

Eye Movement Prediction by Oculomotor Plant Kalman Filter with Brainstem Control

Oleg V. KOMOGORTSEV¹, Javed I. KHAN²

(1. Department of Computer Science,
Texas State University-San Marcos, 601 University Drive, San Marcos, TX, USA;

2. Department of Computer Science,
Kent State University, Kent, OH, USA)

Abstract: Our work addresses one of the core issues related to Human Computer Interaction (HCI) systems that use eye gaze as an input. This issue is the sensor, transmission and other delays that exist in any eye tracker-based system, reducing its performance. A delay effect can be compensated by an accurate prediction of the eye movement trajectories. This paper introduces a mathematical model of the human eye that uses anatomical properties of the Human Visual System to predict eye movement trajectories. The eye mathematical model is transformed into a Kalman filter form to provide continuous eye position signal prediction during all eye movement types. The model presented in this paper uses brainstem control properties employed during transitions between fast (saccade) and slow (fixations, pursuit) eye movements. Results presented in this paper indicate that the proposed eye model in a Kalman filter form improves the accuracy of eye movement prediction and is capable of a real-time performance. In addition to the HCI systems with the direct eye gaze input, the proposed eye model can be immediately applied for a bit-rate/computational reduction in real-time gaze-contingent systems.

Keywords: Eye movement prediction, bio engineering, human computer interaction.

1 Introduction

There has been a substantial amount of research in the HCI community that investigated the use of the eye gaze information as a primary or auxiliary input to computer systems [1] [2]. This research indicates that such input is especially beneficial for target selection due to the following reasons: a) people look at a target prior to selecting it with an input device. Therefore when an eye gaze is used for selection, the selection delay introduced by an auxiliary input device such as a mouse is eliminated [1]. b) Due to the fact that the eye globe rotates in a fluid inside of an eye socket, the eye has the capability of moving much faster than the limbs

that are burdened by bone weight. Additionally the fibers inside of the extraocular muscles are fatigue resistant within certain limits of the eye movement amplitude [3]. It is noteworthy that limb muscles do not possess such properties, hence repetitive limb movements cause fatigue and excess of such movements might cause repetitive stress injury. c) The eye can provide an additional channel of input in situations where the use of limbs is restricted or unavailable, such as surgeries where both hands are occupied (laparoscopy), user interfaces for the handicapped, etc.

Despite all the benefits there are few challenges that limit eye gaze input technology, *i.e.*, cost,

accuracy, sensor and transmission delays [2].

Our work specifically concentrates on compensation of sensor and transmission delays. Our hypothesis is that delay compensation will allow creating more responsive HCI systems and provide higher level of compression in gaze-contingent systems. The delay compensation approach we selected is based on prediction of future eye movements. If the future eye gaze location is predicted accurately, target selection can be pre-made, therefore increasing the responsiveness of the HCI system with direct eye gaze input. In a gaze-contingent compression system the accurate eye movement prediction will allow to minimize the size of the high quality coded *Region of Interest* (ROI), therefore reducing overall bandwidth or computational requirements, [4], [5].

The challenge of the accurate eye movement prediction lies in the fact that the *Human Visual System* (HVS) exhibits a variety of eye movement fixation, saccades, smooth pursuit, optokinetic reflex, vestibulo-ocular reflex, and vergence [6]. In this paper, we concentrate on the first three due to the observation that these eye movements are exhibited when a person works with a computer. With great simplification, their roles are described as follows: 1) fixation – eye movement that keeps an eye gaze stable in regard to a stationary target providing visual pictures with highest acuity, 2) saccade – very rapid eye rotation moving the eye from one fixation point to another, and 3) pursuit stabilizes the retina in regard to a moving object of interest.

In our previous work [7] we created the *Two State Kalman Filter* (TSKF) eye movement prediction model. The TSKF model assumes that an eye has two states, position and velocity. Our tests indicate that the TSKF allows accurate eye movement prediction during fixations and pursuits but has poor performance during saccades. To improve the accuracy of the eye movement prediction during saccades, we have developed an *Oculomotor Plant Mechanical Model* (OPMM) [8], [9]. This model

mimics eye anatomy by considering physical properties of the extraocular muscles and the eye globe. The OPMM has six states that represent eye position, velocity, muscle location and muscle forces. Our model is based on Bahil's model [10] with two major additions: the ability to start a saccade from any point in a horizontal plane and the ability to direct a saccade in any direction in the horizontal plane. To ensure continuous eye movement prediction the mathematical equations guiding the OPMM were transformed into a linear stochastic difference equation required for a Kalman filter. We call this model the *Oculomotor Plant Kalman Filter* (OPKF). The Kalman filter framework is selected as a classical real-time predictor of the system's future state in a noisy environment with the minimization of the error between the estimated and the actual system's state [11].

This paper is a continuation of our previous work [8]. In the current paper we present: 1) new equations for muscular forces during fixations; 2) new equations for brainstem control parameters during saccades; 3) employment of the brainstem control properties in cases when the eye goes through transitions between fast (saccades) and slow (fixation, pursuits) eye movements; 4) accurate initialization of the Kalman filter state vector at the beginning of each saccade; and 5) tertiary (fixation, pursuit, saccade) velocity-based eye movement classification with detection thresholds provided by neurological literature [6]. Our experimental results indicate that above mentioned additions improve the accuracy of prediction achieved by the OPKF model by 2-3%. This improvement in accuracy is supported by simulation results involving 21 subjects and three multimedia tests. It should be pointed out that the prediction model remains linear and can be used in an online system with real-time performance constraints.

2 Human Visual System

2.1 Control

The eye globe rotates in its socket through the use

of six muscles. These six muscles are: medial and lateral recti, superior and inferior recti; superior and inferior oblique – the muscles responsible for eye rotations around its primary axis of sight; and

superior and inferior obliques can participate in intrusion, extortion, depression, elevation, adduction, and abduction of an eye [6].

During saccades, the brainstem control signal

for each muscle resembles a pulse-step function. [12]. The eye position during the onset of a saccade and the saccade's amplitude and direction define pulse and step parameters of the control signal. Once the parameters of the brainstem control signal are calculated by the brain, the control signal is sent as a neuronal discharge at the calculated frequency.

During eye fixations, neuronal discharge is performed at a constant rate that is linearly related to the eye position [12].

