ABSTRACT

WANG, XUEZHONG. Application of Deterministic and Stochastic Components of the Ocular Dynamic System. (Under the direction of Dr. Simon M. Hsiang and Dr. David B. Kaber.)

With the advent of eye tracking technology, gaze control is perceived to be an alternative way of human-machine interaction. Owing to the nature of eye movement, visually localizing a target is faster than by hand control and less likely to cause fatigue. Because of the connection between eye movement and attention, it is plausible to infer what targets may be of interest to people based on their gaze pattern. However, current design paradigms focus on fitting eye movements to user interfaces initially designed for hand movement. For example, many studies attempted to utilize gaze point as a mouse cursor, which resulted in limited success. To date, gaze-based interaction has been prone to mistakes and it is time consuming compared to other conventional non-keyboard input methods. Therefore, commercial eye tracking systems are almost exclusively used by people with severe disabilities or for military purposes.

Although eye movement and hand movement share biomechanical similarities, there are fundamental differences between the two which pose difficulties encountered in gaze-based interaction. First, eye movement is involved in both active interactions and passive information perception. Looking at a menu item indicates the user “probably”, but not necessarily, intends to click on it. Hand movement on the other hand, serves solely as an interaction proxy without ambiguity. Secondary, from the view of control mechanisms, eye movement is governed by open loop control whereas hand movement is closed loop control with visual feedback. All the models developed for goal-directed hand movement assume a feedback control mechanism, which does not apply to eye movement. These models play a key role in understanding the underlying control mechanism and the speed accuracy tradeoff of hand movement. Specifically, motion time is a function of target distance and target size (i.e., Fitts’ Law). Due to the fact that closed loop control is not valid for eye movement,

previous models for hand movement are no longer appropriate for explaining dynamic features of the ocular plant.

By addressing these two differences between eye movement and hand movement, this dissertation aimed to improve the usability of eye tracking interaction and extend the existing goal-directed model for hand movement to three-dimensional eye movement. A prototype interface called the "Hot-Zone" was implemented for gaze-based interaction and assessment. Previous studies focused on utilizing gaze control for one-stage control such as 'click' on an existing interface. The Hot-Zone acts like a context menu designed for gaze, which provides a solution to two-stage control, such as "Call out first, then select". An experiment was conducted to show the performance of this technique compared with a conventional pointing device.

Because eye movement is closely related to information processing, it is possible to infer the task context (e.g., browse a picture on the web or read an article) by analyzing eye movement data. A system with predictive intelligence to evaluate a user's attention based on passive eye movement could be beneficial to support interaction quality and improve productivity. Hidden Markov Models and Support Vector Machine were used to differentiate reading and visual searching gaze patterns based on eye movement.

To explore the control mechanism of eye movement, deterministic and stochastic components of the ocular dynamic system were explicitly constructed based on mechanical principles and neuronal and physiological evidence established so far. The kinematics were verified by simulation and compared with experimental data for validation. This simulation model extends the eye movement from one-dimensional to three-dimensional space. Supplied with a velocity profile as input, the model applies stochastic components during the movement and predicts the final state of gaze. In order to generate the velocity profile with reasonable credibility, an optimal control model is developed. This optimal control model reveals a plausible criterion, which governs the eye movement. The criterion explicitly shows the tradeoff between the competing considerations (e.g., time optimal and minimum effort) in

motion planning. Being a deterministic model by its nature, the present model resembles the minimum variance model numerically from the view of optimization. The latter was considered to be one of the few models capable of accounting for stochastic components of eye movement, as it addresses signal-dependent noise.

Application of Deterministic and Stochastic Components of the Ocular Dynamic System

by
Xuezhong Wang

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Industrial and Systems Engineering

Raleigh, North Carolina

2010

APPROVED BY:

Dr. Simon Hsiang	Dr. David Kaber
Co-chair of Advisory Committee	Co-chair of Advisory Committee
Dr. Peter Bloomfield	Dr. Yuan-Shin Lee

DEDICATION

To my wife Yingjie, who is the initial impetus for me to join the graduate school; and to my parents Zhikun Wang and Yueyun Tang, for their supports along the way.

BIOGRAPHY

Xuezhong Wang was born in Jiangsu province, China. After graduating from South China University of Technology with a dual Bachelor's degree in Civil Engineering and Industrial Engineering, he joined the Department of Industrial and Systems Engineering in North Carolina State University in spring 2005. He worked as a research assistant under the direction of Dr. Simon M. Hsiang. He had his internships as a programmer with Applied Science Laboratories and as a research assistant with Ravin Advanced Imaging Laboratories in 2010. Xuezhong Wang also enjoys playing trumpet, exploring various musical genres ranging from classical to Jazz.

ACKNOWLEDGMENTS

I owe my deepest gratitude to Dr. Simon Hsiang, my advisor and Dr. David Kaber, my co-advisor, for their support and guidance. I am also heartily thankful to my committee members: Dr. Yuan-Shin Lee and Dr. Peter Bloomfield, for their academic counsel. Lastly, I offer my regards and blessings to all my friends, without whom this work would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

J: Moment of inertia for the eye globe

B: Viscous damping constant of the ocular system

K: Effective elasticity constant of the ocular system

f_n Natural frequency of the eye plant

ζ: Damping ratio of the eye plant

ω_m: Peak velocity of saccadic eye movement

ω_0: Optimal saccade velocity constant, 450˚/s

A_0: Optimal saccade amplitude constant, 7.9˚

A: Saccade amplitude

$\{h_1, h_2, h_3\}$: Head-fixed coordinate system

$\{e_1, e_2, e_3\}$: Eye-fixed coordinate system

$\mathbf{v}$: Axis of eye rotation

$(\mathbf{v}, \theta)$: Axis/angle representation of rotation

$q(t)$: Unit quaternion

$\mathbf{H}$: Set of quaternions

$\mathbf{L}$: Listing's plane

D: Distance from the start point to the center of the target

W: Width of the target along the axis of motion

T_1: Duration of the primary submovement

T_2: Duration of the secondary (corrective) submovement

L_i: Mean absolute distance to the center of the target after i^{th} corrective movement

$\mathbf{S}(\bullet)$: Skew symmetric matrix operator

σ_R: Saccade variance in radial direction

σ_ϕ: Saccade variance in tangential direction

$\xi(t)$: Signal dependent noise

$\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$: Hidden Markov Model

$\mathbf{A} = \{a_{ij}\}$: State transition probability distribution matrix

$\mathbf{B} = \{b_j(k)\}$: Observable symbol probability distribution matrix

$\boldsymbol{\pi} = \{\pi_i\}$: Initial state distribution matrix

1 Introduction

1.1 Motivation

The eyes are the windows to the soul. From a practical point of view, since the Stone Age, human history is very much defined by the evolution of hand tools. Nevertheless, the developments of eye tracking techniques for interaction remain in their infancy.

Existing eye tracking systems for human computer interaction are used almost exclusively for military purposes or by users with severe disabilities who do not have alternatives. Current design paradigms for gaze-based interaction treat the gaze end effector as a cursor to imitate manual control. Such practice violates the nature of eye movement, degrades the merits of eye tracking technology, and results in limited success. Such difficulties come from the fact that eye movement has unique characteristics different from hand movement.

Due to the close relationship between visual perception and attention, passive eye movement is ubiquitous, which poses difficulties in determining a user's true intention. This ambiguousness is generally considered as a disadvantage of gaze interaction. As a result, eye movement dynamics are usually treated as noise and filtered out for practical purposes. Nevertheless, passive eye movement may be a rich source of knowledge to help identify gaze patterns such as reading and visual searching.

Eye movement is also different from hand movement in the view of motor control. Indeed, eye movement, an informative component of human observable behavior, still obeys the common neurological and physiological principles of muscle contraction, like limb movement (Meyer, Smith, & Wright, 1982). Kinematic similarities manifest in the

sinusoidal-shape velocity profile and impulse-step neural control signals (Meyer, Kornblum, Abrams, Wright, & Smith, 1988). However, hand movement is closed loop control with visual feedback. Most models developed to account for hand movement assume that the position of the end effector, either a fingertip or a stylus, is known during the movement. This assumption does not apply to eye movement, where the image on the retina is too blurry during the movement. The eye movement is more likely to be preprogrammed before the movement is initiated with no feedback correction during the course.

Eye movements comply with particular neural constraints. Donders' Law states that for each direction of gaze, the eyes assume the same position. Listing's Law indicates that when the eye moves to any position, it rotates about an axis that is perpendicular to the initial and final directions of gaze at the point of their intersection (Cannata & Maggiali, 2008; Martinez-Trujillo, 2005). The models for goal-directed hand movements can neither account for these features, nor explain the underlying control mechanism of eye movement.

This study aimed to improve gaze-based human computer interaction and deepen the understanding of ocular motor control mechanisms by addressing the aforementioned characteristics of eye movement. For gaze-based interaction, a new interface widget was designed to accommodate eye movement. Passive eye movement was utilized to help identify task context. To explore the control mechanism of eye movement, knowledge from both deterministic and stochastic aspects of the ocular dynamic system were assimilated into a comprehensive model. The deterministic component related the kinematic indices, such as position and velocity, to external *stimuli* (e.g., saccade amplitudes). The stochastic component applies variances to the target position acquisition prior to saccade and to the motor control during saccade. Optimal control theory was used to reveal the underlying control mechanism. This approach supports the deterministic model component by providing

realistic dynamic properties as inputs to a simulation model.

1.2 Research scope and deliverable results

1. An interface widget the “Hot-Zone” was developed to facilitate interaction through eye tracking. An experiment was conducted to compare the performance of gaze-based interaction with conventional mouse interaction for a simple selection task with the context menu. This new widget can be integrated with existing software such as Microsoft Office etc. to extend the scope of usability for eye tracking technology.
2. Passive eye movements were studied to extract information about visual task context. Hidden Markov Models and Support Vector Machines were tested for gaze pattern recognition intended to discriminate reading and visual searching.
3. The deterministic and stochastic components of ocular dynamic systems were analyzed based on established neuronal and anatomical evidence. A Simulink model was constructed to represent the mechanical feature of three-dimensional eye rotations. This model was shown to agree with physiological constraints such as Listing’s Law to faithfully represent the eye rotation. Such a model is of practical potential in determining the endpoint variance given the target distance, size and direction. This information can be used to identify design parameters for Hot-Zone widget. An optimal control model was constructed to generate saccade velocity profiles needed for the Simulink model as the inputs. This optimal control model revealed the tradeoff among competing considerations for motor planning. It was shown that this model synthesizes the features of several other optimal control models, including the minimum variance model, which considered signal-dependent noise in its performance measure.

1.3 Organization

This dissertation consists of seven chapters. Chapter 1 addresses the objectives and deliverable results. Chapter 2 provides background on ocular anatomy, gaze-based interaction, pattern recognition algorithms and biomechanical modeling for goal-directed movement. The pilot work (involving all three objectives) is presented in Chapter 3. Preliminary results are presented and discussed, raising questions for further investigation. Chapters 4, 5 and 6 present the formal study of the three objectives with results and discussion included. Chapter 7 draws conclusions and proposes future research.

2 Literature review

The objective of this chapter was to provide background relevant to the model development in terms of eye movement variability and eye tracking based applications. In particular, the following topics are to be discussed (Table 2-1):

Table 2-1 Specific research questions to be answered by this chapter

Question	Related background
What are the components of the ocular system?	Anatomy/Physiology
What is the function of eye movement?	Biomechanics
What are the challenges with existing eye tracking techniques for HCI applications?	Human factor
How can unintentional eye movements be utilized?	Pattern recognition
What approaches are used for pattern recognition in this study?	Pattern recognition
How can the eye movements be represented in the three-dimensional space?	Biomechanics
What are the similarities and differences between eye movements and hand movements?	Neurophysiologic
What is the possible criterion that governs eye movement planning?	Optimal control

2.1 Ocular anatomy and physiology

2.1.1 Globe and retina

Intuitively, eyes act like cameras that form the image of the visual field. Likewise, each eye consists of three elements: a mechanical element (the eye globe, equal to the camera body), an optical element (the lens) and a neuronal/chemical element (the retina, equal to the

film). The eye globe, as the camera body, is a slightly asymmetrical sphere with an approximate sagittal diameter of 24 or 25 mm and a transverse diameter of 24 mm.

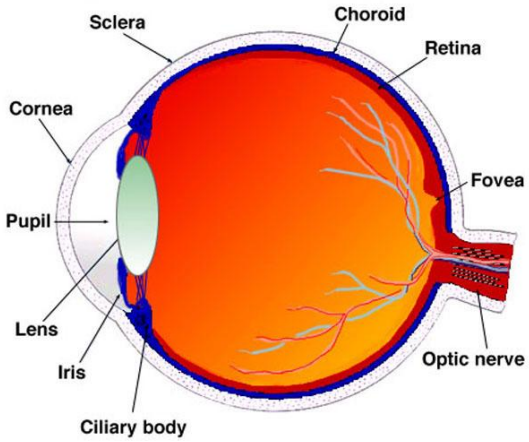


Figure 2-1 Sagittal section of adult human eye (Kolb, Fernandez & Nelson, 2008)

The pupil is an aperture that allows light to enter (Figure 2-1). The dark color comes from the absorbing pigments in the retina. The iris is a circular muscle that controls the size of the pupil so that the amount of light allowed to enter the eye can be adjusted. The cornea is a transparent external surface covering both the pupil and the iris. This powerful lens produces a sharp image at the retina. The sclera is the supporting wall of the eyeball that is continuous with the cornea.

The lens, as the optical component, is a transparent body located behind the iris that is suspended by ligaments (zonular fibers), which are attached to the anterior portion of the ciliary body. Ciliary muscle actions contract or relax these ligaments and change the shape of the lens. This accommodation process is essential to focus an image of an object at distance on the retina (Kolb, et al., 2008).

Of all the parts of the eye, the retina is probably of the most important. It shares the

same function as the film in a camera: to form the final image. Essentially, the retina is a piece of brain tissue that gets direct stimulation from outside light and images. The retina is a circular disc of approximately 22 mm in diameter. The center area of the retina is called the fovea, which is also the central point for image focus or the visual axis that provides the finest detail and directly transfers these detail to the brain for higher order visual perception. The central retina is a circular field of approximately 6 mm in diameter around the fovea while the remainder represents the peripheral retina stretching to the ora serrata (Kolb, 1991; Polyak, 1941; Van Buren, 1963).

The retina has a complex layered organization as shown in Figure 2-2. The innermost layer contains the ganglion cell axons that are the output neurons of the retinal, connected directly to the brain. The outermost layer consists of the pigment epithelium and choroid.

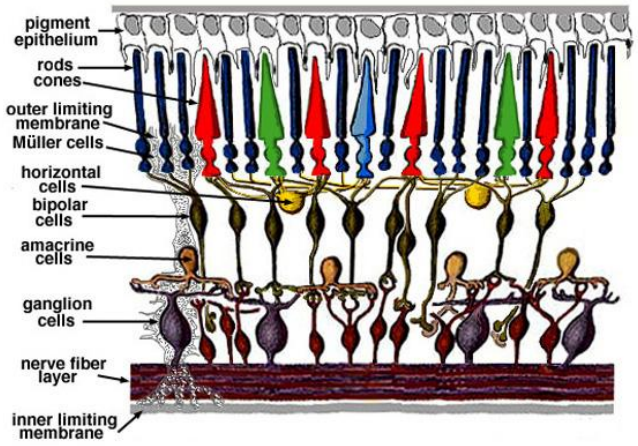


Figure 2-2 Structure of retina (Kolb, et al., 2008)

The light first contacts the outer layer and then travels through the thickness of the retina to reach and activate the photo sensors (the rod cells and the cone cells). The visual pigment of the photoreceptors absorbs the photons and transforms them into a biochemical

message and finally an electrical message that stimulates all neurons of the retina. There is a delicate difference between the retina and film, which could be an important factor in the proposed model for eye movement variability. That is, the film depends on the chemical reaction of the light-sensitive grains to record the pattern of light. Since the grains are sprayed evenly on the film, the developed image has a constant resolution. However, for the retina, anatomical evidence shows that the central retina close to the fovea is much thicker than the peripheral retina (Figure 2-3). The packing densities of photoreceptors, and the associated bipolar and ganglion cells increase in a radial direction toward the fovea. As a result, humans have the sharpest vision only at the central circular region that measures less than a quarter of a millimeter.

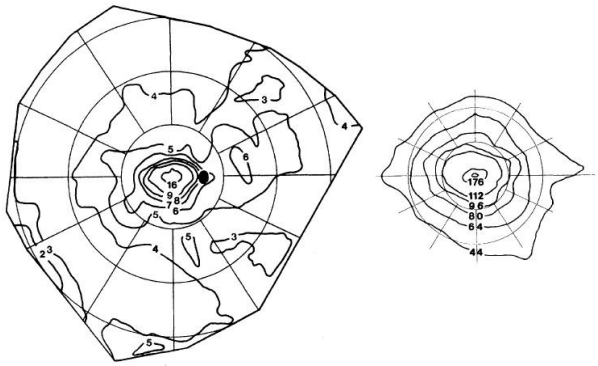


Figure 2-3 Distribution of cones in human retina, reproduced from (Curcio, Sloan, Packer, Hendrickson, & Kalina, 1987)

2.1.2 Extraocular muscles (EOMs)

Each eyeball is held in position in the bony orbit by ligaments, muscles and fascial expansions. To direct the gaze direction to a desired target in three-dimensional space, groups of extraocular muscles attached to the eyeball work together, in a way similar to a three-axis

gimbal set. To describe the eye's orientation, it is necessary to define three orthogonal directions: horizontal, vertical and torsional (Figure 2-4). It is reasonable to approximate the eyeball as rotating around a fixed single point, where three axes intersect. To be specific, several terminologies are commonly used for describing eye rotation in three-dimensional space. Abduction is the horizontal rotation away from the nose, whose complement is adduction, the rotation toward the nose. Elevation and depression are the vertical rotations in opposite directions. Intorsion and extorsion are the rotations of the top of the cornea toward and away from the nose respectively (Table 2-2).

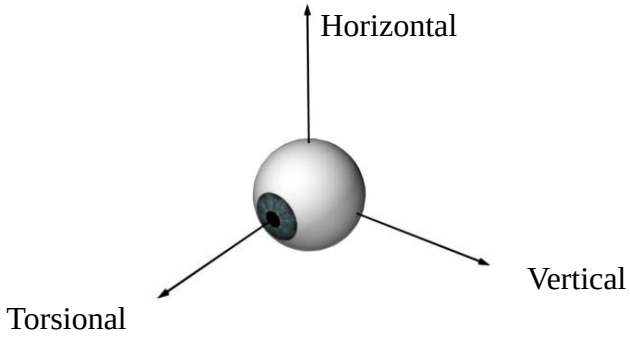


Figure 2-4 Rotation axes of eyeball

Table 2-2 Rotation directions and functions of extraocular muscles (EOMs)

Extraocular muscle	Rotation direction	Primary function
Medial rectus (a)	Horizontal	Adduction
Lateral rectus (b)	Horizontal	Abduction
Inferior rectus (c)	Vertical	Depression
Superior rectus (d)	Vertical	Elevation
Inferior oblique (e)	Torsional	Extorsion
Superior oblique (f)	Torsional	Intorsion

It should be noted that although there are three degrees of freedom, the eye does not assume all possible torsional rotations. Torsional movements are necessary to minimize the

tilt to stabilize perception of horizontal lines, but they only become apparent when they are exaggerated by pathological processes. The eye rotations around the three axes are implemented by three pairs of complementary muscles attached to each eye: superior (Figure 2-5 d) and inferior (c) rectus muscles, medial (a) and lateral (b) rectus muscles, and superior (f) and inferior (e) oblique muscles.

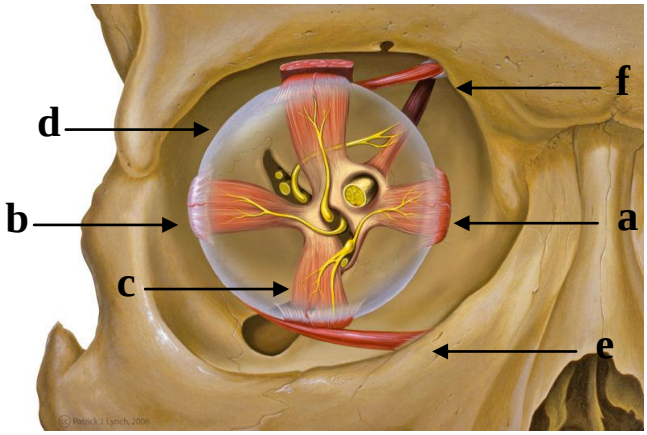


Figure 2-5 Extraocular muscles with motor nerves, anterior (Patrick et al.)

The medial and lateral recti simply produce adduction and abduction, and the actions of the four remaining muscles are complicated because each of these muscles has some torsional component to its action. To reduce complexity, the model is limited under the condition that the inferior and superior recti produce depression and elevation only, and the superior and inferior obliques produce intorsion and extorsion. Experiments indicated that such an arrangement is a reasonable approximation, given that the saccade amplitude is within a certain range, e.g., 15 degrees (Raphan, 1998). It should be pointed out that although EOMs have the same functionality as a gimbal set, there are clear differences between the two regarding rotational kinematics, which will be covered in Section 2.1.4.

2.1.3 Eye movements

Globe, retina and extraocular muscles are the low-level building blocks for the ocular system. When people visually observe the surrounding environment, eye movement, plays an important role in connecting the external stimulus with the low-level components. Based on their functions and neurological differences, eye movements are generally classified into four categories: saccade, pursuit, vergence, and vestibule-ocular (VO) (Robinson, 1968).

Saccades are rapid eye movements with velocities as high as 800˚ per second. They occur frequently when we read or look at a scene, searching for an object. During saccades, sensitivity to visual input is reduced because the eyes are moving so quickly that only a blur image would be perceived on the retina. In other words, we don't receive information during saccades. The peak velocity and duration of the saccade is mostly determined as a function of the distance covered by the eye movements. Pursuit eye movements occur when the eyes follow a slowly moving target. Vergence is a conjugate movement which means the two eyes rotate in the opposite direction to accommodate the change of visual field depth. Vestibule eye movements occur to compensate head movement and keep a target on the fovea (Table 2-3).

Eye movements are highly stereotyped in nature compared to other human body movements (Robinson, 1968). This feature is generally attributed to the fact that the extraocular muscles deal with almost the same mechanical load throughout life, which is different from other skeletal muscles. It is also a well-observed phenomenon that eye movement is usually much swifter than human body movements, such as hand movement.

Table 2-3 Classification of eye movements

Eye movement	Directions	Peak velocity	Function
Saccade	Conjunctive	400- 800˚/s	Switch focus
Pursuit	Conjunctive	Up to 100˚/s	Follow slow motion target
Vergence	Conjugate	20˚/s	Make focus clear
VO	Conjunctive	300˚/s	Compensate head motion
Fixational	Random	60 Hz	Keep firing neuron spikes

There exist three types of small movements: nystagmus, drifts, and micro-saccades, which are also known as fixational eye movements. Due to these micro eye movements, the eye balls are never really still. Nystagmus is the constant tremor of the eyes, whose magnitude is quite small. Nystagmus is generally considered to be related to perceptual activity in that it keeps firing the nerve cells in the retina. Drifts and micro-saccades are often larger movements in magnitude than the nystagmus. Drift is a small and slow movement caused by the less than perfect control of the oculomotor system. Micro-saccades are small rapid movements to compensate for drift (Rayner, 1998). In this study, saccadic eye movements were carefully investigated. Saccades could be utilized for study of real world applications since they reflect visual attention processes. The motor control mechanisms that underlie saccades are also of general interest for both academic and practical purposes.

2.1.4 Noise in eye movements

Like hand movements, eye movement demonstrates a systematic variance pattern that has interested researchers from different fields (Faisal, Selen, & Wolpert, 2008). In fact, variability is a prominent feature (Shadlen & Newsome, 1998), which is embodied in both perception and action, and it permeates every level of the neuronal system. Variability generally arises from two distinct sources. The first source is the deterministic properties of

the system that depend on the initial state and input for each movement. The second source is pure noise that can be characterized by certain probability density functions. In accordance with the information flow of external stimuli, three types of noise - sensory noise, cellular noise and motor noise - affect the system at three levels (Table 2-4).

Table 2-4 Levels of noise in the human body

Noise	Source	Example
Sensory noise	Magnitude of the signal; anatomy and physiology structure	Quality of the image, distance to the focus
Cellular noise	Action potential (AP) variance of neuronal cell	Electrical noise & synaptic noise
Motor noise	Number of firing motoneurons and muscle fibers	Signal dependent noise

Sensory noise is introduced during the perception stage when the external sensory stimulus is converted and amplified into a chemical signal. For visual perception, photons are absorbed by photoreceptors on the retina and converted to electrical signals. Consequently, the photoreceptor density, and its ability to accurately respond to a stimulus, sets a limit for perception. The second category is cellular noise that involves electrical noise and synaptic noise. Electrical noise is related to membrane potential fluctuations in the absence of synaptic inputs (van Rossum, O'Brien, & Smith, 2003). Synaptic noise refers to the trial-to-trial variability in the post-synaptic response (Kleppe & Robinson, 2006).

The last category which is most relevant to this study is motor noise. Direct or indirect control over the muscles with signals from the central nervous system (CNS) is accomplished by motoneurons that essentially act as relays. The amount of force generated depends on the number of muscle fibers innervated by the motoneurons. Henneman’s size principle indicates that, as muscle forces increases, the number of active motoneurons as well as their firing rates

increase, accordingly (Henneman, 1957).

Because of these physiological properties of the motorneurons and muscle fibers, as well as the variance of action potential (AP) timing (Christakos, Papadimitriou, & Erimaki, 2006; Frank, Friedrich, & Beek, 2006), the variability in the generated force is proportional to the average force produced by that muscle (Jones, Hamilton, & Wolpert, 2002). Experimental and theoretical evidence revealed that the proportional variability of force is an inescapable consequence of the organization of the motoneuron and muscle fibers. Noise is an important factor in formulating the principles of the nervous system, which is the key to understanding the control mechanism and modeling realistic dynamic profiles of motor movement.

2.2 Gaze control for interaction

Eyes always fixate at a target before any hand movement is initiated. This indicates that eye tracking can be an alternative with superior potential for human-computer interaction (HCI) (Jacob, 1990). Past studies (Shumin, Carlos, & Steven, 1999) mostly focused on utilizing the gaze point as a mouse cursor to manipulate existing interface widgets. Those studies resulted in limited success attributable to distinct differences between eye and hand movements.

In the early 1990s, in one of the first real-time eye movement-based interaction applications, Jacob studied the utility of the eye tracking technique for some common HCI tasks such as object selection and movement, screen scrolling and menu selection. (Jacob, 1990) In his study, Jacob identified the so-called “Midas Touch” problem. Because normal visual perception involves all types of eye movements, no specific movement can be exclusively used as an indicator for initiating commands. Jacob also emphasized the use of natural eye movements rather than training a user to move his or her eyes in certain ways to

operate the system. In 1999, Zhai et al., developed a Manual And Gaze Input Cascaded (MAGIC) pointing system (Shumin, et al., 1999). By using this system, manual controlled pointing and selection were aided by eye tracking. Although the MAGIC pointing techniques were expected to provide potential advantages, such as reduction of manual stress and fatigue and shorter performance time for large magnitude pointing operations, the benefit gained by using eye-tracking technology was not promising.

Recently, the HCI group at Stanford University developed a gaze-based interaction system {Kumar, 2007 #45} called Gaze-enhanced User Interface Design (GUIDe), which explored the effectiveness of utilizing gaze information as an augmented input besides the keyboard and mouse. The system provided a two-stage object selection sequence that presents a magnified "confident area" in the first stage as a basis for the user to adjust more precisely in the second stage. To fully cover the functions of a mouse, the system designated six hotkeys in the number pad for users to work with. To evaluate the usability of their system, several real world tasks, such as web browsing were tested (Figure 2-6). Qualitative results showed that 75% of the subjects preferred to use EyePoint instead of the mouse for faster and easier control. However, the error rate was much higher than the mouse input.

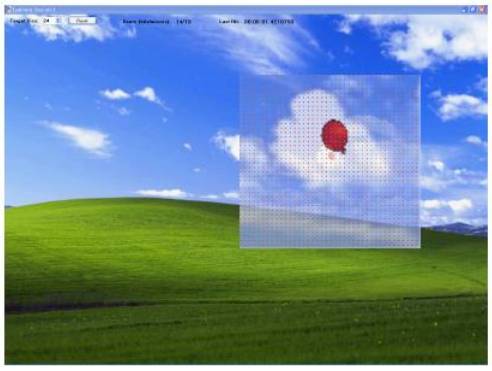


Figure 2-6 EyePoint test application with GUIDe: Point at the balloon

Other techniques have been proposed to improve the usability of eye tracking. Lankford proposed to use zooming to overcome eye tracking accuracy limitations (Chris, 2000). Ashmore and Duchowski et al. present a system using a fish-eye lens to magnify the region where the user is looking (Michael, Andrew, & Garth, 2005). Despite these efforts, eye tracking for interaction has had limited popularity over the past decades. It is used almost exclusively for military purposes or by users with severe disabilities, who have no other choices. Table 2-5 lists some of the requirements for non-keyboard input devices, which are applicable to gaze-based interaction, as well those from ISO 9241. In general, current interfaces developed for mouse operation are not suitable for the eye tracking technique. Widgets, such as buttons and menu items, are simply too small for gaze-based manipulation. New UI widgets must be designed to accommodate the dynamics of eye movements. At present, eye tracking systems are expensive and typically cost around $30,000. This is mostly due to the limited user population. In fact, working video-based eye tracking technology only requires an infrared ray source and a high-speed camera which is relatively inexpensive and can be easily integrated with a LCD monitor and laptop. A laptop can be used for image processing and gaze position calculation, which can be done by a single digital signal processor. There have been studies attempting to build eye tracker systems from low-cost off-the-shelf components, such as an ordinary web camera (Hansen, MacKayu, Hansen, & Nielsen, 2004; Li & Parkhurst, 2006). As eye tracking technology advances, it will become more robust and convenient to use at a lower price. With this in mind, it is plausible that given a larger user population, the manufacturing cost of eye tracking devices may be dramatically reduced.

Table 2-5 ISO 9241, CD Part 9: Non-Keyboard Input Devices requirement (applicable part)

Requirement topic	Mouse	Eye tracking
Fine positioning anchor	√	
Repositioning possible without tools	√	√
Feedback provided	√	√
Resistance to unintended input	√	
Access from work position	√	√
Visual feedback on input	√	

2.3 Gaze pattern recognition

As mentioned in section 2.1.4, the so-called “Midas Touch” problem caused by passive eye movement poses difficulties in determining a user’s true intention. To avoid such a problem, all types of eye movements except fixational movement, except fixational movement, have been disregarded in previous studies. This practice not only ignores the nature of eye movement but also throws away valuable information embedded in unintentional eye movements. Passive eye movement is closely related to concurrent visual task performance. As a vital component of the human perceptual system, the ocular system performs different perceptual tasks throughout life. These tasks range from simple ones as reading and visual search (Duchowski, 2002) to medical image diagnosis (Dempere-Marco et al., 2002) and monitoring of aircraft cockpit information displays, which require a higher level of knowledge. Since the information is presented in different formats, the corresponding eye movements demonstrate different patterns (Figure 2-7).

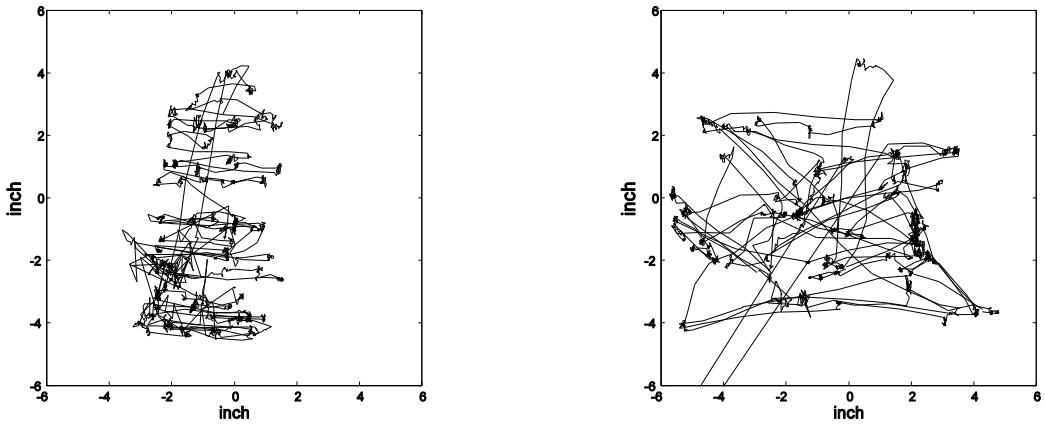


Figure 2-7 Eye movement patterns for reading (left) and visual search (right)

Consider English reading tasks as an example. Eye fixations last about 200-250 ms and mean saccade amplitude is 7–9 letter space; most are made from left to right. Interestingly, it was found that word length has a direct impact on the probability of fixation. For example, 2-3 letter words are fixated about 25% of the time while words of 8 letters or longer are almost always fixated (Rayner & McConkie, 1976). Comparatively, visual search is much less stereotypical. Saccades are made more frequently with larger directional variance, and fixation duration is determined by the nature of currently fixated information (Vaughan, 1982). When searching for words, the word length no longer influences fixation duration (Rayner, Sereno, & Raney, 1996).

