To appear in the proceedings of the ACM International Conference on Advances in Computer Entertainment Technology, 2009

Instantaneous Saccade Driven Eye Gaze Interaction

Oleg V. Komogortsev
Department of Computer Science, Texas State University-San Marcos, USA
+1 512 245 0349
ok11@txstate.edu

Young Sam Ryu
Ingram School of Engineering
Texas State University-San Marcos
+1 512 245 4467
yryu@txstate.edu

Do Hyong Koh
Department of Computer Science, Texas State University-San Marcos, USA
+1 512 245 0349
dk1132@txstate.edu

Sandeep M. Gowda
Department of Computer Science, Texas State University-San Marcos, USA
+1 512 245 0349
sm1499@txstate.edu

ABSTRACT
In this paper, we introduce and evaluate a new Instantaneous Saccade (*IS*) selection scheme for eye gaze driven interfaces where the speed of the target selection is of utmost importance. In the *IS* selection scheme, target selection occurs at the start (onset) of a saccade requiring only constant amount of time to be completed. The *IS* performance is compared to the conventional Dwell Time (*DT*) selection scheme where target selection is triggered when a user fixates on an object for a certain amount of time. The IS method is also compared to the Saccade Offset *(SO)* selection scheme where target selection occurs at the end of a saccade. All three schemes were evaluated in terms of task completion time and the throughput of input performance in horizontal target selection task by six subjects. Results show that the Instantaneous Saccade selection was 57% faster than the *DT* selection to complete a task. In terms of throughput comparison, the throughput of the *IS* selection is 1.9 times greater than the throughput of *DT* selection. We hypothesize that Instantaneous Saccade selection will be beneficial in gaming environments that require fast very interaction speeds.

Categories and Subject Descriptors
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *evaluation/methodology, input devices and strategies, interaction styles.*

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

Keywords
Human Computer Interaction, Kalman Filter, Interaction Technique Evaluation, Eye Tracker, Saccade Trajectory Prediction.

1. INTRODUCTION
Today's video games incorporate sophisticated interaction techniques that give users a more exciting experience. 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 [6]. In addition, Microsoft presented a controller-free interface code-named “Project Natal” for the Xbox 360 game console [6]. Project Natal aims to provide controller-free, full body motion capture, voice recognition, and facial recognition for gaming and entertainment. However, the use of eye movement recognition has not been applied to consumer-oriented video game controls.

In the Human Computer Interaction (HCI) community, eye-guided interfaces and their interaction techniques have recently attracted research interest [9, 11, 13, 16, 20, 22, 29]. An eye-tracking device can be used as an interactive input modality for users with disabilities, or as an additional interaction method for other users [4]. Smith and Nicholas Graham [27] and Komogortsev and Khan [11] explore the use of an eye tracking technique for video game control. Smith and Nicholas Graham [27] applied eye gaze interaction to a first-person shooter game, a role playing game, and an action/arcade game, concluding that an eye gaze interaction technique can produce a new experience for current video game users.

To enable eye-gaze guided interaction, the raw eye position signal must be analyzed by sophisticated algorithms. Parts of the signal are classified into meaningful components such as fixations (movements that occur when gaze is dwelling on objects), saccades (movements between two separate fixations), and pursuits (movements that occur when eyes are tracking moving objects). Fixations (or Dwell Time) are the most common modality for an eye-gaze-guided computer interaction [18, 21, 32]. This modality assumes that a fixated object is selected when the duration of a fixation reaches a predefined threshold. Very little research has employed saccade-based interaction [30]. Pursuit-based interaction seems to be an unexplored topic in the HCI community. In this paper, we consider the fixation-based mode of interaction and introduce a new saccade-based interaction.

Several models exist for eye movement classification, including the most commonly used Velocity-based Threshold (I-VT) model described by Salvucci and Goldberg [24]. The Velocity-based Threshold model is often used because of the ease of its implementation and its low computational cost. However, the model is not robust and is not capable of handling high levels of noise present in eye position data. Thus, in this work we used a Kalman Filter to improve Eye Movement Identification.

The Kalman filter is a recursive estimator that computes a future estimate of the dynamic system state from a series of incomplete and noisy measurements. Eye trackers frequently fail to report eye position data, and the reported data is susceptible to noise due to the individual anatomical properties of users and limited spatial resolution of the equipment. Therefore, a Kalman Filter framework can be used to provide more accurate and robust estimation of the eye position signal. At the same time, the Kalman Filter is capable of classifying eye movements [25]. Komogortsev and Khan [13] were the first ones to discuss the use of a Kalman Filter in a real-time eye-gaze-guided computer interface, and they have indicated that the filter can be successfully used during eye-tracking failures. Kumar et al. [17] presented the case where a Kalman Filter provided smoothing to a raw eye position signal, thereby increasing the stability of the input. Koh at el. [11] provides a comprehensive evaluation of the interface performance when driven by a Kalman Filter eye movement classifier. In their research, using a Kalman Filter for the eye gaze interface provides better performance than using the Velocity Threshold model [11].

