UNIVERSITY OF CALIFORNIA, SAN DIEGO

Investigating Human Movement and Vision in 3-Dimensional Space

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Bioengineering

by

Cory Eli Stevenson

Committee in charge:

Professor Gert Cauwenberghs, Chair
Professor Todd Coleman
Professor Tzyy-Ping Jung
Professor Eduardo Macagno
Professor Gabriel Silva

2018

The dissertation of Cory Eli Stevenson is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2018

DEDICATION

To Grandma Harriet who patiently listened to my ramblings, encouraged the dreaming from which stories came, and unequivocally supported the education needed to realize dreams.

EPIGRAPH

Noise is just signal that you're not interested in at the moment.

Depending on what mood Matlab is in...
A qualifier in the TA's instructions
of my first college math class's
first introduction to Matlab.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

SD	standard deviation
EOG	electroocculography/electrooculogram
EEG	electroencephlography/electroencephlogram
C-ET	camera based eye tracker
EOG-ET	EOG based eye tracker
2D/3D	two/three dimension(al)
LSL	Lab Streaming Layer
IR	infrared
REM	rapid eye movement
VE-HuNT	Virtual Environment Human Navigation Task
VR	virtual reality
HMD	head mounted display
CAVE	cave automatic virtual environment
MCI	mild cognitive impairment
AD	Alzheimer's Disease
MS	mean sensitivity
SAP	strategy automated perimetry
MoCA	Montreal Cognitive Assessment Test
SBSOD	Santa Barbara Sense of Direction Scale

LIST OF SYMBOLS AND DIACRITICS

Symbols

ψ	azimuthal/horizontal angle
θ	elevation/vertical angle
γ	vergence angle
Δ	differential or error

Diacritics

$'$	head/eye relative coordinate
$_L$	left eye
$_R$	right eye
$\hat{}$	estimated/predicted value
$^+$	pseudoinverse

ACKNOWLEDGEMENTS

There are many many people who I would like to recognize for the impact that they made on me and my work throughout my long journey through graduate school.

I could not have, and shouldn't understate that I would not have, made it through my degree without the love and support of my family. Though they drove me crazy asking "What do you do?" and "When will you be done?", their support, in practically every sense of the word, never wavered. This goes doubly for my mother, Nancy, who cared for my wellbeing even when I didn't; resulting in more "Damn, she was right." moments than I'd care to admit.

I also want to thank my friends who were critical in maintaining my sanity and taking part in the experiences that make up life. Of special note are my best friends, Kelvin Liu-Huang and Sandra Ogletree, who I shared an inordinate amount of time and emotions with, and without which I would not be writing this dissertation. I also want to recognize the help and care from my other close friends (in approximate order of meeting them): James Wu, Nick Nolta, Lissette Wilenski, Mitchell Celaya, Eugene Sato, Siddharth Joshi, and Corey Stone; some of which put up with me as roommates, most of which pushed me to be better, and all of which I wish I could have spent more time with.

Attempting a degree is, of course, not possible without assistance from a group of supportive academics and mentors, and there are some which deserve special amounts of recognition. In no particular order, I would like to thank: Mary Boyle, whose class that I took over a decade ago is partially responsible for putting me on the path of studying the brain, and who has since then enthusiastically and optimistically (somewhat contagiously) provided advice on practically every topic I could bring up; Leanne Chukoskie, who would patiently listen when I would drop by randomly, provided advice on how to navigate academia, and reminded me that my work had purpose; Joe Snider, provided help and suggestions on practically any research problem I could bring forward or casually mention

(much of which I barely grasped). I would also like to recognize a few other academics who critically assisted me over the course of the years: my graduate coordinator, Jan Lenington; Dr. Amy Sung; my fellow graduate students and others from the ISN Lab; Shawn *et al* from SCCN; and Howard, Manuel, Jason, and office staff from TDLC and INC.

Chapter 2 is original to the dissertation. Some of its contents are being prepared for eventual publication. I would like to recognize and thank my advisors on this project, Dr. Gert Cauwenberghs and Dr. Tzyy-Ping Jung, who will be included as co-authors when this work is published. I would also like to thank the Swartz Center for Computational Neuroscience (UC San Diego; San Diego, CA) for usage of their facilities for data collection and the support of their technical staff, notably for helping me setup and run the equipment used in the experiments. I would particularly like to thank J.T. and S.H., my labmates who helped pilot the experiment and allowed me to take and use their images.

Chapter 3, covering results from the Virtual Reality Human Navigation Task project, was produced by the efforts of many groups of people over many years and iterations. Parts of this chapter are original to this dissertation and parts are adapted from the article "Wayfinding and Glaucoma: A Virtual Reality Experiment" on which Fabio Daga, Eduardo Macagno, Cory Stevenson, Ahmed Elhosseiny, Alberto Diniz-Filho, Erwin Boer, Jurgen Schulze, and Felipe Medeiros were co-authors and which was published in the ARVO Investigative Ophthalmology and Visual Science in 2017. For this project I developed the experimental design and the implemented virtual environment, performed data extraction and analysis, and developed the data visualizations. The teams at the Visual Performance Laboratory (UC San Diego) and CalIT2/Qualcomm Institute (UC San Digo), including the authors, were instrumental in developing, implementing, and collecting data for this project. I would particularly like to thank the multiple generations of programmers (Lelin Zhang, Cathy Hughes, Scott Runyon, Philip Weber) who worked with me to develop the code necessary for a virtual environment and experiment in the CAVEs. This project

was financially supported by National Institutes of Health/National Eye Institute Grants EY025056 (FAM) and EY021818 (FAM) and a CalIT2 Strategic Research Opportunities (CSRO) seed grant.

And of course I would like to thank my advisor, Gert Cawenberghs, who Additionally, I would like thank Eduardo Macagno and Tzyy-Ping Jung for their continuous advise and assistance on my education and projects, and for serving on my committee, for which I would also like to extend my appreciation to Todd Coleman and Gabe Silva.

VITA

2010	B. S. in Bioengineering: Biotechnology, University of California, San Diego
2012	M. S. in Bioengineering, University of California, San Diego
2017	Ph. D. in Bioengineering, University of California, San Diego

PUBLICATIONS

F. Daga; E. Macagno; C. Stevenson; A. Elhosseiny; A. Diniz-Filho; E. Boer; J. Schulze; F. Medeiros. Wayfinding and Glaucoma: A Virtual Reality Experiment. *ARVO Investigative Ophthalmology and Visual Science.* 2017;58:3343-3349. doi: 10.1167/iovs.17-21849

M.E. Hernandez; J. Snider; C. Stevenson; G. Cauwenberghs; H. Poizner. A Correlation-based Framework for Evaluating Postural Control Stochastic Dynamics. *IEEE Transactions on Neural Systems and Rehabilitation Engineering.* vol.PP, no.99, pp.1-1 doi: 10.1109/TNSRE.2015.2436344

C. Stevenson; T.P. Jung; G. Cauwenberghs, Estimating direction and depth of visual fixation using electrooculography. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).* Milan, 2015, pp. 841-844. doi: 10.1109/EMBC.2015.7318493

M.E. Hernandez; C. Stevenson; J. Snider; H. Poizner. Center of pressure velocity autocorrelation as a new measure of postural control during quiet stance. *Neural Engineering (NER), 2013 6th International IEEE/EMBS, Conference on.* vol., no., pp.1270,1273, 6-8 Nov. 2013

C. Stevenson, A. Elhosseiny, A. Diniz-Filho, G. Cauwenberghs, F. Medeiros, E. Macagno. Virtual environment human navigation task for assessing spatial cognition and navigation. *2016 47th Annual Meeting of the Society of Neuroscience (SfN2016).* San Diego, CA. 12-16 Nov 2016.

A. Elhosseiny; E. Macagno; C. Stevenson; A. Diniz-Filho; J. Shulze; P. Weber; G. Dawe; E. Boer; F. Medeiros. 3D Virtual Environment Human Navigation Task (VE-HuNT) in Glaucoma. *Association for Research in Vision and Ophthalmology (ARVO) Annual Meeting.* Seattle, WA, May 2016

C. Stevenson, C.; Runyon, S.; Schulze, J.; Cauwenberghs, G.; Raffii, M.; Macagno, E., "Virtual Environments to Assess Facility Design for the Cognitively Impaired", *Academy of Neuroscience for Architecture (ANFA), 2014 Biannual Conference.* San Diego, CA. 18-20 Sept 2014.

M.E. Hernandez, C. Stevenson, J. Snider, H. Poizner, "Human Cortical Theta During Quiet Stance Encodes Postural Stability". *2013 43rd Annual Meeting of the Society for Neuroscience.* San Diego, CA. 9-13 Nov 2013.

C. Stevenson, C. Hughes, M. Chi, T.P. Jung, J. Schulze, J. Townsend, G. Cauwenberghs, E. Macagno. "Wireless Monitoring of Human Neurological Activity and Movements in Immersive and Interactive Virtual Environments". *2013 14th Annual University of California Systemwide Bioengineering Symposium.* San Diego, CA. 19-21 Jun 2013.

C. Stevenson, M. Chi, J. Gossmann, E. Edelstein, T.P. Jung, P. Otto, J. Schulze, G. Cauwenberghs, E. Macagno, "Wireless Monitoring of Human Neurological Activity and Movements in Immersive 3D Virtual Reality Environments", *IEEE EMB/CAS/SMC BMBI Workshop, 34th International IEEE EMBS Conference.* 27 Aug. 2012

E. Macagno, L. Zhang, E. Edelstein, J. Schulze, C. Stevenson, G. Cauwenberghs, T.P. Jung, J. Gossman, P. Otto, J. Townsend. "Next-Gen Wireless Monitoring of Electrical Activity and Movement in Virtual Reality: Evidence-based Analysis/Diagnosis in Realistic Environments" *Federation of European Neuroscience Societies (FENS) 8th Forum of Neuroscience.* Barcelona, Spain. 17 July 2012.

L. Zhang, J. Gossmann, C. Stevenson, M. Chi, G. Cauwenberghs, K. Gramann, J. Schulze, P. Otto, T.-P. Jung, R. Peterson, E. Edelstein and E.R. Macagno. "Spatial Cognition and Architectural Design in 4D Immersive Virtual Reality: Testing Cognition with a Novel Audiovisual CAVE-CAD Tool. " On-line contribution to *Spatial Cognition in Architectural Design (SCAD-11) Conference.* New York, NY. 16-19 Nov 2011.

ABSTRACT OF THE DISSERTATION

Investigating Human Movement and Vision in 3-Dimensional Space

by

Cory Eli Stevenson

Doctor of Philosophy in Bioengineering

University of California, San Diego, 2018

Professor Gert Cauwenberghs, Chair

Vision and movement are instrumental components of normal functioning in everyday life. Due to this importance, it is valuable to understand the neurological mechanisms which underlie these processes and how they are employed. Correspondingly, much research is done on deficits in these processes and how people are impacted by such deficits. This dissertation focuses on the development of techniques to better be able to perform investigations into vision, movement, and the interaction between them.

In order to study vision it is often important to know where the eyes are looking and how they are moving; this is often done using a method of eye tracking, typically using a camera. In a first project as part of this work, eye tracking based upon electrooculography

(EOG) was developed as a potentially inexpensive, high temporal resolution eye tracker system, with capabilities of tracking gaze fixation in 3D space. An experimental calibration setup and procedure were created to calibrate and test the capacities of the EOG based eye tracker, including comparison to a standard, camera based eye tracker.

Vision is also involved in movement through space. Navigation within an environment requires the utilization of cues, typically visual ones, to determine location, plan a path, and subsequently execute it. In a second project a virtual reality system and experimental platform, the Virtual Environment Human Navigation Task (VE-HuNT), was created to assess the navigation skills and strategies utilized within conditions applicable to everyday life. VE-HuNT was employed to investigate the effects of visual impairment due to glaucoma on navigation, and the aspects of the environment that would improve functioning in these individuals. Behavioral metrics for wayfinding ability were assessed and the impact of glaucoma on the ability of subjects to utilize environmental cues for wayfinding was determined.

Together the outcomes of these two complementary projects further the ability to perform human cognitive research under conditions representative of typical human activity. As an important tangible benefit of these advances, the improved capabilities for assessment of clinical conditions which impact such activities open up new avenues for neurological understanding of vision and movement.

Chapter 1

Introduction

Movement is a fundamental part of human life. Our bodies and our brains are structured in a way to achieve movement on multiple scales, from tiny movements of one part of the body relative to another, to movement of the body through the entire world. To successfully achieve this vast spectrum of movement within the variety of conditions that occur in the world, senses, planning, and memory are employed to coordinate and execute these movements effectively across the space and time within which they can occur. Correspondingly, movements also impact these processes, creating a system of interactions complex enough to be capable of surmounting this challenge. With such importance and ubiquity, better understanding movement and the processes which enable them is also of importance.

Of particular importance is the interaction of vision and movement. Vision can inform movement before it happens and while it is being planned and provide feedback to stabilize and refine the movement as it is being executed [52]. As vision is carried out by the eyes, which are housed within the moving body and move themselves, movement also has a direct effect on vision. Therefore, to understand this important interaction, movement and vision must be studied together.

1.1 Vision

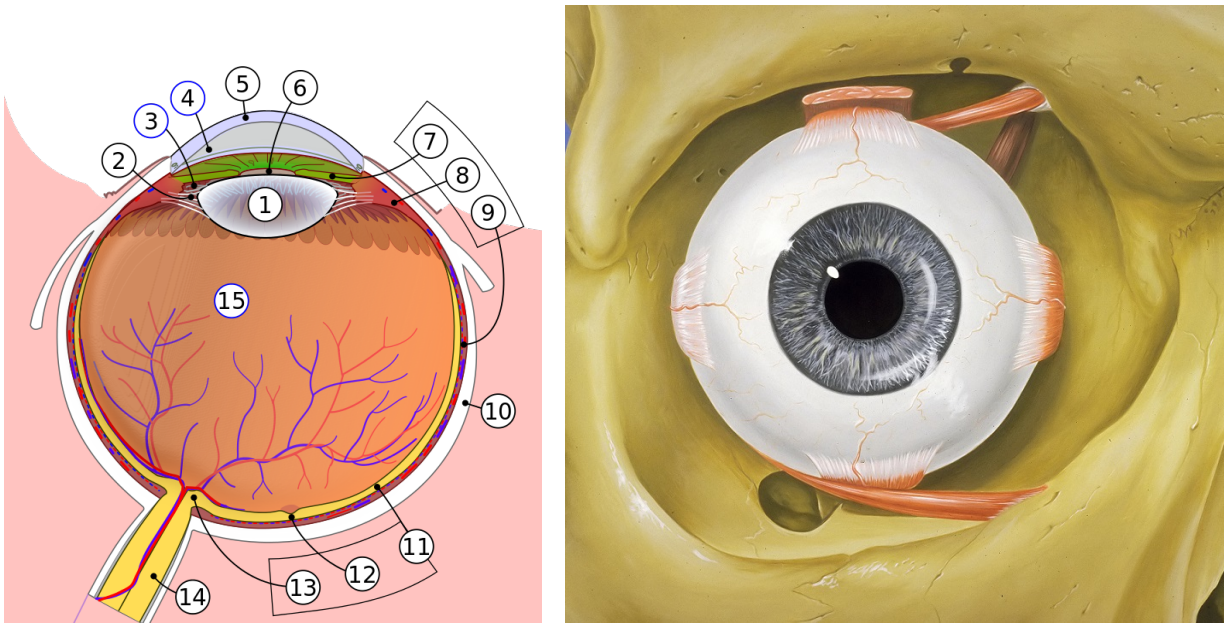
The human visual system consists primarily of the two eyes and the brain. Generally, light containing information from the environment enters each eye, is converted to neural signals, and transmitted to the brain, where the signals from both eyes are further processed, interpreted, and responded to.

1.1.1 Structure of the eye

In order to see, light containing an image passes through the lens and cornea at the front of the eye, onto the retina at the back of the eye (Figure 1.1a). The retina contains interconnected layers of neurons, on top of and connected to, a layer of photoreceptors, which lie upon a layer of pigmented epithelial cells [30]. The photoreceptors respond to the indecent light by hyperpolarizing and releasing glutamate. This neurotransmitter induces a response in the connected interneurons which process and encode the signal, which is sent through the optic nerve to the brain. The visual image produced by the eye is therefore dependent on this structure of the retina. The distribution and types of the photoreceptors are not uniform throughout the retina; in the fovea where the photoreceptors are densest the corresponding visual image's spatial resolution is substantially superior [30].

1.1.2 Binocular Vision

When fixating on a visual object, each eye is pulled by the three sets of extraoccular muscles in order to rotate in the eye socket (Figure 1.1a) and position the corresponding image on the fovea of the eye, therefore generating the best image possible. This movement to a target is called a saccade [36]. Typically, both eyes are used in conjunction, with each eye fixating on the same position in space, allowing for information redundancy and the encoding of depth information. Binocular vision encodes 3D spatial information in two



(a) (b)

Figure 1.1: **Diagram and illustration of the human eye (a)** Diagram of the structure of the eye (dorsal view) with notable structures: lens (1), cornea (5), pupil (6), sclera (10), retina (11), fovea (12), optic nerve (14), and vitreous humour (15). **(b)** Illustration of the human eye depicting an anterior view of the eye in its orbit, along with the six extraoccular muscles.

ways: by direction information for each eye in the form of feedback from the extraoccular muscles and by stereopsis, which is the use of image disparities produced by the discrepancy in distance between the two eyes, and which typically dominates depth estimation in many circumstances [49].

1.2 Movement

1.2.1 Gross Motor Movement

Human movement employs contractions of muscles, coordinated by the brain, to adjust the position of parts of the body and propel the body through space. Gross motor functions are body scale movements of the head, torso, and limbs. They rely on feedback from a combination of visual, vestibular, and proprioceptive senses [12], in order to ensure that the intended motor output is occurring properly and to provide the necessary information for modifying the motion as subsequently needed. Consequently, deficiencies in these senses, or the synthesis of them, can severely impact the motor function of an individual [53]. In order to successfully move within a space, such as opening a door or walking across a room, it is necessary to execute a set of precisely timed, placed, and appropriately balanced gross motor movements.

1.2.2 Movement through Space

Moving through an environment requires planning and combining multiple sets of motor movements over an extended period of time. Because of the relative scales, both in space and time, memory and more complex processing are used by the brain to coordinate the movement in the environment, combining positional information from the past with sensory information from the present. Navigation is the process of or capabilities for moving through space on the scale of an environment. Navigating requires detecting the relevant

features of the environment, both from current senses and memory, determining, from these, the current positions of ones person and these features [26, 46, 65, 54], and planning the requisite movements to proceed to the intended location; it is generally necessary that these requisite processes be executed in a continuous and parallel fashion. Like gross motor function, any deficiencies in the necessary processes can have deleterious effects on the ability to navigate [47, 56, 46, 42, 65].

1.2.3 Investigation

As vision and motion, and their interaction, are complex, it is very difficult to study this in unfettered, realistic conditions, even though understanding their function in these conditions is generally the desired, ultimate objective. Typically, studies in realistic conditions are difficult, in part, due to difficulties in controlling specific aspects of the study, isolating various influencing factors, and monitoring specific physiological processes of interest. Therefore, studies of vision and navigation are often performed in restricted settings, divorced from untenable realistic situations. Ideally, the technology to perform studies in controllable and measurable scenarios, more akin to the realistic ones, can be developed.

1.3 Overview of Dissertation

1.3.1 Objectives

The purpose of the work included in this dissertation is to develop better technologies to study human vision and movement. The ideal objective is to expand the capacities of these technologies for monitoring human subjects in conditions more reminiscent of realistic environments and then employ these technologies to assess aspects of relevant physiology.

1.3.2 Composition

Chapter 2 of this dissertation focuses on the development and characterization of an eye tracking system based upon electrooculogrpahy (EOG). Eye tracking experiments were designed and implemented, and the analysis developed, to produce calibration procedures capable of predicting visual fixation in 3D space using EOG. The capabilities of this eye tracker were analyzed, including a comparison to camera based eye tracking (C-ET).

