Fast Target Selection via Saccade-driven Methods

OLEG V. KOMOGORTSEV

Texas State University-San Marcos

YOUNG SAM RYU

Texas State University-San Marcos

and

DO HYONG KOH

Texas State University-San Marcos

Three fast, saccade driven, target selection methods are explored. The first method selects a target at the beginning of a saccade with the objective of providing target selection in almost constant amount of time regardless of the distance to the target. The second method selects a target at the end of a saccade. The third is a hybrid method combining the speed of the saccade-driven selection with the accuracy of the conventional Dwell-Time selection. Theoretical evaluation of the proposed methods conducted via characteristics of the Human Visual System and a mathematical model of the human eye indicates that the objective is tenable. Practical evaluation of the proposed methods is conducted with the Multi-Directional Fitts' Law task and with a real-time eye-gaze-guided video game designed to simulate gaming environments where selection speed of a target is of outmost importance. The results indicate that proposed methods show an increased throughput and task completion performance compared to the conventional Dwell-Time target selection method.

Categories and Subject Descriptors: H.1.2 [**User/Machine Systems**]: Human Factors

General Terms: Algorithms, Measurement, Performance, Design, Experimentation, Human Factors

Additional Key Words and Phrases: target selection, saccade, oculomotor plant, interaction, eye movement classification, human computer interaction, video game

---

This work is supported in part by funding from Texas State University - San Marcos to Dr. Komogortsev and Dr. Ryu and Sigma Xi Grant-in-Aid of Research program. In addition, we would like to thank Blizzard Entertainment for the World of Warcraft beta account provided to support eye-gaze-guided interaction research. We would like to express our thanks to Sandeep Munikrishne Gowda for help with programming Multi-Directional Fitts' Law task and Balura game.

Authors' addresses: Oleg V. Komogortsev, Department of Computer Science, Texas State University-San Marcos, Texas 78666; Young Sam Ryu, Ingram School of Engineering, Texas State University-San Marcos, Texas 78666; Do Hyong Koh, Department of Computer Science, Texas State University-San Marcos, Texas 78666.

1. INTRODUCTION

The aim of this study is to investigate the performance capabilities of the saccade-driven target selection methods. Saccades are the fastest eye movements that the HVS can exhibit, with eye ball rotation velocities achieving the speeds of up to 700º/s [Duchowski 2007]. Saccades are already parts of the target selection sequence for the conventional mouse and Dwell-Time (DT) selection methods, where an eye dwells/fixates on the target before it is selected. For example, a mouse user looks at the target first using a sequence of saccades and fixations and then drags and clicks the target with the mouse in hand [Jacob 1990]. In the DT selection, the mouse dragging is eliminated, however, the duration of the fixation to trigger the selection prolongs the overall selection time. Therefore, the saccade-driven selection has the potential to eliminate the delays from target selection. We envision that the saccade-driven target selection methods will be applicable widely in action-oriented video games.

Today's immersive gaming environments are pioneering sophisticated interaction techniques that give users more exciting and engaged gaming experiences. The Nintendo Wii game console has gained immense popularity with its novel motion-based remote controller. Recently, Sony Computer Entertainment introduced a prototype motion controller called a “motion-sensing wand” for the Playstation 3 [E3 2009]. In addition, Microsoft presented a controller-free interface Kinect for the Xbox 360 game console [E3 2009]. Project Kinect aims to provide controller-free, full body motion capture, voice recognition, and facial recognition for gaming and entertainment. However, the use of eye-movement-based control has not been applied to consumer-oriented video games.

In the HCI research community, eye-guided interfaces and their interaction techniques have existed for several decades [Majaranta and Räihä 2002], but recently have attracted a significant research interest [Istance et al. 2008; Koh et al. 2009; Komogortsev and Khan 2007; MacKenzie and Zhang 2008; Nakayama and Takahasi 2008; Tien and Atkins 2008]. Eye-tracking devices serve as input devices for users with disabilities, or as additional interaction channels for other users [Douglas et al. 1999]. In our previous work, we have explored the use of eye movements for video game control [Komogortsev and Khan 2007], with users reporting that the “game environment feels more alive” and immersive.

Eye movements are already considered as a more natural tool for target selection than conventional pointing devices such as mouse because users look at a target before they actually "click" it [Jacob 1990]. The DT selection methods were estimated to be faster than the mouse-based selection [Zhang and MacKenzie 2007]. However, the DT method requires data buffering for at least 100 ms, which introduces delays [Kumar et al. 2007; Sibert and Jacob 2000; Zhai et al. 1999]. The two saccade-driven target selection methods we are proposing can eliminate the delays. The first method, called *Saccade Offset* (SO), selects a target at the offset (end) of the saccade. Essentially, the SO is similar to the DT without the dwell time period following after the saccade. The second method, called *Instantaneous Saccade* (IS), selects the target at the onset (beginning) of a saccade. The ultimate goal of the IS method is to provide an almost constant selection time regardless of the distance to the target. In addition, we introduce a target selection scheme called Hybrid Saccade (HS) that combines the IS and DT. The HS overcomes practical challenges associated with the IS, such as direction and amplitude prediction of future saccades, so that it still allows accurate target selection when the IS selection fails. It must be mentioned that saccade-driven target selection methods proposed here do not attempt to solve the Midas Touch problem [Istance, Bates, Hyrskykari and Vickers 2008], necessitating special interface design techniques where erroneous selection is not detrimental to a user's experience. Such design can be noticed in games similar to World of Warcraft [Blizzard 2009] where erroneous selection of a friendly target does not produce negative consequences.