Previous researchers have introduced interaction schemes that go beyond the realm of interaction based on Dwell Time. Blanch and Ortega [1] introduced a rake cursor interaction technique that combines mouse controlled selection and cursor activation by the gaze. Spakov and Miniotas [28] developed a target expansion scheme during real-time eye tracking calibration. In addition, Miniotas et al. [21] evaluated target expansion during tasks controlled by eye-gaze. Salvucci and Anderson [23] present an intelligent gaze-added interface that uses a probabilistic algorithm. One of the goals of the above mentioned methods was to improve the accuracy and ease of selection.

The Instantaneous Saccade (*IS*) selection method that we propose in this paper is designed for the eye gaze driven HCI systems that require extreme interaction speeds. In the most common eye gaze driven systems, interface component selection uses Dwell Time (*DT*) methods that involves data buffering for at least 100 ms [18, 26, 32]. In such interfaces, the duration of the detected fixation initiates a "click". The goal of the selection method proposed in this paper is to make a selection as soon as the eye movement to a new target is detected. Fixation-based selection necessitates data buffering and therefore introduces a delay in the system. Future pursuit-based selection methods may require some data buffering for pursuit detection. Here, we are interested in extreme interaction where the speed of "clicking" a target is of utmost importance. Urbina and Huckauf [30] utilized saccade selection for typing. In that scheme, a component of a pie-like menu was selected when a saccade crossed the outer border of a slice. In that interaction scheme, pie dimensions did not vary, and the landing point of a saccade was not important.

In this work, we are specifically interested in an interaction scheme where the landing point of a saccade is important and indicates the coordinates of a target that has to be selected. Therefore in the IS interaction scheme, target selection is performed at the onset of a saccade. The direction of the future saccade is performed by the model presented in [15]. Saccade Amplitude prediction is performed with the help of the Horizontal Oculomotor Plant Mechanical Model [12], and uses a regression approach similar to [15]. The resulting IS selection method is free of a buffering delay and independent of the distance to the target.

To explore the benefits of saccade based selection schemes we introduce and evaluate an interaction scheme called a Saccade Offset (*SO*) Selection. In the *SO* selection scheme, selection is triggered at the offset of a saccade (the moment when the saccade ends). In other words, *SO* selection can be considered as *DT* selection without Dwell Time threshold.

The main motivation for *IS* selection is to provide a new interactive experience which improves target selection speed. For example, *IS* selection can provide a more exciting experience to users when they are playing first-person shooting game applications, which require extreme reaction speed. More specifically, *IS* selection can provide an exciting experience in multiplayer games such as World of Warcraft [2], where players compete in a hostile environment and each saved millisecond might add to the probability of victory. Based on the findings provided in this paper, we expect that *IS* selection will be used in the eye-gaze-guided interfaces designed for games. Theoretical evaluation of *IS* scheme shows that there is a possibility of 45% improvement in speed, while practical results suggest a possible 57% improvement. Further, pointing device throughput calculations driven by the *IS* scheme suggest a 91% increase in throughput.

The paper is organized as follows: description of the Eye Movement Identification by a Kalman Filter (I-KF) is followed by Saccade Amplitude Prediction by the Horizontal Oculomotor Plant Mechanical Model (OPMM). These discussions are followed by presentation of the Instantaneous Saccade Driven Eye-Gaze Interaction, and Pointing Device Evaluation by Throughput. The paper concludes with a presentation of experimental results, detailed discussion of the effectiveness of the approach, and conclusions based on theoretical and experimental considerations.

2. EYE MOVEMENT IDENTIFICATION BY A KALMAN FILTER (I-KF)

The Kalman Filter is a data processing algorithm that predicts a future estimate of the dynamic system state with existence of incomplete and error signals. A Kalman Filter minimizes the error between the estimated and actual values of a system's state. Only the estimated state from the previous time step and the new measurements are needed to compute the new state estimate. Many real dynamic systems do not exactly fit this model; however, because the Kalman Filter is designed to operate in the presence of noise, an approximate fit is often adequate for the filter to be quite useful [3].

Brown and Hwang [3] describe the general mathematical framework of the Kalman Filter. In our implementation, Identification by the Kalman Filter models an eye as a system with two states: position and velocity. The acceleration of the eye movement is considered to be white noise with known maximum acceleration.

The use of a Kalman Filter allows the generation of a predicted position and velocity signal. Thus, velocity prediction can be applied as a part of Chi-square test for eye movement classification [11, 13, 25] and eye position signal prediction can provide the data during eye-tracking failures. The predicted position signal is employed for fixation parameter calculation, providing a signal during tracking failures (data loss).

To appear in the proceedings of the ACM International Conference on Advances in Computer Entertainment Technology, 2009

Komogortsev and Khan [13] have presented the details of the Kalman Filter parameterization that we have employed in this work.