Chapter 3 of this dissertation focuses on the development and implementation of virtual reality (VR) systems to study human navigation. The Virtual Environment Human Navigation Task (VE-HuNT) developed was designed to study the various cognitive processes and sensory capabilities critical to navigation. VE-HuNT was employed to investigate the effects of visual degeneration due to glaucoma on navigation.

Chapter 2

Electrooculogram Based Eye Tracking

2.1 Introduction

Physiological monitoring is an important facet of studying human cognition. This is particularly true for the recording of eye movements, as they reflect internal state by having functional properties of both motor output and sensory input. Many techniques have been developed for monitoring eye movements and vision in humans. These methods can generally be broken into two categories, electrophysiolgical and visual recording.

2.1.1 Electrooculography

Many physiological processes in the human body generate macroscopically detectable and geometrically localizable electrical potentials, which can be recorded with minimal invasiveness. Electrodes placed in particular locations on the body can be used to study particular electrophyslolgical signals, particularly: electroencephlography (EEG) to measure signals from the processes of the brain, electrocardiography (ECG) to measure signals from the contractions of the heart, or electrooculography (EOG) to measure signals from the movements of eyes. These monitoring methods have a similar technological basis and are

often employed in conjunction with each other to study human neurophysiology.

Ocular Dipoles

The physical basis of EOG is the measuring of apparent changes of an electric field as the source dipole rotates. The retina is maintained in a bath of ions necessary for phototransduction, creating a negative charge at the back of the eye [8]. This results in the corneo-retinal standing potential, an electric dipole aligned from the front to the back of the eye, in the range of $0.4 - 1.0$ mV [36]. This dipole is continuously present under normal conditions, with some amplitude fluctuations over timescales of min to hours, based upon light adaptation and physiological conditions [29, 8].

As the position and orientation of the dipole is fixed to that of the eye, as the eye rotates the orientation of the dipole consequently changes relative to the head [27, 64]. The magnitude of the electric field measured from a fixed location near the eye, such as that from an electrode placed on the face, changes accordingly as it is moved into a different location in the dipole's electric field. When the angular rotation of the eye is constrained to $\pm 30^{\circ}$ and electrodes placed on the face are mostly perpendicular to the neutral dipole position, the magnitude changes to the electric field can be considered approximately linear [36, 27]. Consequently, as the primary change in the measured electric field is dependent on the movement of the eyes the measured electric signal can be used as an analogue for eye movement.

Typical Usage

The existence of EOG has been known for many years, with Mowrer *et al.*making references to the phenomenon as early as the 1930s [41]. Since this time, it has been used primarily as a method of detecting movements of the eyes, where changes in the EOG are used to determine speed and direction of eye movement. This has phenomenon has

been used in polysomnography, monitoring eye movements in sleep [14], and in tests of vestibular function [36]. EOG is also collected incidentally during EEG experiments, due to its relative signal strength, and cleaned from the data as an artifact [28]. More recently there have been attempts at using EOG as a geometric eye tracker [27], developing a method which can tell the absolute position of gaze.

2.1.2 Camera Based Eye Tracking

Much of the human eye tracking done currently in commercial applications and research is camera-based eye tracking (C-ET), where a camera is used to make a visual recording of the eye movements and the resulting images used to determine the location where a subject is looking. C-ET uses light reflected off of the eye to identify the position of the eye, using either visible light, isolating the position of the dark pupil as compared to the white sclera, or equivalent reflections in infrared (IR). In both cases, C-ET requires a steady source of the appropriate spectrum, which in visible light is typically provided by the ambient light in the environment or in IR with a dedicated source of IR illumination, which is often built into the C-ET system and part of the necessary power requirements.

To use a C-ET, the camera, and necessary illumination setup, must be placed so that a clear image of the eye(s) can be acquired. The collected image of the eye is then fed to a computer, which then needs to identify the relevant visible characteristics of each frame. Initially a calibration is performed to determine how the relevant features are related to the extant geometries and subsequently to identify eye positions, usually relative to either a fixation position on a screen of known position or the angular gaze relative to the head. If the line of sight from the camera to the eye is obscured or the required reflection is changed (in amplitude or position) for any period of time, then the C-ET determination of gaze position is distorted or lost for that period of time, or depending on the software and design of the system, an update to the calibration could be required.

Many C-ET designs have been developed, generally able to be split into three categories: desk mounted, computer or tablet mounted, or head mounted. Desk mounted systems are placed between a subject and a screen, above or below the subject's view of the screen (classically on or in a desk). Historically, most research-grade eye trackers were of this category. In many cases subjects place their head on a chin mount in order to stabilize the visual acquisition of the eye and to provide a more fixed geometric basis of calibration, allowing eye movements relative to the head to be synonymous with movements relative to the world. Computer and tablet mounted C-ET, are the extension of desk based systems, using similar principles, but with smaller cameras able to be mounted on or as part of a smaller, more portable electronic device. These suffer from the same geometric limitations of desk mounted systems, where the head position is either fixed relative to the mount or a method is needed to track the position of the head in space in addition to the eye position relative to the head. The third type, head mounted systems, attempts to overcome this problem by attaching the camera mount to the head itself, often in the form of a pair of glasses, so that when the head moves, so does the camera, maintaining the same relative position to the head. However, one of the difficulties in developing head mounted systems has been the development of small cameras with focal lengths short enough to acquire clear images of the eye from a few centimeters away. Head mounted systems typically require a front facing camera necessary to identify the specific scene at which a subject is looking during head movements.

The capabilities of C-ET vary depending on the system used and the situation in which it is employed. High end, research-grade systems are capable of precise angle measurements, with prescribed errors around $\sim 0.5^\circ - 1^\circ$, and are capable of the high frame rates, > 60 Hz, necessary to capture high-speed saccades [50, 57, 58]. These systems, however may be extremely expensive, with research standards such as SR Research's desk mounted EyeLink1000+ or SMI's head mounted glasses, SMI ETG2, ranging into the

tens of thousands of dollars. Other systems, generally intended for commercial purposes, tend to be cheaper, but generally have slower frame rates less suited for capturing faster eye movements. Recent technologies employing the built in cameras in smart phones and tablets as eye trackers, have become more widespread, however are limited to the standard camera frame rate of 30 Hz and are calibrated only to the area of the screen and sensitive to environmental noise [21].

2.1.3 Objectives

The purpose of this experiment is to further the development of human eye tracking based on electrooculographic signals. This will be achieved by developing an experimental calibration procedure, along with the models and calculations necessary to perform this calibration. The results of this eye tracker will then be characterized and compared with typical camera based eye tracking systems.

2.2 Methods

2.2.1 Subjects

A total of ten (10) subjects volunteered to participate in the experiment. Subjects were recruited, compensated, and participated according to the approved IRB, and provided the required consent. Three (3) subjects were eliminated from the analysis; one due to unsuccessful EOG electrode digitization, one completed only a fraction of the experimental trials, and one had a strong, uncorrected, asymmetric astigmatism. The seven (7) remaining subjects had an average age of 23.4 (SD=2.8), three (3) were female, and all were proficient in English and self-reported general good health, with no known motor or other neurological conditions. The visual acuity of each of these subjects was self-reported as normal (n=4) or corrected to normal, in which case they wore their prescribed contact lenses while

participating in the experiment. All images of subjects that were taken and/or published, were done with the permission to do so by the depicted subjects.

2.2.2 Experimental Setup

The experiment was designed to provide a single-point visual target of known location in 3D space, from which an eye tracking calibration could be obtained and against which estimations could be compared.

Sensors were placed on the head to record electrical potentials and head position and movements (Figure 2.1a). To detect EOG signals, eight standard EEG electrodes (Biosemi B.V., Amsterdam Netherlands) were placed around the eyes; two on either side of each eye, one below each eye, and one above each eyebrow. Electrodes were attached as close to the eyes as possible without infringing on the field of view or otherwise interfering with the movement of the eyes or head. The subjects head was fitted for an EEG cap, though no scalp electrodes other than the EEG system ground and reference electrodes were applied. The motion tracking system (PhaseSpace, Inc., Impulse system, San Leandro, CA) used four infra-red LEDs attached to the crown of the snug EEG cap, in a roughly rectangular pattern, to mark the 3D position and orientation of the head. The PhaseSpace system used multiple IR cameras mounted around the room to precisely track the LED marker positions within the room during the course of the experiment. The positions of both EEG electrodes and motion tracking markers, relative to each other and the head, were digitized (zebris CMS20 universal measuring system, zebris Medical GmBH, Isny, Germany) at the beginning of the experiment in order to construct and situate head and eye relative coordinate bases necessary for head-unfixed eye tracking calibrations (Figure 2.1c).

Each subject was seated comfortably facing forwards, towards a projector screen (Figure 2.1b). The EEG system was placed behind the subject with the leads arranged so that they did not fall within the subject's visual field and so that they did not interfere

with the movement of the subject's head, as the subject's head remained unfixed during the experiment. An infra-red, camera-based eye tracker (EyeLink 1000, SR Research Ltd, Mississauga, Ontario, Canada) was placed on a small table in front of the subject. The lens was adjusted so that the C-ET had a clear view of the subject's left eye and a tracking dot placed on the subject's forehead, which as used for the C-ET's internal determination of gaze without a fixed head. The height of the table and subject were adjusted so that the subject was within the range of the C-ET during small head movements and that the C-ET and table were not intruding into the subject's view of the projector screen, placing the C-ET approximately 0.5 m or less away from the subject, depending on the subject's height and angle needed to maintain an optical lock on the eye. The projector screen was placed approximately 3.0 m away from the subject, attempting to maintain a relative visual size and aspect ratio similar to that of a standard sized monitor placed closer to the subject, as a typical desk-based monitor would be positioned, immediately above and behind the C-ET (Figure 2.1d). The precise position of the screen in each experiment was determined by matching PhaseSpace marker positions to the corners of the screen.

2.2.3 Paradigm

After the subject had been seated and sensors were properly attached, the lights to the room were turned off and the EyeLink system's 17-point calibration sequence was performed, using the projector screen in lieu of a standard monitor. After C-ET calibration, the projected image was switched to a solid grey background and recording was started. The data from each modality was captured and stored onto the same computer system via Lab Streaming Layer software (Swartz Center for Computation Neuroscience, UCSD, San Diego, CA). The Biosemi EEG data was collected at 512 Hz, the PhaseSpace position data at 480 Hz, and the Eyelink C-ET data at 66 Hz.

The experiment consisted of multiple trials of smooth pursuit, where the subject

would fixate their vision on a moving target. A visible, red LED, tracked with the PhaseSpace system was attached to the end of a 1.0 m stick, for use as a fixation point with a known 3D location. The LED fixation target was moved as randomly and smoothly as possible within an area between the subject and the projector screen. At the beginning of the experiment, the subject was instructed to follow the target with their eyes, keeping head movements to the minimum necessary to keep the target within visual range. Each subject completed multiple trials in blocks of different paradigms. The first primary block consisted of three trials of 2D, planar movement; where the target was moved along the vertical plane and bounds of the projector screen, as would a target projected on the screen. The second primary block consisted of three trials of 3D movement; where the target was moved throughout the intervening space between the subject and the screen, a range of approximately 0.5 to 3.0 m away from the subject. Each 2D trial lasted for approximately 70 seconds (range 50 to 90) and each 3D trial 100 seconds (range 70 to 130). The subject remained seated and was allowed to rest between each trial, until they indicated they were ready to continue. Additional blocks of alternate paradigms were used to look at the sensitivity of EOG to other conditions or signals, such as eye blinks, but were not used for the determination of position.

2.2.4 Analysis & Modeling

With a known target location and head and eye positions, it is possible to develop a calibration to estimate eye fixations from EOG.

Data Stream Processing

The collected, timestamped data were imported into MATLAB using the EEGLAB toolbox, and separated for processing into its component streams via the Mobilab extension. The eight (8) channels of EOG data were filtered using a low-pass filter with a 50 Hz cutoff,

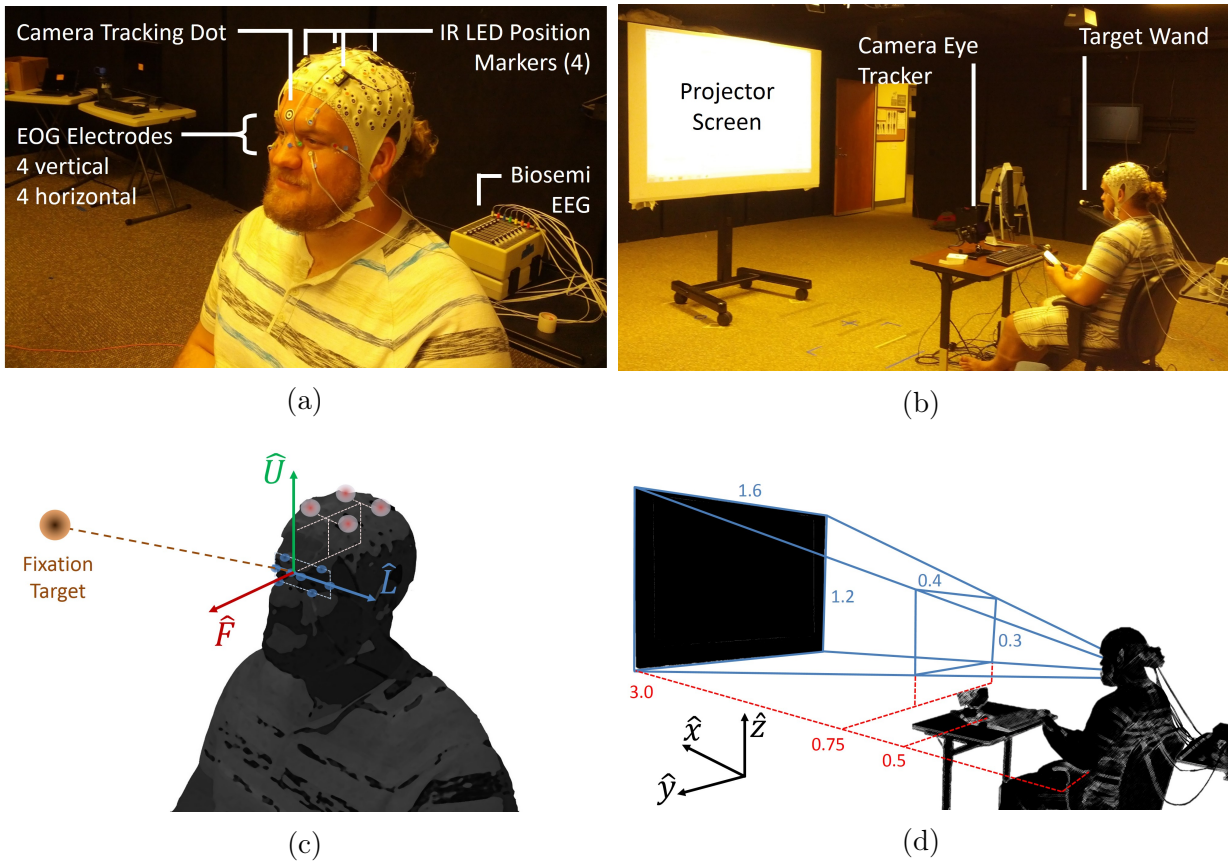


Figure 2.1: **Images and diagrams of the EOG-ET experimental setup.** **(a)** Electrodes for EOG recording were placed around the subject's eyes and IR LEDs for head tracking on the head via the EEG cap, and **(b)** the subject was seated facing the projector screen with the C-ET placed in front of them, in a position that it would typically be at if the screen was a closer, standard monitor. **(c)** A right-handed, relative coordinate system fixed to the head was constructed using the sensor locations as a basis, which was **(d)** converted from the global coordinate system as the target was moved in the intervening space between the subject and the screen.

to eliminate ambient AC line noise and ocular tremor, then with a high-pass filter with a 1/40 Hz (0.025 Hz) cutoff, in order to remove the majority of slow electrode drift. Each data stream was then broken up into independent trials and each stream cleaned of drops (e.g. occluded head markers). Each stream was interpolated to a common time base of 200 Hz, to provide regular sampling across all streams.

The Eyelink C-ET system provided the position of a subjects' gaze on the projector screen, reporting the horizontal and vertical position in pixels. In order to use the C-ET data in a comparison to the global-space target position and EOG estimations, the known position of the projector screen was used to convert the 2D pixel representation of gaze movement into a planar trajectory of $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ coordinates in the common global space. The C-ET data was cleaned by removing points where the system lost track of the eye, typically during blinks, and interpolated to a 30 Hz time base.

Ideal Geometric Model

In order to determine gaze direction, the position of a known object in a global coordinate system, such as the fixation target, needed to be translated into a local coordinate system relative to each eye. As eyes rotate within their sockets, but do not translate significantly relative to each other or the head, a relative coordinate system fixed to the head can be effectively used as a basis for an eye-relative coordinate system. The assumption that the head and face can be treated as a rigid body, as minimal changes to the shape of the head and face occur, allows the digitized head markers and eye electrodes to be used to construct a right handed, orthonormal coordinate basis fixed to the head; with an up vector being generated from the normal vector of a plane of best fit for the four head markers; a left vector being generated from the line passing through the center of both eyes, calculated as the average position of the four electrodes around that eye; and a front vector, pointing out from the face, orthogonal to the other two vectors. The origin of this

coordinate system ($\boldsymbol{P_H}$ in global coordinates) can be placed at each eye or between them to generate eye and head relative positions, respectively.

Any known position in space, $\boldsymbol{P}$, can therefore be translated from the global $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{z}})$ coordinate system into $\boldsymbol{P'}$ in the $(\hat{\boldsymbol{F}}, \hat{\boldsymbol{L}}, \hat{\boldsymbol{U}})$ head-relative system, by a standard Cartesian coordinate transformation, Eq. 2.1. As the head can translate and rotate in a scenario where the head is not fixed, as was the case in the experiment, this translation is done at each timestep of the experiment for each stream of spatial data, including C-ET derived gaze coordinates.

$$\boldsymbol{P'} = \mathbf{R} * (\boldsymbol{P} - \boldsymbol{P_H}), \quad \mathbf{R} = \begin{pmatrix} \hat{\boldsymbol{F}} \\ \hat{\boldsymbol{L}} \\ \hat{\boldsymbol{U}} \end{pmatrix} \tag{2.1}$$

Theoretically, when the eyes fixate on a position in space, each eye rotates independently to move the image to the center of each eye's visual field. Ideally, rays projected from the center of each eye intersect at the fixation target. Therefore, knowledge of the position of the eyes, along with their angular deflections, can determine the position of the fixation, and in reverse, the knowledge of their positions and that of a target can determine their angular deflections. However, applying this ideal model directly in a practical manner has limitations. As the head can move triaxially, the eyes maintain their fixation on a target during pitch and roll movements, in part by compensating with their own pitch and roll (torsional eye movement). While the pitch and roll components are part of the physiological mechanism of fixation, pragmatically the image orientation of a single point-fixation can be neglected and, in normal activity, the roll of the head and eyes is minimal, allowing these components to be reasonably collapsed into solely a pitch relative to the horizontal axis of the eyes.