2.2 Oculomotor Plant Mechanical Model

Horizontal *Oculomotor Plant Mechanical Model* (OPMM) is represented by the eye globe and two extraocular muscles – lateral (LR) and medial (MR) recti. Each muscle can play the role of the agonist (AG) or the antagonist (ANT). The agonist muscle pulls the eye globe in the required direction, and the antagonist muscle resists the pull. Figure 1 presents the OPMM diagram during rightward eye movements. Each individual extraocular muscle is represented by a set of separate components generating a specific force. These components include: 1) active state tension \vec{F}_{LR} , \vec{F}_{MR} modeled as an ideal force generator; 2) series elasticity $\vec{F}_{SE_{LR}}$, $\vec{F}_{SE_{MR}}$, and length-tension components $\vec{F}_{LT_{LR}}$, $\vec{F}_{LT_{MR}}$ modeled as ideal linear strings with coefficients $K_{SE} = 2.5 \text{ grams}/^\circ$ and $K_{LT} = 1.2 \text{ grams}/^\circ$; and 3) force velocity relationship components $\vec{F}_{B_{LR}}$, $\vec{F}_{B_{MR}}$ characterized by damping coefficients $B_{AG} = 0.04 \text{ grams} - s/^\circ$, $B_{ANT} = 0.02 \text{ grams} - s/^\circ$ and the velocity of muscular contraction $\Delta\dot{\theta}_{LT_{LR}}$, $\Delta\dot{\theta}_{LT_{MR}}$. The actual force applied by lateral \vec{T}_{LR} and medial \vec{T}_{MR} recti to the

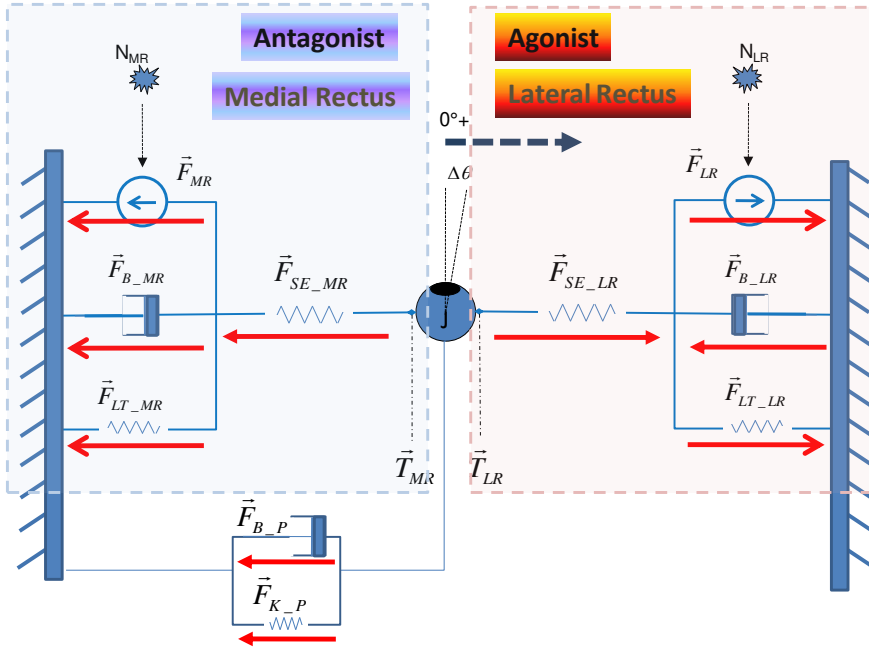


Figure 1. Oculomotor Plant Mechanical Model.

vertical eye movements.

The brain sends a brainstem control signal to each muscle to direct the muscle to perform its work. A brainstem control signal is anatomically implemented as a neuronal discharge that is sent through a nerve to a muscle [12]. The frequency of this discharge determines the level of muscle innervation and results in a specific amount of work that a muscle performs.

Each eye movement trajectory can be separated into horizontal and vertical components. The control signal for the horizontal component is generated by the premotor neurons in the pons and medulla [12] and primarily executed by the medial and the lateral recti muscles [6]. The rostral midbrain generates a brainstem control signal for the vertical eye movement component [12]. The vertical eye movement component is executed primarily by the superior, the inferior recti and the superior, the inferior oblique [6]. Muscle roles are more complex when secondary and tertiary muscle roles are considered, *i.e.*, superior and inferior recti,

eye globe is a summation of all these forces. Passive elasticity components related to the eye globe tissues combined with passive elasticity of extraocular muscles generate force \vec{F}_{K_P} and viscous elements of the orbital tissues generate force \vec{F}_{B_P} . Eye globe inertia is $J = 0.000043 \text{ grams} - \text{s}^2/\text{°}$. More detailed description of the components presented above can be found in [8], [9].

Eye globe rotation occurs as a result of brainstem control signal generating active state tension inside of each muscle therefore allowing the contracting muscle to rotate the eye globe. The next subsection presents brainstem control signal equations for fixations and saccades. The neuronal control signal for pursuits is not presented in this paper.

2.3 Brainstem Control Signal

2.3.1 Fixations

Scalar values for muscle forces \vec{T}_{LR} , \vec{T}_{MR} presented in Figure 1 were measured during a strabismus surgery – a type of surgery in which muscles are detached and then reattached to the eyeball to correct muscle dislocation. Using the values of those forces, Bahill [10] has proposed a linear relationship between the force that each muscle applies to the eye and the eye rotational position $\Delta\theta$ in its socket: $T_{AG_fix}(\Delta\theta) = 14 + 0.8|\Delta\theta|$, $T_{ANT_fix}(\Delta\theta) = 14 + 0.3|\Delta\theta|$ where $\Delta\theta$ is the eye rotation measured in degrees. $|\Delta\theta|$ is the absolute value of $\Delta\theta$. We propose changing the relationship between the antagonist force and the eye rotation to $T_{ANT_fix}(\Delta\theta) = 14 + 0.5|\Delta\theta|$. This modification allows following the trend of the eye position-force relationship identified by Bahill, but allows avoiding negative brainstem control signal values for eye positions $|\Delta\theta|$ up to 45° .

$$T_{AG_fix}(\Delta\theta) = 14 + 0.8|\Delta\theta| \quad (1)$$

$$T_{ANT_fix}(\Delta\theta) = 14 + 0.5|\Delta\theta| \quad (2)$$

During a fixation state, the scalar values of active state tensions F_{LR} , F_{MR} are assumed to be same as the brainstem control signal $N_{AG_fix}(\Delta\theta)$, $N_{ANT_fix}(\Delta\theta)$ sent to the muscle. Knowing the scalar values for the force components inside of each muscle [8], [9] we have:

$$T_{AG_fix}(\Delta\theta) = \frac{N_{AG_fix}(\Delta\theta)K_{SE}}{K_{SE} + K_{LT}} - \frac{\Delta\theta K_{SE}K_{LT}}{K_{SE} + K_{LT}} \quad (3)$$

$$T_{ANT_fix}(\Delta\theta) = \frac{N_{ANT_fix}(\Delta\theta)K_{SE}}{K_{SE} + K_{LT}} + \frac{\Delta\theta K_{SE}K_{LT}}{K_{SE} + K_{LT}} \quad (4)$$

where $K_{SE} = 2.5 \text{ grams}/\text{°}$ and $K_{LT} = 1.2 \text{ grams}/\text{°}$ are force coefficients of the series elasticity and length tension components.

