKenzie, I.S. An error model for pointing based on Fitts' law. In *Proc. CHI 2008*, ACM Press (2008), 1613–1622.
37. Zhai, S., Morimoto, C. and Ihde, S. Manual and gaze input cascaded (MAGIC) pointing. In *Proc. CHI'99*, ACM Press (1999), 246–253.
38. 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.