With this simplification we construct the ideal geometric model (Figure 2.2a), where

the position of the target, $\boldsymbol{T}$, has been translated into relative coordinates $\boldsymbol{T}'$ and used as the basis to determine gaze direction. With this basis as a reference, the gaze vector, bvS, defined by the direction cosines from an eye to the target position relative to that eye, $\boldsymbol{S}(F, L, U)$, can be found ($\boldsymbol{S} = \boldsymbol{T}'/\|\boldsymbol{T}'\|$). The deflections of these vectors can also be represented in angular form by Eq. 2.2, with the azimuthal (horizontal) angle, ψ, measured from the front vector to the left vector and the elevation (vertical) angle, θ, measured from the front vector towards the up vector (similar to the conventions used for yaw and pitch), together providing a heading towards the point of fixation.

$$\tan\psi = \frac{L}{F} \quad \sin\theta = U \tag{2.2}$$

Using such a heading for each eye, two sets of eye-relative direction cosines, $\boldsymbol{S}_R$ and $\boldsymbol{S}_L$, can be defined, providing the necessary degrees of freedom to triangulate a 3D fixation point relative to the head-eye coordinate system. Using the interpupillary distance, d, set between the two eyes along the $\hat{\boldsymbol{L}}$ axis, the fixation point occurs when the extension of the two vectors intersect, as expressed in parametric form in Equation (2.3), with λ as the parametric constant for that line and $\boldsymbol{d}$, as the vector equivalent of d, defined as the difference in position of the left and right eye ($\boldsymbol{E}_L - \boldsymbol{E}_R = (0, d, 0)$) in the head-eye relative coordinate system.

$$\begin{gathered} \boldsymbol{T}' = \boldsymbol{T}'_L = \boldsymbol{T}'_R \\ \boldsymbol{T}'_L(\lambda_L) = \lambda_L \boldsymbol{S}_L + \tfrac{\boldsymbol{d}}{2}, \quad \boldsymbol{T}'_R(\lambda_R) = \lambda_R \boldsymbol{S}_R - \tfrac{\boldsymbol{d}}{2} \end{gathered} \tag{2.3}$$

The fixation point in vector form, $\boldsymbol{T}'(F, L, U)$, can be calculated as:

$$\boldsymbol{T}' = \frac{dL_R(\boldsymbol{S}_L \cdot \boldsymbol{S}_R) - dL_L}{1 - (\boldsymbol{S}_L \cdot \boldsymbol{S}_R)^2}\boldsymbol{S}_L + \frac{\boldsymbol{d}}{2} \tag{2.4}$$

This, being a derivative of the the same as general solution for an intersection as found in

eq. (2.12). This equation has the equivalent angular form of:

$$\boldsymbol{T}' = \frac{dL_R \cos\gamma - dL_L}{\sin^2\gamma}\boldsymbol{S}_L + \frac{\boldsymbol{d}}{2} \tag{2.5}$$

where $\gamma = \arccos(\boldsymbol{S}_L \cdot \boldsymbol{S}_R)$, notable as the angle of vergence for the two direction cosines, and which clearly shows that when the gaze direction cosines are close to parallel $\csc^2\gamma$ becomes undefined and the estimated target jumps to infinity.

For cases in which the target is towards the center of the field of vision and the distance to the target is sufficiently larger than the distance between the eyes ($r \gg d$) then it can be assumed that the direction cosines to the target for both eyes are close to parallel, with $\psi_L \approx \psi_R$, and $\theta_L \approx \theta_R$. This allows the estimated angle of the subjects gaze relative to the head to be calculated as the average of the two eye angles and the depth to the target estimated as a function of the difference in these angles (Equation (2.6)).

$$\psi = \frac{\psi_L + \psi_R}{2}, \quad \theta = \frac{\theta_L + \theta_R}{2}, \quad r \approx \frac{d\cos\psi}{\cos\theta(\psi_L - \psi_R)} \tag{2.6}$$

Target Estimation (Regression & Realistic Model)

With the known fixation position, and the direction cosines to that position indicating the gaze direction of the eye, it is possible to construct a calibration between the EOG signal collected from around the eye, $\boldsymbol{V}$, and the gaze direction, $\boldsymbol{S}$. For small angles, this relationship between the measured EOG and the eye position is close to linear, which can be represented by the calibration matrix $\mathbf{W}$ in Equation (2.7).

$$\boldsymbol{S} = \mathbf{W} * \boldsymbol{V}, \quad \boldsymbol{S} = \begin{bmatrix} F_R \\ L_R \\ U_R \\ F_L \\ L_L \\ U_L \end{bmatrix}, \quad \boldsymbol{V} = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_8 \end{bmatrix} \tag{2.7}$$

With this relationship, the regression matrix can be computed, in Equation (2.8), by taking the product of gaze vectors to the target, $\boldsymbol{S}_0$, for a particular training trial and the pseudoinverse of corresponding EOG voltage signals, $\boldsymbol{V}_0$.

$$\mathbf{W_0} = \boldsymbol{S}_0 * \boldsymbol{V}_0^+ \tag{2.8}$$

This trained matrix can then be applied to EOG signals from a proximal trial to estimate the position of gaze, $\hat{\boldsymbol{S}}$, for that trial (Equation (2.9)). As components of the EOG signal from each eye are present in the EOG measurements across all of the electrodes, the transformation can be done with the direction cosines of both eyes together, agnostic to the distance between the eyes.

$$\hat{\boldsymbol{S}} = \mathbf{W}_0 * \boldsymbol{V} \tag{2.9}$$

With the predicted direction cosines for each eye, it is then possible to use them in the ideal geometric model. However, for practical use, there are some modifications to the model that are needed in order to better reflect realistic conditions. The ideal model assumes that the fixation target of both eyes is a single point in 3D, with the rays depicting gaze intersecting at a point; this is geometrically improbable in 3D, particularly as there is no guarantee of this with estimated gaze directions. Assuming that the gaze is directed

at the target, an estimation of target position should be estimated as a point where the two rays come as close to intersecting one another as possible, as seen in Figure 2.2b. In this case the parametric equations of the lines are the same as the ideal model, but the estimated fixation point, $\hat{\boldsymbol{T}}'$, does not lie on the either line and is the midpoint between $\hat{\boldsymbol{T}}'_L$ and $\hat{\boldsymbol{T}}'_R$, the points where two eye vectors make the closest approach:

$$\hat{\boldsymbol{T}}' = \frac{\hat{\boldsymbol{T}}'_L + \hat{\boldsymbol{T}}'_R}{2}, \quad \hat{\boldsymbol{T}}'_L = \lambda_L \hat{\boldsymbol{S}}_L + \boldsymbol{E}_L, \quad \hat{\boldsymbol{T}}'_R = \lambda_R \hat{\boldsymbol{S}}_R + \boldsymbol{E}_R \tag{2.10}$$

These two points can be found by adjusting λ_L and λ_R as to minimize the intervening distance:

$$\min_{\lambda_L, \lambda_R} : \|\lambda_L \hat{\boldsymbol{S}}_L + \boldsymbol{E}_L - (\lambda_R \hat{\boldsymbol{S}}_R + \boldsymbol{E}_R)\|^2 \tag{2.11}$$

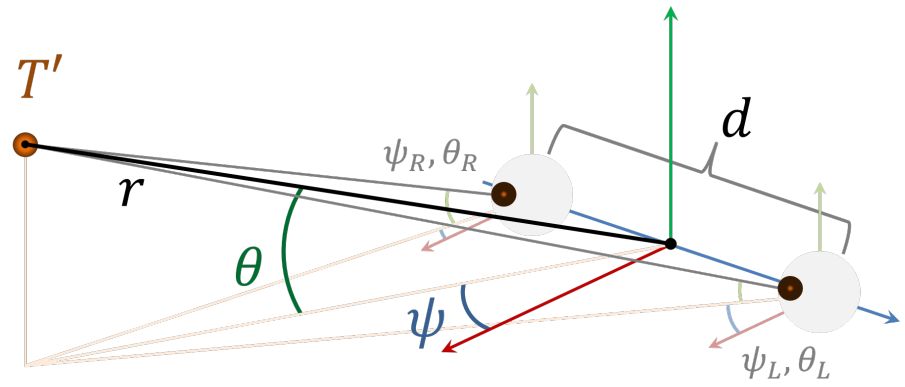
which occurs at:

$$\lambda_L = \frac{dL_R(\hat{\boldsymbol{S}}_L \cdot \hat{\boldsymbol{S}}_R) - dL_L}{1 - (\hat{\boldsymbol{S}}_L \cdot \hat{\boldsymbol{S}}_R)^2}, \quad \lambda_R = \frac{dL_R - dL_L(\hat{\boldsymbol{S}}_L \cdot \hat{\boldsymbol{S}}_R)}{1 - (\hat{\boldsymbol{S}}_L \cdot \hat{\boldsymbol{S}}_R)^2} \tag{2.12}$$

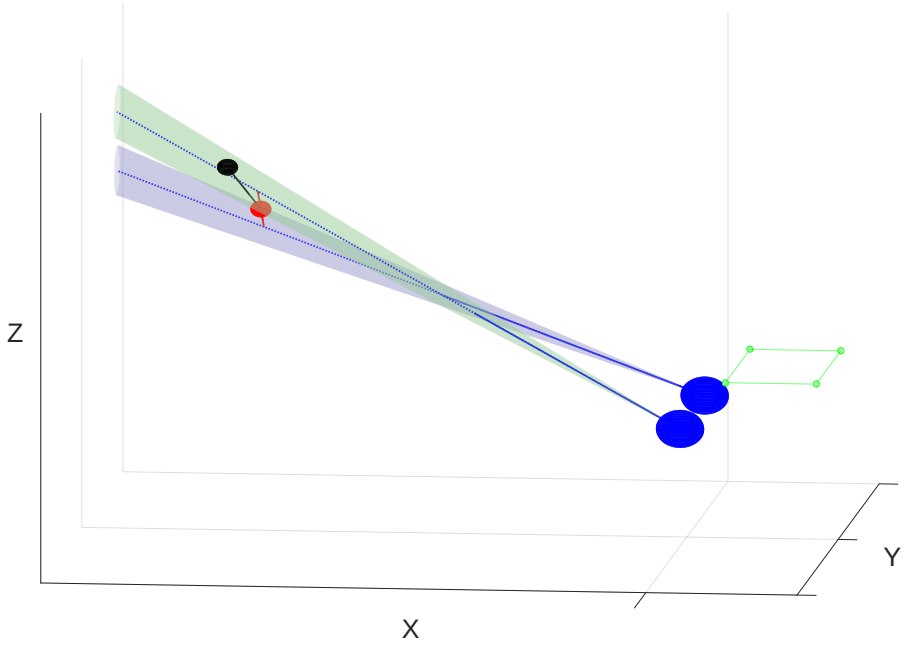
which, similarly to the ideal model, has a virtual vergence angle, $\hat{\boldsymbol{\gamma}} = \arccos(\hat{\boldsymbol{S}}_L \cdot \hat{\boldsymbol{S}}_R)$.

2.3 Results: 2D Capabilities

To examine the capabilities of EOG based eye tracking under conditions comparable to those generally used for C-ET, the EOG eye tracker was calibrated and tested in 2D. During the blocks of three 2D trials, each subject would track the target with their gaze as it moved across the plane of the screen. With the simultaneous recording of EOG and C-ET, the EOG eye tracker's estimation of gaze direction for each trial, based on a calibration performed using the remaining two trials, allowed for the comparison of these modalities against each other and the known fixation target's position.



(a)



(b)

Figure 2.2: **Diagrams of Geometric Models (a)** In the ideal physical model, each eye, set d apart along the $\hat{\boldsymbol{L}}$-axis, fixates on the target at $\boldsymbol{T}'$. The line of gaze to $\boldsymbol{T}'$ (black line) is defined by ψ, the azimuthal angle measured from the $\hat{\boldsymbol{F}}$-axis in the $\hat{\boldsymbol{F}}\hat{\boldsymbol{L}}$ plane, and θ, the elevation angle measured from the $\hat{\boldsymbol{F}}\hat{\boldsymbol{L}}$ plane vertically along $\hat{\boldsymbol{U}}$. A ψ and θ angle can be defined for both eyes; the intersection of the rays extending the direction cosines for each eye (grey line) is at $\boldsymbol{T}'$ and the resulting distance to the target, r, can be calculated. **(b)** In the more realistic model used to determine the estimated target position, there is a fixation target (black) toward which both eyes (blue) should be pointing. The estimated position of the target is the location where the rays extending the estimated direction cosines of each eye come nearest to intersecting, and is defined as the midpoint of the line spanning this minimal distance (red). Projected around the direction cosines are cones depicting regions of high visual acuity, with the area of their intersection being indicative of the uncertainty of the estimated target position.

2.3.1 Effectiveness

The angular direction to the target relative to the head, when decomposed into azimuth and elevation components, produced traces of sinusoidal curves for the position angles over the course of a trial (Figure 2.3). The target typically moved within azimuth and elevation angle ranges of 32.1 ± 3.3 degrees and 21.7 ± 1.1 degrees, respectively. A typical trace of the estimates of the gaze direction angles, shows that both EOG and C-ET estimations follow similar courses to that of the target. Over the course of the trial, the estimated position typically varied around the target, but with no discernible pattern (Figure 2.4). In some trials there was an offset, usually in the elevation, which typically had an error and noise higher than that of azimuth, and was likely due to the position of the vertical electrodes and/or artifacts caused by eyebrow movements. Systematic errors due to experimental design were assessed. No significant lag in estimated angular position and target positions was detected. Correlations between velocities and accelerations of the target against the angular errors were also negligible.

2.3.2 Comparison to Camera Based

Error Mode Comparison

The EOG eye tracker produced viable estimations of angular position. When compared to the known angular position of the target, the mean absolute angle error across subjects was $2.37^\circ \pm 0.84^\circ$ and $2.65^\circ \pm 0.76^\circ$, in azimuth and elevation, respectively (Figure 2.5). The performance in azimuth generally performed better, with a correlation R^2 value 0.96 of estimated to actual azimuthal angle across subjects, compared to an R^2 value of 0.84 in elevation. However, estimations produced by the EOG eye tracker were consistently larger than the C-ET, with a mean average azimuthal error of $0.81^\circ \pm 0.18^\circ$ and elevation error of $1.21^\circ \pm 0.19^\circ$, the comparisons for which can be seen in Table 2.1.

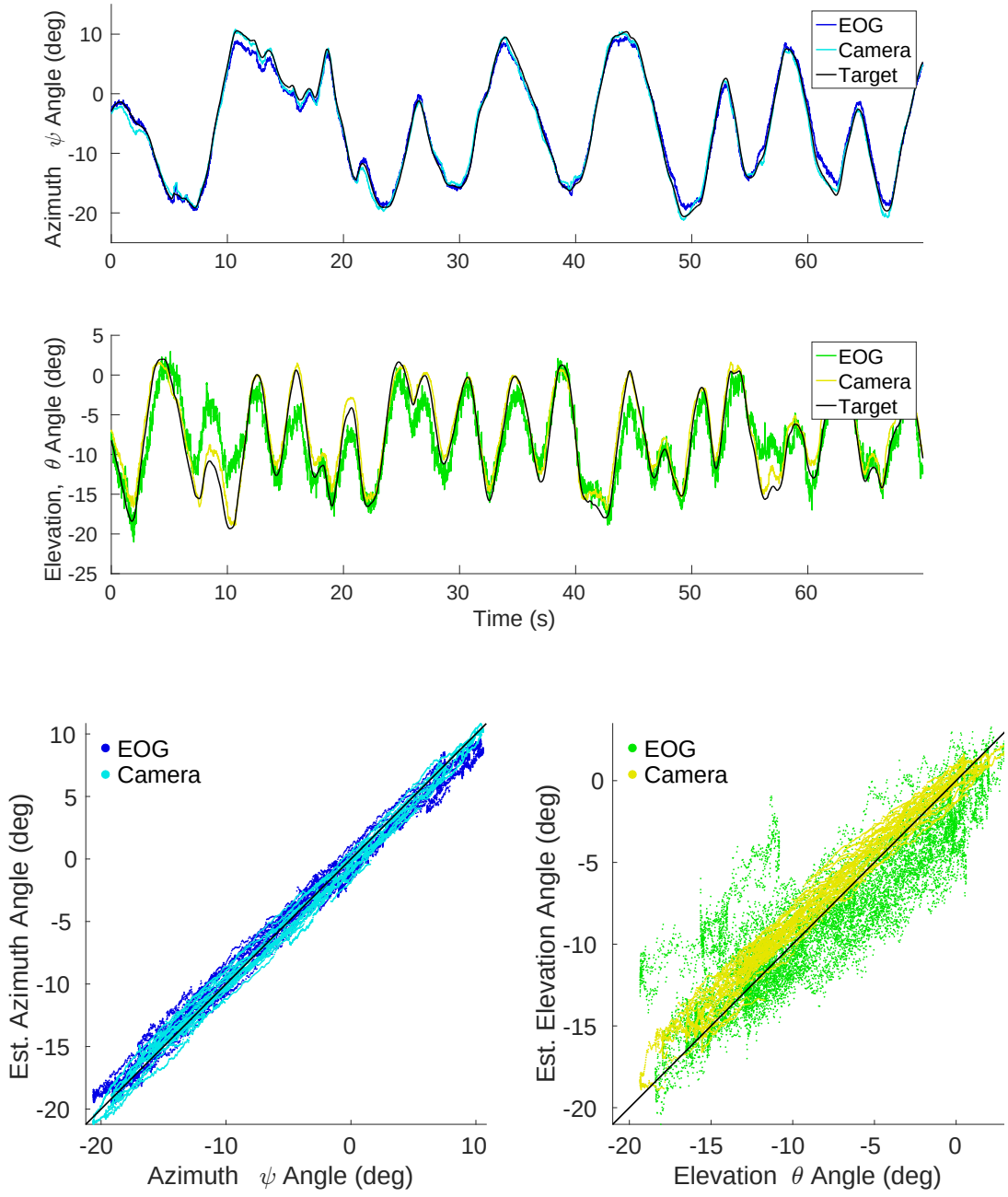


Figure 2.3: **Angular position in typical trial** Traces of angular position (ψ and θ) in a typical 2D trial showing the estimated angular positions from the EOG-ET and C-ET, tracking the known angular target position (black trace) over the course of the trial, along with the corresponding scatter plots of actual versus estimated angle values (with lines of symmetry in black).

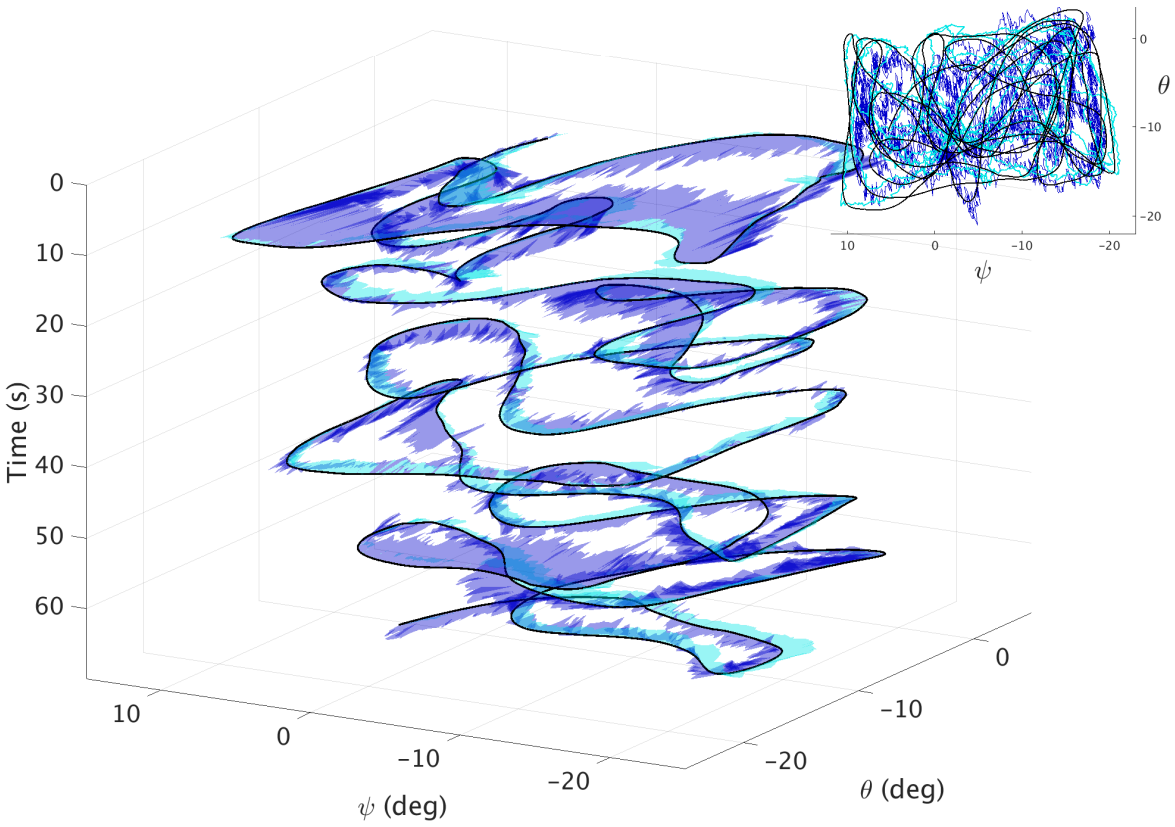


Figure 2.4: **Estimation discrepancies in a typical trial** The traces of the estimated and actual positions in 2D ($\psi\theta$) show the tracking of each modality (blue for EOG-ET, light blue for C-ET) along with the target position (black). To highlight the variable discrepancies between the estimated and actual $\psi\theta$ curves for the entire trial (seen in the inset) the data is displayed separated across the time of the trial and the area between the estimated and actual curves shaded.