Different types of models have been proposed regarding eye movement control during reading and visual search (Kennison & Clifton, 1995; Oregan & Jacobs, 1992) in order to predict when eye movement is initiated and where the eyes will move. However, studies that defined “optimal” or “normative” gaze patterns have focused on theoretical purposes rather than practical applications. No inference regarding current visual tasks can be made from these models. Information on task context has potentially important implications for

designing assistant devices to support user performance in tasks. In this study, Hidden Markov Models (HMMs) and Support Vector Machines (SVMs) were tested to identify reading and search behaviors from eye movement data. As the most popular approaches for pattern recognition, these two methods are briefly introduced in next section.

2.4 Approaches for pattern recognition

2.4.1 Hidden Markov Models (HMMs)

Hidden Markov Models were chosen for this application based on the following methodological requirements: (a) the model must be suitable for real-time implementation, (b) it must have a low computational cost; and (c) it must be robust to noise. Hidden Markov Models, as a statistical classification method, are capable of addressing all the three conditions. In fact, HMMs are an extensively applied tool, which dominant human behavior recognition applications such as speech recognition (both isolated word and continuous) and handwriting recognition (both on-line and off-line). Given a well-trained model, the computational cost is low. In addition, problems such as boundary detection, which commonly occur in other methods (i.e., dynamic time warping) are eliminated.

To use HMMs for pattern recognition (Figure 2-8), a two-phase procedure is generally adopted, which involves feature extraction and conditional probability calculation (Rabiner, 1989). During the first phase, temporal analysis is applied to convert eye tracking data into recognition units. In the second phase, the sequence of recognition units are compared with pre-built HMMs obtained by the training procedure. For each trained HMM, the conditional probability of a given observation sequence based on the HMM is calculated. The one with the highest conditional probability is picked as a matched pattern for that observation

sequence. The two phases are interdependent in that the appropriate feature is closely related to the topology and recognition unit of the HMMs.

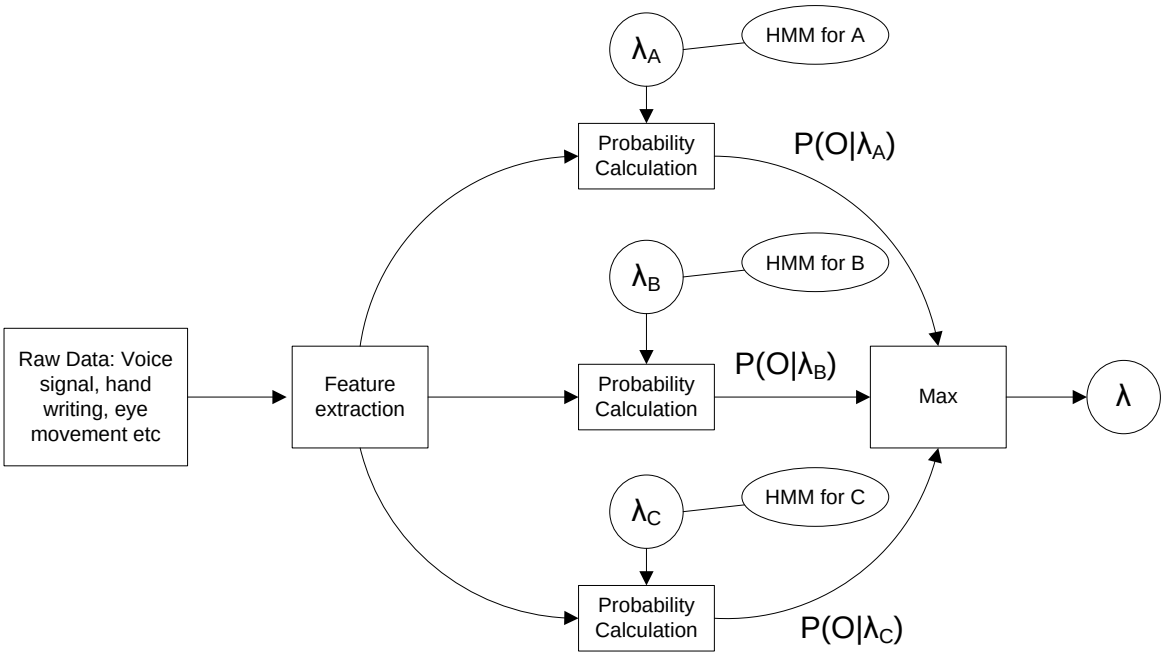


Figure 2-8 Hidden Markov Models for pattern recognition

Definitions and notations

To fully describe a HMM, a compact triple parameter notation is used: $\lambda = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$, Where $\mathbf{A} = \{a_{ij}\}$ is the state transition probability distribution matrix,

$$a_{ij} = P(q_{t+1} = S_j \mid q_t = S_i)$$

$\mathbf{B} = \{b_j(k)\}$ is the observable symbol probability distribution matrix,

$$b_j(k) = P(O_t = V_k \mid q_t = S_j)$$

$\boldsymbol{\pi} = \{\pi_i\}$ is the initial state distribution matrix, in which $\pi_i = P(q_1 = S_i)$ is the probability of

the model starting at state i. To be coherent, the probability matrix needs to satisfy:

$$a_{ij}, b_j(k) \geq 0, \quad \forall i, j, k$$
$$\sum_j a_{ij} = 1 \qquad \forall i,$$
$$\sum_k b_j(k) = 1 \qquad \forall j,$$

The HMM assumes the Markov assumption, which indicates that the state at time $t+1$ only depends on the state at time t. In other words, it's assumed that the current state is conditionally independent of all previous states except for the most recent.

$$P(q_{t+1} = S^{t+1} \mid q_t = S^t, q_{t-1} = S^{t-1}, \cdots, q_1 = S^1) = P(q_{t+1} = S^{t+1} \mid q_t = S^t)$$

Since a HMM is a two stage process, it also assumes the Markov assumption for output symbol generation. Therefore, the probability of a particular symbol generated at time t only depends on the state at that moment, and is independent of the past.

$$P(O_t = V_k \mid q_t = S^t, q_{t-1} = S^{t-1}, \cdots, q_1 = S^1) = P(O_t = V_k \mid q_t = S^t)$$

With the above two assumptions, conditional probability can be easily factorized by the Bayesian equation. Although these assumptions limit the memory of the HMM, they reduce the number of parameters to be determined, and they make it possible for learning and decoding algorithms to be extremely efficient.

HMM topology

The topology of the HMM has a direct effect on both the computational load imposed by the model on resource and model recognition accuracy. Choice of the appropriate topology depends on the nature of the application. Commonly used HMM topologies (Figure 2-9) include the left to right model, the full-connected (ergodic) model, and the parallel path left-right model, etc. Since the number of states and possible output symbols are very limited in the present application, the computation complexity is relatively low. Discrete density was

used for both the state transition matrix and the output matrix. To apply probability calculations, each frame of the extracted feature must be represented by a symbol from a finite alphabet. Vector quantization is usually used during this data compression procedure.

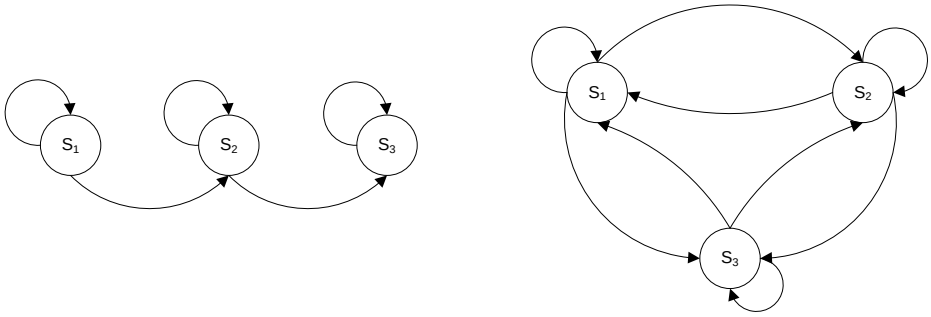


Figure 2-9 Left-right (left) and full connected (right) HMMs

There are also continuous HMMs that use continuous density functions, most commonly multivariate Gaussian density, in which an output probability density function is described by a mean vector and a covariance matrix. This approach provides a distribution that is more accurate and can directly estimate the parameters. However, such advantages are achieved at the cost of a considerably higher computational load. As a result, they are generally inefficient, and inconsistent in performance.

Many researchers have conducted experiments comparing discrete and continuous HMMs. Studies in speech recognition suggest that discrete HMMs perform better because of their efficiency in calculation and capability to represent any distribution. Given the nature of this application, discrete HMMs were used for gaze pattern recognition.

Solutions to the three fundamental problems in HMMs

Three basic problems must be addressed in utilizing HMMs for pattern recognition: a learning problem, a decoding problem, and an evaluation problem (Rabiner, 1989). Learning

problems deal with training the HMMs, and solutions are aimed at determining the parameters in the triple notation to maximize the probability of the observation sequence given the model. This learning problem is usually difficult in terms of obtaining a global optimal solution. The Baum-Welch method (also known as the expectation maximization method) provides a locally optimized solution with an interactive procedure.

$$\begin{aligned}
&\text{Let } \beta_t(i) = P(O_{t+1}O_{t+2}\cdots O_T, q_t = S_i \mid \lambda) \\
&\quad \beta_T(i) = 1 \qquad \beta_t(j) = \sum_{i=1}^{N} a_{ij} b_j(O_{t+1}) \beta_{t+1}(j) \\
&\text{Let } \gamma_t(i) = P(q_t = S_i \mid O, \lambda) \\
&\quad \xi_t(i,j) = P\left(q_t = S_i, q_{t+1} = S_j \middle| O, \lambda\right) \\
&\quad \xi_t(i,j) = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{P(O \mid \lambda)} = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{\sum_i \sum_j \alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}
\end{aligned} \tag{2-1}$$

where $\sum_{t=1}^{T-1} \gamma_t(i)$ is the expected number of transitions from state S_i and $\sum_{t=1}^{T-1} \xi_t(i,j)$ is the expected number of transitions from state S_i to S_j Formally, the Baum-Welch algorithm can be defined as follows:

(i) assume an initial HMM $\lambda_0 = (\mathbf{A}_0, \mathbf{B}_0, \boldsymbol{\pi}_0)$

(ii) re-estimate the parameter matrices with the following equations:

$$\bar{\pi}_i = \gamma_1(i), \quad \bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \quad \bar{b}_j(k) = \frac{\sum_{t=1, O_t = V_k}^{T-1} \gamma_t(i)}{\sum_{t=1}^{T-1} \gamma_t(i)} \tag{2-2}$$

(iii) Check convergence criteria; if satisfied, exit the loop, otherwise, go to step (ii)

To evaluate HMMs, $P(O \mid \lambda)$ must be calculated, given the observation sequence $O = O_1 O_2 \cdots O_T$. The forward algorithm is commonly deployed, which reduces computational

complexity from $2T \times N^T$ to N^2T. The induction can be formulated as follows

$$\text{Let } \alpha_t(i) = P\left(O_1 O_2 \cdots O_t, q_t = S_i \middle| \lambda\right)$$

$$\alpha_1(i) = \pi_i b_i(O_1) \qquad \alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1}) \tag{2-3}$$

$$P(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i)$$

where $\alpha_t(i)$ is the probability of the partial observation sequence $\{O_1, \cdots O_t\}$ $(1 \le t \le T)$ with the hidden state i at time t given the model λ. To determine the most probable internal state sequence, a dynamic programming method called the Viterbi algorithm is employed. Since this method is not used in the present study, it is not discussed here. More detailed information can be found in other papers or books on the HMM topic (Rabiner, 1989).

2.4.2 Support Vector Machines (SVMs)

Introduction

Support vector machine (SVM) is a learning technique originating from statistical learning theory (SLT), which has recently received substantial attention (V.N. Vapnik, 1998, 2000). Vapnik et al. (V. N. Vapnik & Kotz, 2006) showed that SVM generalizes well, which means that method has robust performance on unseen data (given a small training data set), the SVM method is mainly motivated by two ideas: classification with maximal margin and mapping data to higher dimensions. By transforming the original input space into a higher dimension, data from two classes is separated by a hyperplane. The SVM method has been successfully applied to many practical pattern recognition problems, i.e., handwriting recognition, object recognition and text categorization (Cortes & Vapnik, 1995). This section

provides background on applying SVMs to learn from sampling data and solve pattern recognition problems. Consider a binary classification problem with n training samples, given as: $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \ldots, (\mathbf{x}_n, y_n)$, $\mathbf{x}_i \in \mathbb{R}^l$, $y \in \{1, -1\}$. The objective during the learning stage is to find parameters $\mathbf{w} = [w_1\ w_2 \ldots w_l]^T$ and b of a decision function $f(\mathbf{x}, \mathbf{w}, b) = \mathbf{w}^T\mathbf{x} + b$, which labels training samples.

Linear maximal margin classifier for linearly separable data

When the training data set is linearly separable, the hyperplanes obtained perfectly separate training samples into two classes:

$$\begin{aligned} y_i &= \operatorname{sign}\left(f\left(\mathbf{x}_i, \mathbf{w}, b\right)\right) \\ &= \begin{cases} 1 & \mathbf{w}^T\mathbf{x}_i + b \geq 0 \\ -1 & \mathbf{w}^T\mathbf{x}_i + b \leq 0 \end{cases}, \end{aligned} \tag{2-4}$$

Equation (2-4) indicates that the decision plane is not unique. If $f(\mathbf{x}, \mathbf{w}, b)$ is a separation function, any functions of the form $f(\mathbf{x}, k\mathbf{w}, kb), k > 0$ are correct decision functions too. Among the decision planes, the ones with the lowest test error rates are defined as the canonical hyperplanes.

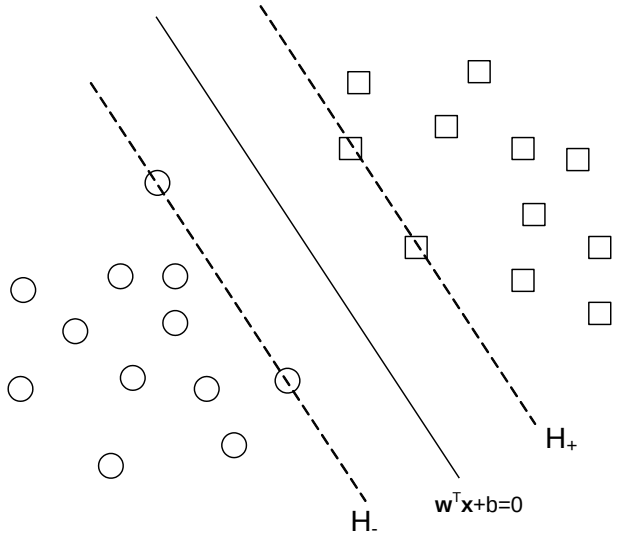


Figure 2-10. Linearly separable data

A canonical hyperplane satisfies $\left|\mathbf{w}^T\mathbf{x}_i + b\right| = 1$ and SVMs search for the optimal canonical hyperplane, which have a maximal margin. Specifically, we denote $H_+ : \mathbf{w}^T\mathbf{x} + b = 1$ and $H_- : \mathbf{w}^T\mathbf{x} + b = -1$ as training samples in the canonical hyperplanes, which are support vectors. Consequently, the margin $M = \frac{2}{\|\mathbf{w}\|}$, which indicates that to find the hyperplane with a maximal margin is equal to minimizing $\|\mathbf{w}\|$. This approach can be formulated as a convex quadratic programming problem:

$$\begin{aligned} \min_{(\mathbf{w},b)} \quad & \tfrac{1}{2}\mathbf{w}^T\mathbf{w} \\ s.t. \quad & y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 \\ & i = 1, 2, \ldots, n \end{aligned} \tag{2-5}$$

This optimization problem is solved by the saddle point of the Lagrange function:

$$L(\mathbf{w}, b, \boldsymbol{\alpha}) = \tfrac{1}{2}\mathbf{w}^T\mathbf{w} - \sum_{i=1}^{n} \alpha_i \left[y_i(\mathbf{w}^T\mathbf{x}_i + b) - 1 \right]$$

The Karush-Kuhn-Tucker (KKT) conditions for this function to be optimum are as follows:

$$\begin{aligned} \frac{\partial L}{\partial \mathbf{w}} &= 0, \quad or \quad \mathbf{w}^* = \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i \\ \frac{\partial L}{\partial b} &= 0, \quad or \quad \sum_{i=1}^{n} \alpha_i y_i = 0 \end{aligned} \tag{2-6}$$

The following complementary conditions must also be satisfied:

$$\begin{aligned} &\alpha_i \left\{ y_i \left[\mathbf{w}^T \mathbf{x} + b \right] - 1 \right\} = 0 \\ &y_i \left[\mathbf{w}^T \mathbf{x} + b \right] - 1 \geq 0 \\ &\alpha_i \geq 0, \quad i = 1, 2, \ldots, n \end{aligned} \tag{2-7}$$

Substituting (2-6) into $L(\mathbf{w}, b, \boldsymbol{\alpha})$, the primal problem is formulated as:

$$\begin{aligned} \min \quad & \tfrac{1}{2} \boldsymbol{\alpha}^T \mathbf{H} \boldsymbol{\alpha} - \mathbf{f}^T \boldsymbol{\alpha} \\ s.t. \quad & \mathbf{y}^T \boldsymbol{\alpha} = 0, \\ & \boldsymbol{\alpha} \geq 0 \end{aligned}$$

Where $\mathbf{H}$ denotes the Hessian matrix of this problem $\mathbf{H}_{ij} = y_i y_j (\mathbf{x}_i^T \mathbf{x}_j)$ and $\mathbf{f}$ is a unit vector $\mathbf{f} = [1 \;\; 1 \;\; \ldots \;\; 1]$. The weighing vector $\mathbf{w}^*$ and bias term b in the primal problem are determined by solving the dual problem for $\boldsymbol{\alpha}$ and substituting it into (2-6). The support vectors are readily seen as samples with $\alpha_i > 0$, which lies on the margin.

Linear soft margin classifier for overlapping classes

For overlapping training data sets, overlapped samples are not classified correctly with a linear hyperplane. The quadratic programming solution introduced in the last section tends to select all training samples as support vectors. To find a proper classifier with maximal margin, we define a soft margin in which all training samples within that margin are neglected.

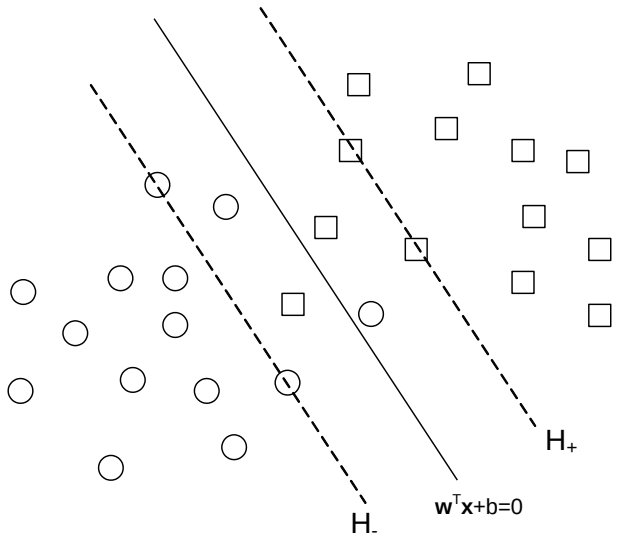


Figure 2-11. Linearly overlapping data

The width of a soft margin is controlled by a penalty parameter C chosen by the user. The constraints are relaxed by introducing slack variables $\xi_i \in \mathbb{R}, \quad i = 1, \ldots, n$. Constraints are rewritten as $y_i \left[\mathbf{w}^T x + \mathbf{b} \right] \geq 1 - \xi_i$. The problem is reformulated with the penalty parameter C:

$$\begin{aligned} \min_{(\mathbf{w},b,\xi)} \quad & \tfrac{1}{2}\mathbf{w}^T\mathbf{w} \\ s.t. \quad & y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i \\ & \xi_i \geq 0, \quad i = 1, 2, \ldots, n \end{aligned} \tag{2-8}$$

The Lagrange function is:

$$L(\mathbf{w}, b, \boldsymbol{\alpha}, \boldsymbol{\beta}) = \tfrac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_{i=1}^{n}\xi_i - \sum_{i=1}^{n}\alpha_i \left[y_i(\mathbf{w}^T\mathbf{x}_i + b) - 1 + \xi_i \right] - \sum_{i=1}^{n}\xi_i\beta_i$$

where α_i and β_i are Lagrange multipliers. The optimal saddle point is found by applying the KKT conditions:

$$
\begin{aligned}
&\frac{\partial L}{\partial \mathbf{w}} = 0, \quad or \quad \mathbf{w}^* = \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i \\
&\frac{\partial L}{\partial b} = 0, \quad or \quad \sum_{i=1}^{n} \alpha_i y_i = 0 \\
&\frac{\partial L}{\partial \boldsymbol{\xi}} = 0, \quad or \quad \alpha_i + \beta_i = C \\
&\alpha_i \left\{ y_i \left[\mathbf{w}^T \mathbf{x}_i + b \right] - 1 + \xi_i \right\} = 0, \quad i = 1, \ldots, n
\end{aligned}
\tag{2-9}
$$

The resulting dual problem is very similar to the separable cases:

$$
\begin{aligned}
&\min \quad \tfrac{1}{2} \boldsymbol{\alpha}^T \mathbf{H} \boldsymbol{\alpha} - \mathbf{f}^T \boldsymbol{\alpha} \\
&s.t. \quad \mathbf{y}^T \boldsymbol{\alpha} = 0, \\
&\qquad 0 \le \alpha_i \le C, \quad i = 1, \ldots, n
\end{aligned}
\tag{2-10}
$$

The final quadratic optimization problem is the same as the separable case, the only difference is that an upper bound C is now applied to each α_i.

Nonlinear classifier

The SVM method can be extended to handle nonlinear classification problems by transforming training samples to a higher dimensional space with a mapping $\phi : \mathbb{R}^l \to \mathbb{R}^m$. Linear SVMs are then applied to separate the classes in the feature space $\mathbb{R}^m$. The dual problem is then formulated as:

$$
\begin{aligned}
&\min \quad \tfrac{1}{2} \boldsymbol{\alpha}^T \mathbf{H} \mathbf{K} \boldsymbol{\alpha} - \mathbf{f}^T \boldsymbol{\alpha} \\
&s.t. \quad \mathbf{y}^T \boldsymbol{\alpha} = 0, \\
&\qquad 0 \le \alpha_i \le C, \quad i = 1, \ldots, n
\end{aligned}
\tag{2-11}
$$

where $\mathbf{K}$ is the kernel function which satisfies $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$. Commonly used kernel functions include the polynomial kernel, Gaussian kernel and Sigmoid kernel functions. The decision function is now in the form:

$$y_j = \text{sign}\left\{\sum_{i=1}^{n} \alpha_i y_i K(x_i, x_j) - b\right\}$$

2.5 Eye rotation model in three dimensional space

2.5.1 Problem description

Eye movement is the direct outcome of extraocular muscle contractions. The forces applied by the EOMs determine the dynamic features of the eye movement such as trajectory and velocity profile. The transition from the force to the final gaze orientation is rather mechanical and complies with the same physical principles as any other rigid body rotation. Information presented in this section is important for modeling the deterministic component of the ocular dynamic system, where the neuronal constraints that specifically apply to the eye movement are also introduced. These rules, along with the physical principles will be developed rigorously in Chapter 3.

2.5.2 Eye movement in one dimension

The analysis of eye movement is relatively standardized, and with less degree of freedom compared to the analysis of limb movement. The relationship between dynamic components, such as amplitude, duration and peak velocity, has been well established by regression and exponential models, as in the following equations (Collewijn, Erkelens, & Steinman, 1988).

$$\omega_m = \omega_0 \times (1 - e^{\frac{A}{A_0}}) \qquad (2\text{-}12)$$

$$d = 2.7A + 23 \qquad (2\text{-}13)$$

where ω_m is the peak velocity (degrees per second), and d is the duration in milliseconds. This relationship is also called the "main sequence" of saccadic eye movement. Both ω_m and d are determined almost exclusively by the amplitude A. ω_0 and A_0 are the constant coefficients determined by the experiment. Although these equations were established for horizontal saccadic eye movements, they are also fair approximations for saccades in oblique directions. Since eye movement, like any other mechanical movement, involves the physical qualities of force and mass, it can be modeled as a differential equation relating displacement to force or torque applied. The first eye movement model (2-14) made connections between eye position θ and the firing rate of ocular motoneurons, which innervate extraocular muscles and generate force (Westheimer, 1954; Young & Stark, 1963). These early models were later confirmed by directly recording the firing rate R of the motoneuron and refined as follows:

$$R = m\ddot{\theta} + r\dot{\theta} + k\theta \tag{2-14}$$

The coefficients k, r and m are the slopes of the linear regression fit to the experimental data (Robinson, 1970; Skavensk.Aa & Robinson, 1973). Although the scale of the three slopes changes with the amplitude, Robinson found that the ratio between them was relatively constant. The existence of such a relationship was observed by neurological experiments (Robinson, 1970). Given that changes in the discharging rate of the motoneuron causes changes in muscle force, Equation (2-14) can be rewritten as

$$T = J\ddot{\theta} + B\dot{\theta} + K\theta \tag{2-15}$$

where T denotes the extraocular muscle forces. J, B and K represent the moment of inertia for the globe, the viscous damping, and the effective elasticity constant associated with tissues, such as tenon's capsule and the optic nerve. (D. A. Robinson, 1975). This model is

compatible with the neurological structure and can be modified to agree with experimental eye movement data by adjusting the ratio of the three parameters (Robinson, Gordon, & Gordon, 1986). For theoretical studies, a second order differential equation has been widely accepted for one-dimensional saccadic eye movement:

$$
\begin{aligned}
&T = \ddot{\theta} + \frac{t_1 + t_2}{t_1 t_2}\dot{\theta} + \frac{1}{t_1 t_2}\theta \\
&J = 1, \quad B = \frac{t_1 + t_2}{t_1 t_2}, \quad K = \frac{1}{t_1 t_2}, \quad \frac{B}{K} = t_1 + t_2 \\
&f_n = \frac{1}{2\pi}\sqrt{\frac{K}{J}} = \frac{1}{2\pi\sqrt{t_1 t_2}}, \quad \zeta = \frac{B}{2\sqrt{KJ}} = \frac{t_1 + t_2}{2\sqrt{t_1 t_2}}
\end{aligned}
\qquad (2\text{-}16)
$$

where f_n is the natural frequency, and ζ is the damping ratio. f_n and ζ are important indices for system stability. The time constants are fixed according to experiments with $t_1 = 224\,ms$ and $t_2 = 13\,ms$, based on prior measurement of human subjects (C. M. Harris & Wolpert, 1998; Robinson, et al., 1986; Tanaka, Krakauer, & Qian, 2006). Equations (2-14) to (2-16) illustrate that to move the eye, the brain must generate three signals proportional to eye position, angular velocity and angular acceleration. The sum of the three components is presented as the input to the motoneurons. Anatomical findings have revealed the sources of the velocity component for saccadic and vestibule-ocular movements, etc; whereas the position component must be determined from velocity commands (D.A. Robinson, 1975). Experimental evidence supports the claim that velocity of all four types of eye movements route through the same type of velocity-to-position integrator (Figure 2-12), which has been referred to as the "final common integrator" (Robinson, 1968; D.A. Robinson, 1975; Waespe & Henn, 1977). Furthermore, well-established observations from neuron-ophthalmology indicate that the location of this common integrator is in the paramedian pontine reticular formation (PPRF), as lesions of PPRF abolish all versions of

eye movements (Leigh & Zee, 1999). This velocity-to-position integrator plays a central role in understanding how the ocular motor system produces kinematically efficient behavior since this transforming process occurs during all kinds of eye movements.

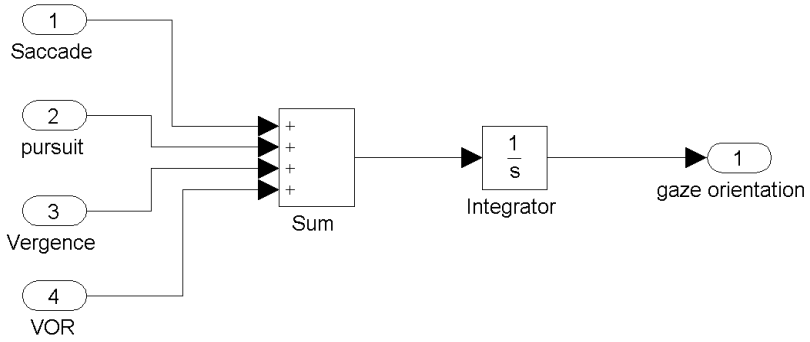


Figure 2-12 The oculomotor integrator in one dimension

2.5.3 Eye movement in three dimensions

When research has attempted to extend the established ocular dynamic model from one dimension to three dimensions, one fundamental principle of rotation is taken into account: rotations of a rigid body are non-commutative (Martinez-Trujillo, 2005; Spong, Hutchinson, & Vidyasagar, 2006; Tweed & Vilis, 1987). Because of the sequence dependent, non-commutative nature of eye rotation, integration of component angular velocity does not necessarily yield angular displacement in three dimensions (Figure 2-13). For one-dimensional eye movement, two rotations do commutate, provided they are on the same axis. This is what happens in the one-dimensional rotation model, since only horizontal rotations are considered. There are several widely accepted tools to represent rigid body rotation and the choice of the most convenient one depends on the nature of the application. A

brief review of these mathematical tools will be covered and one of them will be chosen for the modeling work.

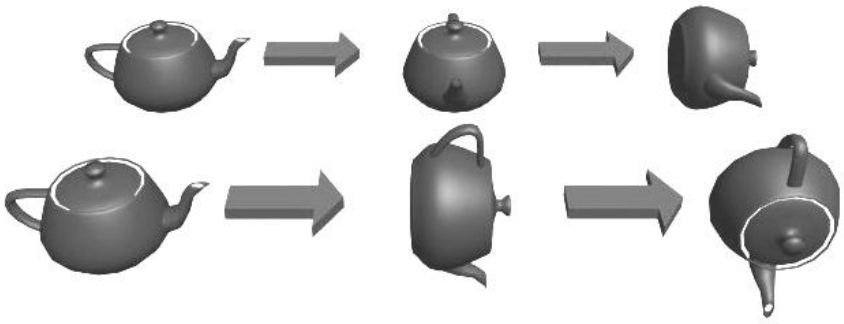


Figure 2-13 Non-commutative rotation in three-dimensional space

To represent the relative position and orientation of the eye with respect to the head, we attach a right-handed coordinate system to each of them (Figure 2-14). Let $\{h_1, h_2, h_3\}$ denote the head-fixed frame such that h_3 coincides with the gaze direction when the eye is in the primary position. Let $\{e_1, e_2, e_3\}$ denote the eye coordinate system such that it coincides with the head coordinate system when the eye is in the primary position.

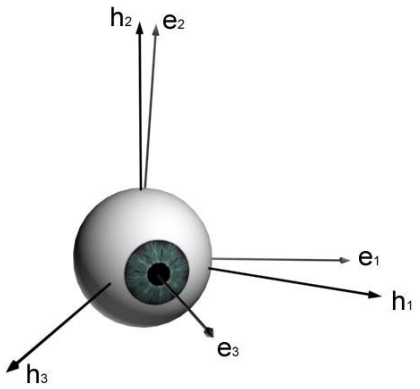


Figure 2-14 Head-fixed and eye-fixed frames

The orientation of the eye-fixed coordinate system in regards to the head-fixed coordinate system could then be described with a rotation matrix R of the form

$$e_i = Rh_i$$
$$R = \begin{bmatrix} e_1 \cdot h_1 & e_2 \cdot h_1 & e_3 \cdot h_1 \\ e_1 \cdot h_2 & e_2 \cdot h_2 & e_3 \cdot h_2 \\ e_1 \cdot h_3 & e_2 \cdot h_3 & e_3 \cdot h_3 \end{bmatrix} \tag{2-17}$$

R belongs to the special orthogonal group of order three (SO(3)), which holds the following properties:

(a) $R^T = R^{-1} \in \mathrm{SO}(3)$

(b) Columns (therefore the rows) of R are mutually orthogonal

(c) Each column (therefore each row) of R is a unit vector

(d) $\det(R) = 1$

These nine elements in a general rotational matrix R are not independent. Rather, it requires only three quantities at most to fully specify the orientation. Accordingly, there are three

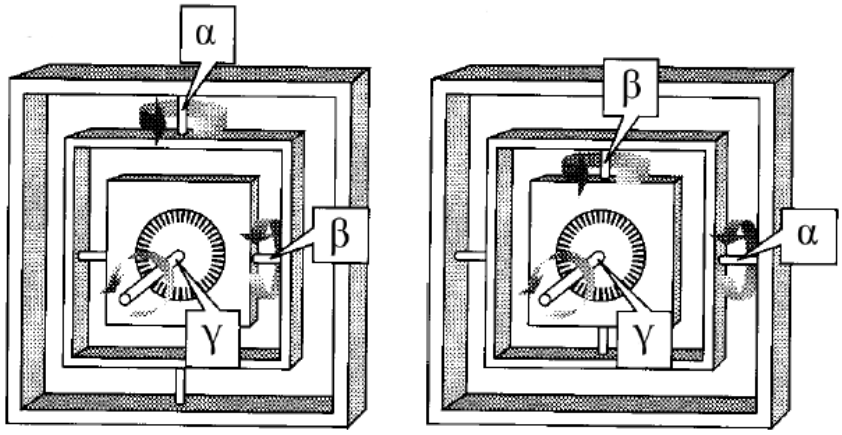
conventional methods to parameterize an arbitrary rotation: the Euler-angle representation, the roll-pitch-yaw representation, and the axis/angle representation.