Significant research has employed the DT selection in the eye-gaze guided interfaces [Jacob 1990; Kumar, Paepcke and Winograd 2007; Zhai, Morimoto and Ihde 1999]. However, very little research has been done using saccade-driven selection. Huckauf and Urbina employed saccade-driven selection, similar to the SO, to select the pieces of a pie-like menu for typing and multiple-choice selection tasks [Huckauf and Urbina 2008; Urbina and Huckauf 2010]. They concentrated on the typing performance of the pie-menu representation and did not evaluate the characteristics of their saccade-driven method for target selection explicitly. The performance of the DT in 2D was evaluated previously by using Multi-Directional Fitts' Law (MD-FL) task [Zhang and MacKenzie 2007], but no such evaluation was conducted for the saccade-driven target selection methods. In our previous work [Komogortsev et al. 2009], we have explored the performance of the SO and IS for target selection with horizontal saccades only. The results indicated that the IS provided 57% faster target selection with 1.9 higher throughput than DT did. Current study investigates the performance of saccade-driven target selection methods on a 2D plane (viz. computer monitor). The results indicate that the proposed methods show an increased throughput and task completion performance compared to the conventional DT selection method.

In summary, this paper presents two major themes: 1) theoretical design and evaluation of the proposed saccade-driven methods and 2) practical evaluation results from a MD-FL task and a real-time eye-gaze-guided video game.

2. THEORETICAL DESIGN & EVALUATION

2.1 Target Selection Methods

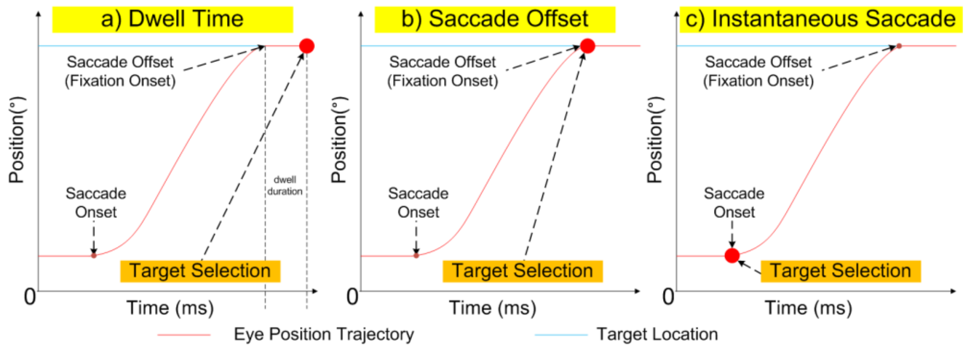

Figure 1. Eye movement driven target selections a) Dwell Time b) Saccade Offset c) Instantaneous Saccade

2.1.1 *Dwell Time*

The DT is designed to select a target when a user fixates on a target (Figure 1.a), for a specified period of time which is usually 100ms. or greater (the value of 100 ms. is employed in our work) [Zhai, Morimoto and Ihde 1999; Zhang and MacKenzie 2007].

2.1.2 *Saccade Offset*

The SO is designed to select a target at the coordinates of the end of a saccade (Figure 1.b), which makes it faster than the DT.

2.1.3 *Instantaneous Saccade*

The design goal of the IS method is to select a target at the very beginning of a saccade (Figure 1.c). Theoretical design of the IS requires two components: amplitude and direction prediction at the onset of a saccade. Following two subsections provide the details of both.

Saccade Amplitude Prediction

Two Dimensional Linear Homeomorphic Oculomotor Plant Model (2D-OP) and Two State Kalman filter (TSKF) were employed for saccades' amplitude prediction.

The 2D-OP models Oculomotor Plant (mechanics of the eyeball, surrounding tissues, and extraocular muscles) as a set of linear components representing major anatomical structures of the plant. The 2D-OP is driven by a neuronal control signal which is sent to each muscle individually. As a result the 2D-OP is capable of simulating accurate saccadic signal on a 2D plane, given the onset and the offset coordinates of a saccade.

The details of the model are published elsewhere [Komogortsev and Khan 2008; Komogortsev and Jayarathna 2008].

The TSKF is a Kalman filter with two states - position and velocity – where acceleration of the eye is approximated by white noise. A chi-square test monitors the difference between predicted and observed eye-velocity, providing the mechanism for the amplitude prediction described below. The details of the TSKF and chi-square test calculation are presented in [Koh, Gowda and Komogortsev 2009; Komogortsev and Khan 2008].

Saccadic signal properties were investigated by simulating 3640 saccades via the 2D-OP at 1000Hz with amplitudes ranging from 1-40° and tilted to 0-90°. Each saccade trajectory was processed by the TSKF deriving a sequence of the chi-square test values. The resulting chi-square test signal had a distinct shape represented by two peaks that existed for every saccade (Figure 2.a). The occurrence of the first peak, counting from the beginning of a saccade, remained in the range of 9-13ms (M=11.89ms, SD=1.01) for the saccade amplitude range of 1-40º. The second peak occurred closer to the end of a saccade. Employing linear regression, a formula was derived connecting the amplitude of the future saccade (A_{sac}) to the first peak in the chi-square test signal (χ^2_{peak}), resulting in R^2=0.98 fit to the simulated data.

$$A_{sac} = -0.002815 \cdot \chi^4_{peak} + 0.7336 \cdot \chi^2_{peak} - 3.494 \qquad \mathbf{1}$$

It is possible to construct a relationship between saccade's amplitude and the time of the first peak occurrence using regression as: $T_{peak} = 0.0876 \cdot A_{sac} + 10.4\, ms$, providing R^2=0.68 and indicating that the IS provides almost constant selection time regardless of the distance to the target.

Saccade Direction Prediction

In the Cartesian coordinate system, the direction between two points can be obtained by finding the direction of the vector $(x' - x) + (y' - y)i$, computed as

$$\mathrm{Dir} = \tan^{-1}\left(\frac{y' - y}{x' - x}\right) \qquad \mathbf{2}$$

In our tests, the same approach is employed where Dir is the saccade direction measured in degrees, (x, y) is the coordinates of the saccade onset, and (x', y') is the coordinates of the point at which saccade direction has to be determined. The direction prediction by the IS occurs when the first peak of the chi-square test signal is detected.

Equations 1 and 2 create the basis for the IS in the case of a 1000Hz eye position signal. However, the actual signal from the eye tracker is frequently produced at a lower sampling rate and is susceptible to noise [Duchowski 2007]. This necessitates an investigation into the lower frequency scenario and the effect of the noise on the prediction accuracy.