Table 2.1: **Table of Angular Errors in 2D Estimations** The average discrepancy of azimuth ($\Delta\psi$) and elevation ($\Delta\theta$) angles and corresponding correlation coefficients between the angular position estimations, $\hat{S}_{EOG}$ and $\hat{S}_C$, of the two modalities, and with the known target position, S.

	$\Delta\psi$	$\Delta\theta$	R^2_ψ	R^2_θ
$\hat{S}_{\text{EOG}} : S$	$2.37 \pm 0.84^\circ$	$2.65 \pm 0.76^\circ$	0.96	0.84
$\hat{S}_C : S$	$0.81 \pm 0.18^\circ$	$1.21 \pm 0.19^\circ$	0.99	0.98
$\hat{S}_{EOG} : \hat{S}_C$	$2.36 \pm 0.92^\circ$	$3.00 \pm 0.96^\circ$	0.95	0.83

While part of the errors in accuracy for the two modalities might be attributed to bias in the subjects gaze, the finding that the errors between the EOG and C-ET estimations were higher and less consistent than those of either mode to the target indicates that a portion of the error was due to estimation errors of the modalities themselves, rather than an exaggeration of that bias.

Timescale Effect on Error

The temporal scale at which the EOG is employed as an eye tracker is important to its accuracy as such. As in most sensor systems, as time progresses after a calibration, the accuracy of the sensor decreases, this is particularly true for electrophysiological recordings, like EOG. In an EOG-ET, this results in increases in the error as time passes from the calibration. Some of this error was diminished by training the calibration using proximal trials and using a high pass filter on the EOG signal. However, electrophysiological sensor drift occurs unpredictably over multiple time scales, and cannot be easily compensated for.

The effect which time scale has on the accuracy of the eye tracking systems was measured by calculating and comparing the errors of each modality at different time scales. The error at a particular time scale can be approximated as the average of all errors for temporal windows of a particular length within a data set.

If a calibration can be considered correct at a particular point in time, t_0, then the relative error ($\Delta S'$) for a time a certain duration, Δt, later would be calculated by

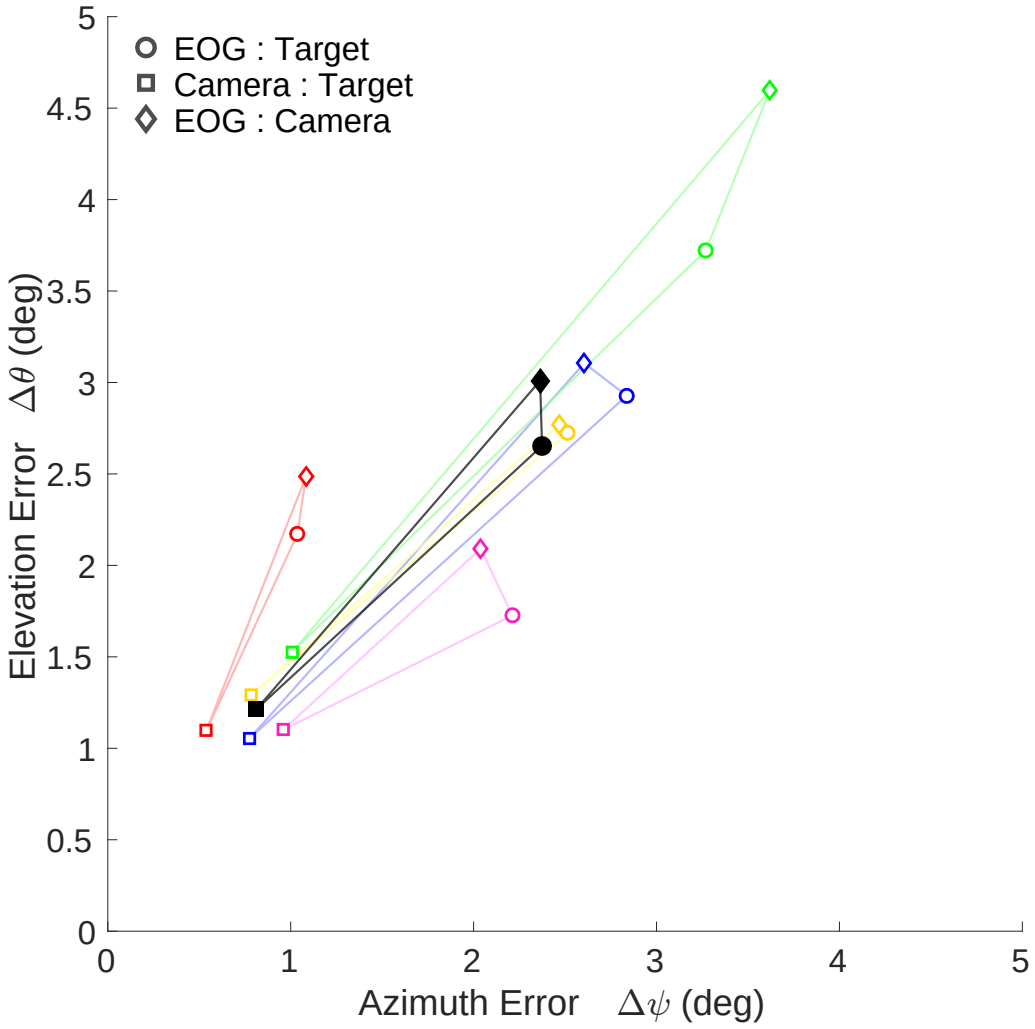


Figure 2.5: **Comparison of angular errors across subjects and modalities.** The angular error results of the two modalities, EOG-ET and C-ET, are shown being compared against the target (circle and square, respectively) and against each other (diamond) for each subject (represented in different colors). The average (black) azimuthal angle error, $\Delta\psi$, for the EOG-ET was $2.37 \pm 0.84^\circ$ compared to the C-ET error of $0.81 \pm 0.18^\circ$. The elevation angle error, $\Delta\theta$, was typically larger than $\Delta\psi$ for both modes, with an average of $2.65 \pm 0.76^\circ$ and $1.21 \pm 0.19^\circ$ for the EOG-ET and C-ET, respectively.

Equation (2.13), with δ_0 removing the offset of the estimation at t_0.

$$\Delta S'(t_0, \Delta t) = \Delta S(t_0 + \Delta t) - \Delta S(t_0) = [\hat{S}(t_0 + \Delta t) - S(t_0 + \Delta t)] - \delta_0 \tag{2.13}$$

This error can be calculated for every time point within the trial. The continuous average across a trial of the relative errors for each point Δt away from a previous time point t_0 is:

$$\overline{\Delta S'}(\Delta t) = \frac{1}{T} \int_0^T |\Delta S'_{\Delta t}(t_0)| dt_0 \tag{2.14}$$

Additionally, for a window of a specific length, Δt beginning at t_0, the error for that window could be written as the continuous average of all absolute errors on the interval of $[t_0, \Delta t]$:

$$\overline{\Delta S'}(t_0, \Delta t) = \frac{1}{\Delta t} \int_{t_0}^{t_0+\Delta t} |\Delta S'(t, \Delta t)| dt \tag{2.15}$$

These can be combined to generate the average error for every window of size Δt:

$$\overline{\Delta S'}(\Delta t) = \frac{1}{T} \int_0^T \frac{1}{\Delta t} \int_0^{\Delta t} |\Delta S'(t)| dt dt_0 \tag{2.16}$$

Which, when written in the discrete form necessary for digital calculations, becomes:

$$\overline{\Delta S'}(\Delta t) = \frac{1}{T} \sum_{n=1}^{T} \frac{1}{\Delta t} \sum_{n=1}^{\Delta t} |\Delta S_0(t_0, t)| \tag{2.17}$$

Using Equation (2.17), the angular errors for a range of timescales were calculated, ranging from the minimum sampling rate of that modality to 30 seconds (Figure 2.6). The error for the EOG-ET was consistently higher and more variable between subjects than that of the C-ET. However, as EOG can be sampled at a significantly higher rate than the C-ET, it was possible to resolve timescales smaller than those resolvable by the C-ET. For both modalities, and as expected, the angular error increased as the timescale increased.

The change in errors as timescale increased were similar in both modalities, until around $1-2$ seconds, where the C-ET errors began to plateau, while the rate of increase of the EOG-ET errors slowed more gradually, plateauing in some subjects more than others. The final values of the errors at the $10-20$ second timescales were generally similar to that of those calculated for the entire trials. This indicates that the viability of EOG for eye tracking is better for short durations, particularly at time scales shorter than the $10-20$ seconds where the error plateaus.

2.4 Results: 3D Capabilities

One characteristic of the EOG-ET is its capability for performing eye tracking in 3D space, with the ability to resolve distance to the fixation point, along with the angular position. Using the model described in Section 2.2.4, it was possible to assess the capacity of the EOG-ET to determine fixation position and depth to target, r. As the C-ET eye tracker was only capable of resolving pixel position on the screen based off of the gaze direction of a single eye, it was incapable of assessing the depth, and as the calibration was constrained to and defined only on the 2D space, the validity of its angular estimations for a target outside of this regime are largely inaccurate. Due to this limitation, the C-ET could not be used as a reference for the estimations of the EOG-ET in 3D.

2.4.1 Effectiveness

Depth to the target was calculated as the Euclidian distance from the target position to the center of the head (midpoint between the two eyes).

$$r = \sqrt{\left(\boldsymbol{T} - \tfrac{\boldsymbol{E}_L + \boldsymbol{E}_R}{2}\right)^2} \approx \|\boldsymbol{T}'\| \tag{2.18}$$

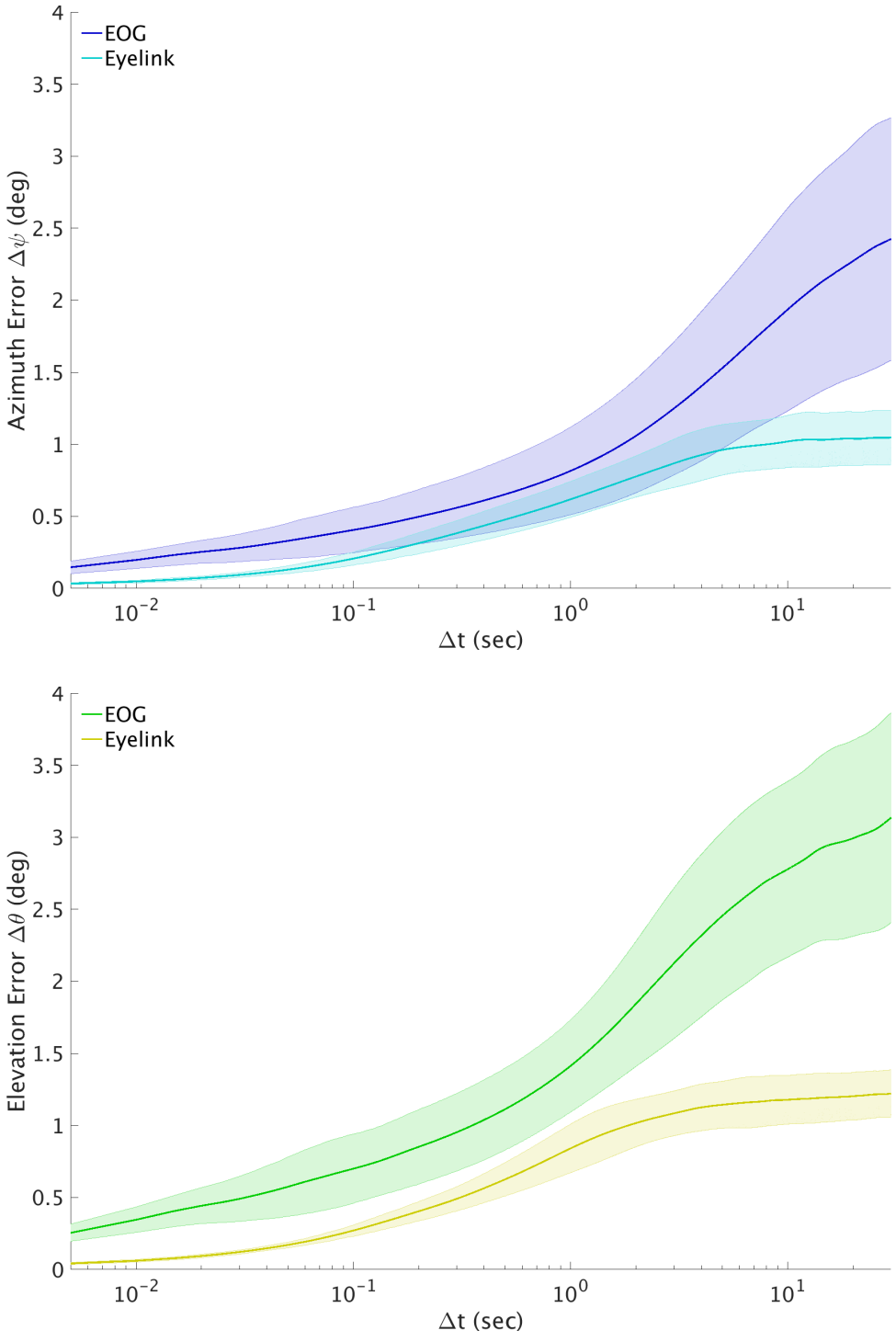


Figure 2.6: **Angular errors at various timescales.** The average relative angular error for all windows of length $\Delta t = (0, 30]$ seconds were calculated relative to the start of each of those windows to compare the accuracy of the EOG and camera based eye trackers at various timescales.

During the 3D trials, the target positions generally filled the space between the screen and the subject (Figure 2.7). On average, the typical target depth ranged from 0.79 meters to 2.9 meters from the subject, and had an angular range of 60° in azimuth and 45° in elevation. The average correlation R^2 value for subjects between estimated, $\hat{r}$, and actual target position was 0.32, corresponding to the average percent error for depth, $\Delta_0 r = |\hat{r} - r|/r$, of 25.3% across subjects, indicating that the EOG-ET was capable of some depth determination.

The behavior seen in Figure 2.7, with the depth predicted by the EOG-ET exhibiting similar movement behavior to the actual target depth, but being more compressed from both the proximal and distal ends, was typical for most subjects. Though compressed, these depth results can be used to indicate whether a fixation target is proximal or distal, which can be sufficient for many purposes, such as clarifying which depth region would be in binocular focus for a subject in augmented reality.

2.4.2 Uncertainty in Depth

Depth is difficult to ascertain, in part, due to uncertainty in the position of fixation. If there is a range of variation in position accuracy within the system, either within the visual system itself or the measurement system, then there is a degree of uncertainty in the predictions of position. Even if nominal in the determination of angular position, the interaction of the uncertainties between the two eyes compounds in the determination of depth.

Both the ideal model and the more realistic model assume that the visual acquisition of the target is a point of fixation defined by a line-of-sight from each eye. Part of this assumption is that each eye is truly pointed directly at the target and that the angle necessary for foviation on a target is precise. If these assumptions are relaxed, then for each eye there is a region of uncertainty around this precise line-of-sight. A cone extending

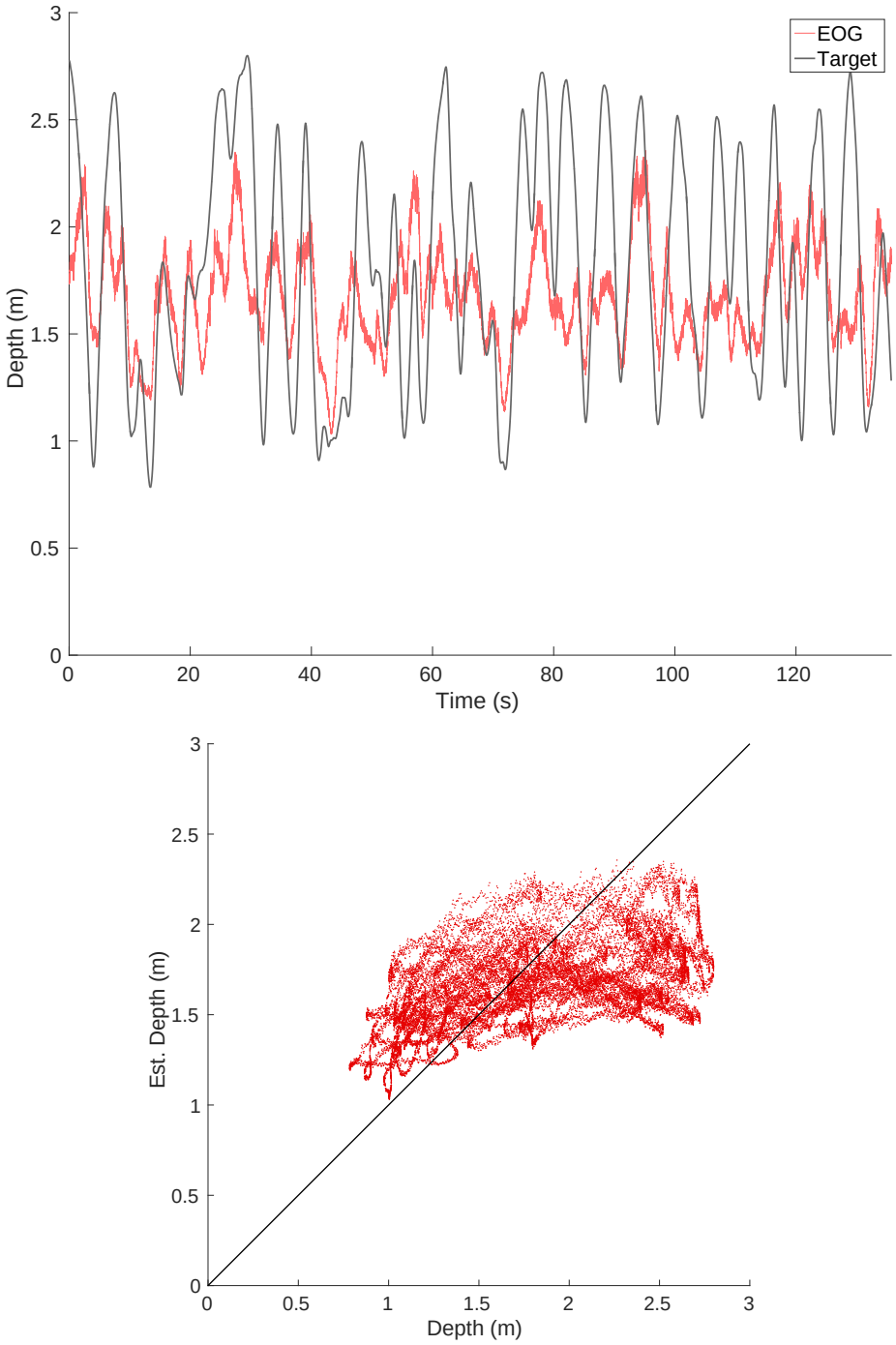


Figure 2.7: **Depth in a typical trial.** During a typical 3D trial the estimated depth (distance from target to subject) predicted by the EOG-ET is shown tracking the depth of the target. **(a)** The estimated depth generally has similar behavior to the actual. **(b)** However, a distal-proximal compression is seen in most trials resulting in significant depth estimation discrepancies, particularly at larger depths, possibly due to limitations in angular accuracy, as described in Section 2.4.2.

out from the eye around each of these rays, can be used to represent a region of uncertainty around the rays, within which it the target could be considered fixated on, as seen in Figure 2.8a. As a point gets further away from the eye the cone of uncertainty gets wider, which geometrically means that even for a fixed angular uncertainty, the euclidean size of the uncertainty grows with depth. In the case of binocular vision, this results in a relationship between the euclidean parameters of depth and the uncertainty of the depth, dependent on the angular parameters of the cone of uncertainty.