The Euler-angle representation is extensively used in classical mechanics of rigid bodies and many other applications. The eye orientation in regards to the head can be specified by three angles (ϕ, θ, ψ), which are referred to as Euler angles. The sequence of the rotational composition is critical to obtain the correct result. For example, we first rotate about h_2 by the angle ϕ. Next, rotate about the e_1 axis by the angle θ and finally rotate about the e_3 axis by the angle ψ. The rotational transformation can be described as the composition of three rotation matrices

$$R = R_{h_2,\phi} R_{e_1,\theta} R_{e_3,\psi} \tag{2-18}$$

It should be noted that there is no uniform standard regulating the order of the axes when applying a sequence of rotations. As a result, there are twelve possible conventions regarding the Euler angles in use and proper definitions should always be stated before employing Euler angles (Spong, et al., 2006).

A rotation matrix *R* can also be described by a product of rotation about head coordinate axes in a specific order. These rotations define the roll, pitch, and yaw angles that are also denoted (ϕ, θ, ψ). A good example to demonstrate this representation is a gimbal system (Figure 2-15).



(a) Fick Gimbal (b) Helmholtz Gimbal

Figure 2-15 Two types of Gimbal sets reproduced from Haslanter (1995)

This sequential style rotation conflicts with the physiological properties of eye movement that the three pairs of extraocular muscles apply their torques simultaneously on three nearly fixed axes. Therefore, the composition of the three components of rotation is mutually interdependent. Such an anatomical arrangement favors a symmetric representation of eye orientation, which will be introduced in the following sections.

According to Euler's theorem, an arbitrary eye position can be achieved from the primary position by a single rotation about a fixed axis. Thus, the angle/axis representation is more efficient since it eliminates the ambiguity of the rotation sequences. Let $\mathbf{\upsilon} = [\upsilon_1, \upsilon_2, \upsilon_3]^T$ be a unit vector defining a rotation axis, and a scalar θ is the angle of rotation about $\mathbf{\upsilon}$. The pair $(\mathbf{\upsilon}, \theta)$ is called the axis/angle representation of R. However, there appears to be four parameters needed in this representation of eye position. Notice that although $\mathbf{\upsilon}$ is a unit vector, the degree of freedom for this representation remains three. There have been several different mathematical tools proposed to characterize single rotation from the primary position. Two will be discussed in the following sections: the quaternion model and the

rotation vector model.

2.5.4 Velocity-to-position integrator in three dimensions

Quaternion model

To develop a three-dimensional analog of the velocity-to-position integrator, a quaternion was introduced as the substitute for the scalar counterpart in the one-dimensional model (Tweed & Vilis, 1987; Westheimer, 1957). Quaternion is a four-component representation of angular position discovered by Hamilton in the mid-19th century. Detailed instruction on and discussion of quaternion and their mathematical properties is beyond the scope of this paper. Here essential background material on quaternion is covered, which is used in the proposed model. Let $\mathbf{H}$ denote the set of quaternion. A quaternion q that uniquely specifies a rotation represented by $(\boldsymbol{\upsilon}, \theta)$ is defined as

$$\begin{aligned} q &\equiv \mathrm{q}_0 + (i\mathrm{q}_1 + j\mathrm{q}_2 + k\mathrm{q}_3) = \mathrm{q}_0 + \mathbf{qI}, q \in \mathbf{H} \\ i^2 &= j^2 = k^2 = ijk = -1, ij = k, ji = -k \end{aligned} \tag{2-19}$$

q_0 is referred to as the scalar component and $\mathbf{q}$ represents the vector component. In general, a unit quaternion with length of 1 describes a pure rotation transformation

$$\begin{aligned} \mathrm{q}_0 &= \cos\frac{\theta}{2} \\ |\mathbf{q}| &= \sqrt{\mathrm{q}_1^2 + \mathrm{q}_2^2 + \mathrm{q}_3^2} = \sin\frac{\theta}{2} \\ \sqrt{\mathrm{q}_0^2 + \mathrm{q}_1^2 + \mathrm{q}_2^2 + \mathrm{q}_3^2} &= 1 \end{aligned} \tag{2-20}$$

The derivative of a quaternion is given by

$$\begin{aligned} \dot{q} &= \frac{1}{2}\bar{\boldsymbol{\omega}} q \\ \bar{\boldsymbol{\omega}} &= (0, \boldsymbol{\omega}) \end{aligned} \tag{2-21}$$

The vector part $\boldsymbol{\omega}$ is the angular velocity of the eye in the head-fixed frame. With the position and velocity components represented by quaternion, Tweed et al. proposed a quaternion integrator (Tweed & Vilis, 1987) to replace the one-dimensional model in Equation (2-15). As for the relationship between the quaternion and rotation matrix, a detailed treatment can be found in other papers (Haslwanter, 1995). The general conclusion is that the two tools are equivalent since they are used to represent the same principles.

Vector model

For a unit quaternion, the scalar component is not independent and provides no additional information once the vector component is given. Thus, a rotation vector $\boldsymbol{r}$ corresponding to a quaternion is given by:

$$\boldsymbol{r} = \frac{\mathbf{q}}{q_0} = \tan\frac{\theta}{2}\cdot\boldsymbol{\upsilon} \qquad (2\text{-}22)$$

Based on this relationship, another approach to model the oculomotor control utilizes rotation vectors (Raphan, 1998; Schnabolk & Raphan, 1994a). The main argument for this effort is that the quaternion model treats the orientation of the eye as the output of the integrator while neglecting the fact that the central nervous system (CNS) activates the muscles to generate the torque to rotate the eyeball.

Following this rationale, Raphan et al. proposed a vector integrator model that converts neural signals, representing an angular velocity, to a state variable $\mathbf{x}_p$. This state variable is linearly related to the muscle torque. The output of this integrator drives the motoneurons to generate the torque needed to rotate the eye globe. The eye plant dynamics convert the applied torque vector into eye orientation. The overall system could be summarized by the following equations

$$\begin{aligned} \dot{\mathbf{x}}_p &= \mathbf{H}_p\mathbf{x}_p + \mathbf{G}_p\boldsymbol{r}_\omega \\ \mathbf{m}_n &= \mathbf{C}_p\mathbf{x}_p + \mathbf{D}\boldsymbol{r}_\omega \\ \mathbf{m} &= \mathbf{M}\mathbf{m}_n \end{aligned} \tag{2-23}$$

$$\dot{\boldsymbol{\omega}} = -\mathbf{J}^{-1}(\mathbf{B}\boldsymbol{\omega} + \mathbf{K}\theta\boldsymbol{\upsilon}) + \mathbf{J}^{-1}\mathbf{m} \tag{2-24}$$

$$\dot{\theta} = \boldsymbol{\omega}\cdot\boldsymbol{\upsilon} \tag{2-25}$$

$$\dot{\boldsymbol{\upsilon}} = \frac{1}{2}(\boldsymbol{\omega}\times\boldsymbol{\upsilon}) + \frac{1}{2}\left(\boldsymbol{\upsilon}\times(\boldsymbol{\omega}\times\boldsymbol{\upsilon})\right)\cdot\cot\frac{\theta}{2} \tag{2-26}$$

$\mathbf{H}_p$ is the system matrix whose eigenvalues determine the time constants of the integrator. G_p and C_p are the coupling matrices applied to input and output neural input. $\mathbf{D}$ is the direct premotor-to-motoneuron coupling matrix. This model has the advantage that it resembles the neural and physiological aspects of eye movements (Raphan, 1998). However, the input to the model is the neural signal which is not easily available. Furthermore, it requires as many as eight parameters, most of which were approximated to be consistent with the dominant time constants associated with Equation (2-14). To justify the model and verify the parameters, a simulation was conducted with an assumed input signal and parameters from the work of Schnabolk and Raphan (1994b).

2.5.5 Donders' Law, Listing's Law and half angle rule

Apart from mechanical restrictions, eye movements are also executed under particular neural constraints. Unlike other human body movements, in which trajectory and velocity can be controlled intentionally, rotation of the eye around the line of sight is generally involuntary. Furthermore, the torsional eye position is not arbitrary; it is governed by the gaze direction. This principle was first discovered by Donders in 1846, which is now referred to as Donders' Law. Listing's Law is an extension of Donders' Law that states that with the head fixed, there

exists a specific eye orientation, called the "primary position", from which other physiological orientations can be reached by a single rotation about an axis. During eye movements, the eyeball assumes a unique torsion for each possible eye orientation (Crawford, Martinez-Trujillo, & Klier, 2003). Listing's Law can be used as a touchstone to test the ocular dynamic model. This neural constraint also indicates that only sequential independent rotation representation should be considered in describing eye movement. One important implication of Listing's Law is the half angle rule, which states that though the eye rotation itself has no torsional component, the angular velocity axes should tilt out of Listing's plane by half the angle of the gaze's deviation from the primary position. Detailed explanation for Listing's Law and the half angle rule will be presented later in the Chapter.

2.6 Motor control strategies for goal-directed movements

2.6.1 The recursive approach

There are resemblances between limb movements and saccadic eye movements. From a physiological perspective, any body movements can be viewed as results of contracting opposing groups of agonist and antagonist muscles. For limb movement, Fitts' Law indicates that motion time is a function of target size and target distance. For eye movement, a similar linear relationship (also known as the main sequence) has also been observed, as mentioned in 2.5.2. Several goal-directed models were developed to explain Fitts' Law for limb movement, which cast insight on eye movement. In this section, we review and point out the limitations of these models before we introduce more sophisticated optimal control models.

The first model to address the control strategy behind hand movement was reported by Woodworth (1899) more than a century ago, which is called the "two-component model".

The two-component model suggests that aiming limb movements consist of an initial impulse phase and a current control phase. The initial impulse phase is rapid and stereotyped and the second phase consists of slow motion adjustments guided by visual feedback. Woodworth examined the contribution of vision to the control loop and suggested that the time lag needed for visual feedback to be effective was approximately 450 ms.

In the 1950s, Fitts (Fitts, 1954) quantified a linear relationship, which governs hand movement as follows:

$$MT = A + B\log_2(2D/W) \qquad (2\text{-}27)$$

MT is the mean time to complete the movement. D is the distance from the start point to the center of the target. W is the width of the target along the axis of motion.

Following Woodworth's work (1899), several processing-based models were developed on the two-component model. In 1968, Keele and Posner substantially reduced the time lag estimate to 190~260 ms in an influential study, which addressed a flaw in Woodworth's experiments (Keele & Posner, 1968). These results provide physiological evidence that eye movement is an open-loop control process because most saccades last less than 200 milliseconds, which is not sufficient for visual feedback to be effective.

Other notable examples include the iterative correction model (Crossman & Goodeve, 1983), single-correction model (Beggs & Howarth, 1970, 1972) and the impulse variability model (Meyer, et al., 1988; Meyer, et al., 1982). All these models attempted to derive a form similar to Fitts' Law based on assumptions. For instance, the iterative correction model (Crossman & Goodeve, 1983) makes three assumptions:

1. There is a minimum time t for an initial movement and each corrective movement.
2. The ratio between the distance to the target after the i^{th} and $(i-1)^{th}$ corrective

movements is constant.

3. The initial movement takes less time (expressed by a constant a) than the corrective movements.

Let L_i denote the mean absolute distance from the center of the target after the i^{th} corrective movement. With the preceding three assumptions, we can obtain the following

$$\begin{aligned} &L_0 = D, \quad L_n = W/2, \quad L_i / L_{i-1} = K \\ &L_n = K^n D = W/2 \\ &\text{solve for } n: \\ &K^n = W/(2D) \\ &n = \log_2 K \cdot \left[-\log_2 (2D/W)\right] \end{aligned} \tag{2-28}$$

Then the mean completion time for the movement can be formulated as

$$\begin{aligned} &MT = (n-1)t + (t-1) = nt - a = -a + b\log_2(2D/W) \\ &b = -t/\log_2 K \end{aligned} \tag{2-29}$$

Based on the iterative correction model and the impulse variability model, Meyer et al. (1982) developed the "optimized dual-submovement model". The key idea is that the motor intelligence controls the movement subject to minimizes (or maximizes) certain performance criterion. For example, in manual positioning tasks, subjects are instructed to choose a control solution to minimize total movement time. This minimizes the average time to reach the target. This model predicts that average total movement time can be approximated by the equation:

$$MT = T_1 + \sum_{i=2}^{n} T_i = k_1 + k_2 \left(\frac{D}{W}\right)^{\frac{1}{n}} \tag{2-30}$$

T_1 is the duration of the primary submovement. T_2 is the duration of the secondary (corrective) submovement. n is the maximum number of submovements.

As a metaphor of the feedback control, this model sheds light on the need for a theory of

the learning process that guides the evolution of the optimal performance seen with well-practiced subjects. However, this model, as well as the iterative correction model, suffers from drawbacks which render it like a mathematical trick rather than a sophisticated model.

First, all these models are derived from rather arbitrary assumptions, some of which do not agree with physiological findings. For example, no physiological evidence is found to show how the second assumption required by the iterative correction model is implemented in the ocular system. Empirical results also indicate that participants always undershoot a target in primary movements (Carlton, 1979; Elliott, Binsted, & Heath, 1999); whereas, the model assumes that movement endpoints are centered on the target. Even if these assumptions were valid, it is not clear why the motor system chooses these rules in planning limb movement.

A more serious problem with these models is that they lack physical and mechanical rigorousness to generate dynamic features, such as a velocity profile, which is important to validate their assumptions. These models provide no quantitative information on movement other than motion time. Furthermore, they cannot explain why the velocity profile of skilled or stereotyped motor behavior is a bell-shaped curve with the peak velocity skewed from the middle of the movement (Baloh, Sills, Kumley, & Honrubia, 1975; Meyer, et al., 1982).

Because the muscles are driven by motoneuron signals, similar dynamic profiles of muscle activities reflect the resemblance of the neural signal patterns. Limb and eye movements generally exhibit a “pulse-step” pattern of motoneuron signals. The pulse part initiates movement towards the target and the step portion is to make correction movements and hold the body segment in place (Raphan, 1998). From the view of evolution, it is not likely that the neuron system generates control signals by accident. It is reasonable to assume

that such control signals have certain advantages compared to other arbitrary signals.

To study the control mechanism behind the observed dynamic features, the framework of optimal control is widely used. The following section will provide background on optimal control theory.

2.6.2 The optimal control approach

Within the framework of optimal control theory, two parts are required to formulate the goal-directed movement problem: (1) a dynamic model that captures the physical and mechanical properties of the system; and (2) a performance measure associated with the particular task. Specifically, the dynamic model usually takes the form of a ordinary differential equation system (e.g., Equation (2-15)); and the performance measure is usually an integral of a cost function over a specified time period.

Of these two components, the performance measure is the key to answering the question of why stereotyped behaviors have evolved as they have. Once the performance measure is defined, the dynamic model acts like a constraint in determining the control signal and the state trajectory (i.e., velocity profile), simultaneously. In this section, it is shown how the optimal control approach accommodates the distinct features of eye movement and the proposed optimal models in the eye movement literature are reviewed.

There are three major concerns in modeling the ocular dynamic system. First, during saccadic eye movements, antagonist muscle activity is often minimal or absent. The reason being that the viscous coefficient of the extraocular muscles is believed to be sufficient for passively decelerating the eye rotation (Robinson, 1964). Such physiological evidence suggests that the ocular system is an over-damped system which is strongly resistant to perturbation.

The second concern involves the saturation property associated with the ocular dynamic system. For saccades over a distance greater than 60 degrees, the muscles typically provide maximum possible force regardless of the distance traveled and the peak velocity appeared to be the same (Collewijn, et al., 1988). This feature implies the existence of the maximum muscle force, which can be formulated as a state constraint of the dynamic model.

The third consideration is that eye movement is open-loop control. Prior to eye movement, the fovea is focused on a start point and the target position is acquired by the peripheral retina with certain sensory noise. No further information is available during eye movement. This feature is particularly appropriate for applying optimal control theory. Various optimal control models proposed in eye movement literature are reviewed as follows.

Time-optimal model

As no information is obtained during saccades, shorter movement duration will increase the chance of survival in natural environments. Therefore, time-optimal control is one of the first performance measures considered for eye movements. The problem is generally formulated as:

$$\begin{aligned} J(u) &= \int_0^{t_f} 1 \mathrm{d}t \\ \dot{\boldsymbol{x}} &= \mathbf{f}(\boldsymbol{x}, t, u) \end{aligned} \tag{2-31}$$

Nelson (1983) used a double-integral plant $m\ddot{x}(t) = u(t)$ to model the ocular system. Since this equation represents the motion of inertial loads in frictionless environments, it is not a good approximation for most biomechanical systems where friction and viscosity are indispensable components. Bahill (A. Bahill, Latimer, & Troost, 1980) and Enderle (Enderle, Wolfe, & Yates, 1984) proposed a 6^{th}-order linear time-varying model instead for time optimal control. The agonist and antagonist muscles were modeled separately as a system of an active state generator, viscosity and elastic elements. This model requires as many as 17

parameters and, being a time-varying system, the computational cost for finding the optimal control is high. Simulation results from the model were compared with experimental data to fine-tune the parameters. This model provides insights into the ocular dynamic system with accurate mechanical explanations. However, with so many parameters to be identified, this model may suffer from the over-fitting problem and high computational demands in solving for the optimal control.

$$\dot{\boldsymbol{x}} = \mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{u}$$

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -P_0 & -P_1 & -P_2 & -P_3 & \delta(K_{st} - B_{ant}/\tau_{ag}) & \delta(K_{st} - B_{ag}/\tau_{ant}) \\ 0 & 0 & 0 & 0 & -1/\tau_{ag} & 0 \\ 0 & 0 & 0 & 0 & 0 & -1/\tau_{ant} \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \delta B_{ant}/\tau_{ag} & -\delta B_{ag}/\tau_{ant} \\ /\tau_{ag} & 0 \\ 0 & /\tau_{ant} \end{bmatrix}$$

Time-optimal control or the "bang-bang" control is not considered to be biologically efficient. Given a linear dynamic system, a piece-wise continuous control function will result in a discontinuity at the point at which the control signal jumps from maximum to minimum. However, the eye movement trajectory is usually smooth, which indicates that the control signal should be continuous and time-optimization is not the only part of performance considered by the ocular system.

Minimum torque change (MTC) model

With the observation of the smoothness of human motor movements, a minimum

torque change (MTC) model was proposed (Uno, Kawato, & Suzuki, 1989) to predict smooth muscle torques from the model. The model is formulated as follows:

$$J(t) = \frac{1}{2}\int_0^{t_f} \dot{\tau}^2(t)dt$$
$$\ddot{x}(t) + a_1\dot{x}(t) + a_0 x(t) = \tau(t) \qquad (2\text{-}32)$$

Tanaka et al. (H. Tanaka, M. Tai, & N. Qian, 2004) showed that with a linear plant, the MTC model will always result in a symmetrical velocity profile regardless of the parameter values. This model also accurately describes the properties of human reaching movements in ordinary situations. However, experimental data shows that as saccade amplitude increases, the peak velocity skews towards the end of the movements which cannot be accounted by MTC model.

Minimum variance (MV) model

Harris and Wolpert (C. M. Harris & Wolpert, 1998) introduced signal-dependent noise and proposed that the motion trajectory is shaped in such a way to minimize the variance of the final eye or arm position. A signal-dependent noise term ξ was added to the system dynamic equations:

$$\dot{\boldsymbol{x}}(t) = \mathbf{A}\boldsymbol{x}(t) + \mathbf{B}[\boldsymbol{\tau}(t) + \boldsymbol{\xi}(t)]$$
$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
$$E[\boldsymbol{\xi}(t)] = 0$$
$$E[\boldsymbol{\xi}(t)\boldsymbol{\xi}(t')] = K\boldsymbol{\tau}^2(t)\delta(t-t')$$

The performance measure to be minimized is defined as:

$$J(\boldsymbol{x}) = K\int_0^{t_f} e^{A(t-s)}B(e^{A(t-s)}B)^T\tau^2(s)ds$$
$$\text{s.t. } E[\boldsymbol{x}(t)] - \boldsymbol{x}(t_f) = 0 \qquad (2\text{-}33)$$

In their original simulations, the terminal time t_f (or duration) was specified to obtain the

best fit for a given velocity trajectory. Thus, the model is heavily dependent on the movement duration, which could not be explained by the model. Tanaka et al. further investigated this model (Tanaka, et al., 2006) and demonstrated that the duration can be derived as a function of amplitude. This complements the MV model and represents a unifying theory for goal-directed movements. The MV model also has the advantage that it predicts the skewness of the peak velocity.

Minimum control effort (MCF) model

Recently, Kardamakis et al. proposed that optimal trajectories are generated to minimize the time-integral of the square of the eye control signals (Kardamakis & Moschovakis, 2009). This resembles the MTC model but with its own feature. The coefficient of control signals in the performance measure is now a weighting function that relates extra-ocular muscle tension with eye position.

$$J = \int_0^{t_f} [\alpha(x_e)u_e^2 + \beta u_h^2]\,\mathrm{d}t$$

$$\boldsymbol{x} = \mathbf{A}\boldsymbol{x} + \mathbf{B}\boldsymbol{u}$$

$$\boldsymbol{y} = \mathbf{C}\boldsymbol{x}$$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 \\ a_{21} & 0 & 0 & 0 \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & a_{43} & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 & 0 \\ b_1 & 0 \\ 0 & b_3 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 0 \\ c_2 \\ 0 \\ c_4 \end{bmatrix}^T$$

The state equations include both eye and head movements. The first two equations represent the ocular system. The latter two represent head movements. The two equation sets are, in fact independent in state space. Both are second order plants with two time constants. The weight function for performance is state-dependent in order to reflect the experimental findings in previous studies (Dean, 1996).

This model has an advantage over the MTC model in that it considers the penalty

caused by eye position on control effort, which agrees with biological observations. However, this model is not without limitations. First, a terminal time is specified for deploying numerical methods. In other words, the terminal time is fed to the model as a known parameter, which is not the case in reality. Second, the state variables x_1 and x_2 represent eye velocity and position; however, the coefficient a_{21} does not equal 1, which means the velocity is the derivative of the position multiplied by a constant. This is not intuitively understandable for such a constant.

2.7 Summary

To summarize, this chapter addressed the questions on different aspects of eye movement listed in Table 2-1. These questions are closely related and have substantial implications in the model developed in Chapter 3.

Table 2-6 Summary of questions for literature review

Questions	Approaches
What are the components of the ocular system?	Eye globe, EOMs (2.1.1 and 2.1.2)
What is the function of eye movement?	Saccade, pursuit, vergence, VOR, fixational (2.1.3)
What are the difficulties with existing eye tracking technique for HCI?	Low resolution of gaze position; unintended input (2.2)
How can unintentional eye movements be utilized?	Gaze pattern recognition (2.3)
What approaches are used for pattern recognition in this study?	Hidden Markov models, Support Vector Machines (2.4)
How to represent the eye movement in 3D space?	Vector and quaternion model (2.5)
What are the similarities and differences between eye movement and limb movement?	Similar biomechanical system; close-loop vs. open-loop control(2.6.1)
What is the possible criterion that governs the eye movement planning?	Time optimal, Minimum torque change, Minimum variance (2.6.2)

3 Pilot work

A pilot study was conducted involving three parts with complementary objectives. The relationships are shown in Figure 3-1. The first two components of the pilot study were aimed at enhancing the usability of existing eye tracking technology in HCI applications. The gaze pattern recognition section provided information to assist gaze control interaction in determining the most likely tools needed for a task. The third component, the ocular dynamic modeling was theoretically constructed on the premise of materials presented in Chapter 2. The endpoint variability was broken down into three noise sources and simulations were conducted to reveal the effects of the noise. This model has the potential to provide information regarding the design parameters for proposed gaze control interaction.

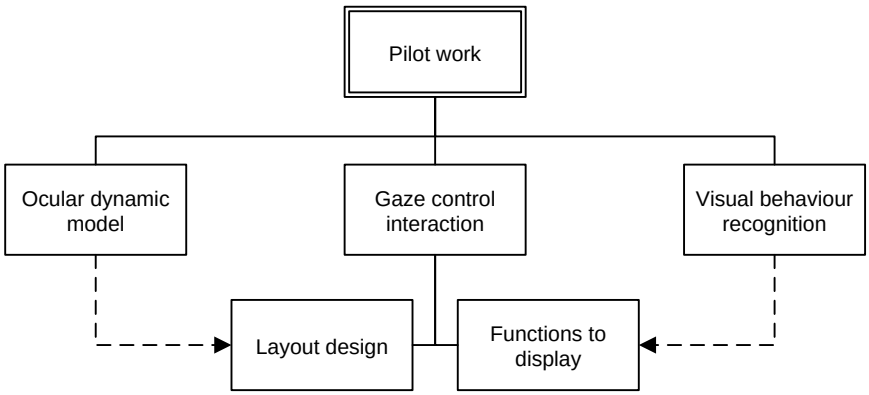


Figure 3-1 Components of pilot work

3.1 Apparatus

In this study, an ASL H6 Head Mounted Optics unit with a head motion tracker was

used for eye movement measurement (Figure 3-2). This is a video-based eye-tracking device dependent upon image processing. Infrared light is imposed on the eyeball and results in a small reflection spot when the light passes through the transparent cornea. One significant feature of this highlighted point is that its location remains the same no matter where the eye looks. On the other hand, the pupil moves along with the eyeball. Therefore, the corneal reflection can be used as an anchor point and the distance of the pupil center from the corneal reflection spot can be used to plot the angle of the eye with respect to the stimuli. To compensate for head motion, a single sensor Flock of Birds motion tracker system was used. With a high-speed camera integrated into the eye tracker headgear, the H6 Head Mounted Optics unit is capable of operating at 120 Hz or a higher sampling frequency. Considering the requirement of this study, the 120 Hz sampling frequency was used.

Figure 3-2 ASL EH 6000 eye tracker

In order to obtain good pupil and corneal reflection recognition, people wearing glasses or contact lenses were excluded from the experiment. For qualified subjects, a 9-point

calibration procedure was performed before the experiment. After that, the subject could move his or her head without affecting the data recording.

3.2 Application of voluntary eye movements – Hot-Zone

3.2.1 Objectives and hypotheses

The application proposed here was intended to enhance the usability of gaze-based interaction for ordinary users by utilizing both active and passive eye movements. Past studies showed that gaze position extracted from fixational eye movement is inadequate for users to interact with real-world applications and eye-based control is usually time consuming and prone to mistakes compared to conventional selective devices (Jacob, 1990; Shumin, et al., 1999). It is proposed here that eye movement dynamics should be used to facilitate human-computer interaction.

3.2.2 Introduction

A major challenge for gaze-based interaction is that existing interface "widgets" such as buttons on a toolbar or items in a menu are too small relative to the resolution of eye trackers. To facilitate a gaze-based control mechanism, a prototype interface called "Hot-Zone" was developed to work with the Windows operating system. Conceptually, the Hot-Zone is very similar to the context menu for the mouse (Figure 3-24). When it is called out, it is centered at the current gaze point for easier access. Several peripheral zones correspond to different context menu items surrounding the central zone. To ensure that users can access the desired zone easily and reliably, design parameters need to be determined, such as the distance from the central to the peripheral zones, the size of the zone and the number of

zones allowed, given certain movement variance tolerance. With the ocular dynamics model proposed in this study, these issues can be tested by simulation without time consuming and expensive experiments.

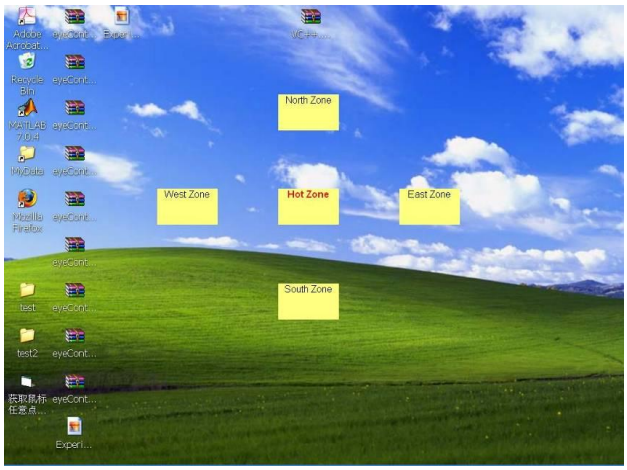


Figure 3-3 Hot-Zone layout

To prevent unintended input, a user presses a hotkey and blinks to call out the Hot-Zone (Figure 3-4). Once the Hot-Zone is displayed, the user is free to look around inside the zone. The zone that the user is currently looking at will be highlighted in red with bold font to serve as feedback. To close the Hot-Zone, the user simply looks outside the Hot-Zone and blinks. Blinks within the Hot-Zone are ignored to prevent unintended command execution. To initiate a command, the user presses the hotkey again and blinks. Researchers have argued that blinking was not as natural as dwell time to be used as a command initiation signal (Jacob, 1990). However, blinking has several advantages over dwell time from several perspectives. First, blinking is a natural behavior and humans blink all the time during the day without feeling fatigue. Fixating at a point requires the extraocular muscles to keep firing to stabilize the eyeball. It consumes much more energy and easily causes fatigue. Second, it is simpler to

identify blinking than fixations; for eye tracking techniques based on pupil and corneal reflection detection, failure to recognize both is a reliable indicator of blinking. The implementation of such a detection mechanism does not require velocity or duration thresholds that are necessary for fixation identification. Finally, blinking from time to time helps reduce eyestrain by moistening the eyeball. On the contrary, dwell time implementation requires users to keep eyes open and stare at the target for a period of time, which can easily cause fatigue.

An underlying ActiveX DLL component (iControl) was developed to support the Hot-Zone for these functions:

1. blink detection
2. gaze direction identification
3. current zone determination
4. a finite state machine to identify gaze state for client program use

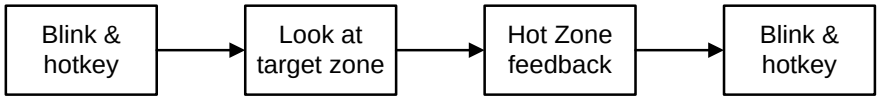


Figure 3-4 Manipulation of Hot-Zone

3.2.3 Methods

To test the usability of the Hot-Zone, a point-and-select task was designed. Performance in gaze control and with the mouse was compared. One male subject with normal vision participated in the pilot experiment for both eye tracking and mouse interaction trials in order to compare performance between the two methods. Two other male subjects participated in the mouse trials to compare performance with the first subject.

On a display screen, there were four labels at each of the four corners. The distance between two neighboring labels was 460 pixels (Figure 3-5). The size of the labels was 120 pixels. At any time, only one label was displayed, which was randomly selected. There were five possible items (North, South, West, East and Center) randomly determined to be displayed on the label, and the subject was required to select the item corresponding to the context menu or Hot-Zone.

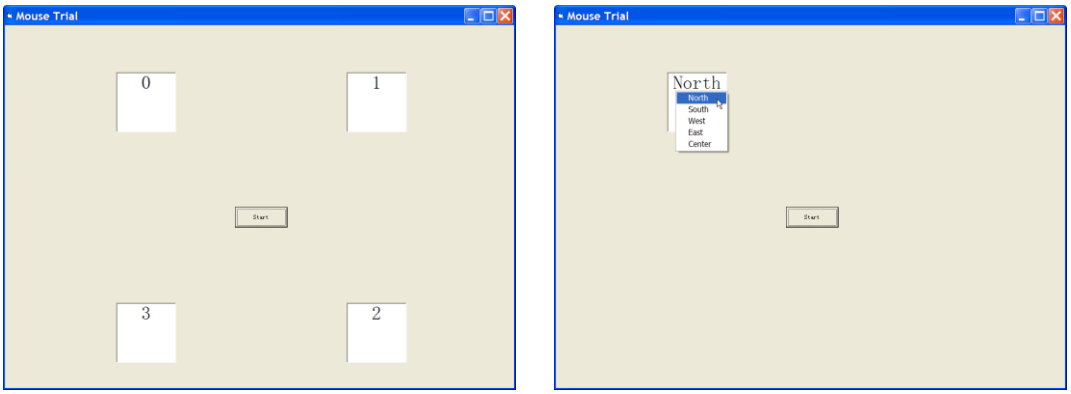


Figure 3-5 Interface for Experiment 2.

When the label was displayed and a specific menu item was selected, the time was marked. The difference between the two time stamps was used as the performance time. There is no correction mechanism for either interaction method and the sequence of displayed items along with the sequence of items that users actually selected was recorded to compute the error rate of the two methods. There were 4 blocks, each consisting of 16 trials for a complete experiment. A rest was provided between two blocks.

3.2.4 Results

Task completion time was the performance index investigated in the pilot study. Given

the small scale of the experiment, the absolute difference between the two interaction methods was assessed instead of the statistical power of the result. This provided an informative and quantitative estimation of the potential for the eye based technique. The researcher, as Subject A, completed two groups of experiments for both interaction methods (Figure 3-6 and Figure 3-7). Subjects B and C only performed mouse trials to serve as a reference for comparison (Figure 3-8). Differences in the mean completion time for the eye tracking and mouse trials for Subject A were within 0.5s. The performances of Subjects B and C in their mouse trials were 1.80s and 1.27s respectively, and shows that Subject A's performance was within normal range (Figure 3-9).

The t-test (Table 3-1 and Table 3-2) showed that task completion time significantly varied with techniques. The-t-test also showed that Subject A's performance was constant in both eye tracking trials ($P = 0.22$) and mouse trials ($P = 0.29$). For the mouse trials, there were significant differences between Subjects A and B ($P < 0.001$) and between Subjects A and C ($P < 0.001$). The difference between subjects was mainly attributed to response agility. The error rate was also considered as a qualitative index for evaluating performance. All three subjects did well in the mouse trials with zero errors. For the eye tracking trials, the error rate was 5% and 2%. Some of the errors were made when part of the "Hot-Zone" was off the screen.