Solving the system of linear equations represented by (1), (2), (3), (4) we come up with the value of brainstem control signal during fixations:

$$N_{AG_fix}(\Delta\theta) = (20.72 + 2.38|\Delta\theta|)$$

$$N_{ANT_fix}(\Delta\theta) = (20.72 - 0.46|\Delta\theta|)$$

2.3.2 Saccades

Each saccade is generated by a brainstem control signal that looks like a pulse step function [13]. This signal can be presented by the following equations:

$$N_{AG_sac}(t) = \begin{cases} N_{AG_sac_onset}, t_{sac_onset} \leq t < t_{AG_sac_pulse_onset} \\ N_{AG_sac_pulse}, t_{AG_sac_pulse_onset} \leq t < t_{AG_sac_pulse_offset} \\ N_{AG_sac_offset}, t_{AG_sac_pulse_end} \leq t < t_{sac_offset} \end{cases}$$

$$N_{ANT_sac}(t) = \begin{cases} N_{ANT_sac_onset}, t_{sac_onset} \leq t < t_{ANT_sac_pulse_onset} \\ N_{ANT_sac_pulse}, t_{ANT_sac_pulse_onset} \leq t < t_{ANT_sac_pulse_offset} \\ N_{ANT_sac_offset}, t_{ANT_sac_pulse_offset} \leq t < t_{sac_offset} \end{cases}$$

t_{name} constants present time parameters for each type of muscle and action phase. t is the time elapsed from the beginning of the saccade. The OPMM developed in this paper uses the following time constants: $t_{sac_onset} = 0$

$$t_{sac_offset} = (2.59 * |\theta_{sac_amp}| + 28.7) \text{ ms}. \quad (5)$$

θ_{sac_amp} is the amplitude of the saccade measured in degrees, $t_{sac_offset} - t_{sac_onset}$ is the duration of the saccade, calculated by the formula proposed by Fioravanti and colleagues [14]. The relationship between saccade duration and the amplitude together with the main sequence relationship (relationship between saccade amplitude and the peak velocity) are representative characteristics of saccade eye movements [6].

The agonist and antagonist muscle related time constants can be found in [8], [9].

The parameters such as amplitude, saccade onset

position, and the direction of the movement should be supplied to the OPMM in terms of brainstem control signal. Therefore, we have created the functions that transform saccade parameters into brainstem control signal: $N_{AG_sac_start}$, $N_{ANT_sac_start}$, $N_{AG_sac_end}$, $N_{ANT_sac_end}$.

$$\begin{aligned}
 N_{AG_sac_start}(\theta_{sac_onset}) &= \\
 &\begin{cases} N_{AG_fix}(\theta_{sac_onset}), & \text{if agonist prior to saccade} \\ N_{ANT_fix}(\theta_{sac_onset}), & \text{if antagonist prior to saccade} \end{cases} \\
 N_{AG_sac_pulse}(\theta_{sac_onset}, \theta_{sac_amp}) &= \\
 N_{AG_fix}(|\theta_{sac_start} + \theta_{sac_amp}|) + 190 \left(1 - e^{-\frac{|\theta_{sac_amp}|}{50}} \right) \\
 N_{AG_sac_offset}(\theta_{sac_offset}) &= \\
 &\begin{cases} N_{AG_fix}(\theta_{sac_offset}), & \text{if agonist after saccade} \\ N_{ANT_fix}(\theta_{sac_offset}), & \text{if antagonist after saccade} \end{cases} \\
 N_{ANT_sac_start}(\theta_{sac_offset}) &= \\
 &\begin{cases} N_{AG_fix}(\theta_{sac_onset}), & \text{if agonist prior to saccade} \\ N_{ANT_fix}(\theta_{sac_onset}), & \text{if antagonist prior to saccade} \end{cases} \\
 N_{ANT_sac_pulse}(\theta_{sac_amp}) &= 0.5 \\
 N_{ANT_sac_offset}(\theta_{sac_offset}) &= \\
 &\begin{cases} N_{AG_fix}(\theta_{sac_offset}), & \text{if agonist after saccade} \\ N_{ANT_fix}(\theta_{sac_offset}), & \text{if antagonist after saccade} \end{cases}
 \end{aligned}$$

The function $N_{AG_sac_pulse}(\theta_{sac_amp})$ is constructed with an assumption that muscle innervations level represented by the brainstem control value can be separated into two components: the first component depends on saccade onset value, and the second component depends on saccade amplitude. Therefore, if saccade amplitude θ_{sac_amp} is zero, $N_{AG_sac_pulse}$ would be equivalent to $N_{AG_sac_start}$. Constants inside of the $N_{AG_sac_pulse}(\theta_{sac_amp})$ function were selected empirically to minimize the error between the model output and the physiological data for large amplitude saccades.

Once the brainstem control parameters are defined, the OPMM model can be transformed into Kalman Filter form.

This section presented the dynamics of the neuronal control signal for the OPMM. The complete description of the OPMM requires six differential equations for leftward and rightward eye rotation. Both sets of those equations are described in [8], [9].

3 Oculomotor Plant Kalman Filter

3.1 Basics of Kalman Filtering

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. A Kalman Filter minimizes the error between the estimation of the system's state and the actual 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 very useful [11].

The Kalman Filter addresses the problem of trying to estimate the state $x \in \mathfrak{R}^n$ of a discrete-time controlled process that is governed by the linear stochastic difference equation [11]:

$$x_{k+1} = A_{k+1}x_k + B_{k+1}u_{k+1} + w_{k+1} \quad (6)$$

with the measurement

$$z_k = H_k x_k + v_k \quad (7)$$

The n-by-n state transition matrix A_{k+1} relates the state at the previous time step k to the state at the current step k+1 in the absence of either a driving function or process noise. B_{k+1} is an n-by-m control input matrix that relates m-by-1 control vector u_{k+1} to the state x_k . w_k is an n-by-1 system's noise vector with an n-by-n covariance matrix Q_k . $p(w_k) \sim N(0, Q_k)$. Not all variables in the state are visible to the measuring instruments. The measurement vector z_k contains state variables that are measured by the instruments. H_k is a j-by-n observation model matrix which maps the state x_k into the measurement vector z_k . v_k is a measurement noise j-by-1 vector with covariance R_k . $p(v_k) \sim N(0, R_k)$.