A Chi-square test monitors the difference between predicted and observed eye-velocity:

$$\chi^2 = \sum_{i=1}^{p} \frac{\left(\hat{\theta}_i^- - \dot{\theta}_i\right)^2}{\delta^2} \qquad (1)$$

where $\hat{\theta}_i^-$ is the predicted eye velocity computed by Kalman Filter and $\dot{\theta}_i$ is the observed eye velocity computed with the eye position signal from the eye tracker. δ is the standard deviation of the measured eye velocity during the sampling interval under consideration. Once a certain threshold of the χ^2 statistic is achieved, a saccade is detected. It was reported that the filter stability improves if δ is selected to be a constant [13]. Empirical evaluation has indicated that values of $\delta^2 = 1000$ and p=5 provide acceptable performance. The value of the χ^2 threshold was empirically selected to be 5.

3. SACCADE AMPLITUDE PREDICTION BY HORIZONTAL OCULOMOTOR PLANT MODEL

Komogortsev and Khan [12] developed a linear horizontal Oculomotor Plant Mechanical Model (OPMM). The OPMM consists of the eye globe and two extraocular muscles: lateral and medial recti. The model accounts for such anatomical properties of the eye as muscle location, elasticity, viscosity, eye-globe rotational inertia, muscle active state tension, length tension and force velocity relationships.

The existence of the OPMM allows examination of the behavior of the Chi-square test over the course of a saccade of any amplitude. The behavior of the Chi-square test presents a unique signature for saccades of any amplitude. Specifically, the Chi-square test signal peaks twice over the course of a saccade. The first peak of the Chi-square test signal can be observed at the moment that is very close to the onset of a saccade. A second peak occurs closer to the end of a saccade. This peaking behavior is consistent for saccades of any amplitude. The onset of the first peak occurs in the range of 12-16ms for a saccade range of 1-30º . Figure 1 presents an example of the Chi-square test behavior during a horizontal saccade with amplitude of 15 º.

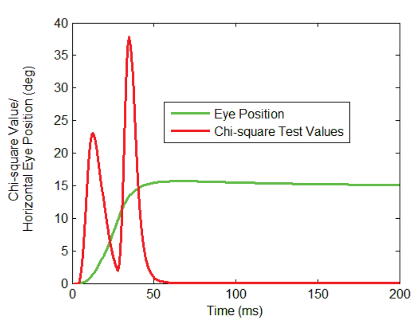

Figure 1. Chi-square Values during a single Saccade

The detection of these peaks provides the necessary means for creation of the IS selection scheme.

By simulating saccade trajectories for amplitudes of 0-40º we can create a graph that shows the dependency between the saccade amplitude and the Chi-square test value during the first peak. For simplicity, we assume that the first peak represents the onset of a saccade. Figure 2 shows this dependency and presents the corresponding regression model.

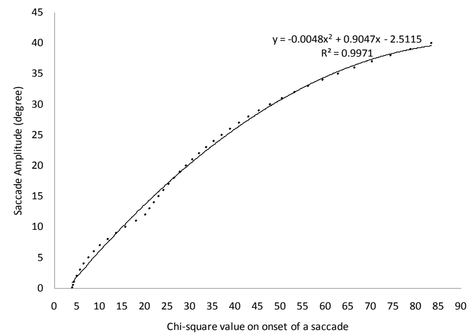

Figure 2. Saccade Amplitude Prediction Model

Based on the data represented by the graph and the corresponding regression formula we can create a saccade's amplitude prediction formula that allows predicting the amplitude of a future saccade based on the Chi-square test value during the first signal peak.

$$A_{sac} = -0.0048 \cdot {\chi^2}^2 + 0.9047 \cdot \chi^2 - 2.5115 \qquad (2)$$

where A_{sac} is an amplitude of a saccade and χ^2 is the Chi-square test value at the first signal peak. Equation 2 serves as a basis for the IS selection discussed next.

It is very interesting to note that the time interval that is required to achieve the first peak does not depend on the amplitude of the resulting saccade. By employing the OPMM model in this theoretical test we found out that the first peak of the chi-square test signal occurs on average at 12.2ms (SD=0.94) from the beginning of a saccade.

4. INSTANTANEOUS SACCADE DRIVEN EYE-GAZE INTERACTION

The main idea behind Instantaneous Saccade (*IS*) interaction is that target selection happens at the very beginning of a saccade. Figure 3 shows conventional Dwell Time (*DT*) selection. In the *DT* selection, selection happens when the user fixates on the target for the amount of time specified as the *DT* Threshold. In our research, we used 100 ms as the *DT* Threshold.

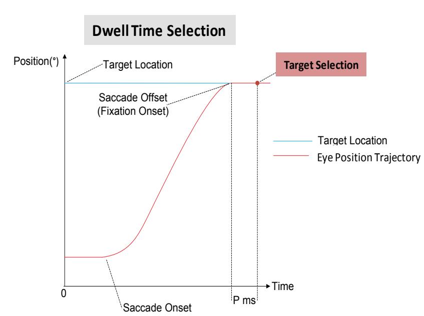

Figure 3. DT Selection

Theoretically, Saccade Offset (*SO*) selection happens as soon as the user's eye gaze lands on the target. Therefore, *SO* selection is triggered faster than *DT* selection. The time saved by the *SO* selection is the same time as the value of *DT* Threshold. Figure 4 shows an example of the *SO* selection.