Lower Sampling Rate and Noisy Signal

To test the lower sampling case, we decided to consider a sampling frequency of

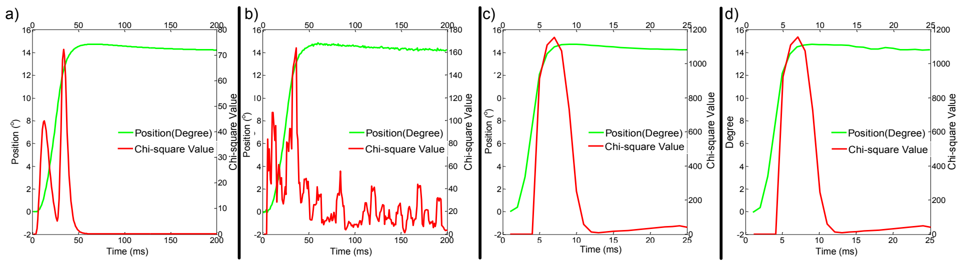

Figure 2. a) Chi-square test behavior during a saccade simulated at 1000Hz with no added noise b) Same as a) but with white noise added c) Chi-square test behavior during a saccade simulated at 120Hz with no added noise d) Same as c with white noise added.

120Hz – the de-facto frequency today for major vendors [SMI 2010; Tobii 2009]. The equation to predict future saccade is amplitude using regression based on the chi-square test with 120Hz sampling frequency is:

$$A_{sac} = -2.011 \cdot 10^{-6} \cdot \chi^4_{peak} + 0.01715 \cdot \chi^2_{peak} + 3.967 \qquad \mathbf{3}$$

with R^2=0.98. It is important to note that only one chi-square test peak exists with 120Hz sampling frequency. The time of the peak, counting from the onset of a saccade, stays in the range of 48-64ms (M=56 ms, SD=8) for the saccade's amplitude range of 1-40°, indicating that it would take 5.3-5.8 times longer to predict saccade's offset coordinates when compared to the case with 1000Hz sampling frequency.

The noise present in the eye tracker can be approximated via precision of the equipment, i.e., minimum amount of the rotation of the eye globe that the eye tracker can recognize. In our test, value of 0.1° was selected, which is equivalent to the equipment with precision of 0.1°. Figure 2 illustrates the cases where saccadic signal was generated via the 2D-OP with the sampling frequency of 1000Hz and 120Hz with and without added noise. Simulation results that discuss the magnitude of the prediction error are presented in Section 7.

2.1.1 *Hybrid Saccade*

The Hybrid Saccade (HS) is the method that is targeted to combine the speed of the IS and the stability of the DT selection. The HS tries to select a target very quickly using the IS on the initial attempt and tries to stabilize the performance in case of failure by switching to the DT until the target is successfully selected. Such factors as noise, low accuracy and the sampling frequency of the eye tracker or individual characteristics of the human visual system such as large number of overshoots/undershoots (see section 6.4 for more details) might contribute to the failure of the initial attempt. If such challenges are present on the initial selection attempt, they are likely to continue to exist. Therefore, to provide a more stable selection performance on the subsequent chain of attempts, the DT method was selected. The alternative choice of using the SO method after the failure of IS was not implemented, because our previous research indicated that the DT is more accurate than the SO [Komogortsev, Ryu, Do and Gowda 2009], therefore providing a better opportunity to stabilize the selection performance.

2.2 Selection Time

The general formula for the estimation of the amount of time saved by a faster selection method over a slower one is given by the following equation:

$$T_{saved} = 100 \cdot \left(1 - \frac{T_{S1}}{T_{S2}}\right) \qquad \mathbf{4}$$

where T_{S1} is the selection time of a faster method (S1) and T_{S2} is the selection time of a slower method (S2).

Target selection time of the DT is given by:

$$T_{\mathrm{DT}} = T_{tar\_acq} + T_{sac\_dur} + T_{dwell\_time} \qquad \mathbf{5}$$

In this equation, target acquisition (T_{tar_acq}) is defined as a time interval between the appearance of a target and the onset of the saccade leading to that target[1]. T_{sac_dur} is the duration of a saccade that lands the eye on the target, and is approximated by [Carpenter 1977]:

$$T_{sac\_dur} = (2.1 \cdot A_{sac} + 22)/1000 \qquad \mathbf{6}$$

where A_{sac} is the saccade amplitude measured in degrees.

The SO target selection time is:

$$T_{\mathrm{SO}} = T_{tar\_acq} + T_{sac\_dur} \qquad \mathbf{7}$$

The IS target selection time can be estimated as:

$$T_{IS} = T_{tar\_acq} + \frac{\mathrm{k}}{f} \qquad \mathbf{8}$$

where f is the sampling frequency of the eye tracker, and k is the number of eye position samples needed for a saccade's amplitude prediction. Therefore, $\frac{k}{f}$ is the amount of time

[1] This time interval is also called saccadic latency [Leigh and Zee 2006]. The name target acquisition is selected to represent its logical meaning for target selection research. Average saccadic latency for jumping targets was reported as 200ms [Leigh and Zee 2006], therefore 200 ms value is employed in this work.

in seconds required to predict the amplitude of a saccade. The results of time saving for all methods are reported in Section 5.

3. PRACTICAL EVALUATION

This section provides the description of the MD-FL Test and the eye-gaze-driven game “Balura” designed to conduct practical evaluation of the target selection methods.

3.1 Multi-Directional Fitts’ Law Test

The MD-FL provides a framework for measuring performance of a target selection method [Zhang and MacKenzie 2007]. The MD-FL presents a sequence of targets displayed at the various eccentricities from the center of the screen initialing each subsequent selection from the center (see supporting video file for an illustration).