The Discrete Kalman filter has two distinct phases that compute the estimate of the next system's state [11].

Predict:

Predict the state vector ahead:

$$\hat{x}_{k+1}^- = A_{k+1}x_k + B_{k+1}u_{k+1} \quad (8)$$

The \hat{x}_{k+1}^- is used as the future eye position

coordinate for predicting eye movement trajectories.

Predict the error covariance matrix ahead:

$$P_{k+1}^- = A_{k+1} P_k A_{k+1}^T + Q_{k+1} \quad (9)$$

The predict phase uses the previous state estimate to predict the estimate of the next system's state.

Update:

Compute the Kalman gain:

$$K_{k+1} = P_{k+1}^- H_{k+1}^T (H_{k+1} P_{k+1}^- H_{k+1}^T + R_{k+1})^{-1} \quad (10)$$

Update the estimate of the state vector with a measurement z_{k+1} :

$$\hat{x}_{k+1} = \hat{x}_{k+1}^- + K_{k+1} (z_{k+1} - H_{k+1} \hat{x}_{k+1}^-) \quad (11)$$

Update the error covariance matrix:

$$P_{k+1} = (I - K_{k+1} H_{k+1}) P_{k+1}^- \quad (12)$$

It should be pointed out that the Kalman Filter maintains the first two moments of the state distribution, $E[x_k] = \hat{x}_k$, $E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T] = P_k$ and

$$p(x_k | z_k) \sim N(E[x_k], E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T]) = N(\hat{x}_k, P_k).$$

The choice of the Kalman gain K_k minimizes the error covariance matrix P_k . The Kalman Filter framework assumes that x_k , z_k are normally

distributed and $E[v_k v_i^T] = \begin{cases} R_k & i = k \\ 0 & i \neq k \end{cases}$, $E[w_k, w_i^T] =$

$$\begin{cases} Q_k & i = k \\ 0 & i \neq k \end{cases}, \quad E[w_k e_i^T] = 0 \forall i, k.$$

3.2 Oculomotor Plant Kalman Filter

Construction

The *Oculomotor Plant Kalman Filter* is defined as the *Oculomotor Plant Mechanical Model*, presented in Section 2.2, transformed into the Kalman filter form.

The specific challenge of this transformation lies in defining the linear stochastic difference equation governing the transition mechanics of the system from one state to another and defining the noise parameters for both system and measurements. Specifically it is necessary to define a state vector x_k , control vector u_k , transition matrix A_k , control matrix B_k . It is also necessary to derive a covariance matrix Q_k for the system's noise w_k and covariance matrix R_k defining the measurement noise v_k . Additionally, to map the actual system's state vector x_k to the measurement

vector z_k , observation matrix H_k is required. The following subsection defines all these parameters.

3.2.1 State Vector

$$x_k = [x_1(k) \quad x_2(k) \quad x_3(k) \quad x_4(k) \quad x_5(k) \quad x_6(k)]^T$$

$$x_1(k) = \Delta\theta - \text{eye rotation}, \quad x_2(k) = \Delta\theta_{LT_LR} -$$

length adjustment of the length tension component of the lateral rectus as a result of $\Delta\theta$ rotation,

$$x_3(k) = \Delta\theta_{LT_MR} - \text{length adjustment of the length$$

tension component for the medial rectus as a result of $\Delta\theta$ rotation, $x_4(k) = \Delta\dot{\theta} - \text{eye velocity},$

$$x_5(k) = F_{LR} - \text{the lateral rectus active state tension,}$$

and $x_6(k) = F_{MR} - \text{the medial rectus active state tension.}$

3.2.2 Transition matrix, control matrix,

control vector: $A_k, B_k, u_k.$

Once a saccade is detected and the amplitude and the direction of the saccade are determined (it is done by a mechanism described in [7]), differential equations defining the OPMM are used to create the transition matrix A_{k+1} , control matrix B_{k+1} and control vector u_{k+1} . This calculation is done using the approximate definition of derivative as

$$\dot{x}(k) = \frac{x^{(k+1)} + x(k)}{\Delta\rho} \quad \text{where } \Delta\rho \text{ is the OPMM internal}$$

sampling clock represented by the time interval between $x(k+1)$ and $x(k)$.

Positive amplitude saccades (rightward movement for the right eye):

The transition matrix A_k is calculated as:

$$A_k = \begin{pmatrix} 1 & 0 & 0 & \Delta\rho & 0 & 0 \\ K_{SE} C_{AG} & D_{AG} & 0 & 0 & C_{AG} & 0 \\ K_{SE} C_{ANT} & 0 & D_{ANT} & 0 & 0 & -C_{ANT} \\ C_J & \Delta\rho \frac{K_{SE}}{J} & \Delta\rho \frac{K_{SE}}{J} & 1 - \Delta\rho \frac{B_p}{J} & 0 & 0 \\ 0 & 0 & 0 & 0 & \left(1 - \frac{\Delta\rho}{\tau_{AG}}\right) & 0 \\ 0 & 0 & 0 & 0 & 0 & \left(1 - \frac{\Delta\rho}{\tau_{ANT}}\right) \end{pmatrix}$$

where $C_{AG} = \Delta\rho \frac{K_{SE}}{(K_{LT} + K_{SE})B_{AG}}$, $C_{ANT} = \Delta\rho \frac{K_{SE}}{(K_{LT} + K_{SE})B_{ANT}}$,

$$C_J = -\Delta\rho \frac{2K_{SE} + K_p}{J}, \quad D_{AG} = \left(1 - \Delta\rho \frac{K_{SE}}{B_{AG}}\right), \quad D_{ANT} = \left(1 - \Delta\rho \frac{K_{SE}}{B_{ANT}}\right).$$

with control vector:

$$u_k = \left[0 \quad 0 \quad 0 \quad 0 \quad \frac{\Delta\rho}{\tau_{AG}} N_{AG_sac}(k) \quad \frac{\Delta\rho}{\tau_{ANT}} N_{ANT_sac}(k) \right]^T$$

where τ_{AG} and τ_{ANT} are activation and deactivation time constants for the agonist and the

antagonist muscles respectively defining the transition between brainstem control signal and active state tension.