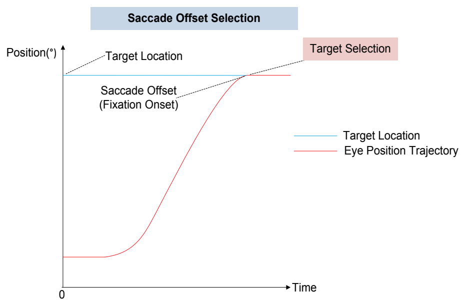

Figure 4. SO Selection

Figure 5 presents *IS* selection. In *IS* selection, saccade direction is determined by analyzing the velocity signature using the mechanism described in [15]. Also, saccade amplitude is predicted by the Chi-square test value mechanism described in the previous section. Thus, the user's intended target location is obtained by predicted saccade trajectory.

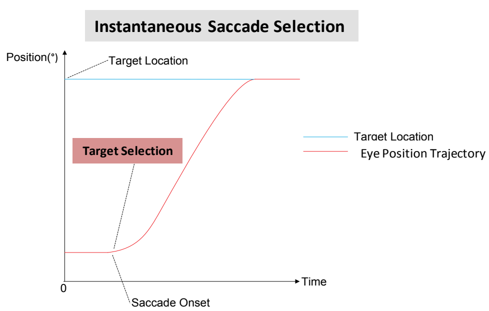

Figure 5. IS Selection

It is important to theoretically evaluate the time-savings which the IS selection provides when compared to the *DT* selection. In the IS selection a "click" occurs immediately after the onset of a saccade is detected. Saccade onset detection is performed by the mechanism described in the previous two sections.

The longer the distance is to the target, the larger the amount of time that can be saved by the *IS* selection. Selection time in any interaction scheme involving eye movements can be connected to the saccade amplitude that is required to reach a target. Saccade duration can be computed by the formula (3).

$$T_{sac\_dur} = 2.1 \cdot A_{sac} + 22 \tag{3}$$

In equation (3), T_{sac_dur} is the duration of a saccade measured in ms, and A_{sac} is saccade amplitude measured in degrees [19].

The amount of time T_{DT} required for the *DT* selection method to "click" on the new target can be computed as in (4).

$$T_{DT} = T_{tar\_acq} + T_{sac\_dur} + T_{dwell\_time} \tag{4}$$

In equation (4), T_{tar_acq} is the amount of time the brain requires to calculate the neuronal control signal for extra-ocular muscles to rotate the eye globe. Usually, the target acquisition time is around 200ms due to the delays in Human Visual System (HVS) [19]. T_{dwell_time} is the duration of the fixation that has to be detected for the selection to take place. In our work, we use the smallest possible selection time of 100ms.

Selection time of the IS T_{IS} can be calculated by equation (5).

$$T_{IS} = T_{tar\_acq} + \frac{\mathrm{k}}{f} \tag{5}$$

In equation (5), f is the sampling frequency of the eye tracker, and k is the number of eye position samples needed for saccade amplitude prediction, therefore, $\frac{k}{f}$ is the amount of time in seconds that is required to predict the amplitude of a saccade.

Time saved can be computed by the formula in (6).

$$T_{saved} = 100 \cdot \left(1 - \frac{T_{IS}}{T_{DT}}\right) \tag{6}$$

Theoretically we wanted to estimate the largest savings in time given our eye tracker sampling frequency of 120Hz. The fastest IS scheme will require just one position sample at the onset of a saccade (k=1). In such case equation (6) indicates that the amount of time saved by *IS* interaction is 35% for very close targets and 45% for the targets which are 30° away from user's current location. When *SO* selection scheme is compared to the *DT* selection the SO allows to save approximately 26% of time. *IS* provides 30% selection time reduction if compared to *SO* selection.

5. POINTING DEVICE PERFORMANCE EVALUATION

The most common evaluation measures for an interaction scheme are speed, accuracy, and throughput [4]. In this paper, speed is equivalent to individual task completion time (selection time of a target). Accuracy is usually reported as an error rate – the percentage of selections outside the target. These measures are typically analyzed over a variety of tasks. Throughput, measured in bits per second, is a composite measure derived from both the speed and accuracy of the selections make as a result of an interaction scheme. Equation (7) explains Throughput calculation for the successful target selection.

$$Throughput = \frac{ID_e}{CT} \tag{7}$$

where CT is the completion time of the successful selection of a target.

Equation (8) calculates the effective index of difficulty.

$$ID_e = \log_2\left(\frac{D}{W_e} + 1\right) \tag{8}$$

The term ID_e is the effective index of difficulty that is measured in "bits." It is calculated from D, the distance to the target, and W_e the effective width of the target. The concept of the "effective" width (W_e) is critical since W_e is the width of the distribution of selection coordinates computed over a sequence of trials which can be obtained by the equation in (9).

$$W_e = 4.133 \cdot SD_x \qquad (9)$$

In equation (9), *SDx* is the standard deviation of the differences between selection and the center of the target coordinates that is measured along the axis of approach to the target. This implies that W_e reflects the spatial variability or accuracy in the sequence of trials. Therefore, throughput captures both the speed and accuracy of the user performance.