If a gaze vector has an angular direction to the target, Ψ, the cone of uncertainty around that vector can be characterized as having an apex angle of $\Delta\Psi$. When the target is directly in front of the subject ($\theta = 0$, $\psi_L - \psi_R = 0$) then Ψ and the distance to target, r, are related by Equation (2.19), in which $\Psi = \psi_L = -\psi_R$.

$$\tan \Psi = \frac{d}{2r} \tag{2.19}$$

The deviation in Ψ can be considered the derivative of this with respect to Ψ and r.

$$\frac{1}{\cos^2 \Psi} \Delta\Psi = \frac{-d}{2r^2} \Delta r \tag{2.20}$$

Resulting in Δr as the uncertainty in the estimation of the depth, which in a form relative to measured depth can be written as:

$$\frac{\Delta r}{r} = \frac{-2r}{d \cos^2 \Psi} \Delta\Psi \tag{2.21}$$

Using the fact that

$$\cos \Psi = \frac{r}{\sqrt{r^2 + (\frac{d}{2})^2}}$$

and

$$\tan\alpha = \tan\Psi + \tfrac{\Delta\Psi}{2}$$

Then the value of relative depth uncertainty, $\frac{\Delta r}{r}$, is determined by:

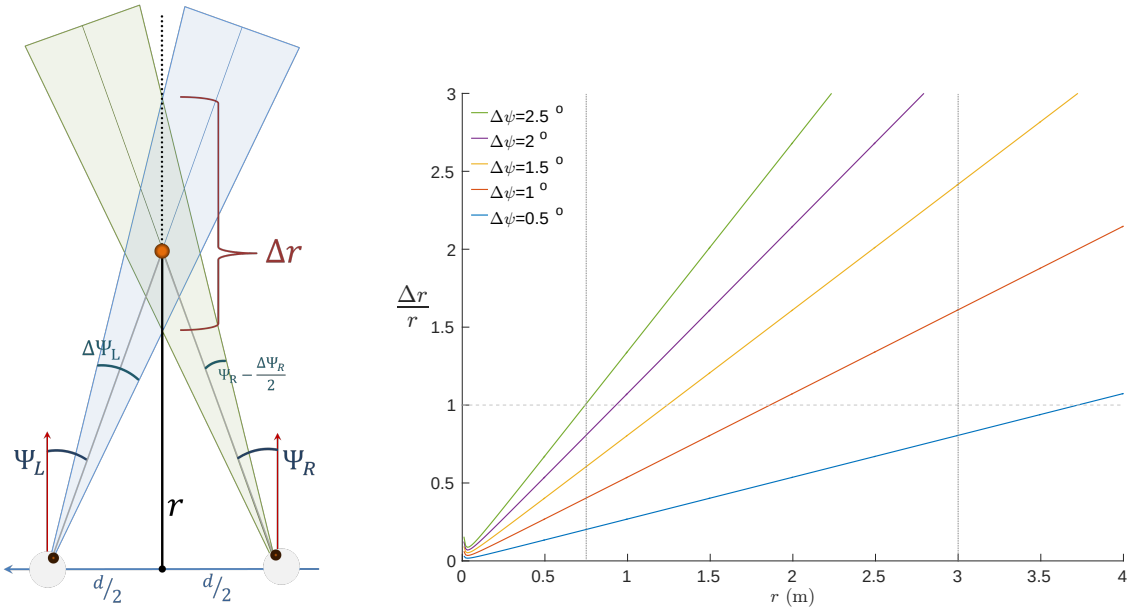
$$\frac{\Delta r}{r} = 2\left(\frac{r}{d} + \frac{d}{4r}\right)\Delta\Psi \tag{2.22}$$

This indicates that $\frac{\Delta r}{r}$ is linearly proportional to r when $r \gg d$, and to Ψ. With this dependency, the importance of $\Delta\Psi$ in being able to determine a depth accurately, is apparent (fig. 2.8b); at distances further away from the subject a small $\Delta\Psi$ is necessary to maintain a Δr small enough to be useful. This would indicate that even with accurate fixation on the target, the ability to determine depth using stereopsis, and the practicality of doing so, is limited for a depth beyond a certain point by moderate values of $\Delta\Psi$. This has physiological relevance as there are limitations on the spatial sensitivity of the visual optics. If the width of $\Delta\Psi$ is considered to be close to the accuracy of foveation, then a standard foveal field of view, of $1 - 2^\circ$ [39], results in a depth accuracy profile indicating that the (corresponding) relative error beyond $2 - 1$ m is greater than 1. A practical consequence of this is that as the depth determination by stereopsis becomes less reliable, other methods for estimating depth should become more dominant. Experimentally, it has been found that use of stereopsis tends to be significant at distances up to approximately $1.5 - 3$ m, after which motion parallax tends to dominate [45].

2.5 Discussion

2.5.1 Improvements

Overall the EOG performs reasonably as a basis for an eye tracking system. While having some advantages over C-ET and is a potentially valuable addition to eye tracking



(a) (b)

Figure 2.8: **Uncertainty of depth estimation (a)** Diagram of target fixation and cones of uncertainty, with target centered. The intersection of the visual field uncertainty for both eyes, each with an angular width of $\Delta\Psi$ yields an area in which the target could be and an uncertainty in depth, Δr. **(b)** Effect of $\Delta\Psi$ on $\frac{\Delta r}{r}$. The relative uncertainty in depth increases linearly with r when $r \gg d$ (approximate interpupilary distance, $d = 0.065$m) and increases faster with higher values of $\Delta\Psi$.

technologies available, the EOG-ET's performance was not superior to the C-ET. Some improvements should be made to make EOG more valuable as an eye tracking technology. Namely, in order to be used as a functional eye tracker the EOG-ET needs to more correctly identify gaze position over a usable duration, such as the course of an experiment.

A few improvements could be made to improve the accuracy of the system. One of the improvements most likely to have a significant impact would be to remove eye blinks. While the frequency of blinks within the data set is relatively low, they produce a large error in the EOG signal, particularly in elevation, which may affect the accuracy of the EOG-ET calibrated in a trial with numerous eye blinks. As the blink signal appears primarily as a sudden vertical shift, similar to that of a quick saccade, this may partly account for the reason why the elevation estimations of the EOG-ET were generally less accurate than those of azimuth. Removing eye blinks may also decrease the error as compared to the C-ET; blinks were reported by the C-ET as a complete loss of signal, necessitating and facilitating their removal, which was not the case with the EOG signal. Another improvement would be a more accurate reporting of head, and corresponding eye, locations during the calibration. Transitory signals in more than one of the motion capturing LEDs leads to a jitter in the determination of head and eye positions, which can distinct effect on position estimation, particularly that of depth. It is also possible that incorporating velocity or some other form of signal memory into the calculations used to develop the calibration could improve the ability to resolve depth, as it would limit the sudden estimation jumps. This however, could potentially limit the ability for a smooth pursuit calibration to be used in the position estimation for conditions in which large, fast eye movements occur.

Accuracy over time is also also an important area which the EOG-ET could be improved. Applying a high-pass filter selected to allow the frequency of the target movements to be seen was effective at removing much of the electrode drift over the course of the experiment. However, there appeared to be shorter periods of drift within trials which were

not well corrected by this. Applying different filters or performing a non-linear calibration may potentially decrease some of the errors caused by such drift.

One of the functional improvements of the EOG-ET would be implementing it for use in real time, allowing for immediate monitoring and feedback based on eye position. The calculations used to perform the EOG-ET calibration were explicitly selected so that it would be adaptable to a real time calibration and target estimation. The capacity to use this real time eye tracker for an extended period of time would be significantly improved by providing a method so that the calibration could be easily updated periodically.

2.5.2 Applications of EOG-ET

While the usage of EOG as a method of eye tracking has its limitations, there are a myriad of potential uses for an EOG-ET. One of the primary potential benefits of an EOG-ET is that systems capable of measuring EOG can be relatively simple, portable, and cheap. While the biopotential recording system used in this experiment was an expensive and non-portable EEG system designed primarily for brain research, a much less complicated system could be used. A simple EOG-ET would require only a few electrodes, a recording system, and a method for performing calibration procedures. The requirements for the processing and transmission of the input signals necessary to produce a gaze estimation from EOG are considerably less intensive than those necessary for the extraction of gaze information from images or the transmission of the corresponding video. This makes an EOG-ET more amenable for implementation with wireless, small, and/or low power technologies. As many EEG systems are capable of detecting EOG, and invariably do during EEG experiments, it would be relatively easy to implement an eye tracking calibration to estimate gaze position in addition to the EEG signal collection, and without the need of additional monitoring technology. This could potentially even be applied to previously collected EEG datasets, which were not designed with eye tracking in mind; as long as there were times during

which the target was known to be fixating on a particular location or object, the EOG signal from these periods could be used as a form of calibration.

Another advantage of an EOG-ET is that EOG is invariant to lighting and does not have a minimal focal depth, which is a challenge for C-ET. An EOG-ET can function in cases where the ambient lighting conditions vary dramatically, such as outdoors or in another visually busy environment. Nor does it have the need to provide its own illumination source, like many C-ET do [64], which reduces an EOG-ET's power requirements. Without such limitations as these, EOG could be a good candidate for an eye tracker in virtual reality systems, a situation where lighting conditions are highly variable, the eyes are generally occluded from where a camera could be mounted, and an apparatus of is often already attached to the head. As the usage and popularity of virtual reality is increasing, both in research and the general public [16], and that they can present a subject with a great variability of visual stimuli, having an eye tracker which can work well in conjunction with these systems would be quite useful.

2.5.3 Other Capabilities of EOG

Beyond using EOG for a general purpose eye tracker, there are other related uses for which EOG could have potential. Just as EOG is agnostic to illumination, it does not need the eye to be visible at all, and could therefore be used to perform "eye tracking" with closed eyes. EOG has been long used in polysomnography, as it is able to detect the eye movements which characterize REM sleep, though primarily for the purposes of identifying this behavior [5], without specific interest in the orientation of the eye movements themselves. An example of a potential usage for eyes closed EOG-ET would be for the investigation of eye movements where visual feedback would be a confounding factor, such as responses to surprising non-visual stimuli.

Due to the ability for EOG to be collected at a relatively high sampling rate it

has the potential to enhance currently existing eye tracking technology. One potential application is using the EOG signal to augment the position estimated by that of a C-ET, allowing for the interpolation of eye movements beyond the temporal resolution of a C-ET's normal frame rate and the recovery of segments where the C-ET lost its optical isolation of the eye, as often happens with eye blinks. The reverse of this is also a potential benefit of EOG incorporation, with EOG enhancing the ability to detect the occurrence of a blink. As some blinks happen very rapidly, a C-ET may not accurately identify the end points of the blink or even register that there was a blink, which would be well within the sensitivity of an EOG recording. Such an incorporation also allows the C-ET's spatial calibration to be used to improve the spatial accuracies of the EOG-ET and decrease the impact which electronic noise has on the determination of position.

2.6 Acknowledgements

While this chapter is original to the dissertation, some of its contents are being prepared for eventual publication. I would like to recognize and thank my advisors on this project, Dr. Gert Cauwenberghs and Dr. Tzyy-Ping Jung, who will be included as co-authors when this work is published. I would also like to thank the Swartz Center for Computational Neuroscience (UC San Diego; San Diego, CA) for usage of their facilities for data collection and the support of their technical staff for helping me setup and run the equipment used in the experiments. I would particularly like to thank J.T. and S.H., my labmates who helped pilot the experiment and allowed me to take and use their images.

Chapter 3

Virtual Environment Human Navigation Task for Studying the Impact of Glaucoma on Navigation

3.1 Introduction

3.1.1 Navigation

Wayfinding is the process of determining and following a path or route between an origin and a destination [26]. As a process humans use to orient and navigate on foot or in a vehicle, wayfinding is an integral part of people's daily life. It is involved in many situations, such as walking inside a home or a building, navigating through a city, or driving across a country. Wayfinding requires the proper encoding, processing, and retrieval of spatial information, during which visual cues are often the main source for perception of the environment. With vision, a traveler can acquire information about the surrounding environment from which an intended travel path can be planned and obstacles detected and avoided.

Egocentric & Allocenteric

Various cognitive strategies are used when performing navigation. These strategies can be categorized into two groups, egocentric strategies and allocentric strategies. The defining characteristic of these strategies is the reference frame which is used to encode the location of oneself and other objects in the environment. An egocentric strategy constructs a representation of the world using the position of proximal landmarks relative to oneself and ones heading and path history [2]. An example of a primarily egocentric navigation method is a series of turn-landmark directions (e.g. drive down Main Street, make a left before the library, continue to until you see the big tree on your right, etc). This contrasts with an allocentric strategy where a more map-like representation of the world is constructed and position is defined based on the relative positions of objects to each other, with the position of oneself being one of these objects [2]. An example of this is converting back and forth between a map and a position (e.g. drawing a map to indicate one's location). Most navigation problems can be solved with either strategy or some combination of the two, though many individuals have a tendency to use a particular strategy under regular conditions [15].

These two navigational strategies have been found to differ by the implicated brain regions [32], and consequently certain navigational situations are significantly more of a challenge when one strategy is used versus the other. The parietal lobes and caudate nucleus have been identified as primary structures for egocentric navigation [32, 65]. Typically, egocentric strategies are more suited and used for well traveled paths. They are effective at navigating along a path repeatedly, but have difficulty when it is necessary to take a new route. Alternatively, allocentric strategies have been found to more heavily utilize the medial temporal lobe and hippocampus [32, 65]. Allocentric strategies typically allow for more flexible route options than egocentric strategies, as the representation of the routes is not dependent on turn-based directions. They are also critical in being able to draw a map

which can be used for a multitude of paths. Allocentric strategies are, however, are more intensive, requiring some capacities of egocentric strategies and more working memory [54], and are consequently more impacted when memory is impaired. Conditions which impact the structures critical to these strategies accordingly affect the ability to utilize the associated strategies.

3.1.2 Virtual Reality

Virtual Reality (VR) is a type of technology which creates an immersive, realistic, and typically, 3D environment virtually for the person using the system. One of the key properties of VR is that it is capable of producing geometrically accurate depictions for the user, combining coverage of a large portion of a subject's field of view and methods for generating the perception of depth. Directional sound, motion tracking, and systems for interacting with the virtual world are often used to augment the virtually generated environment. The important component of VR, a convincing illusion of depth, is generated by taking advantage of the ways that human vision determines distance. This is done through stereopsis, with identical, but slightly shifted, views presented separately to each eye (see Section 1.1.2), and through accurately providing a sense of motion parallax, which is produced by monitoring the movement of the head and changing the images presented accordingly [49].

There are two primary categories of VR systems, differing in the ways in which the virtual environment is presented to the user and the method of providing stereopsis. The first group of VR are head-mounted display (HMD) systems. In HMDs a device is placed on or around the user's head and the two images are provided directly to each eye by screens within the device, generally in conjunction with the complete occlusion of the non-virtual world. The second type of VR are cave automatic virtual environment (CAVE) systems; these use multiple panels of screens and/or projected images, in conjunction with

polarized or frame-synced glasses, to generate depth in a virtual room [10]. Regardless of technology, both types of VR require significant amounts of computational power in order to render the often complex and dynamic geometries at a rate sufficient for real-time human interactions.

The differences in the technological bases of these VR systems yields some important functional and practical differences. HMDs are lighter and easier to transport than even small CAVEs. Recently, due to the commercialized development and popularization of consumer models, the prices of HMDs of increasing quality, have dropped [38]. Historically, the movement of subjects using HMDs was limited, as it was difficult to accurately determine the virtual position and movements of the user's head necessary for presenting a correct view and/or due to constraints of tethering of the HMD to computer generating the virtual world. Technological advances have improved the movement capabilities of HMDs, making them flexible in a variety of scales. CAVEs, on the other hand require a relatively large, fixed amount of space (depending on the size of the CAVE), within which the users are free to move. CAVEs provide the capability of allowing multiple users to view each other and themselves within the same virtual environment, without the need to develop and render virtual avatars of the users [11, 10]. The maintenance of embodiment makes CAVEs useful for collaborative interactions and in cases where the loss of embodiment could be disturbing or detrimental to the user, such as what may occur in aged or individuals with motor impairments.

Regardless of the technology used, VR has great potential for research, especially of cognition and navigation. A critical benefit is the ability to precisely control the environment, so that realistic conditions can be presented while restricting confounding stimuli. Constructing physical apparatuses to present stimuli and employing built, to-scale models of an environment is difficult, costly, and generally limited to the the design of the particular study at hand; VR helps to overcome these research hurdles. In addition VR is

able to provide precisely timed and located stimuli that would be physically impossible to create and has an improved ability to monitor and record meaningful subject data, such as the biopotentials used in E*G.

3.1.3 Health & Aging

In addition to navigation ability varying between individuals, it can also vary throughout an individuals lifetime. The senses and processes necessary for efficient and successful navigation can be impacted by a variety of physiological factors. These factors can be impacted by various conditions such as limited visual capabilities [51] or neurological conditions [56, 32, 33]. Some of these are typically age related conditions such as mild cognitive impairment (MCI), dementia, and glaucoma. However, even normal, healthy aging has been shown to have an impact on the navigation ability and strategies [63, 32, 6].

Glaucoma

Glaucoma is a disease of the eye in which a progressive loss of vision occurs. It is characterized by damage to the optic nerve, which results in the degeneration of cells in the retina and an associated loss of vision [1]. While occurring for a variety of reasons, some factors of which have not been clearly identified , an elevated intraocular pressure has been implicated in the development of glaucoma [24]. During the progression of the disease, namely in open-angle glaucoma, the most common form of glaucoma [9, 59], vision is lost gradually, starting in the periphery and moving towards the center of the visual field [37]. This visual field loss is asymmetric, both between the two eyes and within each eye's visual field. This narrowing of the visual field eventually leads to tunneled vision and complete blindness [37]. Typically, most forms of glaucoma are age related, with open-angle glaucoma, which accounts for three quarters of all glaucoma cases [9, 59], having a worldwide prevalence 3.05% for populations older than 40, with the odds of glaucoma

developing increasing by a ratio of 1.73 for each decade of age [59]

As a progressive optic neuropathy, glaucoma is a prototypical disease causing peripheral visual field loss. The disease is the main cause of irreversible blindness in the world and visual field loss from glaucoma has been shown to impact many daily activities [61, 13, 1, 37, 51]. However, there has been a lack of studies on the impact of glaucomatous visual field loss on efficient navigation. By artificially restricting the extent of the field of view, previous studies demonstrated an association between navigational ability and visual field [47, 22, 34]. Lovie-Kitchin *et al* [34] found that the smaller the solid angle of visual field subtended at the eye, the poorer the performance on a navigation task. Turano and colleagues evaluated mobility performance in glaucoma by investigating the time taken to navigate a established physical travel path [61]. However, their study did not address whether glaucoma subjects had increased difficulty in wayfinding tasks.

3.1.4 Motivation

There are several difficulties in the study of navigation and wayfinding in humans within the context of real-world tasks. Some derive from the large spatial scale of these tasks, such as making the modification or removal of certain visual cues in systematic ways difficult. Real-world navigation experiments may also be difficult for older subjects due to the associated physical effort necessary for repeated observations. New technological advances have enabled researchers to conduct experiments in virtual reality (VR) settings in lieu of a traditional, real-world setup. VR experiments may elicit similar responses from human beings as in real situations [62], with the advantage that experimental conditions can be easily manipulated in systematic and reproducible ways, allowing hypothesis testing of specific factors which influence wayfinding behavior [56]. Furthermore, during VR experiments, participants can have physiological biomarkers and behavior easily recorded throughout multiple experimental observations.