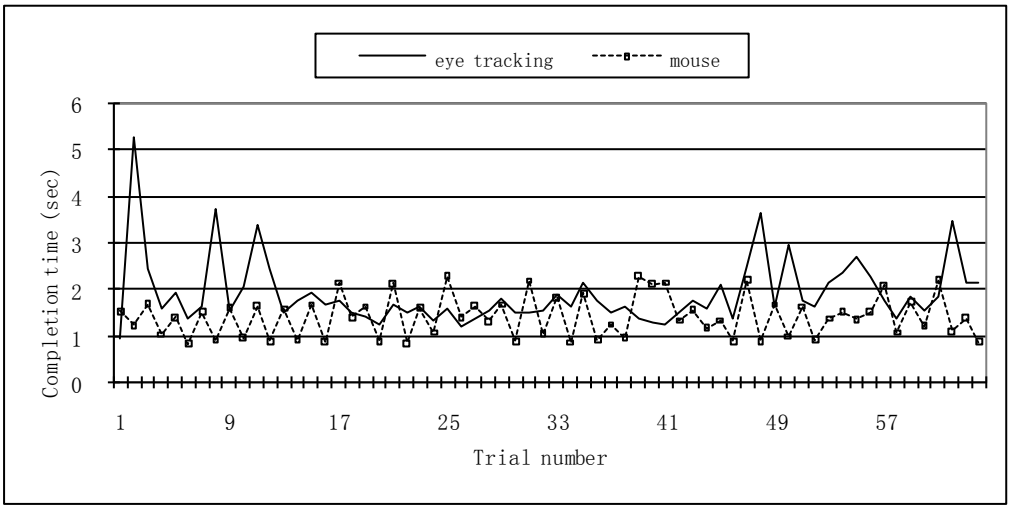


Figure 3-6 Eye tracking and mouse completion time, Test 1

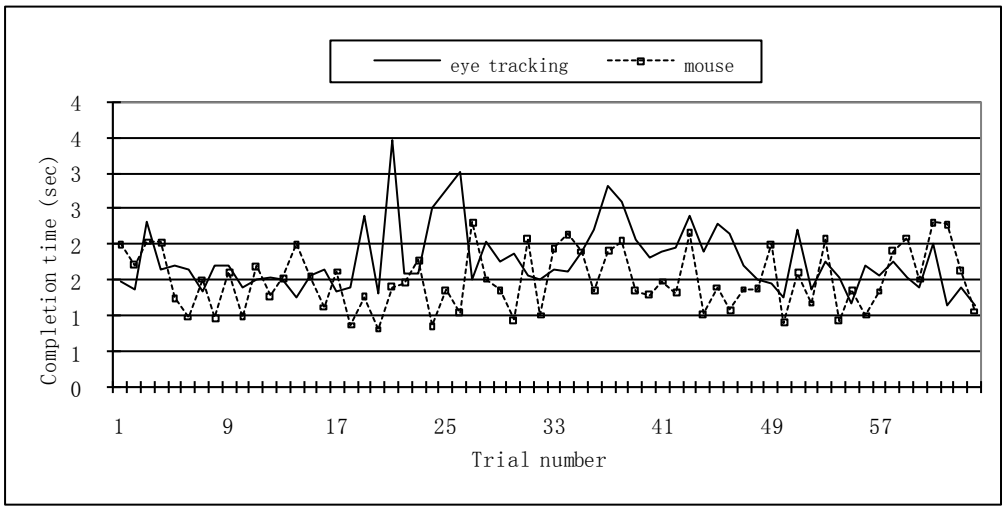


Figure 3-7 Eye tracking and mouse completion time, Test 2

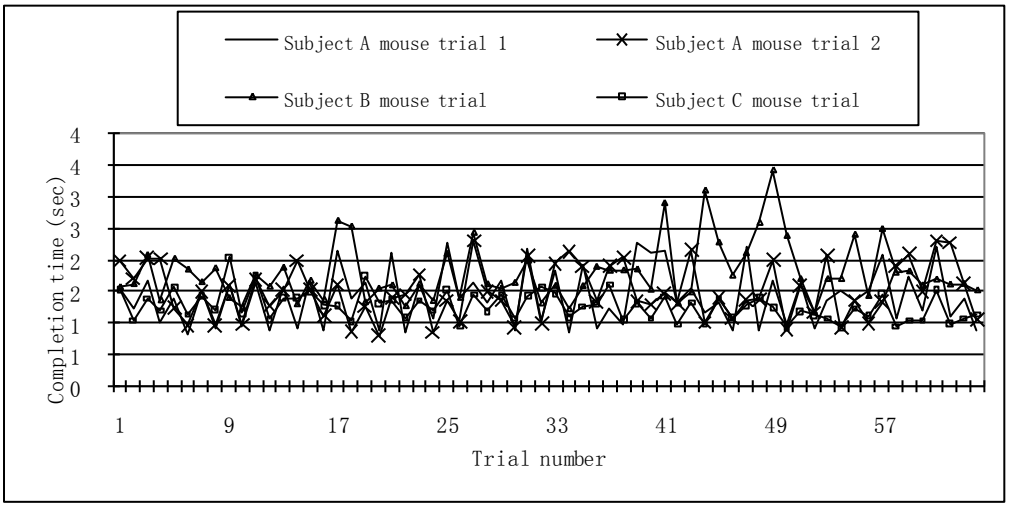


Figure 3-8 Mouse trials completion times for all three subjects

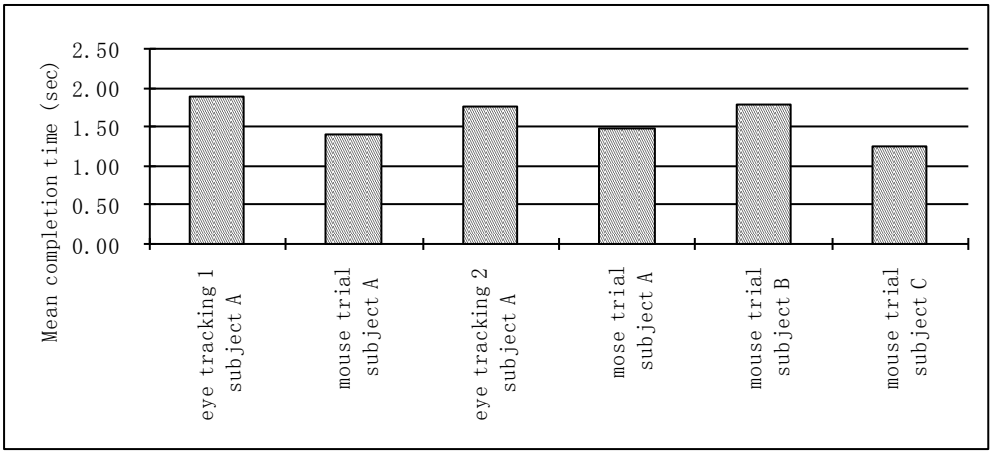


Figure 3-9 Mean completion times for eye tracking and mouse trials

Table 3-1 t-test for eye tracking and mouse, Test 1

	eye tracking I	Mouse I
Mean (second)	1.91	1.42
Variance	0.52	0.20
Observations	64	64
t statistic	105.00	
P(T<=t) two-tail	<0.001	

Table 3-2 t-test for eye tracking and mouse, Test 2

	eye tracking II	Mouse II
Mean (second)	1.77	1.50
Variance	0.22	0.18
Observations	64	64
t statistic	125.00	
P(T<=t) two-tail	<0.001	

3.2.5 Discussion

As a prototype for the pilot test, the Hot-Zone was not fully developed. Some key features were not implemented. First, the appearance of the Hot-Zone could not be modified without editing the code itself. It is desirable for users to customize the Hot-Zone themselves. For example, users may want to change the zone size, color, font, the number of zones and various other properties. Second, the interaction scheme was fixed in the pilot tests. Users had to hold the hotkey and blink to call out the Hot-Zone. However, there are other alternatives worth investigating, such as fixate and press a hotkey to call out Hot-Zone. Current prototype does not provide a means for switching between different manipulations.

The third problem is the range of the display. Sometimes the Hot-Zone could not be fully displayed on the screen because the center was too close to the screen boundary. Thus,

the feedback from the Hot-Zone was lost and the subject could only assume whether the corresponding zone was selected or not. This problem could be avoided by a boundary check and a more flexible display of the Hot-Zone that could change the appearance accordingly.

Finally, the Hot-Zone prototype in the pilot study was developed using Visual Basic 6, which does not support multithread programming. Since the Hot-Zone is required to run with the host program simultaneously, the coordination between the two programs is not optimal. It appeared that sometimes the system did not respond to blinks or hotkey-presses in time, which was an annoying experience for subjects. This difficulty can be solved by using more advanced programming tool with the .NET framework.

3.3 Gaze pattern recognition

3.3.1 Objectives and hypotheses

The Hot-Zone provided easier and reliable access to functions with gaze control, but the number of functions to be displayed at any time was very limited compared to the drop down menu. To make the Hot-Zone more useful, only currently needed functions should be displayed. Two information sources could be used to achieve this goal: (1) the object model of the application, and (2) the users' intension imbedded in the eye movement. This study was focused on the second information source. Gaze pattern recognition was investigated to assist eye-tracking interaction by providing insights on user attention.

3.3.2 Methods

This experiment was designed to test the gaze pattern recognition algorithm and was aimed at distinguishing reading from visual search (Figure 3-10). There were two tasks used

in the experiment, classified as: (1) Read an article, and (2) Search for target ‘T’ among ‘L’s.

In the reading task, a text was displayed in the middle third of a display screen and there were 8 pages in total. The subject pushed the ‘Next’ button to get to the next page. In the visual search task, a randomly generated picture was displayed in the center of the screen. There was only one target ‘T’ among roughly 60 ‘L’s. Once the target was found, subjects pointed the cursor at the target and right clicked to bring out a context menu. By selecting the ‘Next’ menu item, the next picture was drawn on the screen. There were 32 trials in total for visual search. The first 7 pages of the reading task and the first 29 pages of the search task were used for HMM training. The last reading page and last three search pages were used for testing. Since this was the first attempt at such an application, the gaze pattern recognition was conducted off-line.

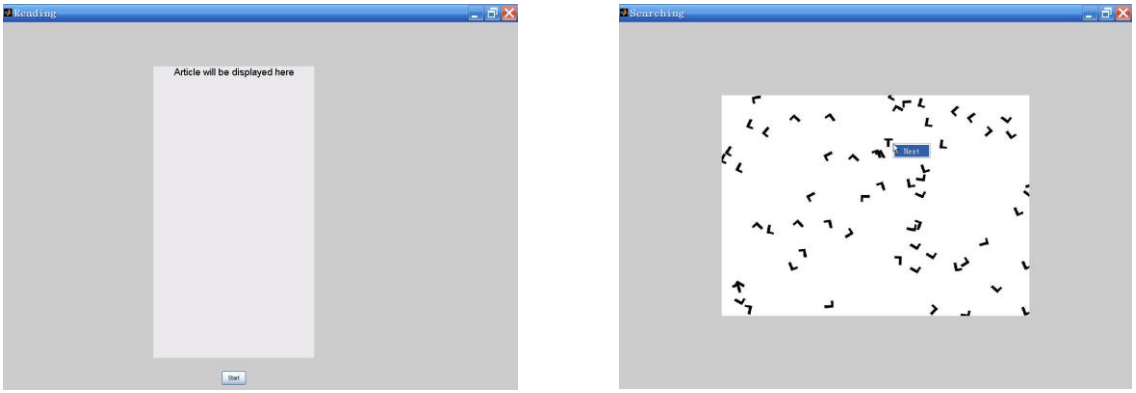


Figure 3-10 Interfaces for reading trial (left) and visual search trial (right)

3.3.3 Results

Feature extraction

To extract the eye movement features from the raw tracking data, a short-time energy

feature was used, which is also commonly used in Voice Activity Detector (VAD) in speech communication. Figure 3-11 shows the velocity profiles for both reading and visual search from the pilot study. The abrupt spikes in the vertical direction (second row) were mainly caused by blinks that appear to be far off the normative velocity values. Since saccades are mostly made from left to right in reading, it was expected that saccades in the horizontal direction (first row) represented the main characteristics of this behavior. Furthermore, saccades are more frequent and less regular in visual search than in reading, as indicated by other studies. Therefore, the magnitude of saccades in the horizontal direction was used exclusively for recognition. In the pilot study, the data packet used to calculate the short-time energy was 6 samples at a 120 Hz sampling rate with a 2-sample overlay. This corresponds to 50 ms for each frame. To reduce the data for HMMs calculation in the later stage, scalar quantization was deployed to assign a symbol from a finite alphabet to each frame. Codebooks with 16, 32 and 64 codewords were designed and applied for the test (Figure 3-12).

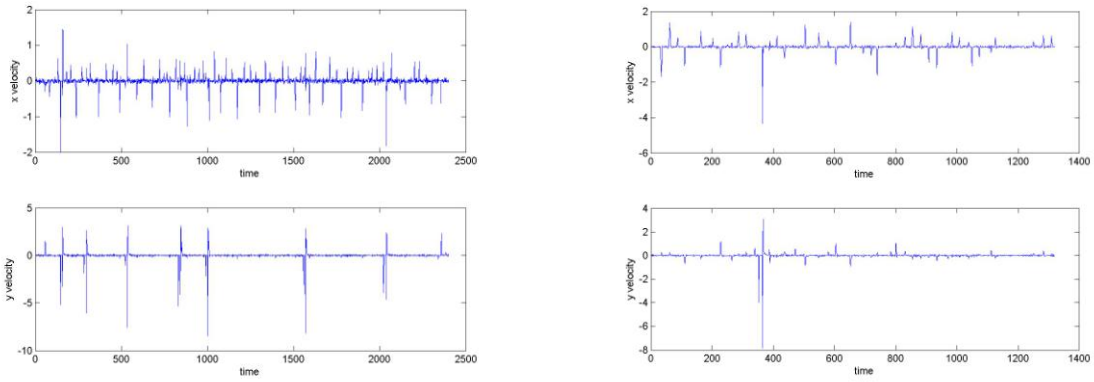


Figure 3-11 Velocity profiles for reading (left) and visual search (right)

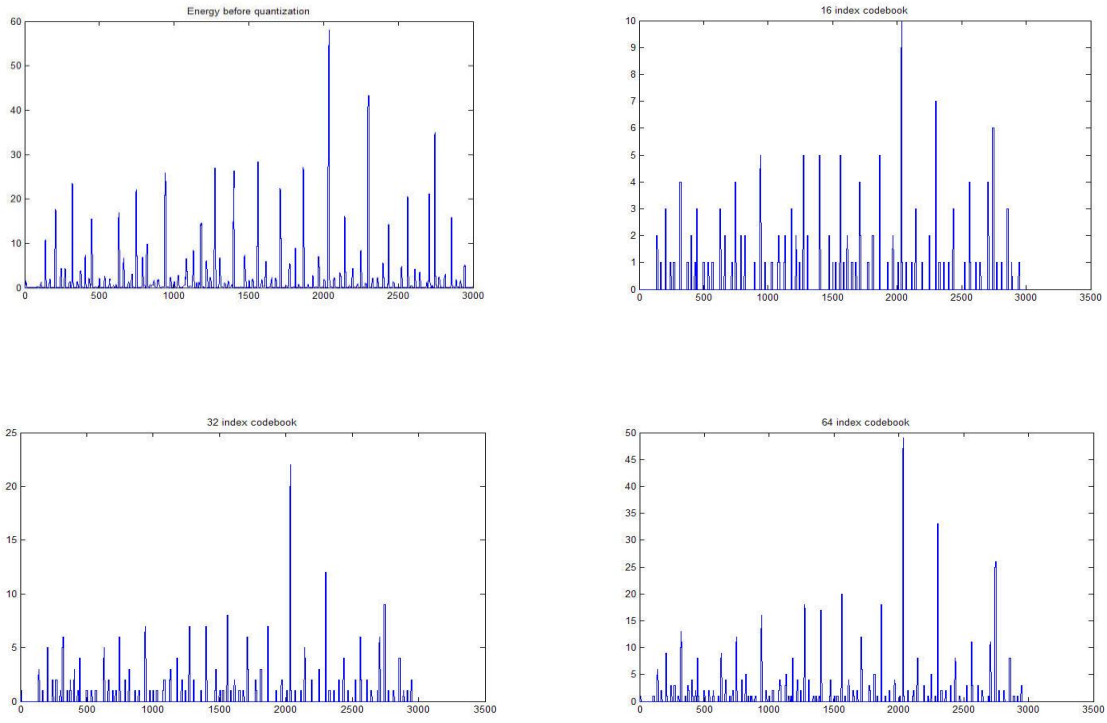


Figure 3-12 Scalar quantization

Recognition unit for gaze patterns

As in other pattern recognition applications, selecting an appropriate unit for HMM modeling is essential. The unit should be a well-defined segment that can be used for training and matching. Another important requirement is that the representation of the unit variance should be bounded. For example, in small-vocabulary speech recognition, words are separated by using a VAD to serve as the recognition unit. However, defining recognition units is not an easy task for gaze pattern recognition. Unlike words, eye movements do not have explicit boundary points for reading and visual search. In this study, we avoided dealing with unit length variance by using 64 fixed samples as recognition units.

Choice of states

In most eye tracking studies, fixations and saccades are the building blocks of various visual activities. Thus, these two are natural candidates for the internal states in HMMs. However, the detection of a fixation or saccade (a binary sequence) alone is not adequate for representing different gaze patterns. By observation, it has been found that the magnitude of the quantized energy is also an important source of information. Accordingly, three internal states of the HMMs are defined, including fixation, small saccade, and large saccade.

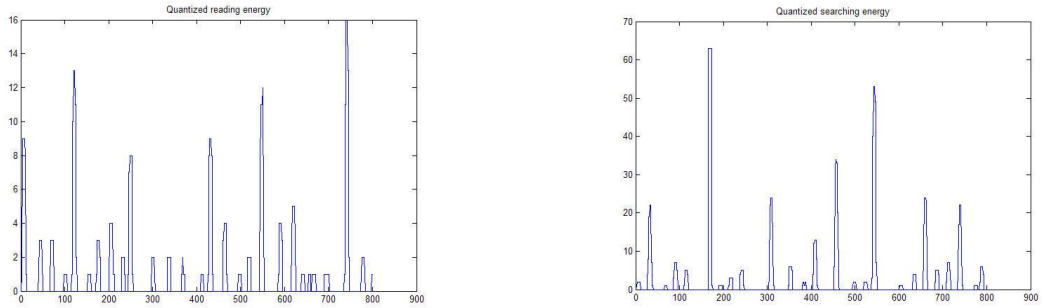


Figure 3-13 Short-time energy for reading and visual search

Recognition rate

For initialization, the following parameter matrices to train the HMMs were assumed for reading and visual search:

$$\boldsymbol{\pi}_0 = \begin{bmatrix} 0.8 & 0.15 & 0.05 \end{bmatrix}; \qquad \mathbf{A}_0 = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.3 & 0.5 & 0.2 \\ 0.4 & 0.4 & 0.2 \end{bmatrix}$$

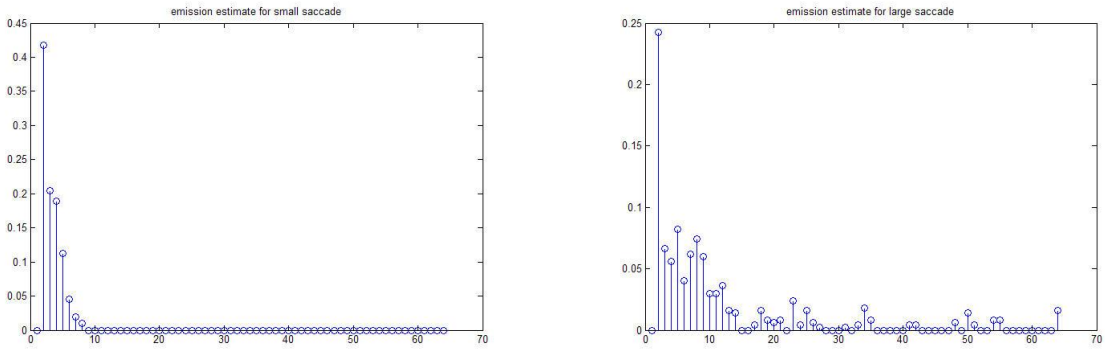


Figure 3-14 Initial emission matrix B_0 estimation for small and large saccades

Past studies have shown that either random or uniform initial estimates of $\boldsymbol{\pi}_0$ and $\mathbf{A}_0$ are usually adequate for giving useful re-estimates in almost all cases. However, a proper estimate for the initial assumed $\mathbf{B}_0$ is important, and needs to be evaluated carefully. Figure 3-14 shows the obtained initial emission matrices for small and long saccades, based on manual counting of several trials. All zero emission probabilities were set to 0.001 because zero values remain unchanged throughout the optimization procedure. Estimation for all three HMMs converged within 12 iterations.

Figure 3-16 shows the HMM representing reading after training with the state transition probabilities marked beside the connections. The preliminary results (Table 3-3) show that the algorithm failed to distinguish reading from visual search. The recognition unit with 128 samples was also tested which made no evident improvement on the results.

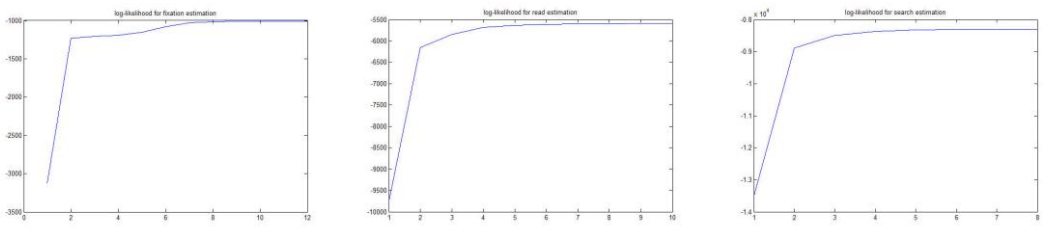


Figure 3-15 Parameter estimation convergence

The final estimated initial matrix for each Hidden Markov Model is:

$$\pi_{fixation} = \begin{bmatrix} 0.9 & 0.0667 & 0.0333 \end{bmatrix}$$
$$\pi_{read} = \begin{bmatrix} 0.7463 & 0.194 & 0.0597 \end{bmatrix}$$
$$\pi_{search} = \begin{bmatrix} 0.7067 & 0.1741 & 0.1192 \end{bmatrix}$$

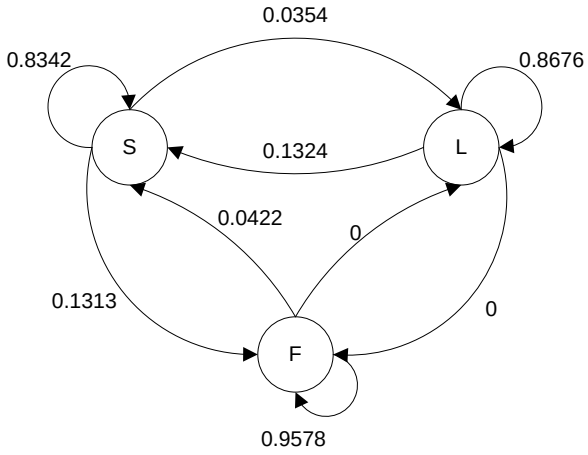


Figure 3-16 Trained HMM for reading. S-small saccade, L-large saccade, F-fixation

Table 3-3 Results of gaze pattern recognition

	Read trials	Search trials
Fixation	20	20
Reading	22	11
Searching	29	12

3.3.4 Discussion

The failure of the HMMs to recognize gaze patterns may attributed to several factors. First, the variance of the fixed length recognition unit is relatively high. This results in the ambiguity among the three HMMs. Second, the feature extraction based on the saccade magnitude in one direction alone is not adequate. Information such as saccade direction needs to be explored as well. The training sequences for the three HMMs also need to be carefully scrutinized to establish templates with consistent initial states and sequential transitions between the states.

3.4 Deterministic and stochastic components for ocular dynamics

3.4.1 Objectives and hypotheses

The objectives of the pilot work were to develop an ocular dynamic model that synthesizes available neuronal-anatomical evidence and behaves in a way compatible with actual eye movements. To use the limb movement models as benchmarks for saccadic eye movements, this study was limited to saccades with small amplitudes such that the physical, physiological and stochastic principles embodied in those models reflect the properties of underlying eye movement mechanisms. Deterministic and stochastic components were explicitly modeled to address the sources of saccade variance. Two factors in the saccadic eye movement variation were proposed: (1) target point inaccuracy prior to eye movements and (2) signal dependent noise during the movements. Figure 3-17 shows the overall organization of the model. Noise from the two sources is applied to the system at different stages before the

deterministic model converted the velocity to eye orientation. This model reveals the uniqueness of eye movements from a biomechanical point of view. It has the potential to provide design guidelines for eye tracking interaction on the basis of the notion that variability in eye movements is related to variability in perception (Coren, 1986).

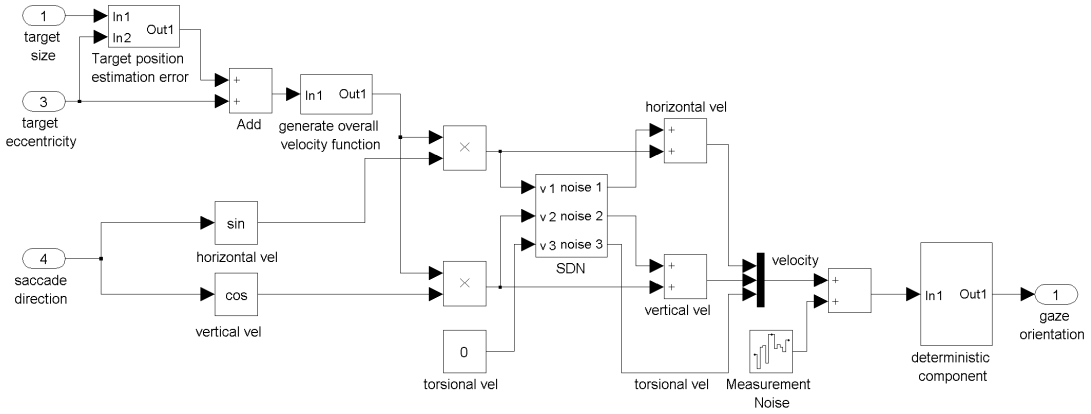


Figure 3-17 Schematic diagram of the ocular dynamic model

3.4.2 Deterministic component modeling

Evidence from recent studies supports the idea of a mechanical implementation of Listing's Law (Martinez-Trujillo, 2005). That is, the vector model is more likely to agree with the physiological organization of the ocular dynamic system. (Background information related to vector and quaternion models was provided in Section 2.5). However, the two-stage transition from neural impulses to muscle forces and finally to eye orientation requires a number of parameters, most of which can only be estimated or assumed based on the time constant of the ocular system. To verify the model, it is necessary to develop a method of generating proper inputs, given the desired output. The performance of the model can then be evaluated by comparing the output of the system with experimental data.

For the vector model, the “impulse-step” input neuronal signals corresponding to a given saccade amplitude are generated according to the velocity profile and stable state requirements of the ocular dynamic system (Raphan, 1998; D.A. Robinson, 1975; Tanaka, et al., 2006). The legitimacy, accuracy and sensitivity of these parameters remain an open question, given that no one has yet pursued these issues explicitly.

The criteria for generating input signals with fidelity may depend on the tradeoff between model complexity and reliability, where the velocity ω seems to be a more suitable candidate than the motoneuron signals. The legitimacy is insured through the relative simplicity of the ocular system, as seen in Equation (2-16), and the movement of the eyeball is a faithful reflection of motoneuron activity. Furthermore, since eye movement in one direction is mostly caused by activities of one single pair of EOMs, variations in component velocity profiles could well represent variability of the corresponding muscle forces.

Past studies also established equations that quantitatively relate the saccade amplitude to peak velocity and duration. With these equations, we can conveniently generate a velocity function and ensure the reliability inherited from the statistical power of those equations. Besides, it is much easier to verify the model since almost all eye-tracking systems provide measurements on eye movement velocity.

Form of the velocity function

In the impulse variability model, it is generally assumed that to produce a desired limb movement, a particular force function is selected. This function specifies the net amount of force generated by the muscles to drive the body part as a function of time (Meyer, et al., 1982). These force functions are assumed to come from a family with the same general shape that could be determined by a force parameter and a time parameter.

Analogically, it is assumed that there is a set of velocity functions with similar

symmetric shapes over time for the oculomotor system to choose from. These functions are characterized by two parameters, including peak velocity and duration, which could be unambiguously determined by (2-12) and (2-13). Figure 3-18 shows the shapes of typical velocity profiles for saccades, which range from 5 to 80 degrees (Collewijn, et al., 1988).

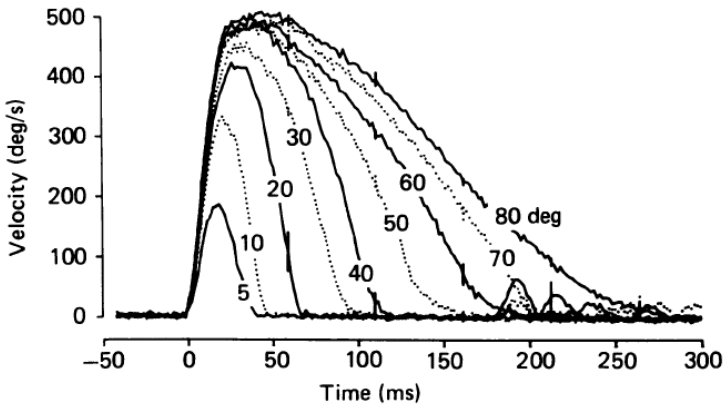


Figure 3-18 Velocity profiles of horizontal saccades (Collewijn, et al., 1988)

The aforementioned assumption arises naturally when we consider the movement dynamics and principles of physics. Past studies indicated that this assumption generally holds for small saccade, whose amplitude is less than 40 degrees (Collewijn, et al., 1988; A. J. Van Opstal & Van gisbergen, 1987). Thus the present experiments were limited to saccades no greater than 10 degrees. Such an arrangement is good enough for most naturally occurring saccades because 86% of saccades are less than 15 degrees (A. T. Bahill, Adler, & Stark, 1975). Past studies in the literature have not reported mathematical functions to represent saccade velocity as a function of time, so the following equation was used in this study

$$\omega(t) = \frac{1}{at^2 + bt + c}; \quad (a > 0, b < 0, b^2 - 4ac < 0) \qquad (3\text{-}1)$$

Figure 3-19 shows the general shape of Equation (3-1), which resembles the velocity profiles obtained from empirical data as seen in Figure 3-18. The choice of the function is not

mandatory but rather practically oriented. The main reason is its simplicity requiring only three parameters to define its shape. The number of parameters can be further reduced to two by shifting the time axis. The remaining two parameters can be derived by peak velocity and duration that in turn are established by existing regression models.

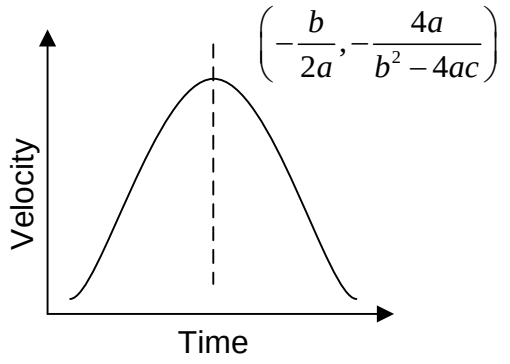


Figure 3-19 Velocity function

The parameter b can be eliminated by shifting the center of the function to be at $t = 0,$ which results in $b = 0.$ Then the following conditions can be satisfied

$$\int_{-d/2}^{d/2} \omega(t)dt = A; \quad -\frac{4a}{b^2-4ac} = \frac{1}{c} = \omega_m \tag{3-2}$$

where ω_m is the peak velocity and A is the saccade amplitude. From Equation (3-2), c is determined by evaluating the peak velocity with Equation (2-12) and a is determined by integrating over the saccade duration d. Figure 3-20 shows the velocity function generated for a 7° saccade by Equation (3-2), with $a = 9.9$ and $c = 0.0048$.

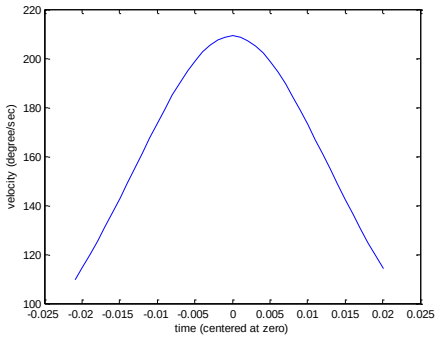


Figure 3-20 The velocity function for a 7˚ saccade

It is also assumed that the three pairs of extraocular muscles work on the eyeball independently with no correlation with each other. Thus, each velocity component could be viewed as the result of activity of a single pair of EOMs'. Therefore, once the overall velocity function was generated, the horizontal and vertical components could be decomposed accordingly based on the saccade direction:

$$\omega_{horizontal}(t) = \omega(t)\cos\theta; \quad \omega_{vertical}(t) = \omega(t)\sin\theta \qquad (3\text{-}3)$$

Validation of the concept

Equations (2-12) and (2-13) were originally derived from experimental data for *horizontal* saccadic eye movements. In the present model, it was assumed that this relationship also apply to saccades at *oblique directions* (Figure 3-21). This assumption is implicitly supported by previous literature, which showed that partial correlation coefficients between duration and amplitude, as well as peak velocity and amplitude, do not vary as saccade direction changes (van Beers, 2007). The pilot study also indicated that dynamic features of oblique saccades can be well approximated by Equations (2-12) and (2-13).