6. METHODOLOGY

6.1 Participants

A total of 7 participants volunteered to evaluate the performance of our algorithms. Participants' ages were from 20 to 25 (*mean* = 23). None of the participants had prior experience with eye tracking. Among these participants, 6 had normal vision and 1 wore glasses or contacts. One subject had an abnormal vision attributed to strabismus and astigmatism.

6.2 Apparatus

The experiments were conducted with Tobii x120 eye tracker, which is represented by a standalone unit connected to a 24-inch flat panel screen with resolution of 1980x1200. 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. The Tobii x120 model allows 300x220x300 mm freedom of head movement. Nevertheless, a chin rest was employed for higher accuracy and stability.

6.3 Procedure

6.3.1 Accuracy test

This procedure involves participants looking at 17 sequentially presented points that are uniformly distributed on the computer screen. When a subject fixates at each point, the raw eye position signal is processed by Identification by a Kalman Filter and corresponding fixation parameters such as location coordinates, the onset time, and the duration are determined. The coordinates of the eye position within the detected fixations are compared to the center of presented stimulus. This allows for the computation of error between reported location of the gaze and the actual gaze point. At the end of the recording, the error values are averaged between all points and presented on the computer screen. Additionally, an accuracy test computes and presents a data loss parameter that indicates the amount of erroneous (not detected) eye position samples provided by an eye tracker for the participant.

6.3.2 Performance Evaluation Task

The experiment we designed was similar to the one dimensional Fitts' Law task described by Douglas et al. [4]. The experiment was separated into several trials with three different selection schemes. Each trial started with a target appearing at a random horizontal location on the screen (vertical coordinate of the target was fixed to the middle of the screen) and subject's goal was to select this target as soon as possible. The target was available for the selection until it was successfully selected, i.e., sometimes a subject had to make several selection attempts before the target was successfully selected. As soon as the initial target was selected (the initial selection was always done by the *DT* method) a second target appeared on the screen at a random location and the timer for the selection completion time was initiated. The width of the target was fixed to 5° of the visual angle. Randomly generated distance to the target was 24° on average. Once the coordinates of the first selection point by each selection scheme were recorded the trial ended and the duration for the target selection completion time was recorded. Each experiment consisted of 30 trials with 15 trials requiring rightward selection and 15 trials requiring leftward selection. In order to avoid a learning effect, the order of the selection schemes was randomly selected. Figure 6 presents an example of a selection trial.

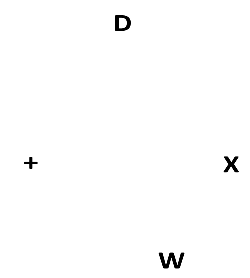

Figure 6. Selection trial example

7. RESULTS

7.1 Accuracy and Data Loss

The data from the subject with abnormal vision was discarded because of an average measured error of 3.10° and 17.5% of data loss. The average accuracy for the remaining six subjects was 1.51° (SD=0.41) and the average data loss was 5.11% (SD=2.58).

7.2 Completion Time - Successful Target Selection

The graph below represents average completion time of the target selection task.

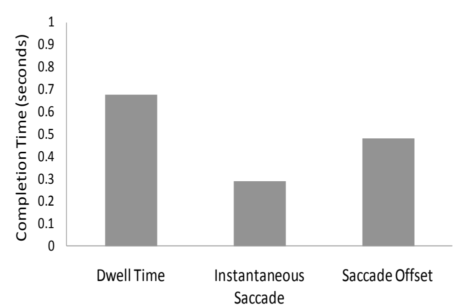

Figure 7. Completion time of the successful target selection of three interaction schemes

Average completion time for the successful target selection in terms of the *DT* selection was 678.41ms (SD=265.17). Average completion time for the *IS* selection was 291.05ms (SD=43.80) indicating 57% decrease in completion time. Average completion time for the *SO* selection was 479.91ms (SD=208.01), which is approximately 30% faster than *DT* selection. The *IS* provides 40% reduction in completion time if compared to the *SO* interaction scheme. The difference in completion time was statistically significant, $F(2,10)=7.82$, $p<0.009$.

It is important to note that practical validation brings higher time savings in terms of the successful target selection than theoretical evaluation provided in Section 4, 57% vs. estimated 45%. This is possible due to physiological differences between the recruited subject pool and the theoretical model presented in Section 4. Another contributing factor was the number of selection attempts before the target was actually selected, i.e., on average a subject had to make two such attempts.