The purpose of the present study was to investigate wayfinding behavior and spatial cognition in patients with glaucoma and healthy individuals using an immersive VR experiment platform, the Virtual Environment Human Navigation Task (VE-HuNT).

3.2 Methods

3.2.1 Virtual Reality Environment

VE-HuNT was implemented using a derivative of the CAVE virtual reality system [11], the 4kAVE. The 4kAVE consists of three vertically oriented 4000 pixel resolution, 80-inch, 3D LED screens, with the lateral screens aligned at 45° relative to the central screen (Figure 3.1) Proprietary software developed at UC San Diego for CAVE systems created spatial displays that engage approximately 150° of a person's field of view. Subjects sat at a narrow table, wore polarizing glasses to discern a 3D image during testing, and navigated through the virtual environment using a steering wheel and an accelerator pedal. A researcher would use a tablet to control the 4kave and monitor the subject's virtual position and experimental progress in real time. VE-HuNT assessed subjects' spatial cognition abilities while performing wayfinding tasks in different sparsely filled virtual rooms with varied arrangements of visual cues, such as colored walls, paintings on the walls, and 3D objects.

Paradigms

Subjects were given an initial training session to familiarize themselves in navigating through the virtual environment. During the training session, subjects were virtually placed near a wall in a square room then asked to "drive" along a green-tiled circular path to a target located in another location in the room (Figure 3.1a). This training paradigm was repeated until subjects were comfortable and adjusted to the virtual environment and the

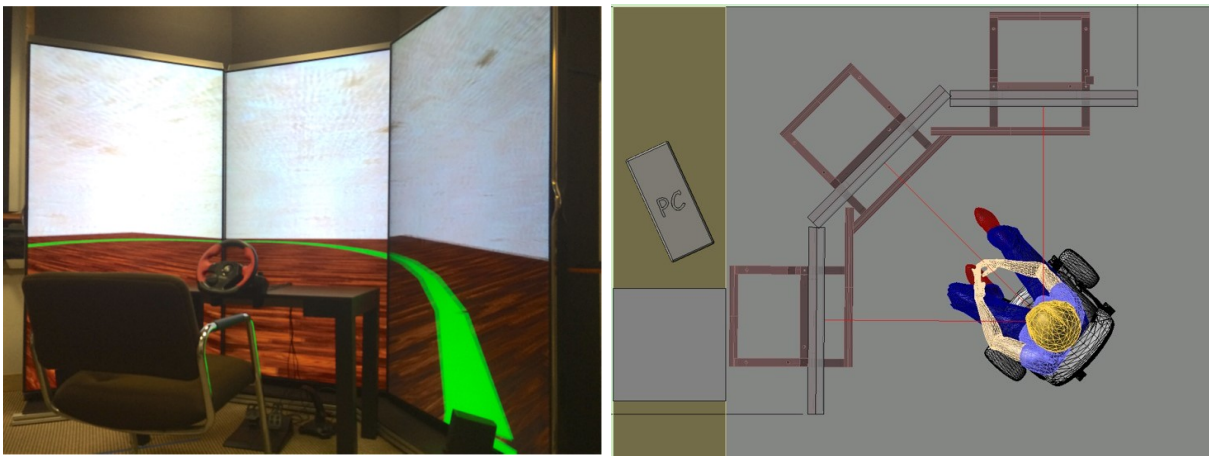
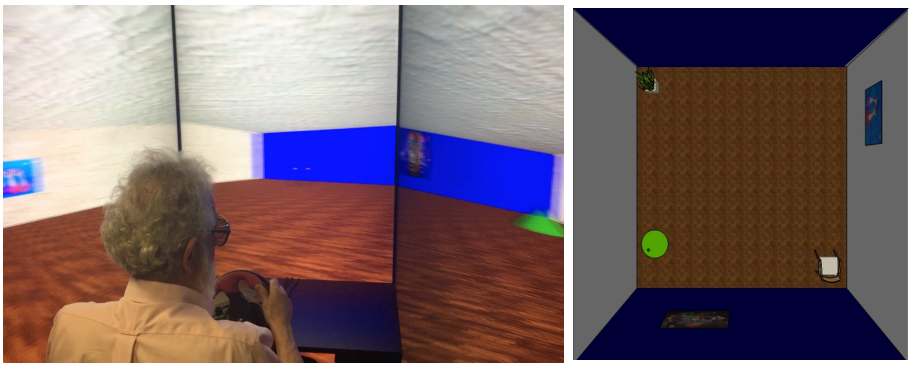


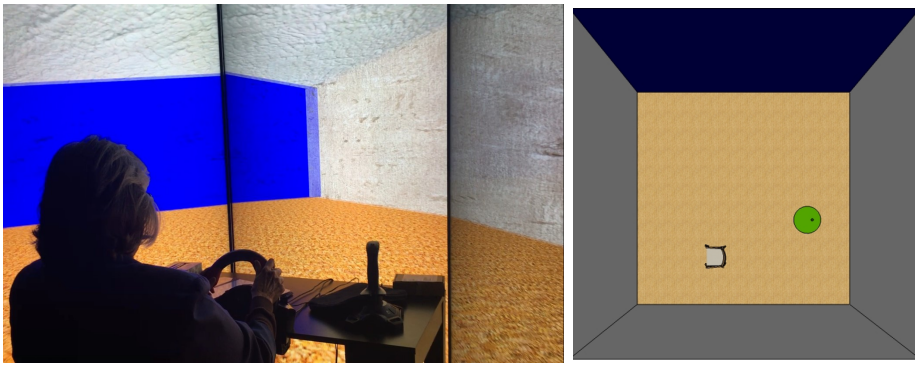
Figure 3.1: **VE-HuNT on the 4kave System (a)** The Virtual Environment Human Navigation Task was a software platform used to study navigation by employing an immersive virtual environment and wayfinding task paradigms. **(b)** VE-HuNT was presented via the 4kave virtual reality system, which consisted of 3 80-inch, 3D LED screens and a steering wheel and pedal interface, which the subject used to move through the virtual environment.

controls and the experimenter was confident that the task would be manageable for them.

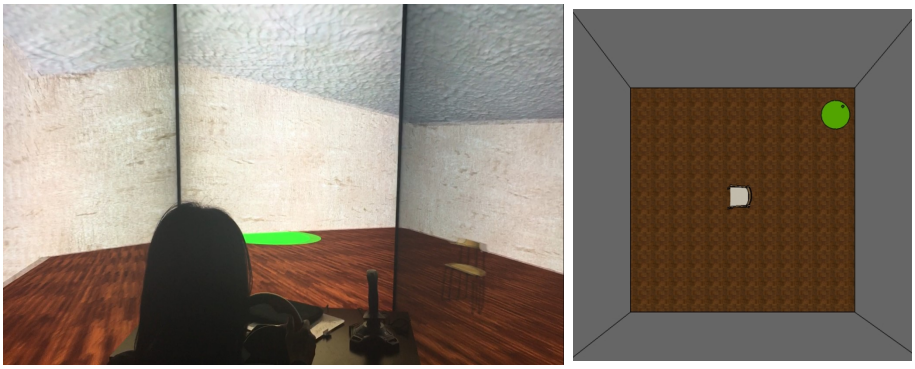
After the training session, subjects were asked to perform the actual wayfinding series of experiments, which were modeled after the Morris water maze classically used to assess spatial cognition in rat models [23]. Three virtual rooms were built. The 3 rooms were geometrically identical (20m x 20m x 3m), but contained a different set of visual cues designed to assess the impact of such visual cues on navigation and spatial cognition. Room 1 contained multiple peripheral visual cues, with 2 large symmetric colored wall cues, 2 asymmetric wall cues (paintings), and 2 asymmetric floor cues (chair and plant, both located peripherally) (Figures 3.2a and 3.2b). Due to the presence of multiple visual cues, this room promoted the use of an allocentric reference frame and cognitive mapping to assist wayfinding (see Section 3.4). The second room (Room 2) contained less visual cues, with 1 large colored wall cue, and 1 ambiguously placed floor cue (a chair, located asymmetrically further from the room) (Figures 3.2c and 3.2d). The third room (Room 3) contained only 1 central cue, a chair near the center of the room with a fixed angle, giving



(a & b) Room 1: Multiple cues emphasizing allocentric navigation, allowing for the use of relational positions. (2 bilateral colored walls; 2 asymmetric wall cues; 2 symmetric, but distinct, floor cues)



(c & d) Room 2: Sparser cues encourage use of egocentric navigation. Multiple cues cannot be seen from most positions. (1 colored wall; 1 spatially offset chair; non-directional floor pattern)



(e & f) Room 3: Single cue enforces usage of egocentric navigation. Positional memory is needed to locate the target. (1 centered chair, where fixed angle gives directional information)

Figure 3.2: **The 3 VE-HuNT Rooms** Each paradigm varied the number and type of wayfinding cues, with the first room being easier to navigate, and subsequent rooms more difficult with sparser cues. Each subject performed multiple trials of driving around each virtual room and performing the location finding task.

a subtle directional cue (Figures 3.2e and 3.2f). The presence of a single cue promoted the use of egocentric reference frame based navigation, decreasing the need of a cognitive map for successful wayfinding.

Subjects were initially instructed to drive to a clearly indicated location in each one of the virtual rooms (visible target), identified by a green cone on the floor. Subjects were instructed to pay attention to the surroundings while driving to the target. Subsequent trials were conducted requiring the patient to revisit the initial target location, but the location was no longer visibly marked (hidden target). The target location would reappear when the subject's path intersected with the target area (the base of the cone), terminating the trial. A maximum time of 120s was alloted to complete each trial, at which point the trial would automatically terminate regardless of whether the target had been found or not. Five trials with the target hidden were obtained for each room with the subjects placed at a different starting position and heading in each. An additional trial with the target visible occurred after the first hidden task in each room, to allow subjects who had not noted the target location during the first exposure to do so. There was a final trial, where the target was removed and therefore the subject was unable to trigger a termination; the maximum time for this trial was abbreviated to 30s to minimize subject frustration. The position of the subject in the virtual room was continuously recorded. The time taken to complete each trial and subject path were extracted and and used as outcome metrics.

3.2.2 Subjects

The study included 31 glaucoma patients and 20 normal control subjects. Table 3.1 shows demographic and clinical characteristics of included subjects. Mean age of glaucoma participants was similar between the two groups (72.8 ± 7.1 years vs. 68.7 ± 8.4 years; $P = 0.069$). There were also no statistically significant differences in gender and race between groups. As expected, glaucoma subjects had significantly worse visual field

Table 3.1: Demographic and clinical characteristics of subjects included in the study. Values are presented as mean ± standard deviation, unless otherwise noted.

	Glaucoma (n = 31)	Control (n = 20)	P-value
Age, years	72.8 ± 7.1	68.7 ± 8.4	0.069
Gender, n (%) female	11 (35)	12 (60)	0.086
Race, n (%)			
White	25 (81)	14 (70)	0.283
African American	4 (13)	6 (30)	
Asian	2 (6)	0	
MD SAP 24-2			
worse eye, dB	-13.5 ± 8.6	-0.8 ± 1.0	< 0.001
better eye, dB	-6.5 ± 6.4	-0.8 ± 1.0	< 0.001
binocular, dB	26.9 ± 5.7	31.0 ± 1.1	< 0.001
MoCA score	29.0 ± 1.3	28.5 ± 1.2	0.202
SBSOD Score	53.6 ± 25.1	54.1 ± 24.8	0.948

sensitivity compared to controls, with average binocular MS of 24.2 ± 5.9 dB versus 30.9 ± 1.0 dB, respectively ($P < 0.001$).

This was a cross-sectional study conducted at the Visual Performance Laboratory of the University of California, San Diego (UCSD). Written informed consent was obtained from all participants and the institutional review board and human subjects committee approved all methods. All methods adhered to the tenets of the Declaration of Helsinki for research involving human subjects and the study was conducted in accordance with the regulations of the Health Insurance Portability and Accountability Act.

Vision Evaluation

All participants underwent a comprehensive ophthalmologic examination including review of medical history, visual acuity, slit-lamp biomicroscopy, intraocular pressure measurement, gonioscopy, ophthalmoscopic examination, stereoscopic optic disc photography, and standard automated perimetry (SAP) using the Swedish Interactive Threshold Algorithm with 24-2 strategy of the Humphrey Field Analyzer II, model 750 (Carl Zeiss Meditec, Inc., Dublin, CA, USA). Visual acuity was measured using the Early Treatment

Diabetic Retinopathy Study chart and letter acuity was expressed as the logarithm of the minimum angle of resolution. Only subjects with open angles on gonioscopy were included. Subjects were excluded if they presented any other ocular or systemic disease that could affect the optic nerve or the visual field. Optic nerve damage was assessed by masked grading of stereophotographs.

Glaucoma was defined by the presence of repeatable abnormal SAP tests (pattern standard deviation with $P < 0.05$ and/or a Glaucoma Hemifield Test outside normal limits) and corresponding optic nerve damage in at least one eye. Control subjects had no evidence of optic nerve damage and had normal SAP tests in both eyes. Only reliable SAP tests were included (less than 33% fixation losses or false-negative errors, and less than 15% false-positive errors). The severity of visual field loss was represented by the binocular mean sensitivity (MS) of SAP 24-2 test. Binocular MS was calculated as the average of the sensitivities of the integrated binocular visual field obtained according to the binocular summation model described by Nelson-Quigg et al.[44] According to this model, the binocular sensitivity can be estimated using the following formula:

$$\text{Binocular Sensitivity} = \sqrt{s_R^2 + s_L^2}$$

where s_R and s_L are the monocular threshold sensitivities for corresponding visual field locations of the right and left eyes, respectively. In order to calculate the binocular sensitivity from the formula above, light sensitivity had to be converted to a linear scale (apostilbs) and then converted back to logarithmic scale (decibels).

Cognitive Assessment

Subjects also completed a general cognitive impairment test, the Montreal Cognitive Assessment Test (MoCA) [43]. The MoCA test is a cognitive screening tool developed

to detect mild cognitive dysfunction. It assesses different cognitive domains: attention and concentration, executive functions, memory, language, visual-constructional skills, conceptual thinking, calculations, and orientations. It consists of a one-page, 30-point test administered by a trained technician in approximately 15 minutes. The total possible score is 30 points, with a score of 26 or above considered normal. Only subjects with normal scores were included in the study.

Subjective Assessment of Spatial and Navigational Abilities

The subjective assessment by each individual of his or her spatial and navigational abilities was evaluated by the Santa Barbara Sense of Direction Scale (SBSOD) questionnaire [46]. This self-report questionnaire analyzes everyday tasks such as finding one's way in the environment and learning the layout of a new environment. Participants answered 15 questions using a 7-point Likert scale. It has been previously demonstrated that the questionnaire is internally consistent and has good test-retest reliability [46]. Rasch analysis of the SBSOD questionnaires was performed to obtain a final Rasch-calibrated score of subjective navigational ability for each participant. Correlations of SBSOD scores with performance on a spatial task have been considered as evidence of the ecological validity of the task [42].

3.3 Results

3.3.1 Behavioral Results

Subject performance on the tasks reflected ability to perform effective navigation, particularly awareness of cues and ability to generate, improve, and recall a representation of the room. The recorded paths and completion times for each subject provided a way of assessing their capabilities.

Most subjects followed a typical behavior pattern, an example of which can be seen in Figure 3.3. At the commencement of a trial, the subject was started at a particular location and heading within the room, and would then rotate around that spot, presumably assessing cues to determine their current position and prepare for moving to the target. The subject would then move in the direction which the subject believed the target to be, refining and correcting their course as needed.

As expected, subjects improved over subsequent trials, with overall times and times initially looking around the room generally decreasing. The performance on the initial hidden trial in each room indicated that it was usually the most difficult for each subject. After being shown the target for a second time (in Trial 3) subjects generally performed better, heading more directly towards the remembered target and with more subjects able to complete the trial. This effect occurred for each room, even though by the second room, they were given instructions identical to the first room and were cognizant of the need to pay attention to cues on the first trial of the room.

Subjects did perform worse in Room 1, as compared to Room 3, as measured by taking longer. However, when taking into account the effect of glaucoma (Section 3.3.2) it was shown that healthy controls did not perform significantly worse in the first room, as the longer completion time was driven mostly by the glaucoma subjects.

One behavioral observation of note is that when some subjects were unable to locate the target in a corner that they had previously believed it to be, some would resort to a search pattern, typically either circling the room in a spiral or exploring the entire area around a corner before moving to another one. Another such observation was while subjects could pass through objects, they were treated as if they were solid, and subjects avoided doing so, preferring to go around the obstacle in almost every case.

In the final trial of each room, where the target was unreachable, subjects were given 30s to attempt to locate the target, not knowing that the target was unattainable.

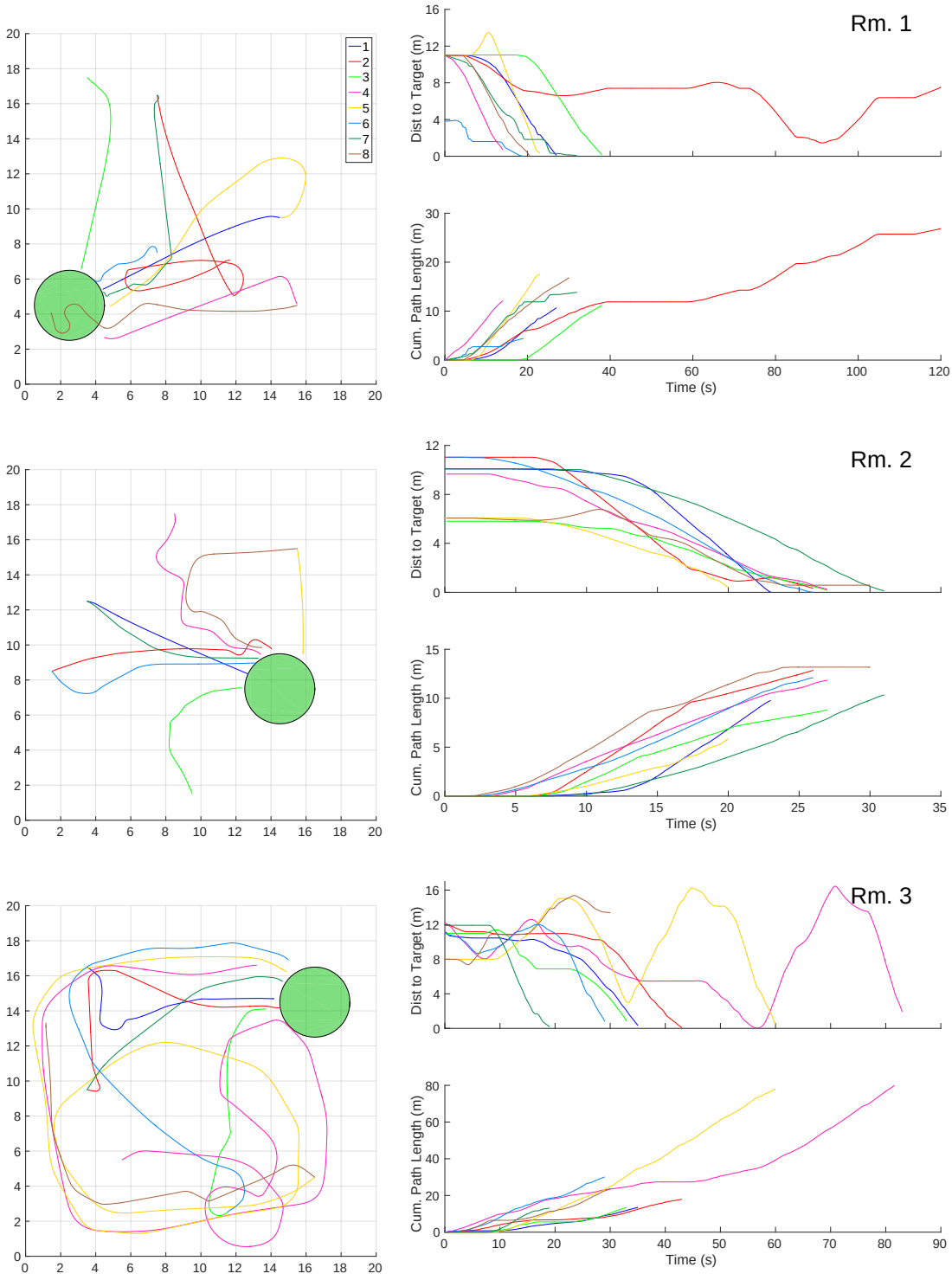


Figure 3.3: **Typical subject paths during the experiment.** The path provides information on how well known is the target location; short, direct paths indicate more precise knowledge while, long, looping paths indicate the target location has been lost.