The control matrix B_k is a 6x6 identity matrix. For detailed calculations of all the components described above please look in [9].

Negative amplitude saccades (leftward movement for the right eye):

The transition matrix A_k , and control vector u_k are as follows.

$$A_k = \begin{pmatrix} 1 & 0 & 0 & \Delta\rho & 0 & 0 \\ K_{SE}C_{ANT} & D_{ANT} & 0 & 0 & C_{ANT} & 0 \\ K_{SE}C_{AG} & 0 & D_{AG} & 0 & 0 & -C_{AG} \\ C_J & \Delta\rho \frac{K_{SE}}{J} & \Delta\rho \frac{K_{SE}}{J} & 1 - \Delta\rho \frac{B_p}{J} & 0 & 0 \\ 0 & 0 & 0 & 0 & \left(1 - \frac{\Delta\rho}{\tau_{ANT}}\right) & 0 \\ 0 & 0 & 0 & 0 & 0 & \left(1 - \frac{\Delta\rho}{\tau_{AG}}\right) \end{pmatrix}$$

$$u_k = \left[0 \quad 0 \quad 0 \quad 0 \quad \frac{\Delta\rho}{\tau_{ANT}} N_{ANT_sac}(k) \quad \frac{\Delta\rho}{\tau_{AG}} N_{AG_sac}(k) \right]^T$$

The control matrix B_k is a 6x6 identity matrix.

3.2.3 Measurement vector, observation matrix: z_k, H_k .

The eye position measurement device is an eye tracker. An eye tracker reports horizontal and vertical eye position coordinates with a time stamp. In this paper, only the horizontal component of the recorded eye movements is considered, thus measurement vector z_k is a scalar that represents horizontal eye coordinates recorded by the eye tracker at the time k .

The angular eye position is the only variable that is observed, thus the observation matrix is $H_k = [1 \ 0 \ 0 \ 0 \ 0 \ 0]$

3.2.4 Measurement noise covariance matrix, system's noise covariance matrix: R_k, Q_k .

By definition, the covariance matrix for the measurement noise is $R_k = E[(v_k - E(v_k))(v_k - E(v_k))^T]$. Because only eye position is measured, v_k is a scalar making $R_k = VAR[v_k] = \delta_v^2$, where δ_v is the standard deviation of the measurement noise. In this paper, it is assumed that the standard deviation of the measurement noise relates to the accuracy of the eye tracker and is bounded by one degree of the visual angle. Therefore, δ_v was

conservatively set to 1° . In cases when the eye tracker fails to detect eye position coordinates, the standard deviation of measurement noise is assigned to be

$$\delta_v = 120^\circ \tag{13}$$

The value of 120° is chosen empirically, allowing the Kalman Filter to rely more on the predicted eye position coordinate \hat{x}_k^- .

By definition, the system's noise covariance matrix is $Q_k = E[(w_k - E(w_k))(w_k - E(w_k))^T]$, where w_k is a 1x6 system's noise vector $w_k = [w_1(k) \ w_2(k) \ w_3(k) \ w_4(k) \ w_5(k) \ w_6(k)]^T$. In this paper, it is assumed that variables $w_i(k)$ are uncorrelated between each other, i.e., $E[(w_m(k)w_n(k))] = E[(w_m(k))E[w_n(k)]]$ for all $n \neq m$. This assumption generates the following system's

noise covariance matrix: $Q_k = \begin{bmatrix} \delta_1^2 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & \delta_6^2 \end{bmatrix}$. Here

$\delta_1^2, \dots, \delta_6^2$ are variances of variables $w_i(k)$. In this paper, it is assumed that the standard deviation of the eye position noise $w_1(k)$ is connected to the characteristics of the eye fixation movement. This is done with the assumption that eye fixation is the most common type of eye movement. Each eye fixation consists of three basic eye-sub-movements: drift, small involuntary saccades, and tremor. Among those three, involuntary saccades have the highest amplitude - about a half degree of the visual angle; therefore, we conservatively set δ_1 to 1° . Standard deviation values for other variables are hard to assess, but the following values performed well in the simulation tests: $\delta_2 = \delta_3 = 1^\circ$ degree, $\delta_4 = 1^\circ$ /s., $\delta_5 = \delta_6 = 1$ gram.

3.2.5 Initial values for state vector and error covariance matrices: x_0, P_0 .

The state vector is:

$$x_0 = [x_1(0) \ x_2(0) \ x_3(0) \ x_4(0) \ x_5(0) \ x_6(0)]^T$$

$x_1(0)$ - horizontal eye position at the onset of a saccade; $x_2(0) = \frac{x_1(0)K_{SE} + N_{LR_fix}(x_1(0))}{K_{SE} + K_{LT}}$ - initial displacement of the length of the lateral rectus

length tension component;

$$x_3(0) = \frac{x_1(0)K_{SE} - N_{MR_fix}(x_1(0))}{K_{SE} + K_{LT}} - \text{initial displacement of}$$

the length of the medial rectus length tension component; $x_4(0) = 0$ - initial value of the eye velocity; and initial active state tension for the lateral rectus

$$x_5(0) =$$

$$\begin{cases} N_{AG_sac_start}(x_1(0)), & \text{if agonist prior to saccade} \\ N_{ANT_sac_start}(x_1(0)), & \text{if antagonist prior to saccade} \end{cases}$$

and medial rectus

$$x_6(0) =$$

$$\begin{cases} N_{AG_sac_start}(x_1(0)), & \text{if agonist prior to saccade} \\ N_{ANT_sac_start}(x_1(0)), & \text{if antagonist prior to saccade} \end{cases}$$

4 Methodology

4.1 Equipment & Test Media

The experiments were conducted with a Tobii 1750 eye tracker. This eye tracker has the following characteristics: sampling rate - 50Hz, accuracy 0.5° , spatial resolution 0.25° , drift less than 1° . During the experiments, every subject was asked to hold his/her head motionless. Before running each experiment, the eye tracking equipment was calibrated for the subject and checked for calibration accuracy.

Three video clips “Car”, “Shamu”, “Airplanes” each with unique perceptual characteristics were selected to test the performance of the eye movement models. The detailed description and the download link can be found in [8].

4.2 Participants

The subject pool consisted of 21 volunteers of both genders and mixed ethnicities, aged 20-40 with normal, corrected and uncorrected vision. The subjects were instructed to watch the video clips in any way they wanted.