7.3 Throughput

Figure 8 shows the performance of the selection schemes in terms of throughput. *DT* selection scheme provided an average throughput for target selection of 5.23 bps (SD=0.24). This result is higher than the result of 3.06 bps presented by Zhang and MacKenzie [27], probably due to much shorter dwell time (100ms vs. 500ms) required for selection. The *SO* selection scheme provided an average Throughput of 7.74 bps (SD=0.88), which is approximately 48% higher than the throughput provided by *DT* selection. The *IS* selection provided an average throughput of 9.98 bps (SD=1.72), which is a 91% increase in throughput when compared to the *DT* selection and approximately 29% higher when compared to *SO* selection. The difference in throughput was statistically significant, $F(2,10)=26.21$, $p<0.0001$.

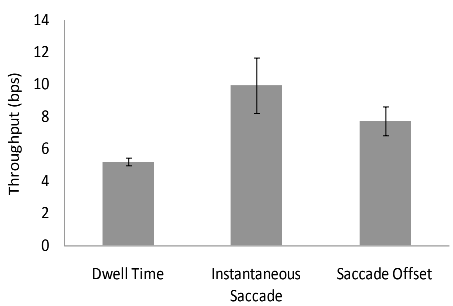

Figure 8. Throughput of three interaction techniques

7.4 Initial Target Selection Time

While data presented by Figure 7 presents individual selection task completion times, the graph does not present the data from the initial target selection attempt. Such data is important because it provides information about possible improvement in completion time provided that all initial selection errors are eliminated. Figure 9 presents the data.

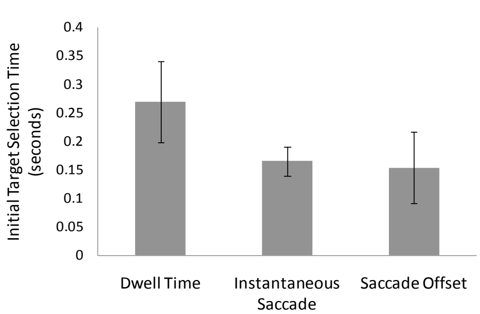

Figure 9. Initial target selection time

Average initial target selection time for the *DT* selection was 269.52ms (SD=71.44). Average initial target selection attempt time for the *SO* selection was 154.10ms (SD=63.11), indicating a 43% decrease when compared to the *DT* method. Average initial target selection attempt time for the *IS* selection was 165.66ms (SD=25.39), indicating 39% reduction when compared to *DT* selection. The difference in initial target selection time was statistically significant, $F(2,10)=14.50$, $p=0.0011$.

It is very interesting to note that for initial target selection, the *SO* method provided slightly shorter times than the IS method. This result can be explained by the fact that the subjects tend to have micro saccades (saccades with an amplitude of less than 0.5º [31]) while fixating on the initial target to start the trial. Such saccades may trigger *OS* selection even before the eyes move to the intended target. Saccades like this are small enough to prevent selection by the IS method (first Chi-square test peak is not detectable) but still detectable to trigger *OS* selection.

7.4.1 Error Rate

The initial selection attempt for each presented interaction scheme provides a high number of selection errors. Specifically, the average error rate for the initial target selection attempt for the *DT* was 29% (SD=5). The *IS* selection average error was 92% (SD=7). The average error rate for the *SO* selection was 87% (SD=11). The difference in average error rate was statistically significant, $F(2,10)=107.38$, $p<0.0001$.

7.5 Target Sequence Total Completion Time

As a result of the experiment each subject completed 30 target selections using each interaction scheme. This sequence of target selections can be treated as a large task where the total completion time would indicate the amount of time required for each scheme to achieve a task. Figure 10 presents these results.

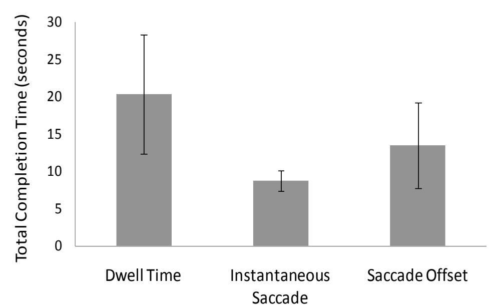

Figure 10. Total completion time

In terms of the total task completion time, the *IS* based interaction was 56% faster than the *DT* scheme, and 35% faster that the *SO* scheme. The *SO* scheme was 34% faster than the *DT* scheme. The difference in target sequence total completion time was statistically significant, $F(2,10)=7.91$, $p=0.0087$. This result indicates that *IS* provides much faster interaction speeds than either *DT* or *SO* for our experimental setup.

8. DISCUSSION

8.1 Error Rate

Based on our observations, participants were able to complete the task faster when Instantaneous Saccade (IS) selection was applied.

This result is remarkable even though the *IS* selection has the highest error rate the initial selection. This result seems peculiar but if we consider the nature of the Human Visual System (HVS) we can find the explanation for this phenomenon. Saccade movements in the HVS are not always precise and are subject to frequent undershoots or overshoots [19]. Such HVS behavior naturally decreases the accuracy of target selection. Nevertheless, such errors do not negate the advantage of the *IS* interaction method which, in terms of the completion time, is almost two times smaller than *DT* selection.