General implementation guidelines for the MD-FL, presented by Zhang and MacKenzie [2007] were followed. The target width was fixed to 64 pixels (1.52°) and only distance to the target was varied following recommendation by Guiard who argued for necessity of reducing the number of degrees of freedom in the Fitts’ Law task by either fixing the target’s width or the distance [Guiard 2009]. The MD-FL was conducted with target distances that started in the most commonly exhibited range of saccade amplitudes (~6-8º [Foulsham et al. 2008; Tatler and Vincent 2009]) and extended to the screen’s periphery where the performance gains of saccade driven-selection are to be more pronounced due to larger saccade amplitudes. Specifically, three eccentricity levels were tested a) short (300 pixels or 7.12º), b) medium (375 or 8.93°), and c) long 450 (10.71°). Each eccentric target distance level consisted of 16 possible selections. Each trial started with an initial target appearing at the center of the screen. As soon as the initial target was selected (the initial selection was always done by the DT method to ensure same baseline for all methods), a target appeared on the screen at a new location and the timer for the selection time was initiated. Participants were instructed to select the new target as soon as possible by looking at the target. The target was available for the selection until it was selected successfully, i.e., sometimes participants had to make several selection attempts to achieve this. Once participants selected the target, the target at the center of the screen appeared again, initiating a new target selection sequence. This sequence was repeated until participants successfully selected 8 targets for each distance level. For each trial, the coordinates of the initial selection attempt and the successful selection together with selection times were recorded. Section 5 provides the performance results.

3.2 Balura Game

Balura is a real-time eye-gaze-guided video game designed to simulate massive battlegrounds where two teams of players compete to achieve a given objective. Balura follows the guidelines suggested for the development of the eye gaze guided applications by Koh and colleagues [Koh, Gowda and Komogortsev 2009] including individual component size and the visual feedback. Visual feedback was provided in the form of a highlighted border when the onset of a fixation was detected to help the user to dwell on the target. Saccade-driven selection methods (except the HS) were not affected by the visual feedback (good discussion on the impact of the feedback on the eye-gaze-guided interfaces is provided by Majaranta *et al.* [Majaranta et al. 2006]). Balura presents 20 red and 20 blue balloons that are randomly moving throughout the screen. Blue balloons represent the players of the friendly team and red balloons represent the opposing team. The main objective is to pop all red balloons as quickly as possible. The selection of the "friendly" balloon does not induce any damage, but highlights the boundaries around the balloon for the visual feedback to the user. The selection of the opposing team's balloon results in its elimination from the game. The target selection experience is designed to be similar to the game World of Warcraft® [Blizzard 2009], therefore to simulate World of Warcraft® player's behavior, all balloons stop at random time intervals, then start moving again in random directions and with random velocity (see supporting video file for an illustration). Participants were instructed to pop the balloons of the opposing team as quickly as possible by looking at them.

4. EXPERIMENT SETUP

4.1 Apparatus

The experiments were conducted with Tobii x120 eye tracker, which is represented by a standalone unit connected to a 19-inch flat panel screen with the resolution of 1280x1024. The eye tracker performs binocular tracking with the following characteristics: accuracy 0.5°, spatial resolution 0.2°, drift 0.3° with eye position sampling frequency of 120Hz [Tobii 2009]. The Tobii x120 model allows 300x220x300 mm freedom of head movement. Nevertheless, a chin rest was employed for higher accuracy and stability.

4.2 Eye Movement Classification & Target Selection

All targets selection methods require real-time eye movement classification performance. Specifically, the DT needs the computation of fixation duration and its coordinates, the SO method requires the detection of the offset of a saccade and its coordinates, and the IS method requires the detection of the onset of a saccade. These requirements were achieved with the help of the real-time eye movement identification protocol [Koh et al. 2010] and a classification algorithm based on a Kalman filter [Koh, Gowda and Komogortsev 2009]. As a result, each target selection method provided coordinates for the selection. If the provided coordinates were within target boundaries (no tolerance area was provided), a successful selection was made.

4.3 Participants & Quality of the Eye Movement Data

Students at Texas State University-San Marcos volunteered to participate in the experiments. The eye-movement data accuracy procedure described in [Koh, Gowda and Komogortsev 2009] was administered for each participant to ensure the quality data. Participants with average positional error greater than 2° and/or data loss greater than 20% did not participate in the experiment. Remaining 14 participants had the average positional error of 1.14° (SD=0.44) and the average data loss of 13.70% (SD=13.09). Participants' ages were from 19 to 45 (M = 27.5, SD=8.31).

4.4 Sequence of Experiments

First, participants performed the accuracy test. Those who passed the accuracy test executed the MD-FL task followed by the Balura game. Each participant was able to complete (i.e. all presented targets were successfully selected without any timeout errors) the MD-FL and Balura game by all target selection methods. The presentation order of the methods was randomized to reduce learning effects.

4.5 Performance Metrics

Completion time: For the MD-FL task completion time represents the time interval between the onset of a trial and the moment when a target is successfully selected. For Balura game completion time signifies the moment when the last opposing team's balloon is eliminated.

Throughput: Throughput is a measurement of performance that envelops both the speed and accuracy of a selection method by a user. The throughput was computed following the methodology described by Zhang & MacKenzie [2007] with movement time represented by completion time metric described above.

Average Magnitude Error: To evaluate positional accuracy of the proposed methods, Average Magnitude Error (AME) computed by the following equation was employed

$$\mathrm{AME} = \frac{\sum_{i=1}^{n} \sqrt{(x_i - \bar{x}_i)^2 + (y_i - \bar{y}_i)^2}}{n} \qquad \mathbf{9}$$

where (x_i, y_i) and $(\bar{x}_i, \bar{y}_i)$ are actual and predicted targets coordinates correspondingly.

Error Rate: The error rate is computed as the number of the initial target selection attempts that fail to land on the target divided by the total number of the initial target selection attempts.

5. RESULTS

5.1 Theoretical results

5.1.1 *Target's Selection Time Savings*

Table I. Theoretical evaluation of time savings for the saccade amplitude range of 1-40º, tilted 0-90º

Method	SO (Ideal)	IS (Ideal)	SO (1000Hz)	IS (1000Hz)	SO (120Hz)	IS (120Hz)
DT	31-25%	38-51%	31-25%	38-48%	31-25%	20-36%
SO	-	10-34%	-	5-31%	-	0.56-16%[2]

Ideal presents the case where saccade's amplitude prediction can be done at the first millisecond of its trajectory (k=1 in the equation (11)). Remaining cases represent sampling frequency of 1000Hz and 120Hz. Theoretical analysis assumes that the selection time for the SO and the DT is not affected by either sampling frequency or the eye movement classification algorithm.