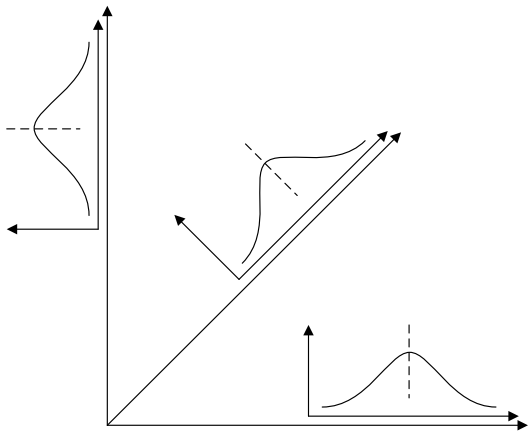


Figure 3-21 Velocity function in oblique direction

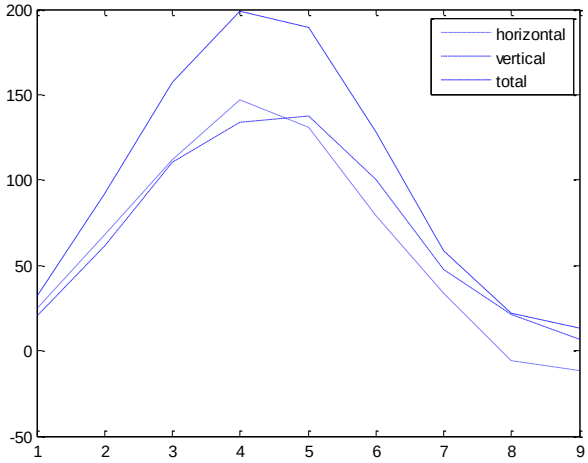


Figure 3-22 Velocity profile of oblique saccade from experiment data

Figure 3-22 shows the velocity profile for a 7˚ saccade in a 45˚ upper right direction, recorded at 120 Hz from the pilot study. The peak velocities in both vertical and horizontal directions were about 130˚/second with an overall velocity of 200˚/second. By comparing this with Figure 3-20, the peak velocity from the experiment is close to the value estimated by the

regression model, which was 209˚/second. Since the saccade amplitude in the pilot study was relatively small (maximum 7˚), this velocity function was considered as a good approximation.

Vector model for the half angle rule

Thus far, the vector integrator model was validated primarily by simulation. Step-by-step, the vector integrator model used in this study was verified for consistency with the quaternion model for the velocity-position transformation in implementing Listing's Law and the half angle rule. A unit quaternion is written as follows:

$$q(t) = \left(\cos\frac{\theta(t)}{2}, \mathbf{v}(t)\sin\frac{\theta(t)}{2} \right) \qquad \text{(3-4)}$$
$$\mathbf{v}(t) \in \mathbf{L}, \|\mathbf{v}(t)\| = 1$$

where $\mathbf{L}$ denotes Listing's plane. At the primary position, the following condition was chosen as the initial state:

$$\upsilon_0 = \frac{\omega_0}{\|\omega_0\|}; \dot{\theta} = \|\omega_0\|$$

Substituting Equation (3-4) into (2-21), the formula for velocity in quaternion form is obtained as

$$\dot{q} = \frac{1}{2}\left(-(\boldsymbol{\omega}\cdot\mathbf{v})\sin\frac{\theta}{2}, \boldsymbol{\omega}\cos\frac{\theta}{2} + (\boldsymbol{\omega}\times\mathbf{v})\sin\frac{\theta}{2} \right) \qquad \text{(3-5)}$$

Since $\mathbf{v}(t) \in \mathbf{L}$ for any t, therefore, $\mathbf{v}(t)^T \in \mathbf{L}$. Because the vector part of (3-5) is a linear combination of $\mathbf{v}$ and $\mathbf{v}^T$, thus, to satisfy the above condition, the following equality must hold:

$$h_3 \cdot \left[\boldsymbol{\omega}\cos\frac{\theta}{2} + (\boldsymbol{\omega}\times\mathbf{v})\sin\frac{\theta}{2} \right] = 0 \qquad \text{(3-6)}$$

Note that

$$\begin{aligned} h_3 \cdot (\boldsymbol{\omega} \times \boldsymbol{\upsilon}) &= h_3 \cdot \mathbf{S}(\boldsymbol{\omega}) \cdot \boldsymbol{\upsilon} = (\boldsymbol{\omega} \times h_3) \cdot \boldsymbol{\upsilon} \\ &= -(h_3 \times \boldsymbol{\omega}) \cdot \boldsymbol{\upsilon} = -\boldsymbol{\omega} \cdot \mathbf{S}(h_3) \cdot \boldsymbol{\upsilon} \\ &= -\boldsymbol{\omega} \cdot (h_3 \times \boldsymbol{\upsilon}) \end{aligned} \tag{3-7}$$

$\mathbf{S}(\bullet)$ is the skew-symmetric matrix. So (3-6) can be rewritten as follows

$$\boldsymbol{\omega} \cdot \left[h_3 \cos\frac{\theta}{2} - (h_3 \times \boldsymbol{\upsilon}) \sin\frac{\theta}{2} \right] = 0 \tag{3-8}$$

Thus, the angular velocity component in h_3 must satisfy

$$\boldsymbol{\omega} \cdot h_3 = \boldsymbol{\omega} \cdot (h_3 \times \boldsymbol{\upsilon}) \tan\frac{\theta}{2} \tag{3-9}$$

Equation (3-9) indicates that the angular velocity must be tilted out of Listing's plane by an angle of $\theta/2$. The plane $\mathbf{P}$ containing all the angular velocity is neither fixed to the head nor eye since its normal depends on both θ and $\boldsymbol{\upsilon}$. This property implied by Listing's Law is the so-called half angle rule. On the other hand, from Equation (2-26), notice that

$$\boldsymbol{\omega} \times \boldsymbol{\upsilon} = \mathbf{S}(\boldsymbol{\omega}) \cdot \boldsymbol{\upsilon} = \begin{bmatrix} -\omega_3\upsilon_2 + \omega_2\upsilon_3 \\ \omega_3\upsilon_1 - \omega_1\upsilon_3 \\ -\omega_2\upsilon_1 + \omega_1\upsilon_2 \end{bmatrix}, \qquad \boldsymbol{\upsilon} \times (\boldsymbol{\omega} \times \boldsymbol{\upsilon}) = \mathbf{S}(\boldsymbol{\upsilon}) \cdot (\boldsymbol{\omega} \times \boldsymbol{\upsilon})$$

The torsional rotation axis vector is obtained as

$$\dot{\upsilon}_3 = \frac{1}{2}(-\omega_2\upsilon_1 + \omega_1\upsilon_2) + \frac{1}{2}(\upsilon_1^2 + \upsilon_2^2)\omega_3 \cot\frac{\theta}{2} \tag{3-10}$$

Listing's Law requires a zero torsional part, i.e., $(-\omega_2\upsilon_1 + \omega_1\upsilon_2) + (\upsilon_1^2 + \upsilon_2^2)\omega_3 \cot\frac{\theta}{2} = 0$.

Therefore, the following equation must hold:

$$\omega_3 = \frac{(\omega_2\upsilon_1 - \omega_1\upsilon_2)}{(\upsilon_1^2 + \upsilon_2^2)\cot(\theta/2)} = (\omega_2\upsilon_1 - \omega_1\upsilon_2)\tan\frac{\theta}{2} \tag{3-11}$$

Rewriting Equation (3-9), we obtain

$$\begin{aligned}\boldsymbol{\omega}\cdot h_3 &= \boldsymbol{\omega}\cdot(h_3\times\boldsymbol{\upsilon})\tan\frac{\theta}{2}\\ &= \begin{bmatrix}\omega_1 & \omega_2 & \omega_3\end{bmatrix}\left(\begin{bmatrix}0 & -1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\upsilon_1\\ \upsilon_2\\ \upsilon_3\end{bmatrix}\right)\tan\frac{\theta}{2} \qquad (3\text{-}12)\\ &= (\omega_2\upsilon_1-\omega_1\upsilon_2)\tan\frac{\theta}{2}=\omega_3\end{aligned}$$

Equations (3-11) and (3-12) show that the same result is obtained for the torsional component of the angular velocity derived from the quaternion and vector integrator models. Because both the quaternion and vector models are used to describe the same motion, the vector integrator model should represent the half angle rule like the quaternion model. To calculate the saccade axis and amplitude from the input angular velocity vector, a Simulink model (Figure 3-23) was developed based on Equations (2-25) and (2-26).

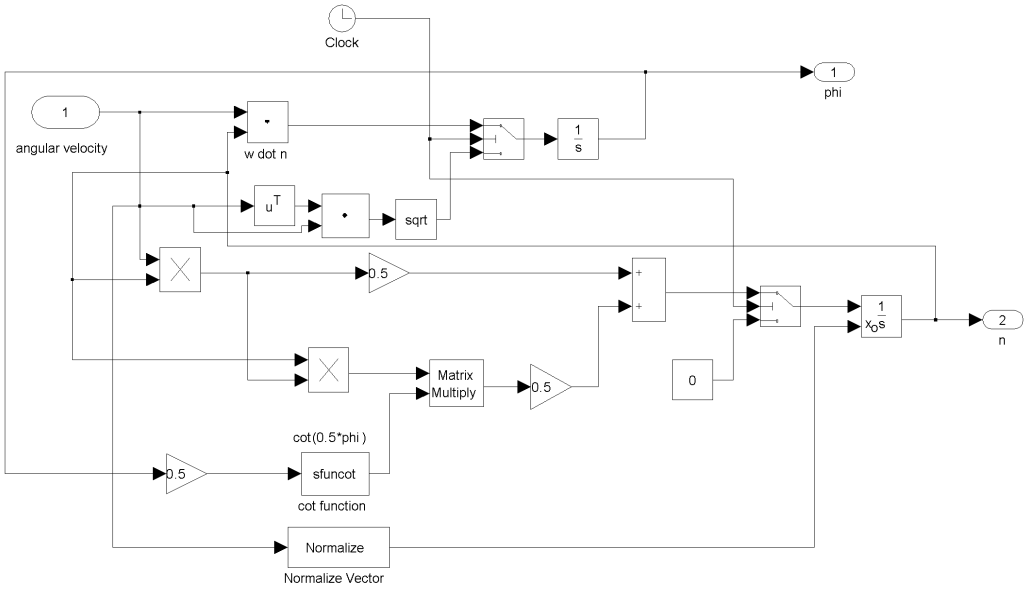


Figure 3-23 Vector rotation model in Simulink

3.4.3 Stochastic component modeling

By comparing the variance of the endpoints, the systematic variability of eye orientation after saccadic eye movements was examed (Figure 3-24). Previous studies showed that corrective movements post-saccade seldom occur and are limited almost exclusively to very small targets (Kowler & Blaser, 1995; D. A. Robinson, 1975). Thus, endpoint variance is conventionally treated as an appropriate representation for saccade variability (van Beers, 2007; A. J. Van Opstal & Van Gisbergen, 1989). Two factors were considered in the stochastic component of the ocular dynamic models (

Table 3-4). First, there are sensory and cellular noises that refer to the inaccuracy of target position perceived by subjects prior to the saccade. Second, there is motor noise that is identical to the signal dependent noise. Other variations such as the device measurement noise were treated as white noise where the magnitude mainly depends on the hardware, and could be estimated through experimental data.

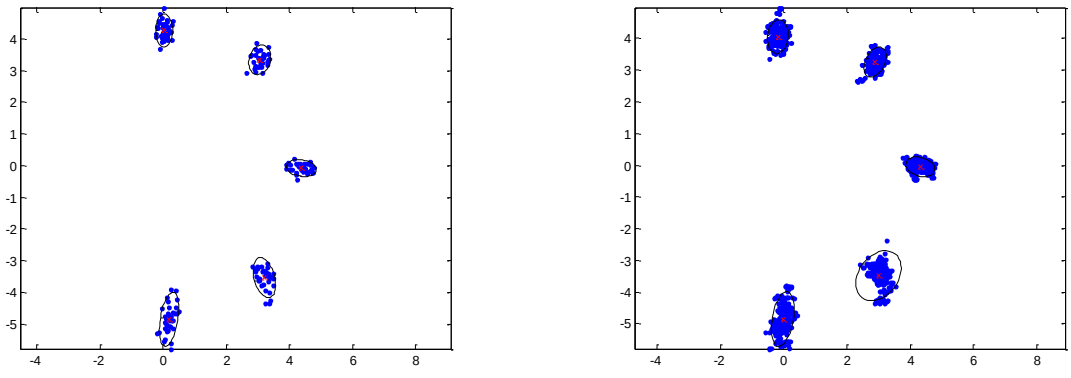


Figure 3-24 Left: End point scatter plot; Right: fixational gaze points after saccades stop

Table 3-4 Stochastic component of ocular dynamic models

Stochastic component	Source of noise	Influence factors
Target position error (prior to eye movement)	Sensory and cellular noise (anisotropy of retina anatomy)	Target distance Target size Saccade direction
Signal dependent error (during eye movement)	Motor noise	Target distance Saccade direction

Variance of target position estimation prior to saccade

As previously mentioned, because of the physiological and mechanical limitations of the photoreceptors, a sharp image can only be acquired in the central retinal area. The peripheral retina also transmits visual stimulus to the brain, but with significantly lower accuracy. Thus, a target with larger distance is located further from the fovea and this contributes to a lower visual resolution. Past studies in horizontal eye movements have shown that given a fixed target size, the variance of target position estimation increases with target distance (van Beers, 2007; A. J. Van Opstal & Van Gisbergen, 1989). On the other hand, given fixed target distances, larger target size means that the variance of target estimation can afford to be larger. For horizontal eye movements, the following equation (Tanaka, et al., 2006; A. J. Van Opstal & Van Gisbergen, 1989) was used to represent the variance of the estimated target points

$$\sigma^2 = (W + sA)^2 \qquad (3\text{-}13)$$

W is the width of the target, A is the target distance, and the slope s is a value between 0.02 to 0.05 in human subjects (A. J. Van Opstal & Van Gisbergen, 1989). It should be noted that the equation was originally developed based on the data of the endpoint post-saccade rather than the target point acquired before the saccade. Therefore, to be exact, this equation already includes the variance caused by the signal dependent noise during the saccade.

Nevertheless, considering that this equation captures the general idea of how the two factors (target distance and size) affect the variance and only requires one parameter for adjustment, it can still serve as a baseline and be used to extend the present work from one dimension to higher dimensions.

When considering the variability of eye rotation in three dimensions, the direction of the saccade starts to play a role. For convenience, we project the saccade trajectory onto a plane. The term radial is used to represent the direction along the radius in the visual field and the term tangential is used to represent the direction orthogonal to the radial direction. Studies have shown that spatial localization of the human fovea is anisotropic, which means the just noticeable difference threshold is different in two orthogonal directions (A. J. Van Opstal & Van Gisbergen, 1989; Westheimer, 2001). To be specific, given a fixed target, variance in the radial direction is larger than that in tangential direction. Again, such phenomenon could be intuitively explained by the retinal anatomy (Figure 2-3): the density of the photoreceptor cells and thickness of the retina decrease along the radial direction while they are relatively constant in tangential direction (Curcio, et al., 1987).

To model the variance caused by the sensory noise, the endpoint variance in the radial direction is evaluated based on Equation (3-13). The challenge is to determine the variance in the tangential direction. One important fact observed in previous studies is that, for a 45° upper right saccade, there is no significant relationship between target distance and the ratio σ_R / σ_ϕ, where σ_R is the variance in radial direction and σ_ϕ is the variance in tangential direction (A. J. Van Opstal & Van Gisbergen, 1989). In this way, the variability caused by signal dependent noise in the horizontal and vertical directions is nullified. The confidence ellipse of endpoints scatter plot should be a circle with σ_R / σ_ϕ (equal to 1) given that only the motor noise is applied. However, experiment showed that this ratio is constantly larger

than 1 with an average of 2 over different subjects (A. J. Van Opstal & Van Gisbergen, 1989). By inference, this constant ratio can only be explained by the target position variance before the saccade and, therefore, the variance in tangential direction was calculated as:

$$\frac{\sigma_R}{\sigma_\phi} = C \tag{3-14}$$

To simplify this complex and asymmetric ocular dynamic, the sensory noise is assumed to be distributed proportionally along the radius of the fovea. The irregular density distribution of the retina may be the source of the asymmetry of the eye movement that is not considered in the model. Asymmetric structures could be incorporated into the existing framework when necessary, at an extra cost of complexity and computational load.

Signal dependent noise during saccades

For limb movement, the variability of the output force is proportional to the mean value of the force, as an analogy of Webber's law in motor behavior (Enoka & Fuglevand, 2001). Signal dependent noise (SDN) has been observed in both voluntary and stimulation-induced muscle contractions (Jones, et al., 2002). As skeletal muscles, the extraocular muscles also show similar noise patterns in output forces. On the other hand, the ocular kinematics that transform muscle force to velocity is dictated by mechanical principles. Therefore, it is plausible to quantify the velocity variability as a function of force variability through Equation (2-16), which is essentially a low pass filter. For simplicity, the SDN of velocity is analyzed according to the same form as the SDN of the force in the pilot study. Equivalently, its variability (standard deviation) increases linearly with its magnitude. This echoes analytically the force pulse variability assumption in the limb movement model (Abrams, Meyer, & Kornblum, 1989; Meyer, Smith, Kornblum, Abrams, & Wright, 1990).

Existence of velocity SDN is implicitly indicated by empirical data, which shows that

the peak velocity magnitude plays a role in saccade variance as the amplitude increases (Boghen, Troost, Daroff, Dellosso, & Birkett, 1974). There have been several mathematical functions proposed to account for the SDN (C. M. Harris & Wolpert, 1998; Jones, et al., 2002). In this study, we followed the practice of treating SDN as a white noise with a mean value of zero and standard deviation proportional to the signal magnitude.

$$E[\xi(t)] = 0; \quad \sigma^2[\xi(t)] = k\omega^2(t) \tag{3-15}$$

Signal dependent noise is applied separately to each of the three angular velocity components. Such an arrangement is based on the previous assumption that the three pairs of EOMs work independently and this could be a reasonable candidate as an estimator. A system usually demonstrates such a phenomenon, when three or more parameters are estimated simultaneously, the combined estimator is more accurate than any method that handles the parameters separately. This is surprising since the parameters and the measurements might be totally unrelated.

3.4.4 Methods

The two sources of variance were applied to the deterministic model at different stages and simulated the output of the ocular dynamics system to generate movement predictions. Then, the predicted saccade variances were compared with the experimental data given the same conditions used in the simulation (target distance, direction, etc.).

One male subject with normal vision participated in the pilot study. Stimuli were presented on a 33 by 24 cm screen with a resolution of 1024 by 768 pixels at a distance of 34.65 inch. To eliminate the effect of head motion caused by fatigue, the subject wore a collar support. After device calibration, the subject was first required to perform a fixation trial in which he fixated at five points on the screen (four corners and a central point) for a short

period. For saccade trials, the central dot (150 pixels) was continuously presented at the center of the screen while the target dot was displayed every 2 seconds and remained for 1.5 seconds. When the target dot disappeared, the subject switched the gaze back to the central point and waited for the next stimulus. Saccades with different amplitudes, target sizes, and directions were tested, with a block design instead of randomly appearing targets to avoid slow changes in saccade performance during an experiment session (Jurgens, Becker, & Kornhuber, 1981).

In each block, the subject changed his gaze 45 times from the central to the target dot. A rest was provided between two blocks. There were two conditions for the target size as well as for the target distance (Table 3-5), and targets were placed in five directions, evenly dividing the right half a circle.

Table 3-5 Experimental conditions (direction is not listed here)

Independent Variable	Condition 1	Condition 2
Target distance	7˚ (large)	3.5˚ (small)
Target size	1.1˚ (large)	0.16˚ (small)

3.4.5 Results

When the target position error prior to each saccade was applied, the length of confidence ellipse axes was determined by the target distance (Figure 3-25). The orientation of the ellipses was determined by the saccade direction. Thus, ellipse size increased as target distance increased and the long axis of the ellipse generally agreed with the radial direction. The size of the ellipse also increased as the target size increased given the same target distance (Figure 3-26). When the SDN alone was applied (Figure 3-27), the length of

confidence ellipse axes depended on both the target distance and direction. Particularly in the horizontal and vertical directions, the ellipses were flattest. In contrast, at oblique directions (45˚ and 315˚), ellipses were more round shaped. Figure 3-28 and Figure 3-29 show the 95% confidence ellipses for the experiment data. As can be seen from the simulation scatter plots, experimental results indicated that both target distance and direction have significant impact on determining the shape and direction of the confidence ellipses while the target size is not as important. Generally, the long axes of confidence ellipses are approximately aligned with the movement direction and the length of axes increases with target distance. These results generally agree with previous studies, as shown in Figure 3-30 (van Beers, 2007; A. J. Van Opstal & Van Gisbergen, 1989).

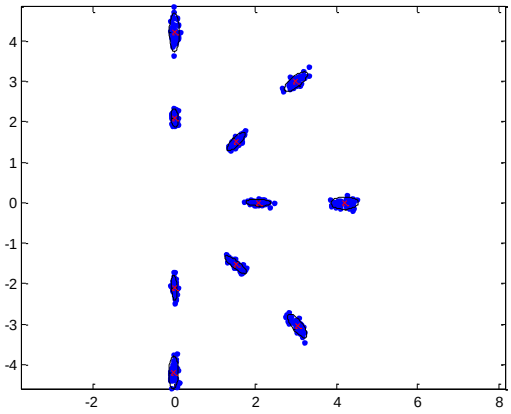


Figure 3-25 Effect of target distance on endpoint variability (small target)

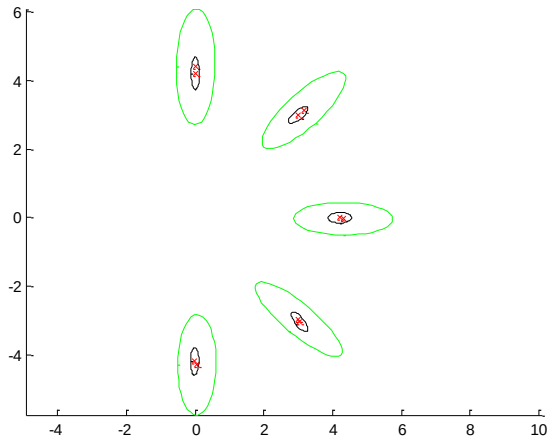


Figure 3-26 Effect of target size on endpoint variability

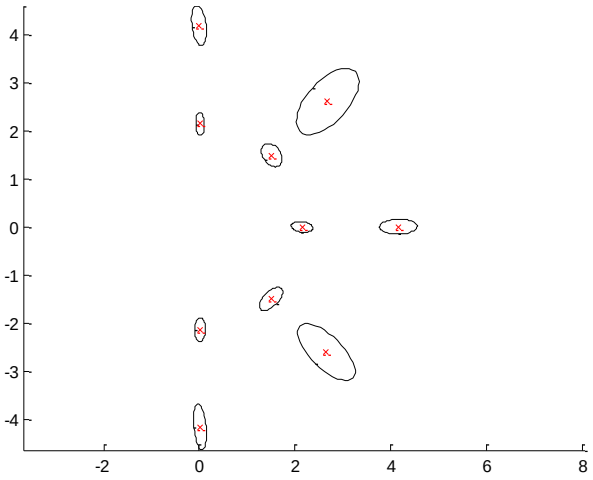


Figure 3-27 95% confidence ellipses of motor noise component prediction

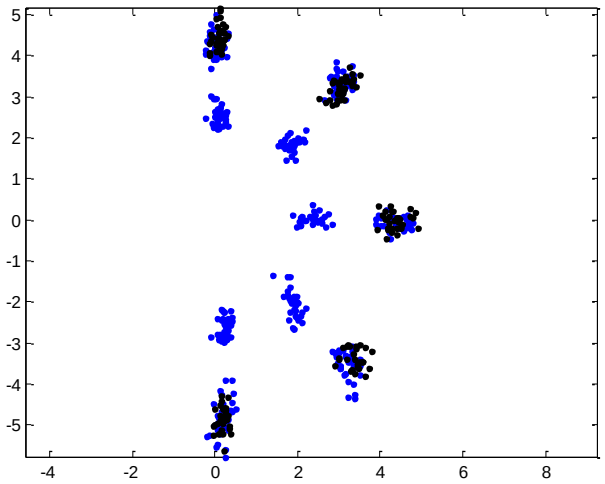


Figure 3-28 Endpoint scatter plot of experiment data

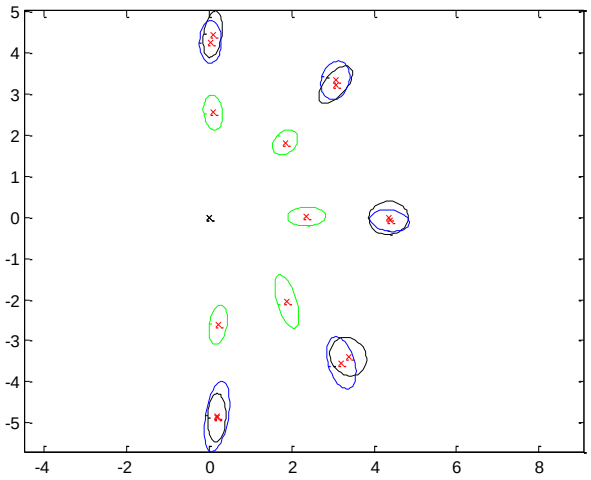


Figure 3-29 95% confidence ellipses on experiment data

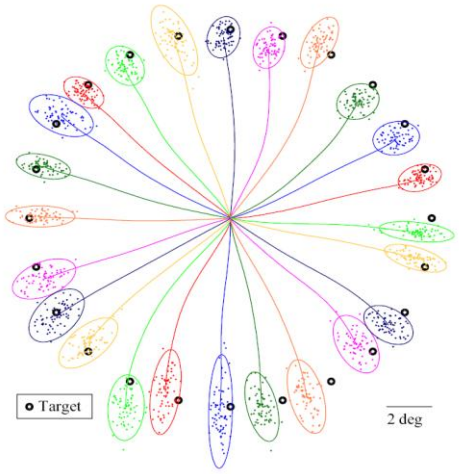


Figure 3-30 Endpoint scatter and 95% confidence ellipses from (van Beers, 2007)

3.4.6 Discussion

By comparing the experimental data with the simulation results, it is noticeable that neither variance source alone fully explains the saccade variability. Target estimation error prior to saccade could not explain the flatter shape of ellipses along orthogonal directions, e.g., horizontal and vertical. Signal dependent noise alone could not account for the ratio of variances in radial and tangential directions (A. J. Van Opstal & Van Gisbergen, 1989).

Since the predictions of neither variance source can fully capture the speed-accuracy tradeoff of saccadic eye movements, it is reasonable to exam if combination of the two may accomplish the goal. The stochastic component of the present model bridges external conditions, such as target distance, saccade direction and target size, with the endpoint variances. However, a few challenges emerged in this study regarding evaluation of assumptions and accumulated computational error in simulation.

First, a simple quadratic function was used to represent the velocity profile in which the

two parameters can be estimated by established linear equations. The velocity function obtained by this means is a coarse approximation of empirical data. There are several obvious flaws of using such a velocity function. For instance, the start and end velocity of the fitted function is much larger (over 100 degrees per second) than the true velocity observed. The shape of the function is always symmetric whereas the empirical data shows the velocity profile is right skewed. More importantly, being merely a mathematical "trick", this function has no physical connection to the ocular dynamic system and lacks the mechanical strength to explain the underlying control mechanism. This problem could be addressed by deriving appropriate optimal control models (C. M. Harris & Wolpert, 1998; Tanaka, et al., 2006) to generate velocity profiles, which agree with the experimental data.

The second problem is the accumulated round-off error in the simulation. Because the parameters of the velocity were determined numerically, and there are limitations on the length of simulation steps, integration of the velocity over the duration does not necessarily amount to the true amplitude. Because the initial state of the system is not defined in the model, there was a perturbation of the rotation axis vector after the simulation initiated. As verification of the simulation, simulated trials with no noise applied were used to demonstrate the effect of round-off error (Figure 3-31). The thickness of the dot cloud in radial direction indicated the amplitude of computational error. The angle covered by the strip-shaped dot shows the rotation axis error.

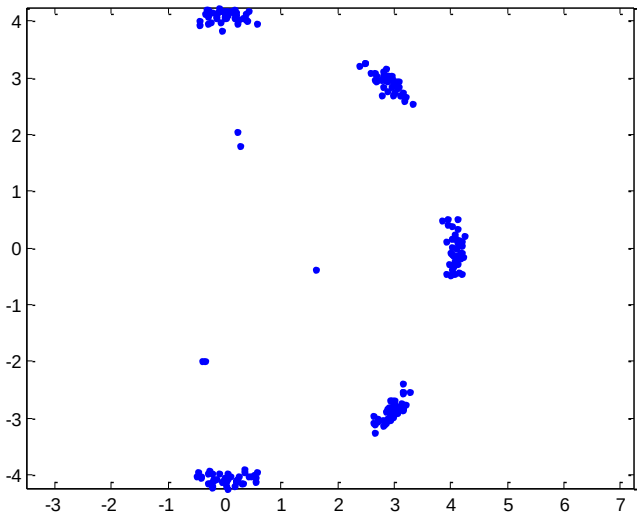


Figure 3-31 Accumulation error of simulation

4 Hot-Zone – a gaze interface

4.1 Research objective

The aim of this formal study was two-fold: (1) develop a gaze-friendly user interface (UI) widget; and (2) explore efficient gaze control schemes for human computer interactions. Given the low resolution of existing eye tracking systems, previous studies focused on fine-tuning the pointing process to acquire small targets. However, eye movement is fundamentally different from hand movement due to the anatomy and functionality involved. It is not plausible to directly apply conventional motor control models to gaze control. This mismatch, i.e., integrating commercially available UI(s) with gaze control, creates a typical problem encountered in human factors: users must constrain their natural behavioral variability to accommodate an existing machine interface. To reverse this design pattern,

developing a new UI widget that accommodates gaze variances is necessary. A solution was proposed to use the Hot-Zone gaze control. Four Hot-Zone operation schemes were designed, implemented, and tested by comparison with the conventional mouse. The Hot-Zone implements an innovative control mechanism called "local calibration", which improves the reliability of gaze control when calibration is poor. A single hotkey is used to achieve multiple goals, including eliminating the "Midas touch" problem, accomplishing the local calibration, and selecting items from the Hot-Zone. The findings of this study shed light on the direction of improving eye tracking techniques in practical applications.

4.2 Introduction

4.2.1 Hot-Zone overview

Conceptually, Hot-Zone acts like the dropdown menu with a different look (Figure 4-1). Users first locate the object to work on, then call out Hot-Zone, look for the right tool, and finally select the tool from Hot-Zone. As shown in Figure 4-1, available tools are located in circular zones (in the form of text for simplicity). In the current configuration, there are five zones in one batch. Users select the arrow on the right side to see more zones available. The zone currently being looked at will be highlighted.

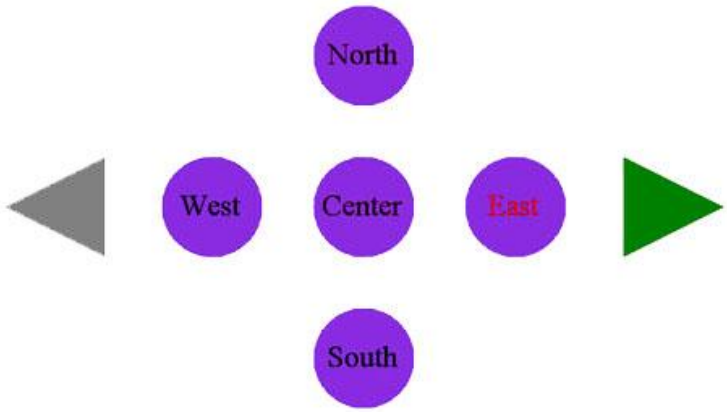


Figure 4-1 Hot-Zone appearance

To eliminate the "Midas touch" problem, a hotkey must be pressed to call out the Hot-Zone. This was a considered to be a reasonable design because the keyboard is often used for many applications; whereas, frequent switches between keyboard and mouse (or other pointing devices such as the touch pad) take time and cause fatigue. By replacing pointing devices with gaze-based interaction, a user's hands will remain on the keyboard without having to switch all the time. This makes the interaction process smoother and more efficient. As to the resolution issue, we designed the work-flow of the Hot-Zone in such a way that a "local calibration" is performed every time the user uses the Hot-Zone.

The Hot-Zone supports two gaze states, fixation and blink, as signals to call out or select commands from the Hot-Zone. When a fixation is used (either for calling out or selecting commands), a hotkey is required to eliminate unwanted actions. On the other hand, blinking can be intentional and causes much less ambiguity in determining a user's intention. Therefore, when blinking is set as the signal, a hotkey is only required for calling out the Hot-Zone, and not for selecting commands. Once a tool is selected, Hot-Zone will be closed. To close it without selecting any tool, the user may simply look outside the Hot-Zone. The

different manipulating schemes for the Hot-Zone are shown in Table 4-1.

Table 4-1 Manipulation schemes for manual and gaze control

<table>
<tr><th>UI widget</th><th colspan="2">Call out</th><th colspan="2">Select tool</th><th>Close without selection</th></tr>
<tr><td>Context menu</td><td colspan="2">Right click</td><td colspan="2">Left click</td><td>Click outside</td></tr>
<tr><td rowspan="3">Hot-Zone</td><td>Gaze state</td><td>Hotkey</td><td>Gaze state</td><td>Hotkey</td><td rowspan="3">Look outside</td></tr>
<tr><td>Fixation</td><td>Yes</td><td>Fixation</td><td>Yes</td></tr>
<tr><td>Blink</td><td>Yes</td><td>Blink</td><td>No</td></tr>
</table>

4.2.2 Local calibration

Ideally, the Hot-Zone is to be centered at the user's current gaze position when called out. However, because eye tracking is not very accurate, it is difficult to precisely map the gaze position back on the screen. Figure 4-2 illustrates this problem. The red dot is the center of the Hot-Zone and the blue dot represents the true gaze position. Instead of overlapping with each other, there is usually a distance between the two. In this study, it was found that the reported gaze position from the eye tracker was at least ¾ inch off the true value even when the calibration was done fairly well. This offset is non-uniformly distributed over the screen and changes in size and direction as time progresses. It typically gets worse because of the "shift" problem commonly encountered in eye trackers. This poses a major difficulty for reliable interaction with the Hot-Zone.