Practical implementation of the *IS* selection involves some technical difficulties. The *IS* selection occurs only when saccade movement is detected by Kalman Filter using a Chi-square test threshold. Since the saccade movement is the most rapid eye movement in the HVS and has very short duration, few eye position samples are available to determine the first peak in the Chi-square test (we are using a 120Hz sampling frequency of the eye tracker). Specifically there are approximately 1.91 eye position samples for making a saccade trajectory prediction. Sometimes the peak in the Chi-square test signal is not correctly identified and the initial target selection attempt does not succeed. A maximum of 5 initial attempts was recorded during all the experiments for the *IS* selection scheme, and the average number of attempts was 1.91 (SD=0.3). Another interesting fact was that 41% of correct predictions occurred at the second prediction attempt. Thus, for the future work, we plan to improve the Instantaneous Saccade selections by incorporating more robust peak detection methods.

When saccade detection fails, e.g. due to noise, *SO* selection is delayed even if the eye correctly lands on the target location. In cases like this the system will wait until the user's eye makes an additional detectable saccade which triggers the selection. In rare cases like this the *DT* scheme is superior to the *SO* method.

8.2 Applications for IS Interaction

In its current state, Instantaneous Saccade-based interaction will be favorable to gaming applications. In this mode of interaction, accuracy is less important than targeting speed. In the action oriented game World of Warcraft [2], players compete in battlegrounds – places where a team of players must overpower the opposite team to complete a task. The interaction between players occurs extremely fast with a player targeting an enemy player and then casting an instantaneous damaging spell. The speed of an enemy player's selection eventually translates into a victory or a loss. In this type of interaction even a saved fraction of a second is a significant achievement. In World of Warcraft, it is impossible to select or damage a friendly player, therefore there is no penalty if instantaneous selection misses the target. To alleviate the impact of misses, the IS selection can be run in parallel with a DT selection scheme allowing the user to perform successful selection in cases when saccade amplitude prediction misses the target.

We are aware of the fact that during the interaction task each subject was presented only with one target for selection at a time. This paper specifically explored potential benefits that can be gained as a result of a new interaction method. The tasks that have multiple target choices for selection will be a part of our future work.

9. CONCLUSION

In this paper, we presented an Instantaneous Saccade (*IS*) selection method that allows target selection at the beginning of the eye movement (saccade) that moves the eye to the target. The new *IS* method was compared to the conventional Dwell Time (*DT*) selection method with dwell time duration of 100ms, and the Saccade Offset (*SO*) selection method (essentially a *DT* selection method without the dwell time threshold). Each method was evaluated through a series of target selection tasks where the completion time of an individual task and the completion time of a series of tasks was recorded. In addition, each selection method was evaluated by calculating pointing device (eye tracker) throughput when driven by each of the presented selection schemes. The results indicate that IS selection is 57% faster than *DT* selection, and 40% faster than *SO* selection. In terms of the sequence of targets selection task the IS method is 57% faster than *DT* and 35% faster than *SO*. The *IS* method provides approximately twice as much throughput (9.98bps vs 5.23bps) when compared to the *DT* method. The throughput achieved by the IS method is approximately 29% higher than the *SO* method.

While providing a significant increase in completion time and throughput, the IS selection method does not address the Midas Touch problem discussed by Jacob [10]. Considering this limitation, we expect that the IS method will be beneficial in virtual environments where accidental selection is not detrimental to the user experience. More specifically we envision that the IS method is applicable in gaming environments where targeting speed is much less important than the accuracy of the initial selection.

10. ACKNOWLEDGMENTS

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.

11. REFERENCES

[1] Blach, R., and Ortega, M. 2009. Rake cursor: improving pointing performance with concurrent input channels. The 27th international conference on Human factors in computing systems (pp.1415-1418). New York: ACM Press.

[2] Blizzard, "World of Warcraft." vol. 2009: Blizzard, 2009

[3] Brown, R. G., and Hwang, P. Y. C. 1997. Introduction to random signals and applied Kalman filtering: with MATLAB exercises and solutions. New York: Wiley.

[4] 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. The ACM Conference on Human Factors in Computing Systems - CHI '99 (pp. 215-222). New York: ACM.

[5] Duchowski, A. T. 2007. Eye Tracking Methodology: Theory and Practice. London, UK: Springer-Verlag.

[6] Electronic Entertainment Expo, E3. 2009. http://www.e3expo.com/

[7] Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices. International Standard, International

Organization for Standardization (2000) ISO. ISO/DIS 9241-9.

[8] Irwin, D. 1992. Visual Memory Within and Across Fixations. Eye Movements and Visual Cognition, Springer Series in Neuropsychology (pp. 146-165). New York: Springer.

[9] Istance, H., R. Bates, et al. 2008. Snap clutch, a moded approach to solving the Midas touch problem. Proceedings of the 2008 symposium on Eye tracking research and applications. Savannah, Georgia, ACM.