The paths during this trial could be investigated and were indicative of whether the subject had truly memorized the target location, analogous to memory testing used in a standard water maze for rats [23]. The path shows whether the subject knew where the target was and gives clues to how clearly they remembered its position (Figure 3.4). Upon the start of the trial subjects would move to where they believed the target to be, a quick, direct path indicating that the subjects knew where they should go. Upon arriving at where the target should be, the subjects certainty in their knowledge could be assessed on how broadly they continued searching for the target. If the subject was sure that they knew the position of the target they would continue to search the local area around the target, less sure subjects would expand their search area or move to another location entirely. Some subjects headed towards where the target should be, but never reached it (16%, 39%, and 37%, in rooms 1, 2, and 3, respectively), indicating that they either were not sure where the target was or believed the target to be in another location and did not correct their course before the trial finished.

3.3.2 Effects of Glaucoma

Completion rates of the experiment differed between glaucoma and healthy subjects. Only 89% of glaucoma subjects completed all trials in Room 2 and 64% in Room 3. Comparatively, all healthy subjects completed the first two rooms, and 94% completed Room 3. This attrition was compensated for in the metrics comparisons between the first and third rooms.

Table 3.2 shows average and 95% confidence intervals of times to complete the wayfinding task for glaucoma and healthy subjects, according to the type of room. Significant differences were seen between glaucoma and healthy subjects for hidden target tasks performed in Room 1 (room with multiple visual cues), but not Room 3 (room with central chair only). For Room 1, glaucoma patients spent an average of 35.0s (95% CI: 31.3s –

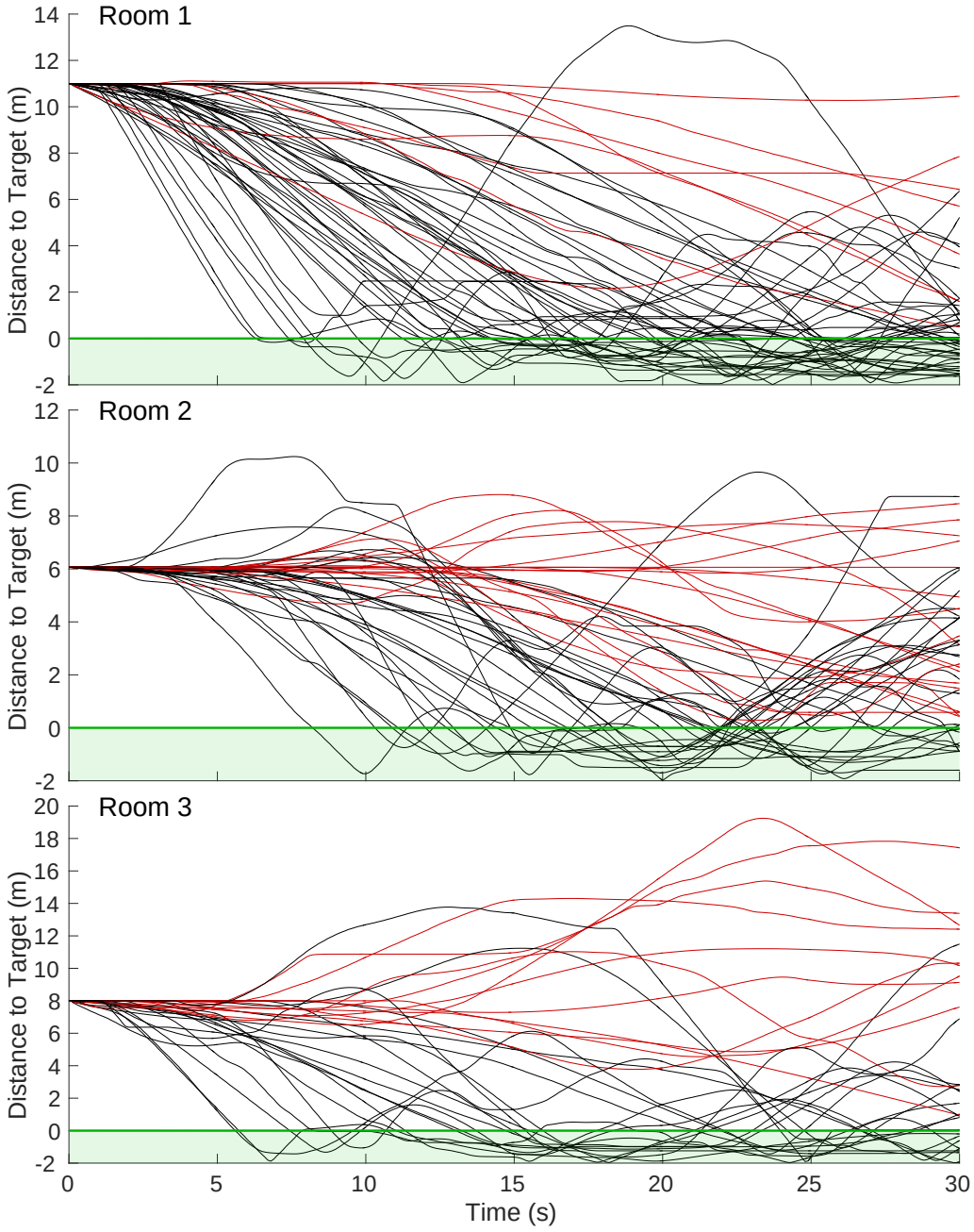


Figure 3.4: **Assessment of subjects knowledge of target position.** The final trial, where the target was unreachable, shows whether the subject knew where the target was, and gives clues to how clearly they remembered its position. Over the course of the trial, the distance to the target decreases steadily if the subject knows the location of the target and moves directly for it. If the subject was confident about their location, the distance to the target would oscillate around 0, with less sure subjects moving to a different area or having larger oscillations. The subjects plotted in red did not intercept the target location, indicated by the shaded green area, within the time alloted.

Table 3.2: **Completion times by room and glaucoma** Mean and 95% confidence interval (CI) of times to complete the wayfinding tasks with hidden targets for glaucoma and healthy subjects according to the type of room.

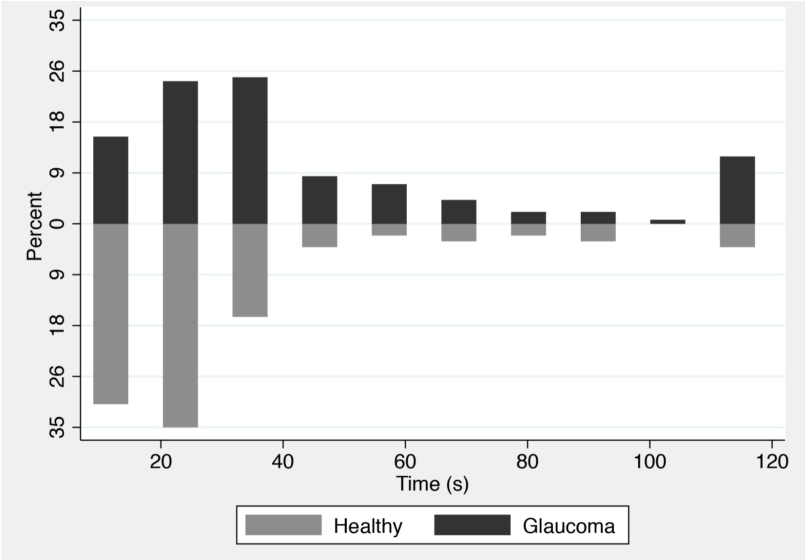
	Glaucoma (n=31)	Healthy (n=20)	P*
Room 1 (time in s)	35.0 (31.3 – 39.1)	24.4 (21.5 – 27.7)	0.001
Room 3 (time in s)	26.2 (22.6 – 27.1)	23.4 (20.3 – 27.0)	0.514

*Estimated from Tobit model. Generalized estimating equation used to adjust for multiple observations per subject.

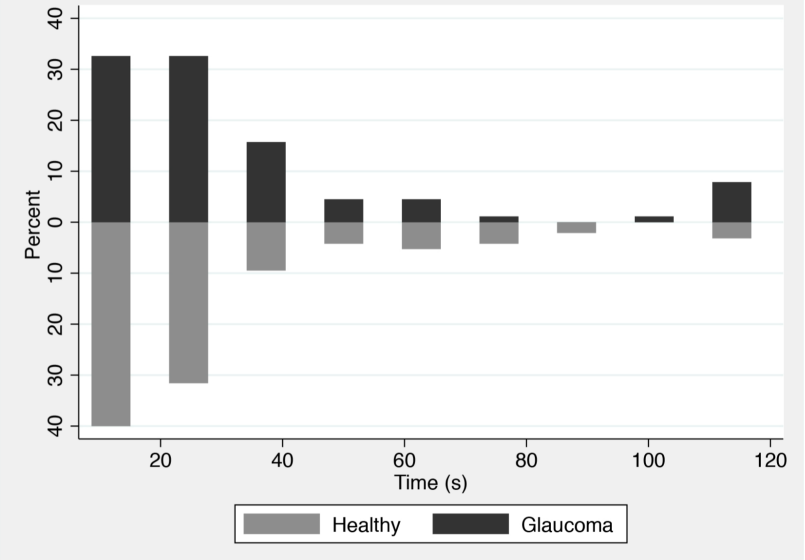
39.1s) performing the wayfinding task, whereas healthy subjects spent an average of 24.4s (95% CI: 21.5s – 27.7s) ($P = 0.001$; Tobit model). For Room 3, no statistically significant difference was seen between glaucoma and healthy subjects on average time to complete the task (26.2s (95% CI: 22.6 – 27.1) vs. 23.4s (95% CI: 20.3 – 27.0), respectively; $P = 0.514$; Tobit model). Figure 3.5 shows histograms of raw times to complete the wayfinding tasks for glaucoma and healthy subjects for Room 1 and 3.

For glaucoma patients, the average time taken to complete the hidden target task in Room 1 was significantly longer than for Room 3 (35.0s vs. 26.2s, respectively; $P = 0.008$; Tobit model). For normal subjects, there were no statistically significant difference in time to complete the task with a hidden target between the two rooms (24.4s vs 23.4s; $P = 0.721$; Tobit model).

We also investigated the relationship between time to complete the task and severity of visual field loss in models adjusting for age, gender and race. Table 3.3 shows results of Tobit multivariable models investigating the relationship between time to complete the hidden target task and binocular MS, for each one of the different rooms used in the VE-HuNT. A significant relationship between time to complete the task and visual field loss was seen for Room 1, but not for Room 3. For Room 1, each 1dB worse binocular MS was associated with 3.4% (95% CI: 1.5% - 5.4%; $P = 0.001$) increase in time to complete the task. Older age and female gender were also significantly associated with longer times to complete the wayfinding task. Figure 3.6 illustrates the relationship between predicted



(a) Room 1



(b) Room 3

Figure 3.5: **Room completion times** Histograms showing the distribution of times to complete the wayfinding tasks in each room for glaucomatous and healthy subjects.

Table 3.3: **Navigation time effects across variables** Results of tobit multivariable models investigating the association between time to complete the wayfinding task in each room and visual field loss assessed by integrated binocular mean sensitivity. Each model for each room adjusted for age, race and gender.

	Coefficient (95% CI)	P	Coefficient (95% CI)	P
Binocular mean sensitivity (per 1 dB lower)	3.4 (1.5 – 5.4)	0.001	1.7 (-0.5 – 4.0)	0.134
Age (per 1 year older)	1.3 (-0.7 – 0.8)	0.035	2.3 (0.1 – 4.6)	0.041
Race (African-American)	22.1 (-1.2 – 45.6)	0.100	33.8 (1.1 – 66.6)	0.043
Gender (female)	20.3 (0.1 – 40.8)	0.050	10.5 (-22.2 – 43.3)	0.528

*Dependent variable on each multivariable Tobit model was log time to complete the task in each room. Coefficients can be interpreted as the approximate percent change in time to complete the task for 1-unit change in the independent variable (for continuous variables) or change in category (for categorical variables).

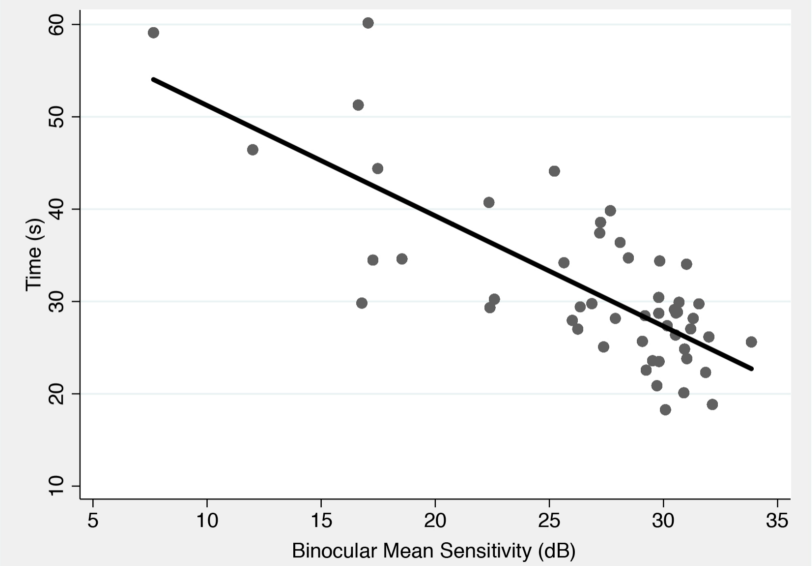
Table 3.4: **SBSOD results across variables** Multivariable model evaluating factors associated with Rasch scores obatined from the Santa Barbara Sense of Direction (SBSOD) questionnaire.

	Coefficient	95% CI	P
VE-HuNT time (per 1 log s slower)	5.69	0.42 – 11.0	0.035
Age (per 1 year older)	0.9	0.1 – 1.7	0.034
Race (African-American)	18.6	0.7 – 36.6	0.042
Gender (female)	16.4	2.9 – 29.9	0.018

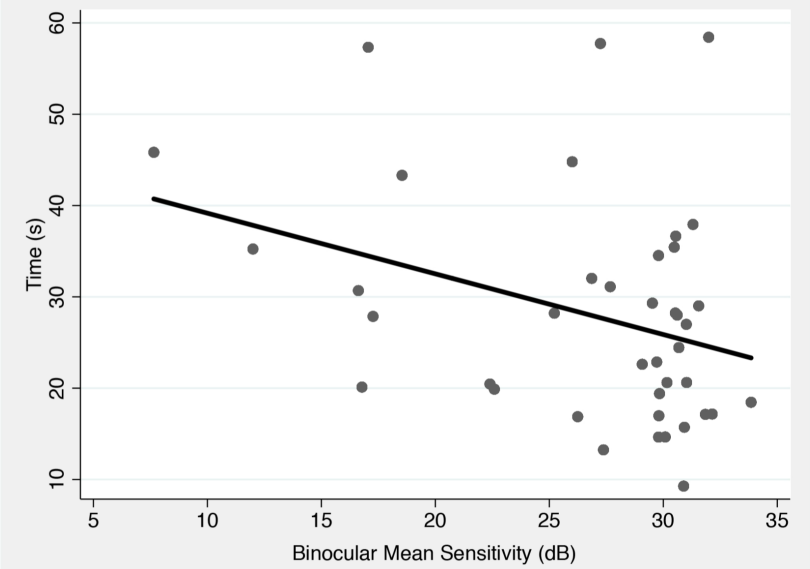
times to complete the hidden target task in the different rooms and binocular MS, adjusting for age, gender and race. The relationship was significantly stronger for Room 1 with $R^2 = 57\%$ (95% CI: 44% – 70%) compared to Room 3 with $R^2 = 11\%$ (95% CI: 0 – 23%).

In the analysis of path length for the hidden target tasks, no significant differences were seen between glaucoma and controls for any of the rooms (P=0.930 for Room 1 and P=0.800 for Room 3).

A statistically significant relationship was observed between Rasch scores from the SBSOD questionnaire and time to complete the hidden target tasks in the VE-HuNT ($P = 0.035$; Table 3.4), even after adjustment for age, race and gender.



(a) Room 1



(b) Room 3

Figure 3.6: **Relationship of binocular MS and completion time** Scatterplots with fitted regression line showing the relationship between times to complete the wayfinding task for each room and integrated binocular mean sensitivity.

3.4 Discussion

Glaucoma patients were found to perform significantly worse than healthy subjects in wayfinding tasks that promoted allocentric-based spatial reference frames to successfully complete the task. This suggests that in the presence of significant visual field loss, glaucoma patients may have difficulty in building accurate representations of the spatial structure of an environment. Our results may have significant implications for a better understanding of how vision loss in glaucoma may affect the ability to perform many everyday tasks that are related to wayfinding, such as driving in a city or walking through a building.

Significant differences were seen in the ability of glaucoma patients to perform wayfinding tasks according to the type of virtual environment used. The virtual rooms which were employed varied primarily in the complexity of the visual scene, the characteristics of which would influence the ability to use particular navigational strategies. In Room 1, times to complete the wayfinding tasks were on average over 40% longer for glaucoma subjects compared to healthy subjects. However, no significant difference between the groups was seen for Room 3. Importantly, when only glaucoma subjects were considered, times to complete the task in Room 1 were significantly longer than in Room 3. This indicates that the longer times for glaucoma patients were not only the result of worse general performance of this group in wayfinding tasks, but depended on the type of visual scene presented. In agreement with this result, the relationship between time to complete the task and severity of visual field loss was much stronger for Room 1 ($R^2 = 57\%$) than Room 3 ($R^2 = 11\%$).

Spatial information is stored in memory according to two general frames of reference: egocentric and allocentric [65, 60]. Egocentric frames of reference use the organism as the central focus point for the representation of the surrounding space. By contrast, allocentric frames of reference are structured around external objects or on the environment itself. By having a central chair placed in a fixed orientation in relation to the target, Room

3 promoted an egocentric-based solution to the wayfinding problem. In contrast, Room 1 promoted an allocentric-based strategy to solve the wayfinding problem. In order to successfully complete the task with the hidden target in Room 1, the subject had to initially build an accurate mental representation of the spatial structure of the environment, or cognitive map, containing the location of the target in relation to the several landmarks or visual cues present in that room [18]. Subsequently, the subject would have to use the cognitive map in order to provide guidance on how to navigate toward the hidden target. Cognitive maps code the spatial relationship between things in the world and can be used to recognize places, compute distances and directions, and help a subject find its way from where it is in a complex environment, to where it wants to be [26].

There is evidence that loss of peripheral vision may prevent building of accurate cognitive maps [17, 48]. Peripheral vision helps provide information about the global structure of the environment within each fixation, such that when an object appears in the periphery, eye and head movements can bring it to the fovea for further processing and accurate localization [25]. Visual information from successive fixations is then bound together to create cohesive representations of the environment in a cognitive map. A loss of peripheral vision may impede the ability to perform effective visual searches in the environment and attend to navigationally relevant objects. In addition, in the presence of significant peripheral vision loss, the visual information acquired from successive fixations may not be coherent enough to allow an accurate representation of the location of objects in space relative to each other and the global environment structure.