4.3 Detection of Basic Eye Movement Types

One of the objectives of the *Oculomotor Plant Kalman Filter* (OPKF) is to continuously predict future eye gaze in real-time. To attain this goal, every eye gaze position sample has to be classified immediately. By employing a *Velocity-Threshold Identification* (I-VT) model [15], we have grouped all position samples into three categories: fixations,

saccades, and pursuits. This oversimplification of eye movement types allowed us to maintain real-time performance of the OPKF without creating any eye position samples detection buffers. The original I-VT model proposed by Salvucci & Goldberg classified eye position samples as fixation when eye velocity was below $100^\circ/s$, and saccades when eye velocity was above $300^\circ/s$. These values seem to be in contradiction with oculomotor research and neurological literature. Leigh and Zee [6] indicate that saccade onset is detected when eye velocity rises to $30^\circ/s$ and smooth pursuit eye movements are detected with eye velocities of $5-30^\circ/s$. A different group of researchers represented by Meyer *et al.* [16] found out that *Human Visual System* (HVS) maintains pursuit motion with velocities of up to $90^\circ/s$. These facts show that it is hard to define an automated instant eye movement detection criterion, especially if eye position sample data is the only source of information. Nevertheless, to test eye movement prediction accuracy of our models, we have adopted velocity values suggested by Leigh and Zee: fixations ($0-5^\circ/s$), saccades (more than $30^\circ/s$), and pursuits ($5-30^\circ/s$). We are aware of the fact that some of the pursuit eye movements could have been classified as a saccade, but the chance of misclassification was quite low because around 90% of the content speed for “Car”, “Shamu”, and “Airplanes” videos was below $30^\circ/s$ threshold.

4.4 Eye Movement Prediction Models

1) *Two State Kalman Filter* (TSKF) models an eye as a system with position, velocity, and white noise acceleration [7]. The TSKF is capable of predicting eye movement trajectories continuously, but it has low accuracy of prediction during saccades. The TSKF is capable of detecting the beginning of a saccade and estimating its duration and amplitude [7].

2) The *Oculomotor Plant Kalman Filter* (OPKF) defined in Section 3.1 predicts eye movement trajectory for the duration of the saccade after it is detected by the TSKF. The duration of the saccade

is calculated by the equation (3). At the end of the saccade eye movement, prediction is switched back to the TSKF. Although part of the trajectory is predicted by the TSKF, this eye movement prediction model is called *Oculomotor Plant Kalman Filter* (OPKF). Figure 2 presents the diagram.

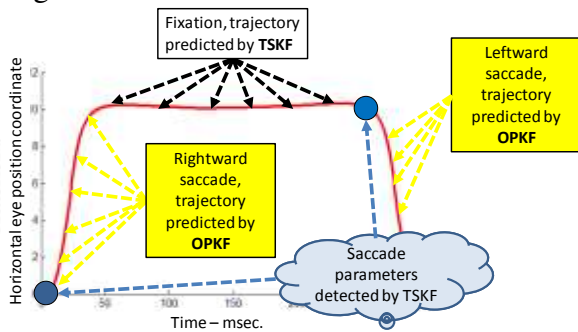


Figure 2. Eye movement prediction by the OPKF.

3) *Oculomotor Plant Kalman Filter Extended* (OPKFE). The OPKFE and the OPKF are almost identical except that the OPKFE continues using *Oculomotor Plant Mechanical Model* for 200 ms. after the end of each saccade. Figure 3 presents the diagram. Figure 4 presents the actual horizontal eye movement trajectory prediction for one of the subjects during viewing of the “Car” video.

There are two reasons for the 200 ms. extension:
 a) The brain needs at least 200 ms. to calculate the parameters of the next saccade after the end of the previous saccade [12]. During these 200 ms., an eye remains in the state of fixation or pursuit. b) If we take the *Oculomotor Plant Mechanical Model* and try to generate a saccade, the eye movement trajectory after the offset of the saccade will closely resemble a fixation. Thus, if the actual eye movement after saccade is a fixation or pursuit, the prediction error between the measured eye position and the eye position predicted by the OPMM will be very small – much smaller if the trajectory was predicted by the TSKF. The experimental results presented in the next section support this logic.

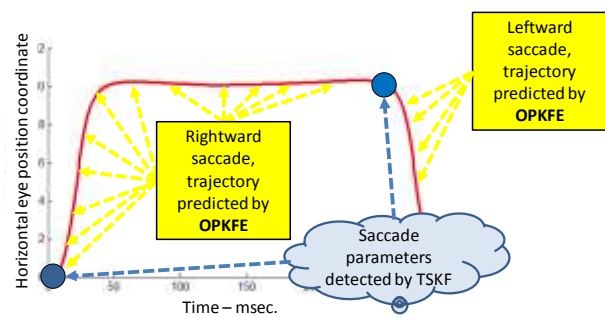


Figure 3. Eye movement prediction by the OPKFE.

4.5 Evaluation & Prediction Accuracy Metric

All models were implemented in MATLAB. From 63 recordings done for 21 subjects, 6 recordings were removed due to excessive noise.

The eye movement prediction was done only for the horizontal movement component of the right eye. Prediction interval was 20 ms., meaning that the future eye gaze coordinate was attained 20 ms. faster in our eye tracker system than in a system without the eye movement prediction.

The *Root Mean Squared Error* (RMSE) between the predicted $\hat{x}_1^-(k)$, Equation (8), and measured z_k , Equation (7), eye position coordinate represents the accuracy of an eye movement prediction during

$$\text{eye movement type } p. \quad RMSE^p = \sqrt{\sum_{k=i}^j \frac{(\hat{x}_1^-(k) - z_k)^2}{j-i}}$$

The ideal eye movement prediction model will have the RMSE of 0°.

The percentage improvement in prediction accuracy (reduction of the RMSE) between two

$$\text{models is calculated as: } \Lambda = 100 \frac{RMSE_{Model_1} - RMSE_{Model_2}}{RMSE_{Model_1}}$$

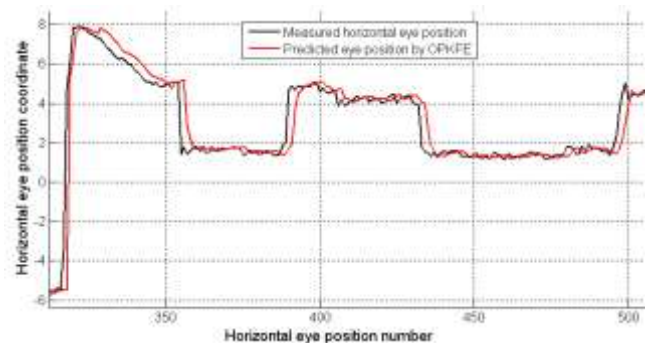


Figure 4. Actual eye movement prediction by the OPKFE.