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

[11] Koh, D. H., S. A. M. Gowda, et al. 2009. Input evaluation of an eye-gaze-guided interface: kalman filter vs. velocity threshold eye movement identification. Proceedings of the 1st ACM SIGCHI symposium on Engineering interactive computing systems, Pittsburgh, PA (pp.197-202). New York: ACM Press.

[12] Komogortsev, O. V. and Khan, J. I. 2008. Eye movement prediction by Kalman filter with intergraded linear horizontal oculomotor plan mechanical model. The Eye Tracking Research & Applications Symposium (pp. 229-236). New York: ACM Press.

[13] Komogortsev, O. V., and Khan, J. 2007. Kalman filtering in the design of eye-gaze-guided computer interfaces. Human-Computer Interaction, Part III, HCII 2007, LNCS 4552, 679–689.

[14] Komogortsev, O. V., and Khan, J. 2008. Eye movement prediction by Kalman filter with integrated oculomotor plant mechanical model. The 2008 symposium on Eye tracking research & applications (pp. 229-236). New York: ACM Press.

[15] Komorgotsev, O. V., Ryu, Y. and Koh, D. 2009. Quick Models for Saccade Amplitude Prediction. Journal of Eye Movement Research, 3(1):1, 1-13.

[16] Kumar, M. and T. Winograd. 2007. Gaze-enhanced scrolling techniques. Proceedings of the 20th annual ACM symposium on User interface software and technology. Newport, Rhode Island, USA, ACM.

[17] Kumar, M., Klingner, J., Puranik, R., Winograd, T., and Paepcke, A. 2008. Improving the accuracy of gaze input for interaction. The 2008 symposium on Eye tracking research & application (pp. 65-68). New York: ACM Press.

[18] Kumar, M., Paepcke, A., and Winograd, T. 2007. Eye point: practical pointing and selection using gaze and keyboard. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, San Jose, CA.

[19] Leigh, R. J., and Zee, D. S. 2006. The Neurology of Eye Movement. UK: Oxford University Press.

[20] MacKenzie, I. S. and X. Zhang. 2008. Eye typing using word and letter prediction and a fixation algorithm. Proceedings of the 2008 symposium on Eye tracking research and applications. Savannah, Georgia, ACM.

[21] Miniotas, D., Spakov, O., and MacKenzie, I. S. 2004. Eye gaze interaction with expanding targets. CHI '04 extended abstracts on Human factors in computing systems (pp.1255-1258). New York: ACM Press.

[22] Nakayama, M. and Y. Takahasi. 2008. Estimation of certainty for multiple choice tasks using features of eye-movements. Proceedings of the 2008 symposium on Eye tracking research and applications. Savannah, Georgia, ACM.

[23] Salvucci, D. D., and Anderson, J. R. 2000. Intelligent gaze-added interfaces. The SIGCHI conference on Human factors in computing system (pp. 273-280). New York: ACM Press.

[24] Salvucci, D. D., and Goldberg, J. H. 2000. Identifying fixations and saccades in eye tracking protocols. The Eye Tracking Research and Applications Symposium (pp.71-78). New York: ACM Press.

[25] Sauter, D., J., Martin B., Di Renzo, N., and Vomscheid, C. 1991. Analysis of eye tracking movements using innovations generated by a Kalman filter. Med Biol Eng Comput, 29(1), 63–69.

[26] Sibert, L. E., and Jacob, R. J. K. 2000. Evaluation of eye gaze interaction. The SIGCHI conference on Human factors in computing systems (pp. 281-288). New York: ACM Press.

[27] Smith, J. D., and Nicholas Graham, T. C. 2006. Use of eye movements for video game control. The 2006 ACM SIGCHI international conference on Advances in computer entertainment technology. ACM International Conference Proceeding Series; Vol. 266 Article No. 20. New York: ACM Press.

[28] Spakov, O., and Miniotas, D. 2005. Gaze-based selection of standard-size menu items. The 7th international conference on Multimodal interfaces (pp.124-128). New York: ACM Press.

[29] Tien, G. and Atkins M. S. 2008, Improving hands-free menu selection using eye gaze glances and fixations. In proceedings of the 2008 symposium on Eye tracking research and applications. Savannah, Georgia: ACM, 2008.

[30] Urbina, M. H., and Huckauf, A. 2007. Dwell time free eye typing approaches. The 3rd Conference on Communication by Gaze Interaction, COGAIN 2007 (pp.65-70). Leicester, UK.

[31] Yarbus, L. 1967. Eye movements and vision. Moscow: Institute for Problems of Information Transmission Academy of Sciences of the USSR.

[32] Zhai, S., Morimoto, C., and Ihde, S. 1999. Manual and gaze input cascaded (MAGIC) pointing. The SIGCHI conference on Human factors in computing systems: the CHI is the limit (pp.246-253). New York: ACM Press.

[33] Zhang, X., and MacKenzie, I. S. 2007. Evaluating eye tracking with ISO 9241 - Part 9. Human-Computer Interaction, Part III, HCII 2007, LNCS 4552, 77