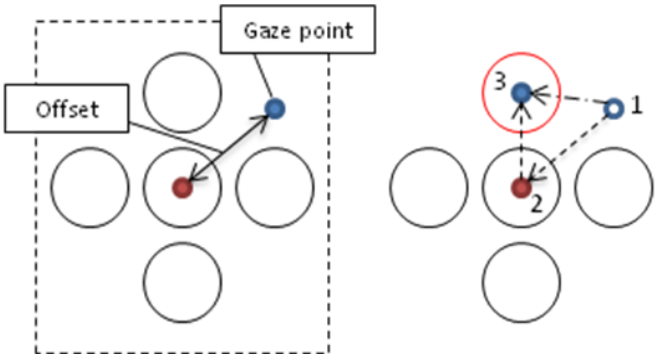


Figure 4-2 Offset (left) and local calibration (right)

To mitigate this problem, an extra step called “local calibration” was designed into the manipulation sequence. As shown in Figure 4-2 (right), the red circle represents the tool the user should select. Instead of directly looking for it (gaze path indicated by dot-dash arrow from Position 1 to 3) from the current gaze position, the user should make a stop at the Hot-Zone center (red solid dot) before moving on to look for the final destination (gaze path indicated by dash-arrow, from Point 1 bypassing Point 2 and finally reaching Point 3). When a user is looking at the Hot-Zone center, the offset is calculated and saved. Gaze coordinates from the eye tracker are later adjusted by this amount of offset to obtain a better estimation of the true gaze position. By this means, the Hot-Zone works well even when the calibration of the eye tracker is poor. Audio feedback is provided to confirm when a tool is successfully selected using the Hot-Zone.

The question remains as to how to inform the computer when a user is looking at the center of the Hot-Zone? This is done by simply releasing the hotkey. Instead of clicking the hotkey to call out the Hot-Zone, the user should hold down the hotkey and look at the Hot-Zone center first. The user then releases the hotkey while looking at Position 2 (the Hot-Zone center) to inform the computer to compensate for the offset. Proper use of the

hotkey is the key to successfully manipulating Hot-Zone. It requires practice for people to coordinate their eyes with their hands. If the Hot-Zone is called out while the gaze position is close to the boundary of the screen, the Hot-Zone will automatically offset towards the screen center to make sure it can be fully displayed. Such a design practice is used to guarantee the user receives feedback when interacting with the Hot-Zone.

4.2.3 Hot-Zone component

The Hot-Zone is organized into a .NET package which consists of three components (indicated by the dashed line in Figure 4-3).

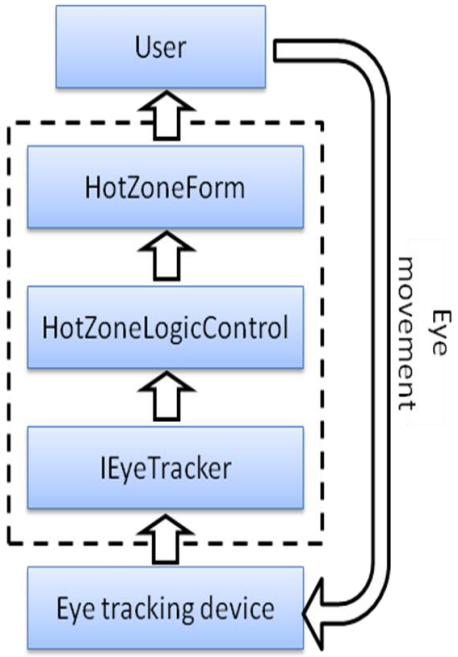


Figure 4-3 Hot-Zone workflow.

HotZoneLogicControl

The HotZoneLogicControl is the core component of the Hot-Zone (Figure 4-4). It is responsible for collecting necessary information and choosing appropriate actions accordingly. Information needed includes gaze state (saccade, blink, or fixation), hotkey state (pressed or not), and Hot-Zone state (whether it's shown on the screen already), etc. These tasks are performed sequentially by calling out different methods in the HotZoneLogicControl.

The data batch size used in this study was 0.05 seconds, or 6 data entries at a 120 Hz sampling frequency. Each data entry has three values: pupil size, and x and y coordinates (in pixels) of gaze point on the screen. Once a batch of data is collected, a simple algorithm is used to determine the gaze state. It first checks the pupil size to see if the user is blinking. If the pupil size is zero, it sets the gaze state to blink. Otherwise, it calculates variances and mean values of the gaze point coordinates. If the variance exceeds a threshold, the gaze state is set to a saccade. If not, the gaze state is set to fixation and the average is saved as the current gaze position.

The hotkey was set as the "Ctrl" (Control) key in this study. Users can use "Alt" or "Shift" as the hotkey by setting the corresponding property of the HotZoneLogicControl. If the Hot-Zone is already shown on the screen, the HotZoneLogicControl goes on and checks which zone or arrow is being looked at. After all information is collected, it is then coded into a state vector that is sent into a finite state machine to determine the proper actions, such as show the Hot-Zone.

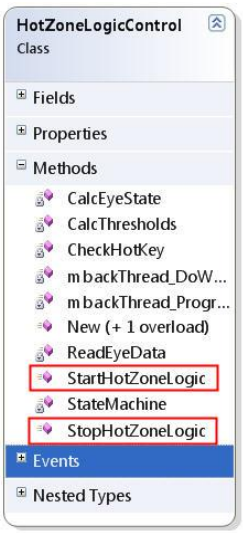


Figure 4-4 Methods of the HotZoneLogicControl.

All the operations above are hidden from users and run in a background thread. Only two methods: StartHotZoneLogic and StopHotZoneLogic are exposed to users to simplify its use. When a user selects a tool from the Hot-Zone, the HotZoneLogicControl generates a message (or raises an event in .NET's language) to the main thread. Users write code to handle this event in the same way as handling mouse events.

IEyeTracker Interface

It is desirable for the Hot-Zone to be hardware independent, which means that eye tracking devices from different manufactures should be able to work with the Hot-Zone properly. To achieve this goal, a programming interface called IEyeTracker was designed (Figure 4-5), which acts like a protocol between eye tracking devices and the HotZoneLogicControl. Methods in the IEyeTracker interface must be implemented with a customized assembly, which is hardware-dependent. The HotZoneLogicControl retrieves eye

movement data via the IEyeTracker instead of directly communicating with the eye tracking devices. Gaze positions are fed into the HotZoneLogicControl as screen coordinates measured in pixels. Proper conversion and mapping should be done by the customized assembly, which implements the IEyeTracker Interface.

Figure 4-5 IEyeTracker interface.

In this study, a function class - ASLEyeTracker – was developed with ASL's SDK (Figure 4-6), which implements the IEyeTracker interface. The gaze position measured by the ASL eye tracker is in inches from a pre-defined calibration plane. It requires two parameters to map the coordinates on the screen: the screen resolution, and the difference between the origins of the two coordinate systems. Eye trackers from different manufactures could work with the Hot-Zone by developing a similar function class, which implements the IEyeTracker interface. The HotZoneLogicControl exposes a property - EyeTrackerDevice - to allow users to apply their customized class and stream eye movement data.

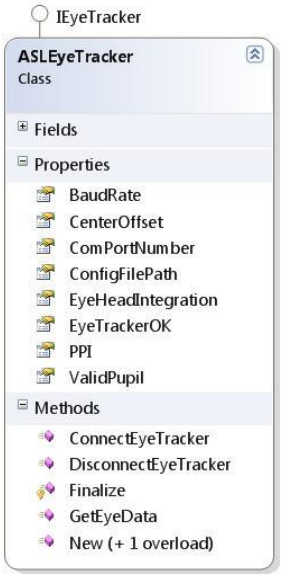


Figure 4-6 ASLEyeTracker class.

Hot-Zone Form

The HotZoneForm is essentially an irregular-shaped window which users directly interact with. Its behavior during interaction is completely controlled by the HotZoneLogicControl and is not exposed to users. However, users can change its appearance, such as the background, color, and overall size, via the HotZoneLogicControl. Figure 4-7 shows a Hot-Zone with 6 peripheral zones instead of four, as seen in Figure 4-1. Though more zones means more tools available in one batch, this configuration requires a higher resolution to achieve a reliable interaction. Changing the number of zones also changes the directions of the peripheral zones since they are evenly distributed around the central zone. Previous studies have shown that the direction of saccade has an impact on eye movement end point variances. In other words, direction will affect how large zones should be to accommodate gaze variance. For simplicity, the number of zones was fixed in this study. The four peripheral

zones are in horizontal and vertical directions, which correspond to the main rotation axes of the eye balls.

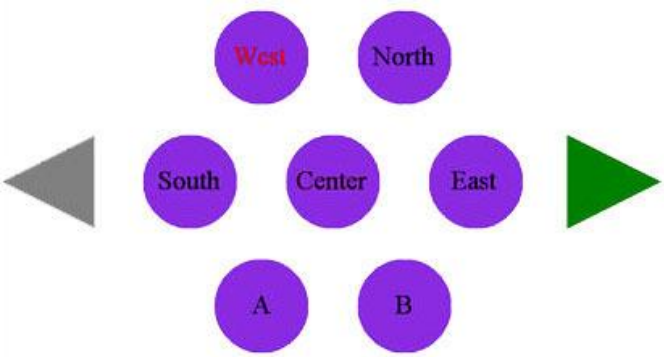


Figure 4-7 The Hot-Zone with 6 peripheral zones.

4.3 Methods

4.3.1 Task and equipment

All four schemes for manipulating the Hot-Zone were tested along with manual interaction. A standard mouse was used for manual trials and a standard keyboard was used to accompany the Hot-Zone trials. In this study, we focused on the differences caused by: (1) different gaze states used for calling out and selection of the Hot-Zone; and (2) the manual and four gaze-based interaction schemes. Design parameters for the Hot-Zone (e.g., number of zones, batch size of the eye tracker data, and the threshold for saccades) were held constant in the experiment. For practical applications, these features can be customized when needed.

The task interface was similar to that used in the pilot study. Four squares resented on a display screen, represented the objects to be worked on. Five items represented available

tools, which include Center, East, North, West, and South. The computer randomly picked one of the items and displayed it on one of the squares, which was also randomly selected.

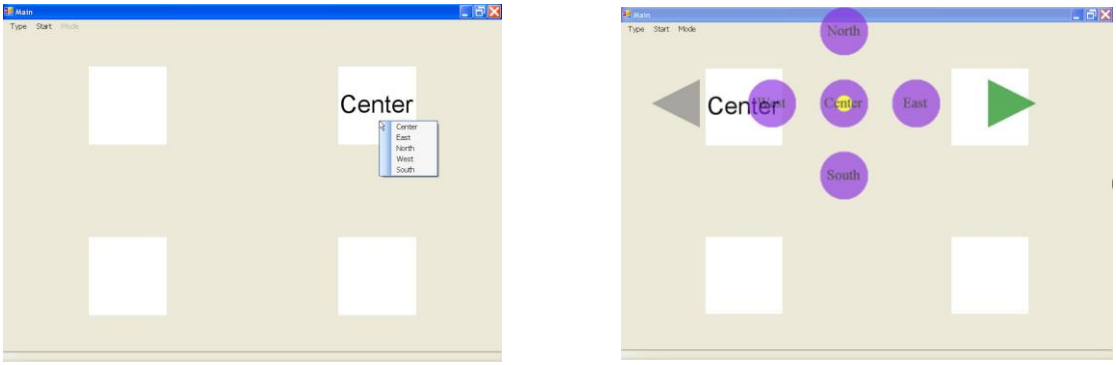


Figure 4-8 Experiment task: Select the item.

In manual interaction trials, these items were presented in a dropdown menu. Subjects worked with the dropdown menu to select the item displayed (Figure 4-8, left). In gaze interaction trials, these items were located in the circular zones of the Hot-Zone; subjects select the correct item by interacting with the Hot-Zone (Figure 4-8, right). The number of tools was set equal to the number of zones provided by the Hot-Zone on one page to simplify the test scenario. Thus, subjects did not need to use the arrow controls to perform the "turn page" operation.

Each subject accomplished the task with all five operation schemes. In manual interaction, subjects were given an 8-trial practice before test trials since the task was fairly simple with the mouse. With the Hot-Zone, 120 training trials were provided before the first testing scheme. Fewer training trials were provided for the following Hot-Zone operation schemes to prevent subjects from feeling bored or fatigued. Under each technique (one manual and four gaze interactions), four 16-trial sessions were performed which resulted in a

total 64 trials for each scheme. Like the pilot study, an ASL H6 Head Mounted Optics with a head motion tracker was used for eye movement measurement. The task was presented on a 19 inch CRT color monitor at a resolution of 1024 by 768 pixels.

4.3.2 Participants

Twelve participants were recruited for the experiment, most of whom were students at North Carolina State University. For the eye tracker to work properly, only subjects without glasses or contact lenses were recruited. It was also required that subjects not have a history of neck injury since the head gear could cause neck fatigue after a long period of time.

4.3.3 Experimental design

A cross-over repeated measures design is used for the experiment. All independent variables are within-subject effects. The dependent variables were the completion time and the correct selection rate. First, any differences caused by gaze interaction schemes were compared. Two gaze states were chosen as the signal to call out or select tools from the Hot-Zone. The resulting four gaze interaction schemes are coded as “Gaze 1” to “Gaze 4” as indicated in Table 4-2. The most efficient gaze interaction scheme was selected that had the fastest completion time and high correct selection rate. Second, we tested the difference between the selected gaze-based and manual interaction schemes. The independent variables therefore were interaction scheme with two levels and subject with 12 levels. The gaze interaction schemes were not randomized during the experiment in order to keep the learning process smooth across subjects.

Table 4-2. Experiment setup

Scheme code	Call out	Select tool
Manual	Right click	Left click
Gaze 1	Fixation	Fixation
Gaze 2	Fixation	Blink
Gaze 3	Blink	Fixation
Gaze 4	Blink	Blink

Table 4-3 Experiment design

Interaction Scheme	Subject	Completion time / Correct rate			
		Session 1	Session 2	Session 3	Session 4
Manual	1-12				
Gaze 1	1-12				
Gaze 2	1-12				
Gaze 3	1-12				
Gaze 4	1-12				

All statistical analysis was performed using SAS, including analyses of variance (MANOVA) with repeated measures applied to the dependent variables. The model for the experiment can be written as follows:

$$y_{ij} = \mu + \alpha_i + \pi_j + \varepsilon_{ij} \qquad (4\text{-}1)$$

where:

y = the response variable (i.e., completion time or correct selection rate)

u = the overall mean response

α_i = the effect of the i^{th} interaction scheme

π_j = the effect of the j^{th} subject

ε =experimental error

Finally, the relationship between completion time (T_i) and the rate of correct selection

(R_i) within the gaze interaction schemes was investigated. The purpose of this test was to see whether increased completion time resulted in higher correct selection.

4.4 Results

The Box plots revealed (Figure 4-9 and Figure 4-10) that the completion time under manual control was significantly smaller than all other gaze-based control methods with small variances. The correct selection rate for manual control was higher with smaller variances compared to gaze-based control. The mean values and standard deviations are listed in Table 4-4.

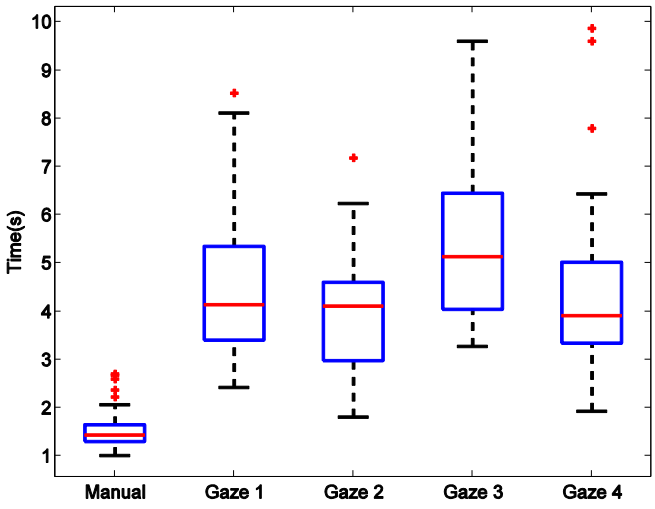


Figure 4-9 Box plot of Completion time

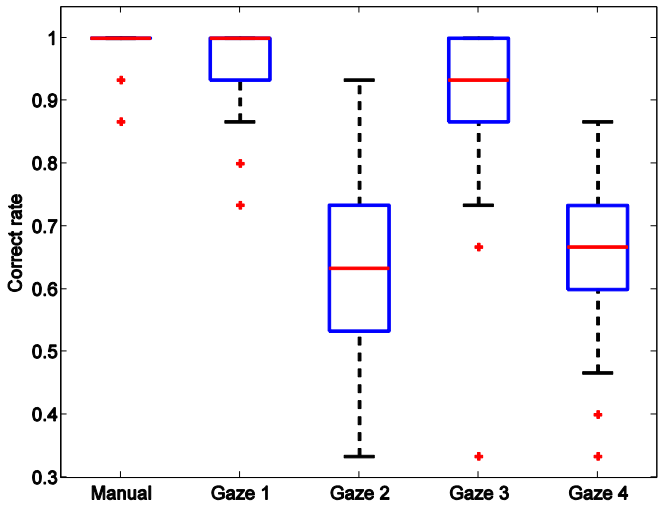


Figure 4-10 Box plot of Correct rate

Table 4-4 Descriptive statistics

Interaction Scheme	Mean completion time (ms)	SDT of completion time	Mean correct rate	SDT of correct rate
Manual	1568	407	99%	3%
Gaze 1	4408	1460	95%	8%
Gaze 2	3898	1144	63%	13%
Gaze 3	5439	1665	92%	12%
Gaze 4	4275	1631	65%	13%

It is suspected that the large variances in the response were due to a flaw in the Hot-Zone control logic. Because the duration of fixation is much longer than the sampling rate of eye tracker device, the Hot-Zone is designed in such a way that it only updates when it detects changes in the gaze state. In other words, only when the transition between gaze states (fixation to saccade or vice visa) is captured will the Hot-Zone update its current state.

Since the HotZoneLogicControl runs as a background thread, it is not guaranteed that

every gaze state change will be captured. This problem is exacerbated by the fact that saccades generally have a very short duration - about 200 ms. If Hot-Zone missed the gaze state transition (e.g., when subjects switch their gaze and look at a different zone), it may fail to highlight the item in that zone. When this happens, subjects are inclined to stare at the target tool and wait for the Hot-Zone to follow up. Quite contrary to what they expect, this would never cause the Hot-Zone to update. The correct manipulation under such a situation is to look around and go back to the target tool again. However, even though subjects were told to follow such a practice, they forgot to do so from time to time. When they failed to select the tool on the first try, it typically took a much longer time to recover and redo the process to acquire the right tool.

It is also clear from the Box-plot that when blink was chosen as the gaze command for selecting tools, the correct selection rate dropped significantly. This was because when blinking, eye movements increased dramatically. Since the gaze position was simply the average of the 0.05 second eye movements, one or two "jumping" points would be enough to create a false selection. This problem could be avoided by improving the Hot-Zone implementation to rule out those outliers when calculating gaze position. On the other hand, blink-to-select resulted in a shorter completion time and subjects also reported that they felt it was easier and more comfortable to select by blinking.

Based on such observations, gaze-based control method Gaze 1 was chosen to be compared with the manual control. The univariate tests revealed that only the scheme factor has significant influences on the completion time (Table 4-5 and Table 4-6). Table 4-7 shows that the overall model is not significant for the correct selection rate ($p=0.292$). However, the scheme factor is still significant at level of 0.05 (Table 4-8). We postulate that this paradox is due to the fact that subject factor has a much stronger impact on the overall model than the

scheme factor. Because the subject factor is clearly not significant, the model shows no differences between the two groups. However, the scheme factor is still significant at level of 0.05, which suggests that the correct selection rate for the two interaction groups is also different. For more strict analysis, both completion time and correct selection rate need to be transformed, as Figure 4-9 and Figure 4-10 show that the variance is dependent on the mean.

Table 4-5 Univariate analysis for completion time

Source	Degree of freedom	Sum of squares	Mean square	F value	Pr > F
Model	12	58.185	4.849	6.55	0.002
Error	11	8.149	0.741		
Corrected Total	23	66.334			

Table 4-6 Variance structure estimates on completion time

Source	Degree of freedom	Sum of squares	Mean Square	F value	Pr > F
Subject	11	8.367	0.761	1.03	0.483
Scheme	1	49.820	49.820	67.25	<.0001

Table 4-7 Univariate analysis for correct selection rate

Source	Degree of freedom	Sum of squares	Mean square	F value	Pr > F
Model	12	0.202	0.002	1.40	0.292
Error	11	0.014	0.001		
Corrected Total	23	0.036			

Table 4-8 Variance structure estimates on correct selection rate

Source	Degree of freedom	Sum of squares	Mean square	F value	Pr > F
Subject	11	0.101	0.001	0.75	0.681
Scheme	1	0.011	0.011	8.59	0.014

It can be noted from the results that trials with longer completion time sometimes resulted in wrong selections. Due to the implementation flaws mentioned above, re-initiating Hot-Zone for selection was not only time consuming, but also frustrating for subjects. In such circumstances, subjects selected randomly, trying to finish the trial as soon as possible. The correct selection rate was plotted against 10%, 25%, 50%, 75% and 100% completion time to show this point (see Figure 4-11).

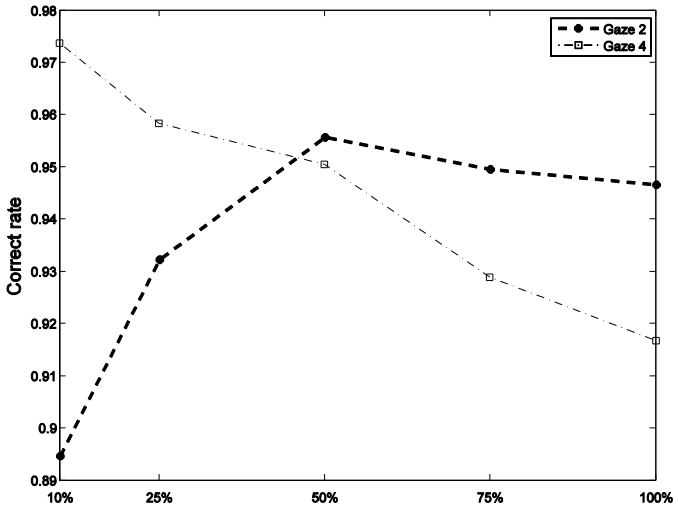


Figure 4-11 Correct selection rate vs. completion time percentile

4.5 Discussion

Although the current Hot-Zone manipulation was slower than manual control, this study provides insight for future research, particularly when the benefit of accessing dropdown menus with gaze is weighted more than the delay (about 3 seconds). The study also identified both objective and subjective obstacles causing longer completion times and larger error rates of with the Hot-Zone manipulation.

At this stage, the Hot-Zone is a software-based solution. The HotZoneLogicControl runs as a background thread sharing the CPU with other processes running on the computer. There is no guarantee that the Hot-Zone will be fully responsive all the time. This problem manifests when the Hot-Zone fails to update and provide feedback to users or fails to select a tool when a user presses a hotkey. If users do not succeed in selecting tool on the first attempt, they typically wait for a while before initiating a second attempt.

Although debugging was undertaken, there are still limitations in the code. For instance, it was observed that when a user tried to select a tool by blinking, although the Hot-Zone detected the blink and correctly de-highlighted the text, it failed to close itself. Instead, Hot-Zone remained on the screen and highlighted again once when the subject opened his/her eyes. Because of these flaws, the subjects did not trust in the Hot-Zone as much as they did the mouse. They were more prudent when working with Hot-Zone, tending to check each step of the operation, which made the whole work-flow much less smooth and more time consuming.

These technical issues had a major negative impact on subject performance, but they could be eliminated by integrating the HotZoneLogicControl with the eye tracker and making it a hardware-based solution. Using a mouse does not require a computer to map hand movement to the screen, nor to monitor whether buttons are clicked. All this work is done by

the mouse itself; the computer just handles the event raised by the mouse. Likewise, if we let the eye tracker does all the background work currently done by the computer, the Hot-Zone could be made more responsive to the user's actions.

Furthermore, training trials provided for gaze control interaction were no comparison with years of experience using a mouse. It is not enough for novices to become accustomed to the Hot-Zone's manipulation since it is quite different from what they are used to. Specifically, "local calibration" takes time to memorize and work with efficiently. It was observed that subjects unconsciously looked for the desired tool directly instead of aiming at the center first when Hot-Zone showed up, even when the training was conducted. It takes practice to overcome previous operation patterns of behavior before subjects become comfortable with the Hot-Zone. The four types of Hot-Zone schemes also caused confusion in some trials as subjects forgot whether to blink or fixate. With a properly designed training procedure, this issue could be alleviated as people pick-up the new operation "rhythm" with their gaze. In fact, experienced users such as researchers are quite comfortable using the Hot-Zone.

Despite these limitations, this study proposed a new direction for utilizing gaze-based interaction. The Hot-Zone is of practical value in HCI, especially for heavily graphic-dependent applications such as PowerPoint. For such applications, the interface objects used are typically large enough to be selected directly by the eyes. Even if direct object acquisition is not satisfying, object could simply be put names in Hot-Zone for users to select from given that the number of objects is small. The Hot-Zone also provides the means of selecting tools to be applied to accomplish the whole interaction process in a unified manner.

The Hot-Zone requires a single hotkey to perform the local calibration and tool selection compared to multiple hotkeys needed to mimic the mouse operation in previous

studies (Manu, Andreas, & Terry, 2007). The workload for manipulating the Hot-Zone is much lower once a user gets the timing right. The Hot-Zone also has the merit of reducing fatigue and potential injury that extended use of a mouse can cause. If it is set to blink-to-select mode, it could even prevent eye strain as blinking moistens the eye balls.

Commercially available eye tracking systems are quite expensive compared to a mouse. However, there have been studies focused on developing cheap eye tracking systems with commonly used web cameras to make the technology affordable to a larger user population (Dan Witzner, David, John Paulin, & Mads, 2004; Li & Parkhurst, 2006). There is also a trend to integrate such systems into computers so users do not need to wear any gear. With enhanced image processing algorithms, video-based eye tracking is expected to be more reliable and convenient to use in the future.

Finally, the Hot-Zone as a prototype is open to modification regarding its appearance, such as the overall size, zone size, and number of zones, etc. Subjects provided feedback regarding how to make the Hot-Zone more user friendly. One subject suggested that the Hot-Zone's location should be anchored, so that the user knows where the center of Hot-Zone should be when they need to do "local calibration" instead of looking for it every time. All these considerations need to be tested and evaluated with formal experimentation. Biomechanical models of ocular dynamic systems may also help in determining design parameters. A better understanding of gaze pattern could also help in developing better operation schemes, which are suitable for gaze control.

5 Differentiating reading and searching gaze patterns

5.1 Research objective

With eye tracking techniques, it is possible to use passive eye movements for facilitating human computer interaction. Many studies have shown that eye movements follow different patterns in specific tasks, e.g., reading and visual search. However, few studies have attempted to take advantage of such differences to improve gaze-based interaction. Discriminating the tasks users perform can enable the computer to provide the right tools for the right needs at the right time.

This chapter is aimed at addressing the gaze pattern recognition problem to discriminate reading and visual search based on user's eye movements. An eigenvalue-based feature extracted from a saccade series was proposed. The performance of this feature for gaze pattern recognition was evaluated with a linear support vector machine (SVM), which resulted in an average recognition rate of 85%. It was demonstrated that saccades, which are usually overlooked in interaction, convey important messages regarding a user's attention. Two issues unique to gaze pattern recognition are explored: (1) the saccade block size (e.g., the chunk length of the saccade series), which plays an important role in the speed-accuracy tradeoff of the discrimination; and (2) recognition rates for reflecting the reading fluency. For example, the recognition rate for native English speakers is significantly higher than for foreign language speakers.

5.2 Feature selection

Given the gaze position reported from eye trackers, there are several measurements available for feature extraction. In this study, velocity, fixation duration, and saccade were considered. These have been the main dependent variables in eye movement studies. It was demonstrated that saccade sequences represent key characteristics of reading and visual search which can be used for gaze pattern recognition.

Velocity

For simplicity, the velocity v_i is calculated as the moving difference between two consecutive gaze positions ($v_i = P_{i+1} - P_i$) at a 120 Hz sampling frequency. It serves as the determining index for separating fixations and saccades as well. Considering the impact of smaller eye movements such as tremors, drifts and flicks (Ditchburn, 1973), the acceleration profile is much noisier and does not provide more information. Thus, acceleration and higher derivatives are not considered in this study.

Fixation duration

To determine the fixation duration, the start and end time of a fixation must be identified first. Since fixation by definition is the interval during which eyes are relatively still, a velocity threshold was used for fixation identification (Figure 5-1).

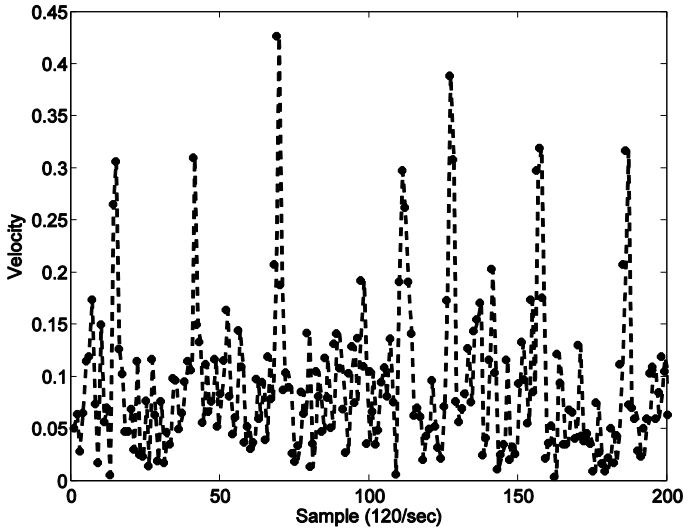


Figure 5-1 Velocity profile

The threshold was calculated based on eye movement data recorded when subjects fixated on the screen during calibration:

$$v_{threshold} = \left| v_{fixation} \right| + 3\sigma_{fixation} \tag{5-1}$$

Then, the eye movement state S_i is determined as follows:

$$S_i = \begin{cases} -1, & \left| v_i \right| \le v_{threshold} \\ 1, & \left| v_i \right| > v_{threshold} \end{cases}$$

This velocity threshold method serves the purpose of state identification well in the present study. There are other advanced methods for fixation identification, such as the use of HMMs or deploying a dispersion threshold. Although more accurate, these methods require more parameters to specify, which have large variances between subjects or require subjective judgments (Salvucci & Goldberg, 2000). Thus, these methods were not considered in this study.

It is also possible to extract information about what a user is looking at on the screen (e.g., a text range or a picture on a slide) from the computer system. This study did not leverage such information because the objective was to design a context-independent pattern recognition scheme. Therefore, the only characteristic that defines a fixation is its duration, which ranges from 100 ms to over 1 second depending on the assigned tasks.

Saccade

Saccades, in switching gaze focus, provide both magnitude and directional information of eye movement. Because saccades only depend on the relative differences between two fixations, it is a reliable indicator of focus changes even when gaze positions reported from eye trackers are shifted. Such "shift" problems usually cause difficulties in utilizing fixation for interaction. It is conceivable that no information can be encoded during saccades. However, saccade sequences say a lot about what users are doing.

With the eye state S_i from fixation identification, the first gaze point ($i \in \{i \mid \Delta S_i = -2\}$) of a fixation is identified as the switch point:

$$\Delta S_i = S_i - S_{i-1} = \begin{cases} -2 \\ 2 \; or \; 0 \end{cases}$$

Then, the saccade magnitude c_j is calculated as the difference between two consecutive switch points:

$$c_j = \Delta P_i = [\Delta x_i, \Delta y_i]^T, \; i \in \{i \mid \Delta S_i = -2\}, \; j = 1, 2 \ldots$$

This method is simple to implement and works fine even when fixation identification is not accurate. By setting a distance threshold to filter out false or small saccades, which have less weight in later recognition stages, the negative impact of false fixation identification can be further reduced.

5.3 Methods

The same eye tracking devices, as used in the pilot study were used here. The same 12 subjects, who participated in the Hot-Zone test, also performed the tasks in this study. The tasks were the same as in the pilot study (Figure 3-10). Subjects were required to perform a reading task and a searching task on a 19 inch CRT color monitor at a resolution of 1024 by 768 pixels. For the reading task, an article consisting of 16 pages, with 106 words average on each page, was presented. Subjects read through each page and clicked the "next" button to see the next page. For the visual search task, a randomly generated image with about 80 letter "L"s rotated at arbitrary angles, was presented to the subjects. Among the letter "L"s, an up-right letter "T" was presented as the target for subject to locate.

In both tasks, there was no time limitation or other constraints. Only eye movements were recorded for analysis. The images with letters "L" and "T" were randomly generated during formal experiments, whereas, the images in the pilot study were pre-constructed in other software.

5.4 Results

Velocity

To mitigate data distortion, six gaze point samples, or 50 ms of eye movement data at a frequency of 120 Hz before and after blinks, were discarded. Because fixation duration (about 200 to 400 ms) is about 10 times longer than saccade duration (about 30 to 50 ms), and eye movement velocity during fixation is small, the distribution of velocity in both tasks overlapped within a small area close to the origin (Figure 5-2). The confidence ellipses of velocity have similar shapes in both tasks, which are difficult to separate. These findings

suggest that velocity data is redundant and inefficient for recognition.