5.1.2 *Accuracy of Saccade Prediction*

Table II shows theoretical accuracy of the IS method in case of the single saccade leading to the target, measured by the AME when saccades are simulated by the 2D-OP model at 1000Hz and 120Hz. Direction of a saccade is computed based on equation (2) and the amplitude is predicted based on equations (1) or (3) depending on the sampling frequency.

Table II. Saccades' landing point prediction error

Sampling frequency, noise condition	AME
1000Hz, no noise	3.88º
1000Hz, white noise added	13.03º
120Hz, no noise	5.71º
120Hz, white noise added	6.4º

A sampling frequency of 1000Hz with no noise allowed better prediction accuracy of saccade offset coordinates. The prediction accuracy decreased substantially when noise was added. However, the amount of increase in the AME with 120Hz was smaller than with 1000Hz, indicating larger impact of noise on prediction accuracy in higher frequencies. It is important to note that the 120Hz case with noise, while yielding smaller AME than the 1000Hz case with noise, takes approximately 5.3-5.8 times longer to estimate the coordinates of the saccade's landing point.

[2] At 120Hz the IS starts outperforming the SO selection when saccade amplitude exceeds 18º.

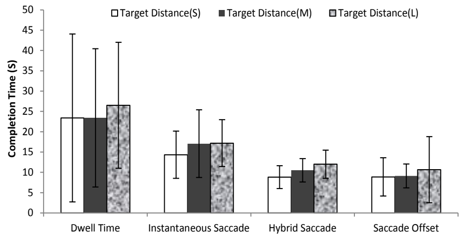

Figure 3. Average completion time for the successful individual target selection in the MD-FL task.

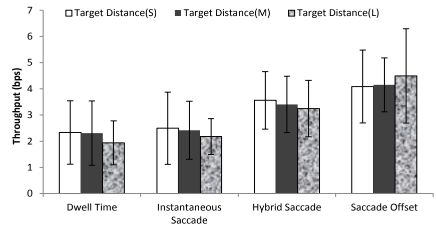

Figure 4. Average throughput of the target selection computed as the result of the MD-FL task

5.2 Practical Evaluation

Practical implementation of the saccade-driven methods employed 120Hz amplitude prediction model by equation (3) and direction prediction model by equation (2) based on the acceptable performance of the simulated data with added noise.

5.2.1 *Multi-Directional Fitts' Law Task*

Completion Time

Figure 3 presents the results. The difference in completion times between methods was statistically significant for each distance category: short - $F(3,36)=5.78$, p=0.0025, medium - $F(3,36)=6.57$, $p=0.0012$, and long - $F(3,36)=8.91$, $p=0.0002$. Post hoc analyses using the Scheffé post hoc criterion for significance indicated that DT completion times were significantly longer than HS and SO for all distance categories, $p<0.05$, but not significantly longer than IS, p>0.05. Among saccade-driven methods, the SO was the fastest and the IS was the slowest, however, there was no significant difference using Scheffé post hoc criterion for all distance categories, p>0.05. The large difference between the DT and SO can be explained, in part, by the challenge of an eye-fixation classification in real-time when parts of the signal contain invalid data (average data loss of 13.7% in our setup).

Table III provides comparative performance in terms of the selection time savings computed by equation (4) for short, medium, and long target distances respectively. Faster methods (second row) are compared to the slower methods (first column).

Table III. MD-FL: completion time saving

Distance	*Short*			*Medium*			*Long*		
Method	IS	HS	SO	IS	HS	SO	IS	HS	SO
DT	39%	62%	62%	27%	55%	61%	35%	55%	60%
IS	-	52%	19%	-	38%	46%	-	30%	38%
HS	-	-	-1%	-	-	13%	-	-	11%

The completion time savings provided by the IS over the DT method were similar to those reported in the theoretical evaluation results in Table I. The SO selection was faster than the IS, contrary to the theoretical evaluation. This phenomenon can be explained by the fact that each target was frequently selected not by a single saccade-fixation movement, but by a sequence of such saccades and fixations due to the presence of complex oculomotor behavior such as undershoots/overshoots (see Section 6.4 for details), inaccuracies associated with amplitude/direction prediction, and the variability in the eye movement classification algorithms [Komogortsev et al. 2010]. Therefore, the SO method which is void of the prediction errors (but still susceptible to the eye movement classification inaccuracies) provided better performance. The HS performance was better than the IS, indicating the advantage of combining the IS and the DT methods together when prediction difficulties are present from the noise and/or sampling frequency.

Throughput

Figure 4 presents the results. The difference in throughput between methods was statistically significant for each distance category: short - $F(3,36)=9.56$, $p<0.0001$, medium - $F(3,36)=9.89$, $p<0.0001$, and long - $F(3,36)=19.57$, $p<0.0001$. Post hoc analyses using the Scheffé criterion indicated that DT throughput was significantly smaller than HS and SO for all distance categories, but not significantly smaller than IS, $p<0.05$. Among saccade-driven methods, the SO was the largest and the IS was the smallest, however, there was no significant difference in throughput using Scheffé post hoc criterion for all distance categories, $p>0.05$

Table IV. MD-FL: comparative performance for throughput increase between target selection methods for short, medium, and long distances respectively

Distance	*Short*			*Medium*			*Long*		
Method	IS	HS	SO	IS	HS	SO	IS	HS	SO
DT	7%	53%	76%	5%	47%	80%	12%	67%	132%
IS	-	42%	64%	-	40%	71%	-	49%	106%
HS	-	-	15%	-	-	22%	-	-	39%

Initial Selection Attempt

Information about initial attempt for target selection allows assessing the initial spatial accuracy of a target selection method and allows comparing actual performance of the IS selection to the theoretical estimation.