5 Results

5.1 Eye Movement Data

Figure 5 presents results. The “Not Reported” category consists of eye position samples for which the eye tracker failed to report the eye position coordinates.

Fixations: two subjects’ recordings did not have any fixations detected by the IV-T model. Average fixation duration in all the video clips was approximately 120 ms., deviating ~10 ms. from the mean. **Saccades:** Average saccade amplitude was the 2.6° for the “Car” and “Shamu” videos, 1.8° for the “Airplanes” video.

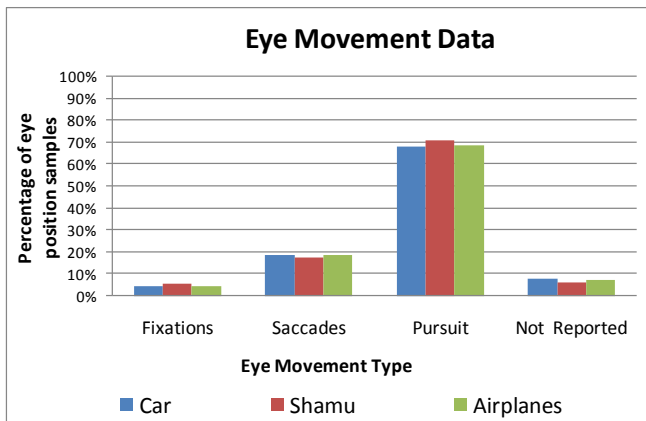


Figure 5. Percentage of eye position samples of each eye movement type detected during a recording of each video.

5.2 Eye Movement Prediction Results

It should be noted that eye movement prediction results reported in our previous work [8], [9] assume that the distance between the subject and Tobii eye tracker is reported in pixel units. In reality, this distance is reported in millimeters. Eye movement prediction results presented in this paper are calculated with the correct distance values. Additionally we have changed velocity threshold values employed for eye movement detection to the numbers specified in Section 4.3. Therefore, the RMSE values were affected by all these changes.

5.2.1 Fixations & Pursuits

Table 1. Prediction data during fixations

Fixations	TSKF - RMSE	OPKF - RMSE	OPKFE - RMSE	Accuracy improvement TSKF vs OPKF	Accuracy improvement OPKF vs OPKFE
"Car"	0.34	0.34	0.34	0%	3%
"Shamu"	0.31	0.31	0.30	0%	4%
"Airplanes"	0.24	0.24	0.17	0%	29%

Table 2. Prediction data during pursuits

Pursuits	TSKF - RMSE	OPKF - RMSE	OPKFE - RMSE	Accuracy improvement TSKF vs OPKF	Accuracy improvement OPKF vs OPKFE
"Car"	0.83	0.83	0.77	0%	7%
"Shamu"	0.76	0.76	0.70	0%	9%
"Airplanes"	0.77	0.77	0.72	0%	6%

The TSKF and the OPKF are essentially the same model during fixations and pursuits, hence identical RMSE values. The OPKFE reduced prediction error by 3-29% during fixations and by 6-9% during saccades. Essentially, this improvement in prediction accuracy was achieved by employing OPMM during fixations/pursuits for 200 ms. if that fixation/pursuit succeeded a saccade.

It should be noted that an eye fixation is often selected as a “click” trigger in eye-gaze-aided computer interfaces. Therefore, the RMSE value during a fixation will indicate the accuracy of such a “click”.

5.2.2 Saccades

Table 3. Prediction data during saccades

Saccades	TSKF - RMSE	OPKF - RMSE	OPKFE - RMSE	Accuracy improvement TSKF vs OPKF	Accuracy improvement OPKF vs OPKFE
"Car"	2.83	2.53	2.53	11%	0%
"Shamu"	2.78	2.53	2.53	9%	0%
"Airplanes"	2.07	1.91	1.91	8%	0%

The OPKF model was substantially more accurate than the TSKF model due to more accurate modeling of the human eye. There is no improvement between the OPKF and the OPKFE models because they are identical for the duration of a saccade.

As we can see from the table above, the “Car” and the “Shamu” videos were the most challenging for the eye movement prediction due to the larger saccade amplitudes. In the case of the “Airplanes”

video, the improvement in the accuracy prediction was the smallest due to the smaller saccade amplitudes exhibited by subjects during this video. Therefore, we can conclude that the accuracy improvement achieved by mechanical modeling of the eye is most effective for saccades of larger amplitudes.

5.2.3 All Eye Movements

The RMSE for “All Eye Movements” category was measured by calculating the RMSE for all eye position samples except “Not Reported”.

Table 4. Prediction data for all eye samples

All Eye Movements	TSKF - RMSE	OPKF - RMSE	OPKFE - RMSE	improvement TSKF vs OPKF	improvement OPKF vs OPKFE
"Car"	1.35	1.26	1.24	7%	2%
"Shamu"	1.23	1.16	1.13	6%	3%
"Airplanes"	1.07	1.02	0.99	4%	2%

The “Car” video had the highest prediction error overall, but prediction accuracy improvement was the largest as well. The “Airplanes” video with smaller saccade amplitudes had the smallest accuracy improvement.

Paired samples T test was conducted to test the level of statistical significance between the RMSE values for the TSKF and the OPKF models for each video test. The differences were significant with the level of significance of 0.05. The same was true for the OPKF and the TSKF models.

5.2.4 Eye Movement Prediction during Eye Tracking Failures

An eye position sample was classified as “Not Reported” when the eye tracker failed to report the proper eye position coordinates. Usually the failure to identify the correct eye position coordinates occurs due to the subject’s blinking, jerky head movements, changes in the content’s lighting, excessive wetting of the eye, and squinting. The amount of “Not Reported” eye position samples varied between 3-33% per experiment. In the cases in which the tracking failed for a brief period of time, all models – TSKF, OPKF, TSKFE were capable of predicting eye movement trajectories by using the measurement noise covariance matrix defined by Equation (13).

Figure 6 presents an example of eye movement

prediction by the OPKFE during eye tracking failures.