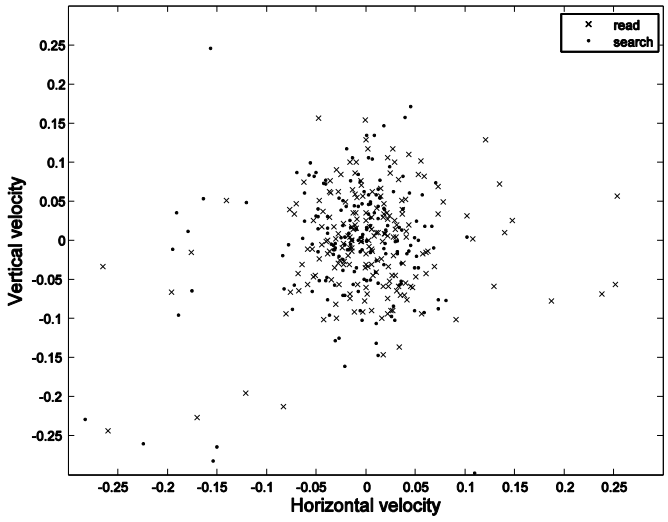


Figure 5-2 Velocity for reading and visual search

Fixation duration

The fixation duration measured in this study agreed with the results from other studies, most of which are within the range of 200 to 400 ms (Figure 5-3). The data suggests that the velocity threshold fixation method works properly for identifying fixation. An unbalanced ANOVA test ($F=46.76$, $p<0.0001$) revealed that fixation duration was significantly different in the two tasks, which has also been demonstrated by other studies (Rayner, 1998). Fixation duration shows an exponential distribution for reading and a skewed normal distribution for visual search.

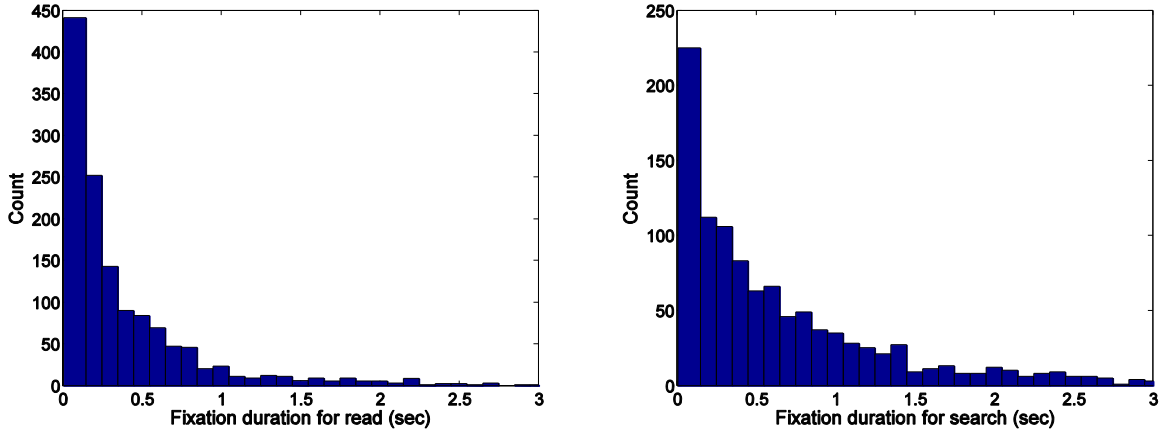


Figure 5-3 Duration histogram of reading and searching

While the difference in distributions may be feasible for gaze recognition, such a method requires substantial eye movement data given the large variance in fixation duration. This requirement will cause long system delays for real-time interactions. Furthermore, the accuracy of the measured duration is limited by the sampling rate of the eye tracker. Higher sampling rates require more expensive eye trackers and produce noisy eye movement data, which is error prone in identifying fixations.

Saccade

The saccade sequences extracted from the two tasks demonstrated significant differences in both direction and magnitude. For the reading task, most saccades were horizontally oriented and grouped into two clusters (Figure 5-4 left). The one on the right of Figure 5-4 represents the saccades between words within the same line, which have a smaller magnitude (less than 2 inches). The one on the left represents saccades that occur when a subject switches to the next line, with a magnitude range from 2 to 4 inches.

It should be noted that the left dot cluster is slightly above instead of below the right

cluster. This is caused by the coordinate system used by the ASL eye tracker, which uses a vertical axis pointing down. Such arrangements have no effect on the recognition results, thus the arrangement was left as is to simplify calculations.

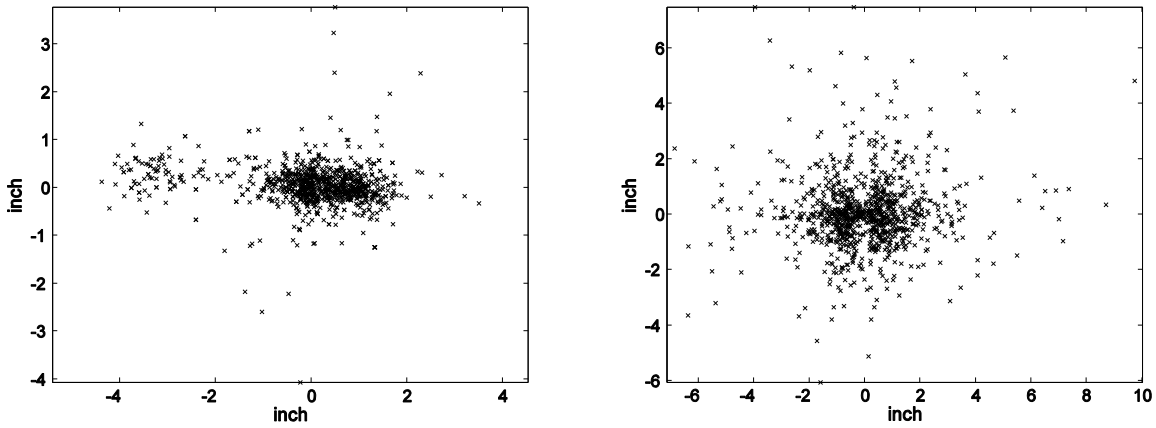


Figure 5-4 Saccades for read and search

Saccades for visual search are more or less scattered symmetrically pointing in various directions (Figure 5-4, right). The area and flatness of the confidence ellipses are significantly different from that of the reading saccades. Such differences revealed the eigenvalues of the saccades covariance matrix as a useful feature for recognition. The feature extraction is calculated as:

$$\begin{aligned} M_k &= \left[c_{nk+1}, c_{nk+2}, \cdots, c_{nk+n}\right]^T \\ \Phi_k &= \mathrm{cov}(M_k^T M_k) \\ A_k \Phi_k &= \lambda_k \Phi_k, \quad \lambda_k = \left[\lambda_{k1}, \lambda_{k2}\right]^T \end{aligned} \tag{5-2}$$

Here, n (the block size) is the number of saccades used for calculating the covariance matrix M_k. A_k and λ_k are eigenvector and eigenvalue vectors. The block size is a very important parameter in extracting the eigenvalue feature. With a larger n, the covariance

matrices for reading and visual search show larger differences from each other. However, a large block size requires a longer time for collecting sufficient saccades. This speed-accuracy tradeoff for recognition rate is a unique characteristic of gaze pattern recognition compared to other modality recognition applications.

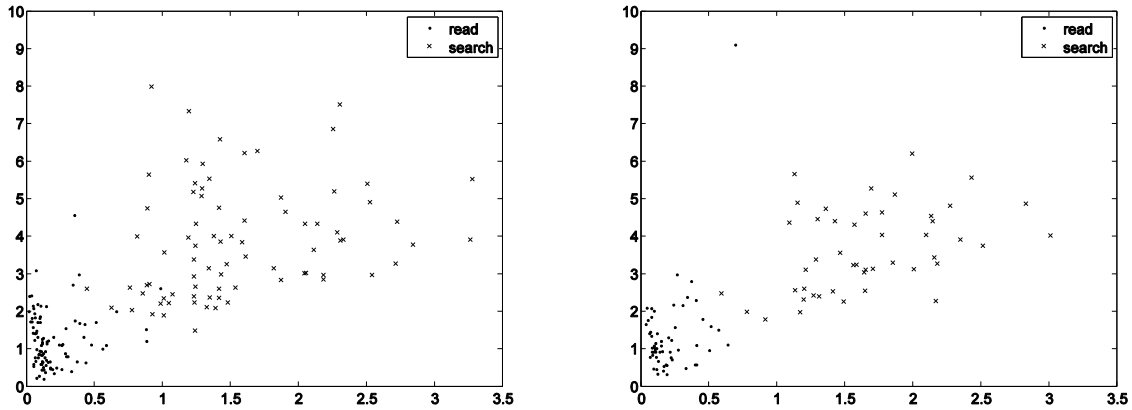


Figure 5-5 Feature extraction (left: n =30, right: n =50)

With eigenvalues extracted from saccade sequences, the read and visual search patterns are projected into the feature space, which intuitively shows the differences of the two. In Figure 5-5, dots represent eigenvalue pairs extracted from reading saccade blocks, while the cross marks stand for search saccade blocks. In the feature space, it is clearly shown that reading saccades aggregated close to the origin in a linear fashion, since they are mostly oriented in a horizontal direction with a small vertical magnitude. However, the visual search saccades are farther away from the origin with various slopes.

Because the total number of saccades from the eye movement data is fixed in this study, the larger n is, the fewer blocks available for training and testing (e.g., fewer dots and cross marks in Figure 5-5(left) than in Figure 5-5(right) as n increases from 30 to 50). With a small block size there are more overlaps between the two patterns; increasing the block size

effectively separates the two farther away from each other.

Gaze Pattern Recognition with Support Vector Machine

For this binary recognition problem, the Support Vector Machine (SVM) with a linear kernel function as the classifier was used. Training and testing were carried out with the Matlab bioinformatics toolbox. Figure 5-6 shows the recognition result of one subject's eye movement data with n=50.

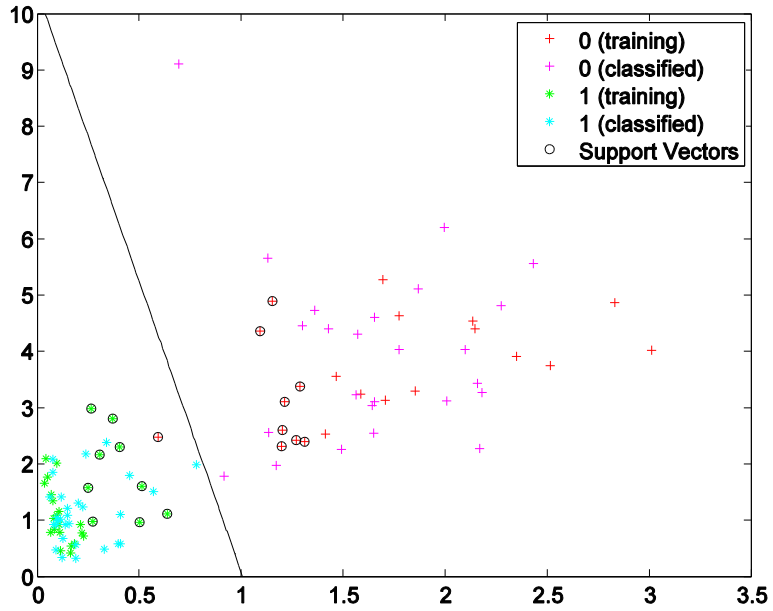


Figure 5-6 Recognition results with Support Vector Machine

Based on interest in investigating the effect of block size on recognition rate, the saccade block size was increased from 5 to 50 in increments of 5 saccades. For each subject, the evaluation was iterated 12 times. Half of the saccade blocks were randomly selected to hold out for testing in each iteration. For each block size, the mean correct selection rate

across the 12 subjects is reported for comparison. As shown in Figure 5-7, the mean correct selection rate for recognition increases as the block size increases (Table 5-1 presents the summary of the recognition rates).

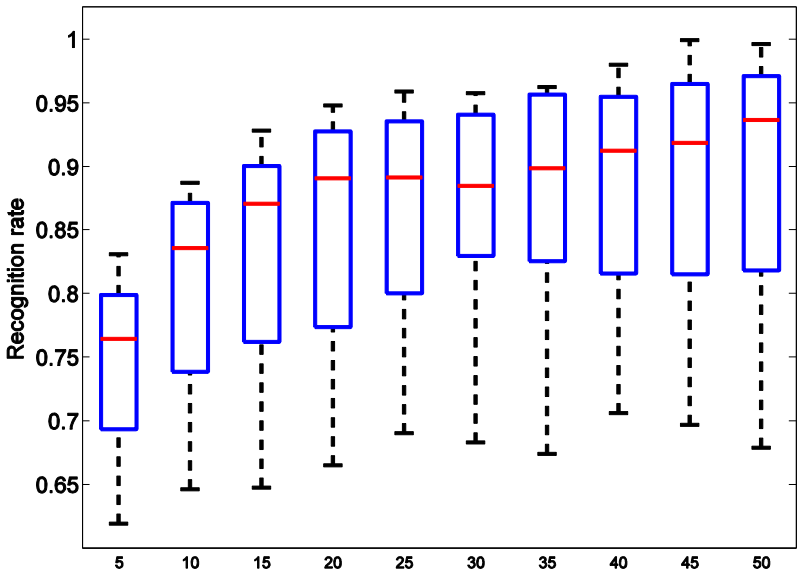


Figure 5-7 Effect of batch size on recognition rate

However, comparatively less improvement was obtained when the block size was over 20. These results indicate that in general, the eigenvalue feature extracted from saccade sequences is appropriate for gaze recognition. The optimal length of the saccade sequence is about 15 to 20 saccades, which is about 40 to 50 seconds, given that the average saccade frequency is 25 per minute (Luo, Vargas-Martin, & Peli, 2008).

Table 5-1 Recognition rate statistics

Block size	Mean recognition rate	Std recognition rate
5	0.7436	0.0704
10	0.8073	0.0795
15	0.8321	0.0891
20	0.8525	0.0901
25	0.8630	0.0900
30	0.8710	0.0803
35	0.8790	0.0873
40	0.8831	0.0862
45	0.8920	0.0922
50	0.8938	0.0974

The recognition rate for subjects varied from 67% to 94% after averaging all block sizes. Two major causes for such differences between subjects were postulated. First was the relative motion between the head gear and the subjects' head. This problem is more severe for female subjects because longer hair causes the head gear to slip more frequently. Second, cognitive factors such as concentration and the fluency of reading English play a role in recognition rate. Figure 5-8 shows the eye movement of a subject who was much less organized in reading compared to others. The left-to-right pattern typically seen in reading is not clearly demonstrated in the subject's eye movements, which resulted in a poor recognition rate.

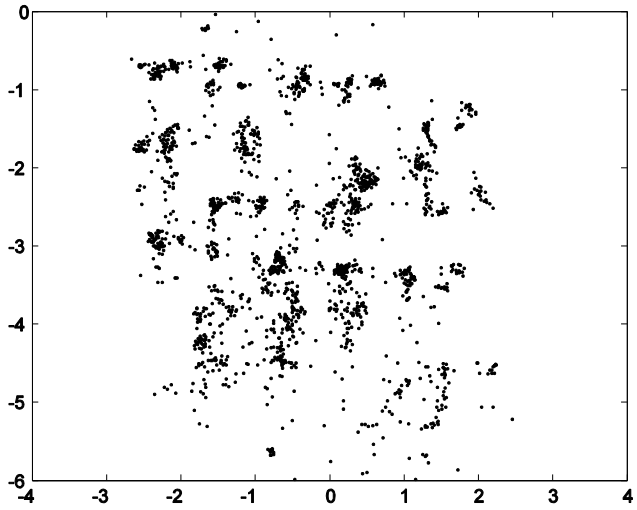


Figure 5-8 Reading pattern with low recognition rate

Although such cognitive factors were not considered in the experiment design, the potential effects with the existing data was evaluated. Among the 12 subjects, 7 were native English speakers and 5 were foreign graduate students. The recognition rate data was grouped by this categorical variable. An unbalanced ANOVA test ($F=19.49$, $p < 0.0001$) was conducted, which showed a significant difference between the two groups. This result suggests that it is plausible to assess reading ability with gaze pattern recognition.

5.5 Discussion

This study provides a first step to address this new paradigm of gaze pattern recognition with an example of discriminating reading and visual search. The results of the experiment demonstrate the discriminating power of saccade sequence, which was generally overlooked for gaze-based interaction in previous research (Kowler, 1990). By extracting the eigenvalues

of the saccade covariance matrix, the reading and searching patterns were transformed to the feature space, in which geometric differences between the two patterns were readily seen.

The overall recognition rate with SVM (V. N. Vapnik & Kotz, 2006) was above 85% on average, which confirmed that the proposed feature extraction is efficient for gaze recognition. The impact of the block size for the speed-accuracy tradeoff was also investigated, which is unique to gaze pattern recognition. A sequence length of 15 to 20 yields satisfying results, which is a critical parameter needed for applying gaze recognition in real-time.

It should be noted that the linear classifier used in this study does not consider the sequential relationship between consecutive saccades; however, internal structures do exist in some gaze patterns (e.g., the saccade in reading is usually left-to-right). This sequential characteristic may provide a way to further reduce the block size required for recognizing specific gaze patterns. However, the challenge is that not all gaze patterns have such well-defined structures, for example, saccades in searching are more or less random in either direction or magnitude. The amount of uncertainty in their structures poses difficulties in defining consistent state transitions for each pattern, which indicate that context-dependent methods such Markov Model may not yet be ready for applications.

More gaze patterns are expected to be included to make this technique useful for broader applications. Visual activities that require highly trained specialists (e.g., medical image diagnosticians (Beutel, Horii, & Kim, 2000) are of interest. Applying pattern recognition on these tasks may help to reveal superior search schemes for successful diagnosis. Furthermore, it was found that reading fluency plays a role in recognition rate, indicating the possibility of using gaze pattern recognition as a complement for assessing learning disabilities such as dyslexia. It has been know for decades that there are

developmental trends in eye movements during reading. As reading skill improves, fixation duration decreases, saccade magnitude increases, and the number of fixations decrease (McConkie et al., 1991). To date, the lack of specific identification criteria for learning disabilities remains a persistent problem in the field (Lyon, 1996). Eye movement analysis has the advantage of being a more objective and consistent method compared to existing diagnostic tools that require a qualified person to carry out various reading tests and IQ tests (Bannatyne, 1968).

6 Modeling tradeoff between time-optimal and minimum energy in saccade main sequence

6.1 Research objective

Saccadic movement is highly stereotyped and commonly believed to be governed for both eyes simultaneously, by a fast open-loop control mechanism. A principle combining time-optimal and minimum control energy criteria was proposed to account for the saccade main sequence as observed from empirical data. The model prediction revealed that the weighting factor of the energy conservation principle was more dominant than the time-optimal, when the saccade amplitude is getting large. It was demonstrated that the model is a general form synthesizing the time-optimum, minimum torque change, and minimum control effort models. In addition, the connection was shown between the model and the stochastic minimum variance model from the aspect of optimization.

6.2 Introduction

The saccade main sequence is the consistent relationship between the duration, the peak velocity and the amplitude of saccadic eye movements (A. T. Bahill, et al., 1975). To simulate the dynamic mechanism for stereotyped trajectories of saccades, optimal control theory is frequently used and debated. This approach originates from Euler's famous statement: "nothing arises in the universe in which one cannot see the sense of some maximum or minimum." In practice, optimal control models have successfully explained the arm movements (Uno, et al., 1989), and walking and posture (Kuo, 1995). In the eye movement

literature, the dynamics of the ocular plant are usually modeled as a linear second or third order system with constant coefficients (C. M. Harris & Wolpert, 1998; Robinson, 1973; A. Van Opstal, Van Gisbergen, & Eggermont, 1985). These linear time-invariant (LTI) systems are generally sufficient to obtain realistic saccade velocity profiles, given a reasonable performance measure to optimize (Uno, et al., 1989). The key issue, therefore, is to choose a performance measure (or performance index) to be applied on the given LTI system. For example, Enderle et al. considered the time-optimal principle, $J = \int_0^{t_f} 1 \, \mathrm{d}t = t_f$, with a sixth order time-varying model (A. Bahill, et al., 1980; Enderle & Wolfe, 1987; Enderle, et al., 1984). Uno (Uno, et al., 1989) proposed a minimum torque change (MTC) model $J = \frac{1}{2}\int_0^{t_f} \dot{\tau}^2(t) \, \mathrm{d}t$, where $\dot{\tau}$ is the change of muscle torque. With a second order LTI system, the velocity profile from the MTC is always symmetric (H. Tanaka, M. H. Tai, & N. Qian, 2004). In contrast, the empirical data shows that the velocity becomes right-skewed as the amplitude increases. More recently, Kardamakis et al. proposed a state dependent minimum-effort principle to predict both head-fixed and head-free saccade velocity profiles (Kardamakis & Moschovakis, 2009). The performance measure used was similar to the MTC model, except for a quadratic weighting function assigned to the control signal: $J = \int_0^{t_f} \boldsymbol{\alpha}(x_e(t)) \, u^2(t) \, \mathrm{d}t$. Here, $x_e(t)$ is the gaze position and $\boldsymbol{\alpha}(x_e) = \alpha_0 + \alpha_1 x_e + \alpha_2 x_e^2$ is a polynomial approximating the extraocular muscle tension as a function of eye position (Dean, 1996). This state-dependent minimum-effort (ME) model has an advantage over the MTC model because it generates a right-skewed velocity profile as the saccade amplitude increases. Such an energy conservation principle may relate to the fact that extraocular muscles, unlike other skeletal muscles, are remarkably resistant to fatigue (Fuchs & Binder, 1983). This fatigue resistance may be seen as a result of the energy conservation

criterion, which governs the saccadic eye movements. With a similar dynamic plant, Harris et al. proposed a minimum variance (MV) model that assumes the saccade velocity profile is the result of minimizing variance in the final gaze position (C. M. Harris & Wolpert, 1998). This model will be discussed in details in Section 6.5.

The current literature indicates the multi-objective nature of saccade movement. Rather than considering these criterions separately, it is assumed that a combined performance measure may be advantageous to gain insights regarding the speed-energy or speed-accuracy tradeoff.

In this Chapter, a weighted linear combination of minimal time and control effort was proposed as the performance measure. This model is capable of representing both time-optimal and minimum energy models by adjusting the weighting factor representing the tradeoff between speed and energy conservation. A critical difference between this study and the aforementioned models is the role of the saccade duration t_f. Typically, t_f is specified in advance and used as a parameter for simulation. In the present model, t_f is a "free" variable to be determined, prior to the state and control variables, in a 2-phase optimization. The weighting factor is first identified by fitting t_f to the empirical data. Once the weighting factor is identified, the saccade main sequence is optimized and the duration-peak velocity relationship and the velocity profile are generated.

To validate the prediction, the output differences between the model and the stochastic MV model are compared in the final section of this chapter. It is demonstrate that the dissimilarity between the two is merely due to the weighting function in the performance measure. As such, it is argued that the model proposed in this chapter synthesizes the aforementioned models (Enderle, et al., 1984; C. Harris & Wolpert, 2006; Kardamakis &

Moschovakis, 2009; Uno, et al., 1989) and provide insights on the speed-accuracy tradeoff from previous studies. Therefore, three out-of-four fundamental components in movement costs are addressed, including time, accuracy and energy (C. Harris & Wolpert, 2006). Stability, the fourth component is not considered, since the dynamic model used is highly over-damped, as seen from the impulse response of the model.

6.3 Methods

The performance measure considered in this chapter is formulated as follows:

$$J(u) = \int_0^{t_f} \left[\lambda + R \cdot \boldsymbol{\alpha}\left(x_e(t)\right) u^2(t) \right] \mathrm{d}t \qquad (6\text{-}1)$$

where $u(t)$ is the control signal; $x_e(t)$ is eye position; t_f is the saccade duration; λ and R are weighing factors attached to time and control energy; and $\boldsymbol{\alpha}(x_e)$ is the same polynomial used in ME model (Kardamakis & Moschovakis, 2009). The dynamics model used in this study is the same one as in Harris and Wolpert (2006) and Tanaka, et al. (2006). This third-order plant fits best to trajectories when using the sensitive technique of Fourier analysis (Harwood, Mezey, & Harris, 1999). The model is stated below:

$$\dot{\boldsymbol{x}} = \begin{bmatrix} 0 & 1 & 0 \\ -a_1 & -a_2 & 1 \\ 0 & 0 & -a_3 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 0 \\ 0 \\ a_3 \end{bmatrix} u \qquad (6\text{-}2)$$

where $\boldsymbol{x} = [x_e \quad \dot{x}_e \quad \tau] = [x_1 \quad x_2 \quad x_3]$ represents the eye position, the eye velocity and the extraocular muscle torque. The last term τ is modeled as a low-pass filtered control signal $u(t)$. Since $\dot{\tau}$ is a function of control u, the term $u^2(t)$ in (6-1) is equivalent to the minimum torque change, as in the MTC model. The coefficients of the state equation and the polynomial function are from Harris and Wolpert (1998) and Tanaka et al. (2006):

$$\alpha_0 = 9.1 \quad \alpha_1 = 0.36 \quad \alpha_2 = 0.014$$
$$a_1 = \frac{1}{t_1 t_2} \quad a_2 = \frac{t_1 + t_2}{t_1 t_2} \quad a_3 = \frac{1}{t_3}$$
$$where: \quad t_1 = 224, \quad t_2 = 13, \quad t_3 = 10$$

To find the optimal control based on the given dynamic system and performance measure, a costate $\boldsymbol{p} = \begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix}$ is first introduced and the *Hamiltonian* is formed as:

$$\begin{aligned} H &= \lambda + R\boldsymbol{\alpha}(x_1)u^2 + \langle \boldsymbol{p}, \dot{\boldsymbol{x}} \rangle \\ &= \lambda + R\boldsymbol{\alpha}(x_1)u^2 + p_1 x_2 + p_2(-a_1 x_1 - a_2 x_2 + x_3) \\ &\quad + p_3(-a_3 x_3 + a_3 u) \end{aligned} \tag{6-3}$$

By applying the necessary conditions for optimality, the costate equations and optimal control are then derived in terms of $\boldsymbol{x}$ and $\boldsymbol{p}$:

$$\dot{\boldsymbol{p}} = -\frac{\partial H}{\partial \boldsymbol{x}} = \begin{bmatrix} \dot{p}_1 \\ \dot{p}_2 \\ \dot{p}_3 \end{bmatrix} = \begin{bmatrix} a_1 p_2 - u^2 R(\alpha_1 + 2\alpha_2 x_1) \\ a_2 p_2 - p_1 \\ a_3 p_3 - p_2 \end{bmatrix} \tag{6-4}$$

$$\begin{aligned} \frac{\partial H}{\partial u} &= 2R\boldsymbol{\alpha}(x_1)u + a_3 p_3 = 0 \\ u &= -\frac{a_3 p_3}{2R\boldsymbol{\alpha}(x_1)} \end{aligned} \tag{6-5}$$

Substituting (6-5) into state Equation (6-2) and costate Equation (6-4), a reduced ordinary differential equation (ODE) system is obtained as follows:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{p}_1 \\ \dot{p}_2 \\ \dot{p}_3 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_3 - a_1 x_1 - a_2 x_2 \\ -a_3 x_3 - (a_3^2 p_3) / \left[2R\boldsymbol{\alpha}(x_1)\right] \\ a_1 p_2 - a_3^2 p_3^2 (\alpha_1 + 2\alpha_2 x_1) / \left[4R\boldsymbol{\alpha}(x_1)^2\right] \\ a_2 p_2 - p_1 \\ a_3 p_3 - p_2 \end{bmatrix} \tag{6-6}$$

Accordingly, the optimal control problem is converted to a two point boundary value

problem (BVP) with following boundary conditions:

$$\boldsymbol{x}(0) = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{x}(t_f) = \begin{bmatrix} \theta & 0 & a_1\theta \end{bmatrix}$$

It is assume that the system starts from a rest position with no initial velocity and no muscle force applied, which coincides with the experiment setup in previous studies (Collewijn et al., 1988). When a saccade is completed, the gaze position is at θ with zero velocity, and $a_1\theta$ is the torque required to maintain the gaze on the target. Since the duration t_f is not specified, an extra boundary condition is needed, which can also be derived from the minimum principle:

$$H\left(\boldsymbol{x}(t_f), \boldsymbol{p}(t_f), u(t_f)\right) = 0$$

This BVP is numerically solved in MATLAB by supplying the seven conditions above. Given a saccade amplitude θ, the duration t_f is determined entirely by the parameter pair (λ, R) in Equation (6-3). Alternatively, there is a simple linear relationship between the amplitude and duration from experimental data (Collewijn et al., 1988): $t = 2.7\theta + 23$. So (λ, R) can be identified by fitting the model to generate t_f, which satisfies the regression model. It should be noted that (6-1) can be further simplified to a single proportionality parameter:

$$\begin{aligned} \tilde{J}(u) &= \int_0^{t_f} \left[\frac{\lambda}{R} + \boldsymbol{\alpha}(x_1)\, u^2(t) \right] \mathrm{d}t \\ &= \int_0^{t_f} \left[\tilde{\lambda} + \boldsymbol{\alpha}(x_1)\, u^2(t) \right] \mathrm{d}t \end{aligned} \tag{6-7}$$

All pairs of (λ, R) with the same ratio $\tilde{\lambda} = \lambda / R$ will produce the same velocity profile and control trajectory. The difference is the scale of final cost $J = R\tilde{J}$, which is not critical for this system identification problem. By comparing (6-7) with the performance

measures of the time-optimal and ME models, it is shown that as $\tilde{\lambda} \to \infty$, (6-7) resembles the time-optimal model (Enderle & Wolfe, 1987). When $\tilde{\lambda} \to 0$, (6-7) is the same as the minimum control effort model (Kardamakis & Moschovakis, 2009).

Although Equation (6-7) more intuitively reveals the relationship between the proposed and previous models, Equation (6-1) is a better option for enumerated solutions. For a given θ, instead of trying to find a single $\tilde{\lambda}$ in one step, one can try several values of λ and find the corresponding R that will generate the duration t_f, which fits best to the regression model. Then, the ratio $\tilde{\lambda} = \lambda / R$ is determined by the linear regression over the series of (λ, R) pairs. Furthermore, it should be noted that $\tilde{\lambda}$ is not necessarily a constant. It is possible to be a function of θ, indicating the satisficing solution between time and control energy as the amplitude changes. This argument will be elaborate this in the next section.

6.4 Results

The values of λ and R were tested within the range of $[0.04,\ 1] \times [4,\ 20]$ and this square region was subdivided evenly. As shown in Figure 6-1, each mesh grid represents a pair of (λ, R), which is supplied to the model to produce a duration t_f. The z-axis is the distance between the output duration and predicted duration from regression (Collewijn et al., 1988). It is clear that for the λ and R, which generate the best t_f for a fixed amplitude θ, the ratio of the two remains constant. This ratio also changes as θ changes. The spikes in Figure 6-1 indicate no solution is found or the solution is lacking accuracy (i.e., if fails to meet the required residual tolerance, which was 10^{-3} in this study). The region of good

solutions shrinks as the value of θ increases.

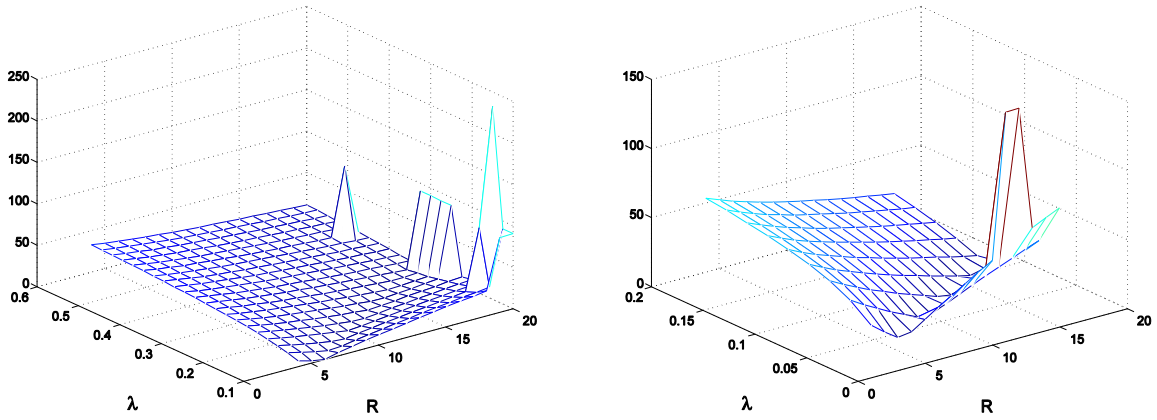


Figure 6-1 Prediction error of duration. Left:$\theta = 25^{\circ}$, Right: $\theta = 45^{\circ}$

Figure 6-2 shows the dispersion result of fitting R to λ with a linear regression model $R = c_1\lambda + c_2$. The slope and intercepts as a function of θ are listed in

Table 6-1. The weighting factor $\tilde{\lambda} = \lambda / R = 1/c_1$ is fitted with a second order polynomial, as shown in Figure 6-3, which shows that the time-optimal weights are much less than the energy conservation weights as the amplitude increases.