<u>*Average Magnitude Error*</u>

Table V presents spatial accuracy of the initial selection attempt via the AME metric. In general, the AME contains four possible types of errors a) calibration error of the eye tracking equipment for a given screen area, b) human error as a result of the complex oculomotor behavior such as undershoots, overshoots, and etc. (more discussion is provided in Section 6.4) c) prediction algorithm's error, and d) eye movement classification error due to the selection of the eye movement classification algorithm and its parameters (for more details see [Komogortsev, Gobert, Jayarathna, Koh and Gowda 2010]).

Table V. Average magnitude error between the center of the target and the initial selection coordinates

Method	AME
DT	3.14º
IS	8.11º
HS	8.22º
SO	7.01º

The DT selection yielded the lowest AME. The SO was the second most accurate, but the difference between the DT and SO was quite substantial (~132%). We hypothesize this difference was specifically due to the classification algorithm that selects accurate coordinates for the center of fixation based on a Kalman filter [Koh, Gowda and Komogortsev 2009], but can potentially dampen saccadic behavior [Komogortsev, Gobert, Jayarathna, Koh and Gowda 2010]. The increase in the AME for the IS method over the SO method was small (<15%). The AME for the IS was larger than the one by the theoretical evaluation (Table II, 120Hz, noise added) by 27%. This might be due to the difference in noise characteristics of the eye tracking equipment and the white noise used in the theoretical evaluation, which can allow prediction algorithm producing larger errors than expected.

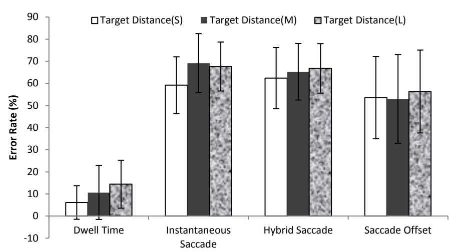

Figure 5. Average error rate during initial target selection attempt in the MD-FL task

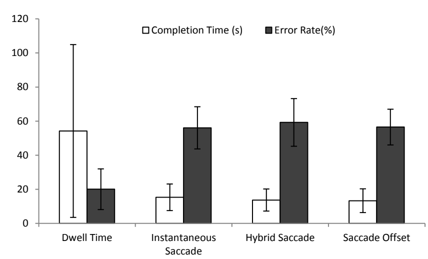

Figure 6. Average completion time and error rate for the initial target selection attempt for Balura game

Error Rate

The error rate was computed by taking into the account the number of the initial target selection attempts that failed to land on the target and the total number of the initial target selection attempts.

For the DT, error rates increased with the targets' distance range (Figure 5), but did not exceed 14%. Both results are comparable to the study conducted by Zhang and MacKenzie [2007]. The IS yielded highest error rates peaking at 69% for the medium distance targets. The SO yielded lowest error rates peaking at 56% for the long distance targets. The highest error rate for the HS was 67% for the long distance targets. Large differences in the error rates between the DT and saccade-driven methods can be attributed to the presence of complex oculomotor behavior discussed in more detail in Section 6.4. The linear relationship between distance to the target and error rates was less prominent for the saccade-driven methods than for the DT.

The difference in error rates between methods was statistically significant for each distance category: short - $F(3,36)=81.96$, $p<0.0001$, medium - $F(3,36)=65.25$, $p<0.0001$, and long - $F(3,36)=65.25$, $p<0.0001$. Post hoc analyses using the Scheffé criterion indicated that DT error rates were significantly smaller than IS, HS and SO for all distance categories, $p<0.05$. Among saccade-driven methods, there was no significant difference in error rates using Scheffé post hoc criterion for all distance categories, $p>0.05$.

5.2.2 *Balura*

Table VI. Balura: completion time savings for various target selection methods

Method	IS	HS	SO
DT	72%	75%	75%
IS	-	10%	13%
HS	-	-	3%

The DT required approximately 72-75% longer to complete the objectives of the Balura than saccade-driven methods (Figure 6, Table VI). Completion time savings were close to the short range targets during the MD-FL task, with saccade-driven methods showing even more advantage in the selection time savings. The performance of the saccade-driven methods was quite similar among each other. This can be explained by the fact that the average saccades' amplitude for those methods was 3.91º while the amplitude for the DT was approximately twice as high. The differences between the amplitudes could be attributed to the quick selection of close by targets by saccade-driven methods, while selection attempts by the DT do not result in a fast selection (balloons are moving in random directions while randomly stopping), making the participant wander to the targets that are further away.

The error rates for Balura indicate the failure of selecting an opposing team's balloon on the initial attempt. The error rates were similar to the MD-FL task. The DT yielded slightly higher rates (20.1% vs. 14%), possibly due to the increased difficulty of selection caused by moving targets. The error rates between the saccade-driven methods were very similar (the difference did not exceed 6%) and close to the highest error rate of the SO in the MD-FL task.

The difference in completion time was statistically significant, $F(3,39)=10.81$, $p<0.0001$. Post hoc analyses using the Scheffé criterion indicated that DT completion time was significantly longer than the others, $p<0.05$. The difference in error rate was statistically significant, $F(3,39)=50.17$, $p<0.0001$. Post hoc analyses using the Scheffé criterion indicated that DT error rate was significantly smaller than the others, $p<0.05$.

6. DISCUSSION

6.1 Target Selection Performance 2D vs. 1D

The results from Komogortsev *et al.* [2009] that investigated saccade-driven methods for horizontal target selection using 1D Fitts' Law (1D-FL) task were compared to that from the MD-FL of this study. **Completion time:** the MD-FL increased target selection time for all methods a) DT –125% b) IS –248% c) SO – 25%. **Throughput:** the MD-FL decreased throughput for all methods a) DT - 58%, b) IS - 76%, c) SO - 45%. **Error rate:** the MD-FL decreased error rates for all methods a) DT - 30.7%, b) IS - 39.04%, c) SO - 35%. Thus, target selection in a 2D plane, represented by MD-FL, seems to be more challenging than a horizontal (1D) selection, represented by 1D-FL, for both the HVS and target selection methods. Specifically, the MD-FL was more challenging to the HVS due to the random direction of target appearance and three possible distances to the target. Algorithmic saccade direction and amplitude prediction were more challenging due to the additional dimension, with its own noise characteristics. Prolonged selection time resulted in smaller error rates during the initial target selection attempt. The results indicate better tolerance of the SO method to 2D plane transition which is indicated by smaller selection time increase, throughput decrease, and moderate decrease of the error rates as compared to the other methods.