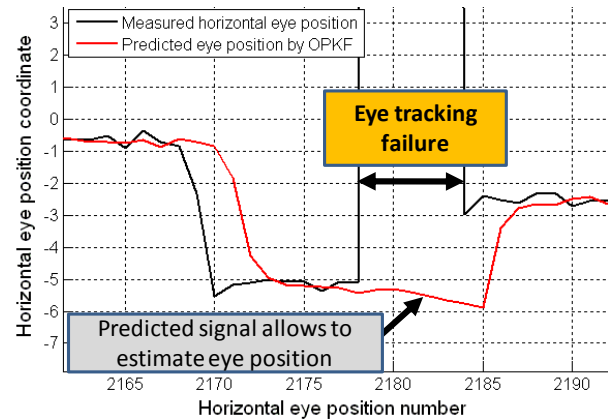


Figure 6. Eye movement prediction during eye tracking failure.

5.2.5 Real-time Performance

The real-time performance comes from the fact that most computationally expensive calculations performed by the OPKF and OPKFE are matrix operations required by the Kalman filter framework. The OPKF and OPKFE are six order systems requiring 6x6 matrixes. Considering the processing capabilities of the modern computers, all necessary types of operations for 6x6 matrixes can be performed in real-time.

6 Conclusion

Eye movement prediction is one of the methods that can compensate for the sensing/transmission delays in the HCI systems with direct eye gaze input. In cases when delay is small, the performance of the system can be increased by pre-selecting visual targets based on the predicted fixation spots. Accurate eye movement prediction can be also beneficial to gaze-contingent compression systems by reducing the size of the high quality coded Region of Interest, therefore lowering the bandwidth or/and computational requirements [5].

In the effort to devise an accurate eye movement prediction algorithm, we have designed a mathematical model of the human eye. This model tries to resemble the neuro-anatomical eye structure by modeling properties of the eyeball suspensory tissues and individual extraocular muscles, *i.e.*, passive elasticity, viscosity,

eye-globe rotational inertia, muscle active state tension, length tension characteristics, and force velocity relationships. The eye model was enhanced by the Kalman filter paradigm where the equations describing the eye movement are transformed into a linear stochastic form and system/measurement noise characteristics are incorporated into this form. Such an approach allows us to achieve a real-time continuous eye movement prediction and to prevent eye tracking failures. When compared to a *Two State Kalman Filter* [7], the model developed in this paper improves the accuracy of prediction by 7-9% overall, considering 20 ms. prediction interval and the pool of 21 subjects. There is also a 2-3% improvement in accuracy of prediction when compared to the previous implementation of the *Oculomotor Plant Kalman Filter* [8].

We hypothesize that the model proposed in this paper will provide the highest accuracy of prediction in a 0-250 ms. prediction interval.

7 References

- [1] K. R. J. Jacob [M]//. Eye tracking in advanced interface design. New York : Oxford University Press, Inc., 1995.
- [2] S. Zhai, C. Morimoto, S. Ihde. Manual and Gaze Input Cascaded (MAGIC) Pointing[C]//. *Proceedings of the 1999 ACM Conference on Human Factors in Computing Systems*: ACM Press, 1999: 246-253.
- [3] D. R. Wilkie. Muscle[M]//. London, 1976.
- [4] D. J. Parkhurst, E. Niebur. Variable resolution displays: A theoretical, practical, and behavioral evaluation[J]. *Human Factors*, 2002, 44: 611-629.
- [5] O. V. Komogortsev, J. I. Khan. Perceptual Multimedia Compression based on the Predictive Kalman Filter Eye Movement Modeling[C]// *Proceedings of the 2007 SPIE Conference on Multimedia Computing and Networking (MMCN 2007)*: SPIE Press, 2007: 1-12.
- [6] J. R. Leigh, D. S. Zee, The Neurology of Eye Movements[M]//. Oxford University Press, 2006.
- [7] O. V. Komogortsev, J. I. Khan. Kalman Filtering in the Design of Eye-Gaze-Guided Computer Interfaces[C]//. *Proceedings of the 2007 12th International Conference on Human-Computer Interaction (HCI 2007)*: 2007: 1-10.
- [8] O. V. Komogortsev, J. I. Khan. Eye Movement Prediction by Kalman Filter with Integrated Linear Horizontal Oculomotor Plant Mechanical Model[C]//. *Proceedings of the 2008 ACM Symposium on Eye Tracking Research & Applications (ETRA 2008)*: ACM Press, 2008: 229-236.
- [9] O. V. Komogortsev, Eye Movement Prediction by Oculomotor Plant Modeling with Kalman Filter[D]//. Kent, OH, USA : Kent State University, 2007.
- [10] A. T. Bahill, Development, validation and sensitivity analyses of human eye movement models[J]. *CRC Critical Reviews in Bioengineering*, 1980, 4: 311-355.
- [11] R. Brown, P. Hwang. Introduction to Random Signals and Applied Kalman Filtering[M]//. New York : John Wiley and Sons, 1997.
- [12] D. L. Sparks. The brainstem control of saccadic eye movements. *Nature Reviews Neuroscience*, 2002, 3 (12): 952-964.
- [13] D. A. Robinson, D. M. Omeara, A. B. Scott, C. C. Collins. Mechanical components of human eye movements. Robinson[J]. *Journal of Applied Physiology*: 1969: 548-553.
- [14] F. Fioravanti, P. Inchingolo, S. Pensiero, M. Spanio. Saccadic eye movement conjugation in children[J]. *Vision Research*, 1995, 35: 3217-3228.
- [15] D. D. Salvucci, J. H. Goldberg. Identifying fixations and saccades in eye tracking protocols[C]//. *Proceedings of the 2000 ACM Symposium on Eye Tracking Research & Applications (ETRA 2000)*: ACM Press, 2000: 71-78.
- [16] C. H. Meyer, A. G. Lasker, D. A. Robinson. The Upper Limit of Human Smooth Pursuit Velocity[J]. *Vision Research*, 1985, 25: 561-563.



Dr. Oleg V. Komogortsev was born in 1979. He received his M.S. and Ph.D. degrees from Kent State University, Kent, OH, USA in 2003 and 2007 respectively. He is

now an Assistant Professor in the Department of Computer Science at Texas State University-San Marcos, TX, USA. His present research interests include human computer interaction, bio-engineering, gaze-contingent multimedia compression. E-mail: ok11@txstate.edu.



Dr. Javed I. Khan has received his Ph.D. from the University of Hawaii and B.Sc. from Bangladesh University of Engineering & Technology (BUET). He is now a Professor in the Department of Computer Science at Kent State University, Kent, OH, USA. His research interest includes extreme networking, complex systems, and perceptual engineering. E-mail: javed@kent.edu.