Table 6-1 Linear regression of $R = c_1\lambda + c_2$

R	c_1	c_2
5	20.0	0.0
10	25.3	-0.2
20	42.4	0.8
30	75	-0.2
40	122.6	-.4
50	255.0	0.0

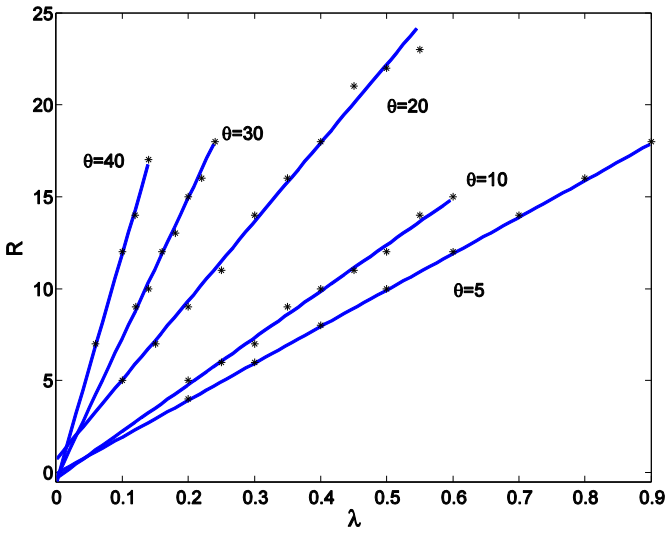


Figure 6-2 Linear relationship between λ and R

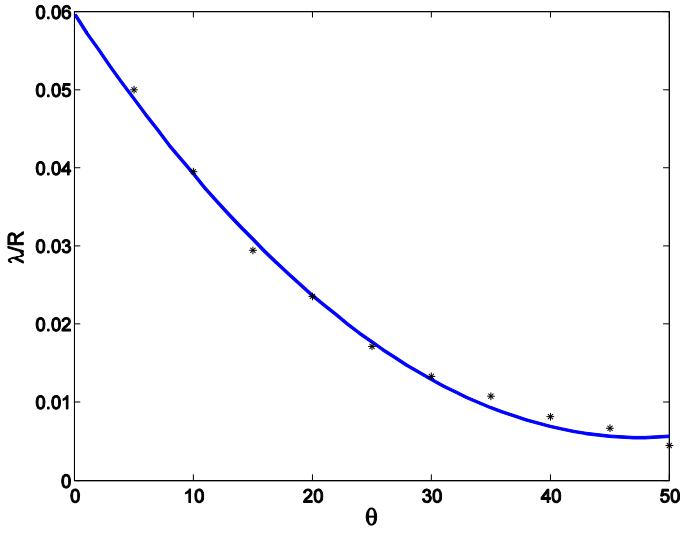
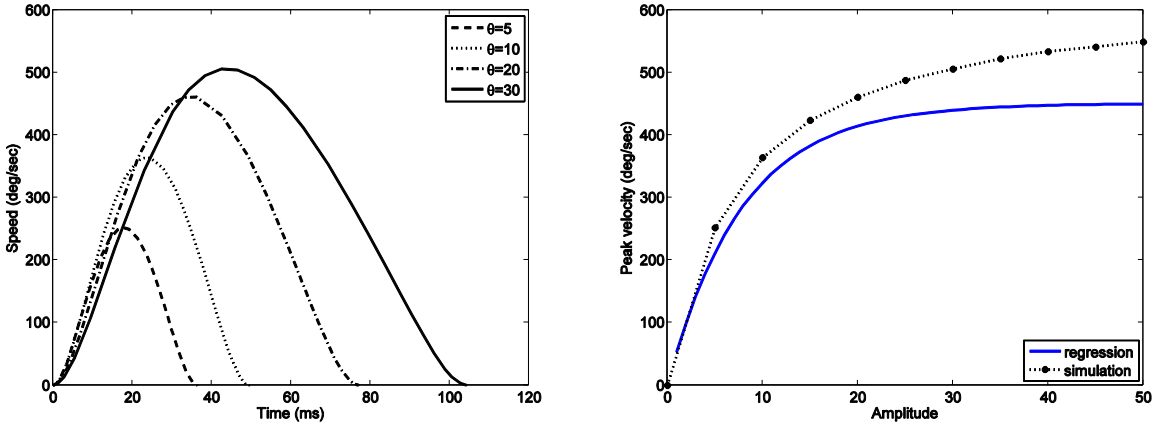


Figure 6-3 Weighting factor (λ / R) as a function of the amplitude θ

The main sequence generated by the proposed model is plotted in Figure 6-4. The velocity profile shows a good fit to the empirical data (Collewijn et al., 1988), which is right-skewed as the amplitude increase. The model captured the compressive nature of the peak velocity-amplitude relationship; however, the peak velocity from the model is a higher than the empirical data. The difference between the two becomes larger as the amplitude increases. Since no constraints were applied to the control (such as saturation), this reflects the pure effect of the performance measure.



(a) (b)

Figure 6-4 (a) Velocity profile; (b) Peak velocity vs. the amplitude

6.5 Discussion

By applying a performance measure that combines the time-optimal and minimum energy principles, it was shown that as amplitude increases, conserving control energy is the dominant factor rather than moving as fast as possible. The ratio of $\tilde{\lambda}$ in (6-7) is as small as 0.0044 when $\theta = 50^{\circ}$, which is negligible compared to other conditions. Therefore, the

proposed model approximates the ME model, which keeps $\tilde{\lambda} = 0$ in the performance measure and generates a satisfying saccade main sequence as well (Kardamakis & Moschovakis, 2009).

It should be noted that the term "minimum energy" in this study, as well as in other eye movement literature, is not the same as in classical optimal control theory (Kirk, 2004). There is a time-dependent weighting function attached to the control signal in the performance measure, which is not common in a classical minimum energy control design. It is well known that without such a weighting function, the minimum energy principle would result in "bang-off-bang" control (Athans & Falb, 1966; Kirk, 2004), which is not accurate for describing biomechanical systems.

The present model, as a solution to an open-loop optimal control problem, is deterministic without considering noise or other stochastic properties of the ocular system. However, it was found that the model was similar to the MV model by Harris et al. (C. Harris & Wolpert, 2006; C. M. Harris & Wolpert, 1998), from the view of optimization. The eye plant used in this study was the same as in Harris and Wolpert (2006), thus, the difference between the two models was the performance measure. In Harris and Wolpert (2006), the performance measure was:

$$\hat{J} = \gamma + \alpha T + k^2 \beta' \int_0^T u^2(t)\, q_T(\tau) \mathrm{d}\tau + \kappa \qquad (6\text{-}8)$$

where $\gamma, \alpha, \beta', k, \kappa$ are all constants. $u(t)$ is the control signal, $q_T(\tau) = \int_T^{T+F} p^2(t-\tau)\mathrm{d}t$ where $p(t)$ is the impulse response of the eye plant, and F is the assumed fixation duration post-saccade. To test dependency, F was varied over a range from 100 to 1000 ms to show that the MV model is not sensitive to F (C. Harris & Wolpert, 2006; Tanaka, et al., 2006), particularly when $F > 300$. By assuming that γ and κ are the same for all possible

trajectories, the performance measure can be further reduced to:

$$\hat{J} = \int_0^T \left[\alpha + k^2 \beta' q_T(\tau) u^2(t) \right] \mathrm{d}\tau \qquad (6\text{-}9)$$

By comparing (6-9) with (6-8), it is clear that the only difference between the two models is the weighting function attached to the control, i.e. $q_T(\tau)$ and $\boldsymbol{\alpha}(x_e)$. $q_T(\tau)$ is the integral of $p^2(t)$ with a translation in the time axis and Figure 6-5 geometrically shows the relationship between the two. The integral is taken from T to $T+F$ which is indicated by the two vertical dash-dot lines. Graphically, $q_T(\tau)$ represents the shaded area under the function $p^2(t)$ between the two vertical boundaries. As $p^2(t)$ moves along the time axis from 0 to T (as (6-9) is integrated over $[0,\ T]$), the shape of the shaded area changes. When the time shift is τ, $p^2(t-\tau)$ is integrated to give the value of $q_T(\tau)$.

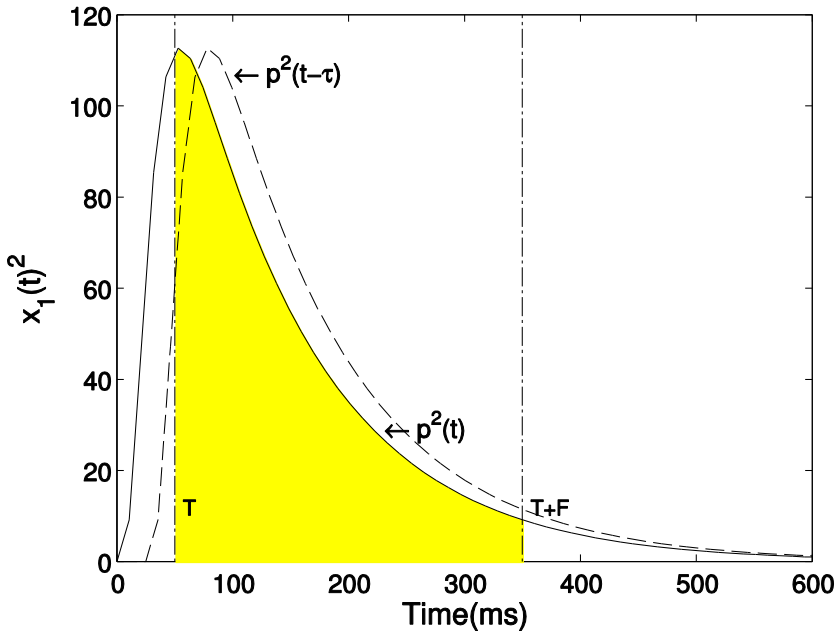


Figure 6-5 Relation between $q_T(\tau)$ and $p^2(t)$.

Since the eye plant is highly over-damped, the impulse response as well as its squared function $p^2(t)$ approaches zeros rapidly. As a result, $q_T(\tau)$ will change little if the right boundary is perturbed further to the right, i.e., increase the fixation duration F. This explains why the MV model is not sensitive to F. Figure 6-6 shows both weighting functions over different amplitudes. The Figure of $q_T(\tau)$ (left, $F = 300$ ms) is adjusted such that it starts at the same point for different amplitudes. The absolute scale of the function does not play a role in optimization since the constant coefficients, $k^2\beta'$ and R, will absorb the differences. The critical adjustment occurs according to the slope of the function. When the amplitude is small, the shape of $q_T(\tau)$ approximates the second order polynomial $\boldsymbol{\alpha}(x_e)$ closely. However, as the amplitude increases, $\boldsymbol{\alpha}(x_e)$ tends to grow faster over the time course. As can be seen from Figure 6-6, when $\theta = 40$, $\boldsymbol{\alpha}(x_e)$ grows more than four times while $q_T(\tau)$ only doubles over the saccade duration. Despite this difference, the weight function $q_T(\tau)$ in the MV model plays a similar role to the state dependent weight function in the present model. Although $q_T(\tau)$ origins from the fixation variance post-saccade, once F is fixed, $q_T(\tau)$ can be readily calculated without any stochastic components. It was also shown that $q_T(\tau)$ is not sensitive to the chosen F because of the highly over-damped nature of the eye plant.

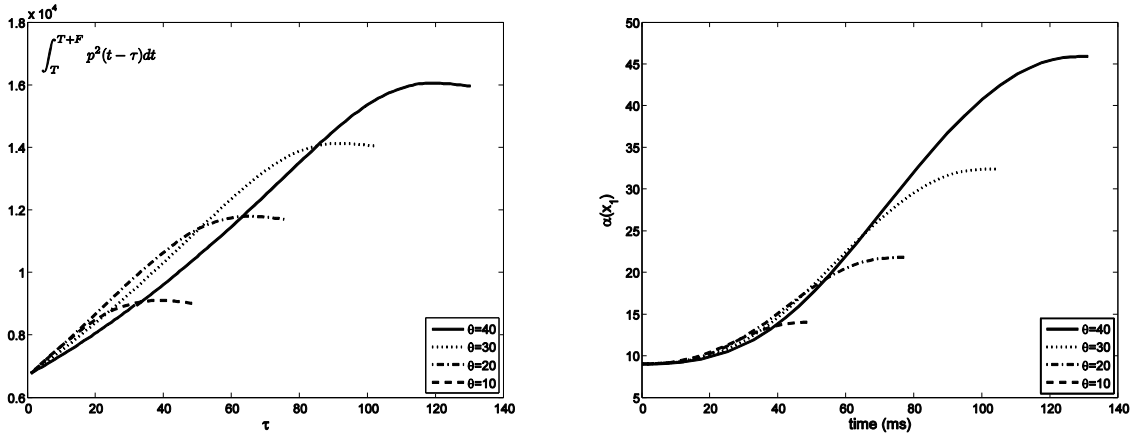


Figure 6-6 Comparison of control weight function.

In summary, it is proposed that the ocular system may evolve to shift the balance between time-optimal and minimum control energy in planning due to saccade amplitudes. The present model is a general form capable of representing most optimal models considered previously. The result from this model integrates and agrees with those from previous studies. It was shown that energy conservation is the dominant consideration for planning as saccade amplitude increases. In addition, there is a transformation mapping between the deterministic model and stochastic MV model from the view of optimization.

It is interesting to notice that, both empirical experiments (Collewijn et al., 1988) and simulations models have been carried out in a discrete fashion, i.e., tested for predetermined levels of amplitudes. All models apply optimal control theory and generate control as a continuous function of the input amplitude. That is, specific amplitude will generate a unique control trajectory. However, the distance between two amplitudes is decreased, it is not likely that the ocular system will be able to keep fine-tuning the control to achieve an optimum. Thus, physiological limitations such as the noise in the motor system and just noticeable differences in the sensory system will possibly result in a resolution threshold for control.

7 Conclusions and future research

7.1 Conclusions

This study addressed unique features of eye movement that are key to enhancing the usability of existing eye tracking technology and understanding the control principles of the ocular system. Respectively, the following contributions were made regarding the ocular dynamic literature.

1. First, an interface widget called the "Hot-Zone" was implemented to accommodate gaze control for human computer interaction. Unlike previous approaches, the Hot-Zone provides for the first time, a two-stage interaction scheme resembling the context menu selection with a mouse. The possibility of using natural gaze state sequence was explored to accomplish complicated tasks; whereas, previous studies solely focused on using fixation for "clicking" on existing interfaces. With its "local calibration" manipulation, the Hot-Zone is more resist to the noise embedded in the eye movement dada. Although the Hot-Zone manipulation is still much slower than manual mouse control, it shed lights on the direction of future gaze control for HCI; that is, designing new interface widgets to work with gaze control.
2. Gaze interaction has the advantage of providing additional information regarding current task performance. Like the context menu, the Hot-Zone must be "filled" with tools related to the user's current task when working with practical applications. Conventionally, this is done by retrieving information on the running program from the operations system. This study explored the

possibility of classifying visual tasks (i.e., reading or visual search) by analyzing user eye movement. The SVM (Support Vector Machine) approach achieved 85% correct on classifying the two different gaze patterns.

3. Finally, this study developed a full-fledged simulation model that faithfully represented the eye rotation in three-dimensional space. It was shown that this model agrees with observed physiological constraints and physical principles. An optimal control model synthesizing features of several previously proposed models was also developed. This model generates countable velocity profiles needed as input to the simulation model. More importantly, it explicitly demonstrates the tradeoff of competing considerations in movement planning. Combining the optimal control model with the simulation model, the deterministic and stochastic components of the ocular dynamic system were constructed. Although the model development is rather theoretical in this study, it can be applied to determine design parameters for Hot-Zone, such as size, distance and direction of the zone.

7.2 Future research

For the Hot-Zone development:

1. The usability of the Hot-Zone was only tested for various gaze control interaction schemes. Many other properties of the Hot-Zone are subject to testing: e.g., the size, the number and the distance of the zones, etc.
2. The Hot-Zone was developed as an independent control object to work with other applications. It would be of interest to test it when working with real world applications such as PowerPoint.

3. The “blink to select” gaze command has the potential to achieve shorter completion times; However, the control logic of the Hot-Zone needs to be modified to deal with “jumps” when blinking, which resulted in poor correct selection rate.
4. It would be desirable to separate the Hot-Zone from the computer and integrate the data processing and logic analysis components with the eye tracking system. Thus, no eye movement data would be missed and the interaction would be more reliable.

For gaze pattern recognition:

1. It would be of interest to investigate other gaze patterns for classifications.
2. Support Vector Machines, as a two-class classifier, may not generalize well for applications with more groups to discriminate. Other pattern recognition algorithms should be tested.
3. The possibility of diagnosing dyslexia by evaluating gaze patterns in reading merits further investigation.
4. Integrating gaze pattern recognition with the Hot-Zone for HCI should be studied.

For ocular dynamics modeling:

1. There were no constraints applied on control signals and state variables in this study. However, for saccades with large amplitudes, there is a “saturation” effect on the control signal. To extend the model to account for saccades with large amplitudes, such constraints should be considered.
2. The control signal considered in this dissertation was a smooth, continuous function, which is only an approximation of the real situation. More general

forms of control should be considered, such as piece-wise continuous and relaxed control.

3. Direct methods will be necessary for solving general optimal control problems, as defined by this study.
4. The round-off error of the simulation model needs to be eliminated for more accurate model prediction of saccadic movement variability.

REFERENCES

Abrams, R. A., Meyer, D. E., & Kornblum, S. (1989). Speed and Accuracy of Saccadic Eye-Movements - Characteristics of Impulse Variability in the Oculomotor System. *Journal of Experimental Psychology-Human Perception and Performance, 15*(3), 529-543.

Athans, M., & Falb, P. (1966). *Optimal control: an introduction to the theory and its applications*: McGraw-Hill New York.

Bahill, A., Latimer, J., & Troost, B. (1980). Linear homeomorphic model for human movement. *IEEE Transactions on Biomedical Engineering*, 631-639.

Bahill, A. T., Adler, D., & Stark, L. (1975). Most Naturally Occurring Human Saccades Have Magnitudes of 15 Degrees or Less. *Investigative Ophthalmology, 14*(6), 468-469.

Baloh, R. W., Sills, A. W., Kumley, W. E., & Honrubia, V. (1975). Quantitative Measurement of Saccade Amplitude, Duration, and Velocity. *Neurology, 25*(11), 1065-1070.

Bannatyne, A. (1968). Diagnosing Learning Disabilities and Writing Remedial Prescriptions. [Article]. *Journal of Learning Disabilities, 1*(4), 242-249.

Beggs, W. D. A., & Howarth, C. I. (1970). Movement Control in a Repetitive Motor Task. *Nature, 225*(5234), 752-&.

Beggs, W. D. A., & Howarth, C. I. (1972). Accuracy of Aiming at a Target - Some Further Evidence for a Theory of Intermittent Control. *Acta Psychologica, 36*(3), 171-&.

Beutel, J., Horii, S. C., & Kim, Y. (2000). *Handbook of Medical Imaging* (Vol. I): SPIE Press.

Boghen, D., Troost, B. T., Daroff, R. B., Dellosso, L. F., & Birkett, J. E. (1974). Velocity Characteristics of Normal Human Saccades. *Investigative Ophthalmology, 13*(8), 619-623.

Cannata, G., & Maggiali, M. (2008). Models for the design of bioinspired robot eyes. *Ieee Transactions on Robotics, 24*(1), 27-44.

Carlton, L. G. (1979). Control Processes in the Production of Discrete Aiming Responses. *Journal of Human Movement Studies, 5*(3), 115-124.

Chris, L. (2000). *Effective eye-gaze input into Windows*. Paper presented at the Proceedings of the 2000 symposium on Eye tracking research \& applications.

Christakos, C. N., Papadimitriou, N. A., & Erimaki, S. (2006). Parallel neuronal mechanisms underlying physiological force tremor in steady muscle contractions of humans. *Journal of Neurophysiology, 95*(1), 53-66.

Collewijn, H., Erkelens, C. J., & Steinman, R. M. (1988). Binocular co-ordination of human horizontal saccadic eye movements. *Journal of Physiology, 404*(1), 157-182.

Coren, S. (1986). An Efferent Component in the Visual-Perception of Direction and Extent. *Psychological Review, 93*(4), 391-410.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning, 20*(3), 273-297.

Crawford, J., Martinez-Trujillo, J. C., & Klier, E. M. (2003). Neural control of three-dimensional eye and head movements. *Current Opinion in Neurobiology, 13*(6), 655-662.

Crossman, E., & Goodeve, P. J. (1983). Feedback-Control of Hand-Movement and Fitts Law. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 35*(MAY), 251-278.

Curcio, C. A., Sloan, K. R., Packer, O., Hendrickson, A. E., & Kalina, R. E. (1987). Distribution of Cones in Human and Monkey Retina - Individual Variability and Radial Asymmetry. *Science, 236*(4801), 579-582.

Dan Witzner, H., David, J. C. M., John Paulin, H., & Mads, N. (2004). *Eye tracking off the shelf*. Paper presented at the Proceedings of the 2004 symposium on Eye tracking research \& applications.

Dean, P. (1996). Motor unit recruitment in a distributed model of extraocular muscle. *Journal of Neurophysiology, 76*(2), 727.

Dempere-Marco, L., Hu, X. P., MacDonald, S. L. S., Ellis, S. M., Hansell, D. M., & Yang, G. Z. (2002). The use of visual search for knowledge gathering in image decision support. *Ieee Transactions on Medical Imaging, 21*(7), 741-754.

Ditchburn, R. W. (1973). *Eye-movements and visual perception*: Oxford University Press.

Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods Instruments & Computers, 34*(4), 455-470.

Elliott, D., Binsted, G., & Heath, M. (1999). The control of goal-directed limb movements: Correcting errors in the trajectory. *Human Movement Science, 18*(2-3), 121-136.

Enderle, J., & Wolfe, J. (1987). Time-optimal control of saccadic eye movements. *IEEE transactions on biomedical engineering*, 43-55.

Enderle, J., Wolfe, J., & Yates, J. (1984). The Linear Homeomorphic Saccadic Eye Movement Model ¿ A Modification. *IEEE Transactions on Biomedical Engineering*, 717-720.

Enoka, R. M., & Fuglevand, A. J. (2001). Motor unit physiology: Some unresolved issues. *Muscle & Nerve, 24*(1), 4-17.

Faisal, A. A., Selen, L. P. J., & Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience, 9*(4), 292-303.

Fitts, P. M. (1954). The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. *Journal of Experimental Psychology, 47*(6), 381-391.

Frank, T. D., Friedrich, R., & Beek, P. J. (2006). Stochastic order parameter equation of isometric force production revealed by drift-diffusion estimates. *Physical Review E, 74*.

Fuchs, A., & Binder, M. (1983). Fatigue resistance of human extraocular muscles. *J Neurophysiol, 49*(1), 28–34.

Hansen, D., MacKayu, D., Hansen, J., & Nielsen, M. (2004). *Eye tracking off the shelf.* Paper presented at the Eye Tracking Research & Application, San Antonio, Texas.

Harris, C., & Wolpert, D. (2006). The main sequence of saccades optimizes speed-accuracy trade-off. *Biological cybernetics, 95*(1), 21-29.

Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature, 394*(6695), 780-784.

Harwood, M., Mezey, L., & Harris, C. (1999). The spectral main sequence of human saccades. *Journal of Neuroscience, 19*(20), 9098.

Haslwanter, T. (1995). Mathematics of 3-Dimensional Eye Rotations. *Vision Research, 35*(12), 1727-1739.

Henneman, E. (1957). Relation between Size of Neurons and Their Susceptibility to Discharge. *Science, 126*(3287), 1345-1347.

Jacob, R. J. (1990). *What You Look at is What You Get: Eye Movement-Based Interaction Techniques.* Paper presented at the Human Factors in Computing Systems: CHI '90.

Jones, K. E., Hamilton, A. F. D., & Wolpert, D. M. (2002). Sources of signal-dependent noise during isometric force production. *Journal of Neurophysiology, 88*(3), 1533-1544.

Jurgens, R., Becker, W., & Kornhuber, H. H. (1981). Natural and Drug-Induced Variations of Velocity and Duration of Human Saccadic Eye-Movements - Evidence for a Control of the Neural Pulse-Generator by Local Feedback. *Biological Cybernetics, 39*(2), 87-96.

Kardamakis, A. A., & Moschovakis, A. K. (2009). Optimal Control of Gaze Shifts. [Article]. *Journal of Neuroscience, 29*(24), 7723-7730.

Keele, S. W., & Posner, M. I. (1968). Processing of Visual Feedback in Rapid Movements. *Journal of Experimental Psychology, 77*(1), 155-&.

Kennison, S. M., & Clifton, C. (1995). Determinants of Parafoveal Preview Benefit in High and Low Working-Memory Capacity Readers - Implications for Eye-Movement Control. *Journal of Experimental Psychology-Learning Memory and Cognition, 21*(1), 68-81.

Kirk, D. (2004). *Optimal control theory: an introduction*: Dover Pubns.

Kleppe, I. C., & Robinson, H. P. C. (2006). Correlation entropy of synaptic input-output dynamics. *Physical Review E, 74*(4).

Kolb, H. (1991). *The neural organization of the human retina*. St. Louis: Mosby Year Book Inc.

Kolb, H., Fernandez, E., & Nelson, R. (2008). WEBVISION The organization of the retina and visual system. from http://webvision.med.utah.edu/index.html

Kowler, E. (1990). The role of visual and cognitive processes in the control of eye movement. *Reviews of oculomotor research, 4*, 1.

Kowler, E., & Blaser, E. (1995). The Accuracy and Precision of Saccades to Small and Large Targets. *Vision Research, 35*(12), 1741-1754.

Kuo, A. (1995). An optimal control model for analyzing human postural balance. *IEEE transactions on biomedical engineering, 42*(1), 87-101.

Leigh, R. J., & Zee, D. S. (1999). The Neurology of Eye Movements *The Neurology of Eye Movements* (pp. 499). Oxford, England: Oxford University Press.

Li, D. H., & Parkhurst, D. J. (2006). openEyes: an open-hardware open-source system for low-cost eye tracking. *Journal of Modern Optics, 53*(9), 1295-1311.

Luo, G., Vargas-Martin, F., & Peli, E. (2008). The role of peripheral vision in saccade planning: Learning from people with tunnel vision. [Article]. *Journal of Vision, 8*(14).

Lyon, G. R. (1996). Learning disabilities. *The future of children*, 54-76.

Manu, K., Andreas, P., & Terry, W. (2007). *EyePoint: practical pointing and selection using gaze and keyboard*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.

Martinez-Trujillo, J. C. (2005). Noncommutativity of eye rotations and the half-angle rule. *Neuron, 47*(2), 171-173.

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). Children's eye movements during reading. *Vision and visual dyslexia, 13*.

Meyer, D. E., Kornblum, S., Abrams, R. A., Wright, C. E., & Smith, J. E. K. (1988). Optimality in Human Motor-Performance - Ideal Control of Rapid Aimed Movements. *Psychological Review, 95*(3), 340-370.

Meyer, D. E., Smith, J. E. K., Kornblum, S., Abrams, R. A., & Wright, C. E. (1990). Speed Accuracy Tradeoffs in Aimed Movements - toward a Theory of Rapid Voluntary Action. *Attention and*

Performance(13), 173-226.

Meyer, D. E., Smith, J. E. K., & Wright, C. E. (1982). Models for the Speed and Accuracy of Aimed Movements. *Psychological Review, 89*(5), 449-482.

Michael, A., Andrew, T. D., & Garth, S. (2005). *Efficient eye pointing with a fisheye lens*. Paper presented at the Proceedings of Graphics Interface 2005.

Nelson, W. (1983). Physical principles for economies of skilled movements. *Biological Cybernetics, 46*(2), 135-147.

Oregan, J. K., & Jacobs, A. M. (1992). Optimal Viewing Position Effect in Word Recognition - a Challenge to Current Theory. *Journal of Experimental Psychology-Human Perception and Performance, 18*(1), 185-197.

Polyak, S. L. (1941). *The Retina*. Chicago: University of Chicago Press.

Rabiner, L. R. (1989). A Tutorial on Hidden Markov-Models and Selected Applications in Speech Recognition. *Proceedings of the Ieee, 77*(2), 257-286.

Raphan, T. (1998). Modeling control of eye orientation in three dimensions. I. Role of muscle pulleys in determining saccadic trajectory. *Journal of Neurophysiology, 79*(5), 2653-2667.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372-422.

Rayner, K., & McConkie, G. W. (1976). What Guides a Readers Eye-Movements. *Vision Research, 16*(8), 829-837.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology-Human Perception and Performance, 22*(5), 1188-1200.

Robinson, D. A. (1964). Mechanics of Human Saccadic Eye Movement. *Journal of Physiology-London, 174*(2), 245-&.

Robinson, D. A. (1968). Eye Movement Control in Primates. *Science, 161*(3847), 1219-&.

Robinson, D. A. (1970). Oculomotor Unit Behavior in Monkey. *Journal of Neurophysiology, 33*(3), 393-&.

Robinson, D. A. (1973). Models of Saccadic Eye-Movement Control-System. *Kybernetik, 14*(2), 71-83.

Robinson, D. A. (1975). Oculomotor control signals. In P.Bach-y-Rita & G. Lennerstrand (Eds.), *Basic Mechanisms of Ocular Motility and Their Clinical Implications* (pp. 337-374). Oxford UK: Pergamon.

Robinson, D. A. (1975). Quantitative-Analysis of Extraocular-Muscle Cooperation and Squint. *Investigative Ophthalmology, 14*(11), 801-825.

Robinson, D. A., Gordon, J. L., & Gordon, S. E. (1986). A Model of the Smooth Pursuit Eye-Movement System. *Biological Cybernetics, 55*(1), 43-57.

Salvucci, D. D., & Goldberg, J. H. (2000). *Identifying fixations and saccades in eye-tracking protocols.* Paper presented at the Proceedings of the 2000 symposium on Eye tracking research & applications.

Schnabolk, C., & Raphan, T. (1994a). Modeling 3-Dimensional Velocity-to-Position Transformation in Oculomotor Control. *Journal of Neurophysiology, 71*(2), 623-638.

Schnabolk, C., & Raphan, T. (1994b). Modeling three-dimensional velocity-to-position transformation in oculomotor control. *J Neurophysiol, 71*(2), 623-638.

Shadlen, M. N., & Newsome, W. T. (1998). The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *Journal of Neuroscience, 18*(10), 3870-3896.

Shumin, Z., Carlos, M., & Steven, I. (1999). *Manual and gaze input cascaded (MAGIC) pointing*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit.

Skavensk.Aa, & Robinson, D. A. (1973). Role of Abducens Neurons in Vestibuloocular Reflex. *Journal of Neurophysiology, 36*(4), 724-738.

Spong, M. W., Hutchinson, S., & Vidyasagar, M. (2006). *Robot modeling and control*: John Wiley & Sons, Inc.

Tanaka, H., Krakauer, J. W., & Qian, N. (2006). An optimization principle for determining movement duration.

Journal of Neurophysiology, 95(6), 3875-3886.

Tanaka, H., Tai, M., & Qian, N. (2004). Different predictions by the minimum variance and minimum torque-change models on the skewness of movement velocity profiles. *Neural Computation, 16*(10), 2021-2040.

Tanaka, H., Tai, M. H., & Qian, N. (2004). Different predictions by the minimum variance and minimum torque-change models on the skewness of movement velocity profiles. *Neural Computation, 16*(10), 2021-2040.

Tweed, D., & Vilis, T. (1987). Implications of Rotational Kinematics for the Oculomotor System in 3 Dimensions. *Journal of Neurophysiology, 58*(4), 832-849.

Uno, Y., Kawato, M., & Suzuki, R. (1989). Formation and control of optimal trajectory in human multijoint arm movement. *Biological Cybernetics, 61*(2), 89-101.

van Beers, R. J. (2007). The sources of variability in saccadic eye movements. *Journal of Neuroscience, 27*(33), 8757-8770.

Van Buren, J. M. (1963). *The retinal ganglion cell layer*. Springfield, Illinois: Charles C. Thomas.

Van Opstal, A., Van Gisbergen, J., & Eggermont, J. (1985). Reconstruction of neural control signals for saccades based on an inverse method. *Vision research, 25*(6), 789.

Van Opstal, A. J., & Van gisbergen, J. A. M. (1987). Skewness of Saccadic Velocity Profiles - a Unifying Parameter for Normal and Slow Saccades. *Vision Research, 27*(5), 731-745.

Van Opstal, A. J., & Van Gisbergen, J. A. M. (1989). Scatter in the Metrics of Saccades and Properties of the Collicular Motor Map. *Vision Research, 29*(9), 1183-1196.

van Rossum, M. C. W., O'Brien, B. J., & Smith, R. G. (2003). Effects of noise on the spike timing precision of retinal ganglion cells. *Journal of Neurophysiology, 89*(5), 2406-2419.

Vapnik, V. N. (1998). Statistical learning theory.: Wiley, New York.

Vapnik, V. N. (2000). *The nature of statistical learning theory*: Springer Verlag.

Vapnik, V. N., & Kotz, S. (2006). *Estimation of dependences based on empirical data*: Springer-Verlag New York Inc.

Vaughan, J. (1982). Control of Fixation Duration in Visual-Search and Memory-Search - Another Look. *Journal of Experimental Psychology-Human Perception and Performance, 8*(5), 709-723.

Waespe, W., & Henn, V. (1977). Neuronal-Activity in Vestibular Nuclei of Alert Monkey during Vestibular and Optokinetic Stimulation. *Experimental Brain Research, 27*(5), 523-538.

Westheimer, G. (1954). Mechanism of Saccadic Eye Movements. *Ama Archives of Ophthalmology, 52*(5), 710-724.

Westheimer, G. (1957). Kinematics of the Eye. *Journal of the Optical Society of America, 47*(10), 967-974.

Westheimer, G. (2001). Relative localization in the human fovea: radial-tangential anisotropy. *Proceedings of the Royal Society of London Series B-Biological Sciences, 268*(1471), 995-999.

Young, L. R., & Stark, L. (1963). Variable Feedback Experiments Testing a Sampled Data Model for Eye Tracking Movements. *Ieee Transactions on Human Factors in Engineering, Hfe4*(1), 38-&.