6.2 Fastest Saccade-driven Target Selection Method

Practical results obtained in this work via MD-FL show overall faster selection times and higher throughput of the SO over the IS method. Theoretical evaluation results (Table I) indicate that for 120Hz eye tracker the IS will outperform the SO for selection amplitudes exceeding 18º. Practical tests involving Balura game allow analyzing such occurrences. The results indicate that the initial selection attempt for the IS was on average 87% faster than the SO (the part of this substantial increase is attributed to shorter saccadic latencies among recorded during the IS selections at this amplitude). Therefore, the IS method would be the most beneficial in cases when the distance to the targets is going to be large, e.g., an action game where the targets are coming from the periphery.

In general, theoretical evaluation results indicate that the IS has potential to outperform the SO in the cases of the higher sampling frequency and low noise. Two areas have to be investigated to make this possible from the implementation side: 1) software/hardware platform that allows performing both REMI and the Kalman filter processing of 1000Hz signal in real-time, and 2) noise reduction algorithms that do not break chi-square peak to saccade amplitude relationship illustrated by equation (1). In addition, models of oculomotor movement biases such as presented by Tatler and Vincent [2009] can be incorporated to improve prediction performance in most common saccadic movement directions.

6.3 Applications of Saccade-driven Target Selection

Proposed saccade-driven methods do not attempt to solve the Midas Touch problem [Istance, Bates, Hyrskykari and Vickers 2008; Jacob 1990]. The speed of the proposed saccade-driven selection makes it difficult to select an auxiliary input method that would allow effectively resolving problem. Previously proposed solutions such as Snap Clutch [Istance, Bates, Hyrskykari and Vickers 2008] and EyePoint [Kumar, Paepcke and Winograd 2007] would prolong the selection time by removing the speed benefit of the saccade-driven methods. The solution to Midas Touch problem lies in a careful interface design. One of the successful examples is Balura game where an erroneous selection is not detrimental to the interaction experience and practical evaluation indicates substantially faster task completion times (up to 132%). We also hypothesize that hierarchical pie menus [Urbina and Huckauf 2010] would allow saccade driven methods to be applicable for the menu-driven navigation and typing.

6.4 Impact of Complex Oculomotor Behavior on Presented Outcomes

Complex oculomotor behavior (COB) which is represented in part by undershoots, overshoots, dynamic and express saccades [Komogortsev et al. 2010] can provide a substantial impact on overall completion times (Figure 3), spatial accuracy (Table V), and error rates (Figure 5). The COB results in a case where initial movement to the target, e.g. saccade, does not land within the boundaries of the target and subsequent correction occurs via a corrective small amplitude saccade or a glissade (post-saccadic drift). In cases when the amount of COB becomes large, the completion times for all target selection methods are extended due to the sequence of the initial and corrective movement leading to the target's selection. Among the target selection methods considered in this work, the DT selection is affected the least by the COB, because of the filtering behavior represented by minimum fixation duration of 100ms and merging techniques employed by eye movement classification algorithms [Koh, Gowda and Komogortsev 2010; Komogortsev, Gobert, Jayarathna, Koh and Gowda 2010].

10. CONCLUSION

In this paper, we have proposed several saccade-driven target selection methods. The performance of the methods was compared to the conventional fixation-based selection method called Dwell-Time (DT) using the two tests - the Multi-Directional Fitts' Law (MD-FL) Task and a real-time eye-gaze driven Balura application created to simulate interaction environments of the action oriented computer games.

Theoretical evaluation results indicate that the Instantaneous Saccade (IS) method, designed to select a target at the very beginning of a saccade, is the fastest selection method, providing almost constant relationship between the distance to the target and the time of selection.

The MD-FL task indicated that all saccade-driven methods outperformed the DT by 35-62% in terms of the completion time savings and increased the throughput by 5-132%. Saccade Offset (SO) method, designed to select a target at the end of a saccade, was the best performer among saccade-driven methods. It was hypothesized that the performance of the IS method was degraded in this practical implementation due to the challenge of estimating saccade parameters in the cases of noisy signal and lower sampling frequency of the eye position signal.

The Balura game indicated even higher time savings due to saccade-driven methods (72-75%) over the DT selection. Following results of the theoretical analysis, target selection performance with the IS and the saccade-driven methods was analyzed for target selection when the amplitudes exceeded 18º during Balura. The performance of the IS was superior to the SO by 87%, supporting the results of the theoretical evaluation.

Saccade-driven methods deliver very fast completion times for target selection applications. Future work lies in the areas of (1) addressing the implementation challenges that arise at high sampling frequencies of the eye movement signal, where saccade-driven methods provide higher performance, and (2) developing interaction schemes where erroneous selection does not negatively impact user's experience. Currently, saccade-driven methods would be beneficial to action-oriented video games.

REFERENCES

BLIZZARD 2009. World of Warcraft Blizzard.

CARPENTER, R.H.S. 1977. *Movements of the Eyes*. Pion, London.

DOUGLAS, S.A., KIRKPATRICK, A.E. AND MACKENZIE, I.S. 1999. Testing pointing device performance and user assessment with the ISO 9241, Part 9 standard. In *Proceedings of the Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*, Pittsburgh, Pennsylvania, United States1999 ACM.

DUCHOWSKI, A. 2007. *Eye Tracking Methodology: Theory and Practice*. Springer.

E3 2009. Electronic Entertainment Expo.

FOULSHAM, T., KINGSTONE, A. AND UNDERWOOD, G. 2008. Turning the world around: Patterns in saccade direction vary with picture orientation. *Vision Research 48*, 1777-1790.

GUIARD, Y. 2009. The problem of consistency in the design of Fitts' law experiments: consider either target distance and width or movement form and scale. In *Proceedings of the Proceedings of the 27th international conference on Human factors in computing systems*, Boston, MA, USA2009 ACM, 1518980, 1809-1818.

HUCKAUF, A. AND URBINA, M.H. 2008. Gazing with pEYEs: towards a universal input for various applications. In *Proceedings of the 2008 symposium on Eye tracking research & applications* ACM, Savannah, Georgia, 1-4.

ISTANCE, H., BATES, R., HYRSKYKARI, A. AND VICKERS, S. 2008. Snap clutch, a moded approach to solving the Midas touch problem. In *Proceedings of the Proceedings of the 2008 symposium on Eye tracking research & applications*, Savannah, Georgia2008 ACM.

JACOB, R.J.K. 1990. What you look at is what you get: eye movement-based interaction techniques. In *Proceedings of the Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people*, Seattle, Washington, United States1990 ACM.

KOH, D.H., GOWDA, M. AND KOMOGORTSEV, O.V. 2010. Real Time Eye Movement Identification Protocol. In *ACM Conference on Human Factors in Computing Systems (CHI)* ACM, Atlanta, GA, 1-6.

KOH, D.H., GOWDA, S.A.M. AND KOMOGORTSEV, O.V. 2009. Input evaluation of an eye-gaze-guided interface: kalman filter vs. velocity threshold eye movement identification. In *ACM Symposium on Engineering Interactive Computing Systems* ACM, Pittsburgh, PA, USA, 197-202.

KOMOGORTSEV, O., V. AND KHAN, J. 2008. Eye Movement Prediction by Kalman Filter with Integrated Linear Horizontal Oculomotor Plant Mechanical Model. In *ACM Eye Tracking Research & Applications Symposium*, Savannah, GA, 229-236.

KOMOGORTSEV, O.V., GOBERT, D. AND DAI, Z. 2010. Classification Algorithm for Saccadic Oculomotor Behavior Texas State University-San Marcos http://ecommons.txstate.edu/cscitrep/18/, San Marcos, 1-4.

KOMOGORTSEV, O.V. AND JAYARATHNA, U.K.S. 2008. 2D Oculomotor Plant Mathematical Model for eye movement simulation. In *IEEE International Conference on BioInformatics and BioEngineering (BIBE)*, 1-8.

KOMOGORTSEV, O.V. AND KHAN, J. 2007. Kalman Filtering in the Design of Eye-Gaze-Guided Computer Interfaces. In *12th International Conference on Human-Computer Interaction (HCI 2007)*, Beijing, China, 1-10.

KOMOGORTSEV, O.V., RYU, Y.S., DO, H.K. AND GOWDA, S.A.M. 2009 Instantaneous Saccade Driven Eye Gaze Interaction. In *ACM International Conference on Advances in Computer Entertainment Technology*, 1-8.

KOMOGORTSEV, V., GOBERT, V., JAYARATHNA, S., KOH, D. AND GOWDA, S. 2010. Standardization of Automated Analyses of Oculomotor Fixation and Saccadic Behaviors. *Biomedical Engineering, IEEE Transactions on 57*, 2635-2645.

KUMAR, M., PAEPCKE, A. AND WINOGRAD, T. 2007. EyePoint: practical pointing and selection using gaze and keyboard. In *Proceedings of the Proceedings of the SIGCHI conference on Human factors in computing systems*, San Jose, California, USA2007 ACM.

LEIGH, R.J. AND ZEE, D.S. 2006. *The Neurology of Eye Movements*. Oxford University Press.

MACKENZIE, I.S. AND ZHANG, X. 2008. Eye typing using word and letter prediction and a fixation algorithm. In *Proceedings of the Proceedings of the 2008 symposium on Eye tracking research & applications*, Savannah, Georgia2008 ACM.

MAJARANTA, P., MACKENZIE, S., AULA, A. AND RÄIHÄ, K. 2006. Effects of feedback and dwell time on eye typing speed and accuracy. *Univers. Access Inf. Soc. 5*, 199-208.

MAJARANTA, P. AND RÄIHÄ, K. 2002. Twenty years of eye typing: systems and design issues. In *Proceedings of the Proceedings of the 2002 symposium on Eye tracking research & applications*, New Orleans, Louisiana2002 ACM, 507076, 15-22.

NAKAYAMA, M. AND TAKAHASI, Y. 2008. Estimation of certainty for multiple choice tasks using features of eye-movements. In *Proceedings of the Proceedings of the 2008 symposium on Eye tracking research and applications*, Savannah, Georgia2008 ACM.

SIBERT, L.E. AND JACOB, R.J.K. 2000. Evaluation of eye gaze interaction. In *Proceedings of the Proceedings of the SIGCHI conference on Human factors in computing systems*, The Hague, The Netherlands2000 ACM.

SMI 2010. Senso Motoric Instruments, 2010.

TATLER, B.W. AND VINCENT, B.T. 2009. The prominence of behavioural biases in eye guidance. *Visual Cognition 17*, 1029 - 1054.

TIEN, G. AND ATKINS, M.S. 2008. Improving hands-free menu selection using eyegaze glances and fixations. In *Proceedings of the Proceedings of the 2008 symposium on Eye tracking research and applications*, Savannah, Georgia2008 ACM.

TOBII 2009. Tobii technology.

URBINA, M.H. AND HUCKAUF, A. 2010. Alternatives to single character entry and dwell time selection on eye typing. In *Proceedings of the Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*, Austin, Texas2010 ACM, 1743738, 315-322.

ZHAI, S., MORIMOTO, C. AND IHDE, S. 1999. Manual and gaze input cascaded (MAGIC) pointing. In *Proceedings of the Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*, Pittsburgh, Pennsylvania, United States1999 ACM.

ZHANG, X. AND MACKENZIE, I.S. 2007. Evaluating Eye Tracking with ISO 9241 - Part 9. In *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments* Springer Berlin / Heidelberg, 779-788.