In our study, the loss of peripheral field in glaucoma patients may have made it difficult for them to build an accurate cognitive map of the virtual environment. This translated later into difficulties in performing the wayfinding tasks with hidden targets for Room 1. It is possible that the longer times taken by glaucoma patients for Room 1 may represent an increased amount of fixations, and consequently time, necessary to capture

spatial relations and build the cognitive map, as a result of ineffective visual searches. It is also possible that due to their restricted field of view, some glaucoma patients may never have had the opportunity to code the target relative to a feature in the scene, namely when the target was initially visible. However, this seems unlikely, as subjects were free to move their heads and eyes as they navigated through the virtual environment during the initial task with a visible target. Future studies employing eye tracking methodology could help clarify these issues and explore the compensatory behaviors associated with visual field loss. The improvement in times and the comparable times to healthy subjects in Room 3, indicates that loss of peripheral field in glaucoma is not as crucial when the relevant cues are restricted. This could result from the fact that only one source of information is functionally needed; scanning for multiple objects and constructing a visual relationship between them is unnecessary, therefore a deficit in this ability would have minimal impact. Importantly, no differences were found in the times to complete the tasks among the different rooms for healthy subjects. This shows that the delay in completing the task in the more complex room for glaucoma patients was due to the loss of peripheral vision, rather than some other inherent characteristic making the task longer in the more complex room.

To the best of our knowledge, our study is the first to assess wayfinding performance in glaucoma patients in a VR environment. The use of VR allows for highly controllable and repeatable experiments. Wayfinding and other metrics derived from VR platform experiments could potentially be used as endpoints for assessment of functional vision in clinical trials of emerging new therapies. The VR platform developed in our research could potentially be used in a variety of eye conditions other than glaucoma. As a more ecologically valid, objective, and reproducible assessment, VR-based platforms could overcome limitations of laboratory-based physical mobility courses [62].

The use of VR for studying clinically relevant navigation tasks has limitations. It is

difficult to directly compare in good accuracy the deficits of subjects seen within the VR world to those seen in real-world wayfinding tasks since standardized, robust tests of real-world wayfinding are not readily available. However, several previous studies have shown the validity of VR for assessing wayfinding in a variety of conditions, such as traumatic brain injury and stroke [62, 56, 33, 7, 55, 31]. The benefits of using VR to study wayfinding in emergency evacuations, such as buildings on fire, has also been shown. The significant association shown in our study between wayfinding performance in the VE-HuNT and the subjective assessment of navigational ability by the SBSOD questionnaire is an indication of its clinical validity. Further, we were able to replicate well-known results derived from conventional wayfinding studies, such as a significant effect of gender on allocentric, but not egocentric-based navigation, [20, 35, 63] as well as a significant effect of age on wayfinding performance [40].

In conclusion, glaucoma patients performed significantly worse than healthy subjects on allocentric-based wayfinding tasks conducted in a VR environment. Our findings suggest that glaucomatous visual field loss may affect the construction of spatial cognitive maps relevant to successful wayfinding, which may translate into significant difficulties in many daily tasks that are wayfinding dependent. VR environments may represent a useful approach for assessing functional vision endpoints for clinical trials of emerging therapies in ophthalmology.

3.5 Acknowledgments

Parts of this chapter are original to this dissertation and parts are adapted from the article "Wayfinding and Glaucoma: A Virtual Reality Experiment" on which Fabio Daga, Eduardo Macagno, Cory Stevenson, Ahmed Elhosseiny, Alberto Diniz-Filho, Erwin Boer, Jurgen Schulze, and Felipe Medeiros were co-authors and which was published in the ARVO

Investigative Ophthalmology and Visual Science in 2017. For this project I developed the experimental design and the implemented virtual environment, performed data extraction and analysis, and developed the data visualizations. The teams at the Visual Performance Laboratory (UC San Diego) and CalIT2/Qualcomm Institute (UC San Diego), including the authors, were instrumental in developing, implementing, and collecting data for this project. This project was financially supported by National Institutes of Health/National Eye Institute Grants EY025056 (FAM) and EY021818 (FAM) and a CalIT2 Strategic Research Opportunities (CSRO) seed grant.

Chapter 4

Conclusions

4.1 Summary of Contributions

This dissertation described two of the major components of my research on human movement and vision. Much of this research consisted of the advancement of technological capabilities for collecting cognitive and related biomarkers, followed by characterization and usage.

In Chapter 2 the development of an electrooculogram based eye tracker was described. It detailed the principles of the technology, the design and implementation of eye tracking experiments, and the subsequent analysis. Ultimately a method of calibration was developed, producing an eye tracking system based on principles other than standard camera based eye tracking. The capabilities of this system were measured and described, indicating that the eye tracking system is capable and useful under certain conditions, and these were compared against C-ET capabilities. Some theoretical limitations of eye tracking were discussed in the context of the EOG-ET, along with practical and planned applications of such a system.

In Chapter 3 the development and implementation of the Virtual Environment

Human Navigation Task was described. A virtual reality system was created on which to run the VE-HuNT software platform. Experiments to test the capabilities and cognitive processes involved in human navigation were developed and VE-HuNT was employed to administer them. The detrimental effects of glaucoma on the ability of subjects to perform wayfinding tasks was investigated using this system.

Overall, many facets of engineering, biomedical sciences, and cognitive neuroscience, were employed to make this research possible. The technology employed was just as varied: eye tracking, electrophysiology, virtual reality, and opthalmologic testing. To perform this work advancing the capabilities of research, it was required to bring these varied aspects together.

4.2 Future Directions

The next stages for the research presented in the dissertation is the extension of the development described into more integrated and application based work. Beyond the further developments for each project mentioned in their individual chapters, the next major step in this research is the combination of the two projects. Applying EEG recording and eye tracking via EOG to cognitive and behavioral research under more realistic environments and relevant clinical conditions.

One such expansion is the application of the eye tracking for testing visual field in ocular degeneration. Visual field loss in glaucoma is typically asymmetric, where parts of the visual field are degraded earlier than others and differentially between the two eyes [37]. Detecting saccades to and from locations based upon specific stimuli presented in space, could elucidate the shape and extent of the visual field. Importantly, this would be within the context of a virtual environment more reminiscent of real-world conditions, and consequently the results are measures of "functional" visual field. This would continue as

investigations into functional consequences and compensatory mechanisms in visual field loss.

Additional directions of the combined work is the application of the technology and techniques to the study of navigation and cognitive processes in other conditions. Normal healthy aging has been shown to have effects on wayfinding [6, 32, 40] and on reflex speed [19]; various age related conditions, such as mild cognitive impairment and Alzheimer's disease, have also been shown to have effects [32]. This expansion allows for the study of these conditions in a dynamic changing environment, with EOG and EEG signals allowing for assessments of critical high-speed reactions, such as being able to quickly gather the information in a passing road sign. The work is also applicable to studying development and Autism, which has been shown to have effects on eye movements [53]. The developed systems would allow recording in an engrossing and interactive environment, which is critical for the collection of data from young individuals. Being able to monitor in such a way allows for the study of things like learning and the effects of environment or tasks on it, such as spatial learning and determining how best directions can be conveyed. Additionally, such systems facilitate the investigation of navigation while minimizing the limitations of mobility and motor control problems, which could be especially beneficial for investigations in the navigation capabilities of autistic individuals.

An additional avenue in which I would particularly like to see the technology applied, is on the function of vestibular sense in navigation. The vestibular sense originates in the vestibular organs of the inner ear and detects translational and rotational accelerations of the head, which is critical in balance [4]. A well known interaction between this and the visual system is the vestibuloocculo reflex, in which signals from the vestibular system quickly modify eye movements in order to compensate for changes in head orientation [3]. These eye movements could be detected and examined via EOG in navigation tasks, to help explore the interplay between these processes.

4.3 Final Thoughts

I had entered into my doctoral program with the intention of studying human motor control and an interest in brain-computer interfaces. Much of my research ended up being related to those, but as things go in research, the path is often a little windy and more than a bit rocky, and therefore a bit more tangential than I would have liked (I had even, ironically, expressed explicit disinterest in studying anything to do with vision). However, the windy path did lead to me contributing to principles and technology aligned with my intentions. Ultimately, I believe I did produce research that expands the fields within which I worked and the capabilities of what could have been done previously. I am proud about what I have done to make developments in places which I feel are relatively unique, even if much of it will never be practical or go anywhere. I do believe that some of what I developed does have a purpose, and perhaps even commercial applications.

Throughout my graduate education there have been discoveries beyond, but instrumental to, the results and projects described in this dissertation. My research spanned across multiple fields and numerous collaborations; I acted as an interface, an intermediary between computer programmers, cognitive neuroscientists, engineers, and clinicians. When invariably (and repeatedly) hit by "impostor syndrome", one thing that was undeniable, just due to the nature of my research, was this expertise; which I believe was critical, not just to my sanity, but also to the general advancement of interdisciplinary research. Another discovery, just as important, is the balance between detailed, incremental progress along a well planned path and entertaining long-shots and the value of "fishing expeditions". Without the first, things move forward sporadically and unpredictably, or more likely, nothing comes of it; without the second, little important does, and nothing particularly interesting. Maintaining both of these, knowing when and how to oscillate between them, is critical to good science. There have, of course, been other discoveries, but I think these are among the most applicable to my experience and investigations of movement and vision.

Bibliography

[1] R. Y. Abe, A. Diniz-Filho, V. P. Costa, C. P. Gracitelli, S. Baig, and F. A. Medeiros. The impact of location of progressive visual field loss on longitudinal changes in quality of life of patients with glaucoma. *Ophthalmology*, 123(3):552–557, mar 2016.

[2] V. Aginsky, C. Harris, R. Rensink, and J. Beusmans. Two strategies for learning a route in a driving simulator. *Journal of Environmental Psychology*, 17:317–331, 1997.

[3] R. S. Allison, M. Eizenman, and B. S. K. Cheung. Combined head and eye tracking system for dynamic testing of the vestibular system. *IEEE Transactions on Biomedical Engineering*, 43(11):1073–1082, 1996.

[4] D. E. Angelaki and K. E. Cullen. Vestibular system: The many facets of a multimodal sense. *Annual Review of Neuroscience*, 31(1):125–150, 2008. PMID: 18338968.

[5] C. Armon, K. G. Johnson, A. Roy, and W. J. Nowack. Polysomnography. https://emedicine.medscape.com/article/1188764, Feb 2016. Published by Medscape.

[6] V. D. Bohbot, S. McKenzie, K. Konishi, C. Fouquet, V. Kurdi, R. Schachar, M. Boivin, and P. Robaey. Virtual navigation strategies from childhood to senescence: evidence for changes across the life span. *Frontiers in aging neuroscience*, 4(November):28, jan 2012.

[7] L. Carelli, M. Rusconi, C. Scarabelli, C. Stampatori, F. Mattioli, and G. Riva. The transfer from survey (map-like) to route representations into virtual reality mazes: effect of age and cerebral lesion. *Journal of NeuroEngineering and Rehabilitation*, 8(1):6, 2011.

[8] P. A. Constable, M. Bach, L. J. Frishman, B. G. Jeffrey, and A. G. Robson. ISCEV Standard for clinical electro-oculography (2017 update). *Documenta Ophthalmologica*, 134(1):1–9, 2017.

[9] C. Cook and P. Foster. Epidemiology of glaucoma: What's new? *Canadian Journal of Ophthalmology*, 47(3):223–226, 2012.

[10] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti. Surround-screen projection-based virtual reality. *Proceedings of the 20th annual conference on Computer graphics and interactive techniques - SIGGRAPH '93*, pages 135–142, 1993.

[11] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart. The CAVE: audio visual experience automatic virtual environment. *Communications of the ACM*, 35(6):64–72, jun 1992.

[12] H.-C. Diener and J. Dichgans. Chapter 22 on the role of vestibular, visual and somatosensory information for dynamic postural control in humans. In O. Pompeiano and J. Allum, editors, *Vestibulospinal Control of Posture and Locomotion*, volume 76 of *Progress in Brain Research*, pages 253 – 262. Elsevier, 1988.

[13] A. Diniz-Filho, E. R. Boer, C. P. Gracitelli, R. Y. Abe, N. van Driel, Z. Yang, and F. A. Medeiros. Evaluation of postural control in patients with glaucoma using a virtual reality environment. *Ophthalmology*, 122(6):1131–1138, jun 2015.

[14] E. Estrada, H. Nazeran, J. Barragan, J. R. Burk, E. A. Lucas, and K. Behbehani. Eog and emg: Two important switches in automatic sleep stage classification. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2458–2461, Aug 2006.

[15] N. Etchamendy and V. D. Bohbot. Spontaneous navigational strategies and performance in the virtual town. *Hippocampus*, 17(8):595–599, 2007.

[16] A. Faisal. Computer science: Visionary of virtual reality. *Nature*, 551:298, Nov. 2017.

[17] F. C. Fortenbaugh, J. C. Hicks, L. Hao, and K. A. Turano. Losing sight of the bigger picture: Peripheral field loss compresses representations of space. *Vision Research*, 47(19):2506–2520, sep 2007.

[18] F. C. Fortenbaugh, J. C. Hicks, and K. A. Turano. The effect of peripheral visual field loss on representations of space: Evidence for distortion and adaptation. *Investigative Opthalmology & Visual Science*, 49(6):2765, jun 2008.

[19] J. M. Furman and M. S. Redfern. Effect of aging on the otolith-ocular reflex. *Journal of Vestibular Research*, 11(2):91–103, 2001.

[20] L. A. Galea and D. Kimura. Sex differences in route-learning. *Personality and Individual Differences*, 14(1):53–65, jan 1993.

[21] George. Top 10 phones of 2017: Best video recording. https://www.gsmarena.com/top_10_phones_of_2017_best_video_recording-news-28894.php, 12 2017 (accessed 2018-02-28). Published by GSMArena.

[22] S. E. Hassan, J. C. Hicks, H. Lei, and K. A. Turano. What is the minimum field of view required for efficient navigation? *Vision Research*, 47(16):2115–2123, jul 2007.

[23] M. Hegarty. Development of a self-report measure of environmental spatial ability. *Intelligence*, 30(5):425–447, oct 2002.

[24] A. Heijl, M. C. Leske, B. Bengtsson, L. Hyman, B. Bengtsson, and M. Hussein. Reduction of intraocular pressure and glaucoma progression: Results from the early manifest glaucoma trial. *Archives of Ophthalmology*, 120(10):1268–1279, 2002.

[25] S. G. Jacobson, A. V. Cideciyan, R. Ratnakaram, E. Heon, S. B. Schwartz, A. J. Roman, M. C. Peden, T. S. Aleman, S. L. Boye, A. Sumaroka, T. J. Conlon, R. Calcedo, J.-J. Pang, K. E. Erger, M. B. Olivares, C. L. Mullins, M. Swider, S. Kaushal, W. J. Feuer, A. Iannaccone, G. A. Fishman, E. M. Stone, B. J. Byrne, and W. W. Hauswirth. Gene therapy for leber congenital amaurosis caused by rpe65 mutations: Safety and efficacy in 15 children and adults followed up to 3 years. *Archives of Ophthalmology*, 130(1):9–24, 2012.

[26] K. A. Jellinger. Wayfinding behavior, cognitive mapping and other spatial processes reginald g. golledge, ed. the johns hopkins university press, baltimore & london, 1999. 428 pp., ISBN 0-8010-5993-x. *European Journal of Neurology*, 7(5):590–591, sep 2000.

[27] C. A. Joyce, I. F. Gorodnitsky, J. W. King, and M. Kutas. Tracking eye fixations with electroocular and electroencephalographic recordings. *Psychophysiology*, 39(5):607–618, 2002.

[28] T.-P. Jung, C. Humphries, T.-W. Lee, S. Makeig, M. J. McKeown, V. Iragui, and T. J. Sejnowski. Extended ICA removes artifacts from electroencephalographic recordings. *Advances in Neural Information Processing Systems*, 10:894–900, 1998.

[29] N. Kikawada. Variations in the corneo-retinal standing potential of the vertebrate eye during light and dark adaptations. *The Japanese Journal of Physiology*, 18(6):687–702, 1968.

[30] H. Klob. *Webvision*, chapter Simple Anatomy of the Retina. University of Utah, 2011.

[31] R. Li. Why women see differently from the way men see? a review of sex differences in cognition and sports. *Journal of Sport and Health Science*, 3(3):155–162, sep 2014.

[32] S. Lithfous, A. Dufour, and O. Després. Spatial navigation in normal aging and the prodromal stage of alzheimer's disease: Insights from imaging and behavioral studies. *Ageing Research Reviews*, 12(1):201 – 213, 2013. Special Issue: Invertebrate Models of Aging.

[33] S. A. Livingstone and R. W. Skelton. Virtual environment navigation tasks and the assessment of cognitive deficits in individuals with brain injury. *Behavioural Brain Research*, 185(1):21–31, dec 2007.

[34] J. E. Lovie-Kitchin, J. C. Mainstone, J. Robinson, and B. Brown. What areas of the visual field are important for mobility in low vision patients? *Clinical Vision Sciences*, 5(3):249–263, 1990.

[35] J. C. Malinowski and W. T. Gillespie. Individual differences in performance on a large-scale, real-world wayfinding task. *Journal of Environmental Psychology*, 21(1):73–82, mar 2001.

[36] J. Malmivuo and R. Plonsey. The electric signals originating in the eye. In *Bioelectromagnetism : principles and applications of bioelectric and biomagnetic fields*, chapter 28. Oxford University Press, 1995.

[37] F. A. Medeiros, C. P. Gracitelli, E. R. Boer, R. N. Weinreb, L. M. Zangwill, and P. N. Rosen. Longitudinal changes in quality of life and rates of progressive visual field loss in glaucoma patients. *Ophthalmology*, 122(2):293–301, feb 2015.

[38] T. Merel. The reality of vr/ar growth. https://techcrunch.com/2017/01/11/the-reality-of-vrar-growth/, Jan 2017. Published by TechCrunch.

[39] L. Mitrani and G. Dimitrov. Retinal location and visual localization during pursuit eye movement. *Vision Research*, 22(8):1047 – 1051, 1982.

[40] S. D. Moffat and S. M. Resnick. Effects of age on virtual environment place navigation and allocentric cognitive mapping. *Behavioral Neuroscience*, 116(5):851–859, 2002.

[41] O. H. Mowrer, T. C. Ruch, and N. E. Miller. The corneo-retinal potential difference as the basis of the galvanometric method of recording eye movements. *American Journal of Physiology-Legacy Content*, 114(2):423–428, 1935.

[42] L. Nadel and O. Hardt. The spatial brain. *Neuropsychology*, 18(3):473–476, 2004.

[43] Z. S. Nasreddine, N. A. Phillips, V. BÃ©dirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings, and H. Chertkow. The montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4):695–699, apr 2005.

[44] J. M. Nelson-Quigg, K. Cello, and C. A. Johnson. Predicting binocular visual field sensitivity from monocular visual field results. *Investigative Ophthalmology & Visual Science*, 41(8):2212–2221, 2000.

[45] M. E. Ono, J. Rivest, and H. Ono. Depth perception as a function of motion parallax and absolute-distance information. *Journal of Experimental Psychology: Human Perception and Performance*, 12(3):331–337, 08 1986.

[46] L. Palermo, G. Iaria, and C. Guariglia. Mental imagery skills and topographical orientation in humans: A correlation study. *Behavioural Brain Research*, 192(2):248–253, oct 2008.

[47] D. G. Pelli. The visual requirements of mobility. In *Low Vision*, pages 134–146. Springer New York, 1987.

[48] M. S. Peterson, A. F. Kramer, and D. E. Irwin. Covert shifts of attention precede involuntary eye movements. *Perception & Psychophysics*, 66(3):398–405, apr 2004.

[49] T. Pfeiffer, M. E. Latoschik, I. Wachsmuth, and J. Herder. Evaluation of binocular eye trackers and algorithms for 3d gaze interaction in virtual reality environments. *Journal of Virtual Reality and Broadcasting*, 5(16):16, 2008.

[50] PupilLabs. Platform for eye tracking and egocentric vision research: Technical specifications and performance. https://pupil-labs.com/pupil/, 2018 (accessed 2018-02-28). Published by Pupil Labs GmBH.

[51] P. Ramulu. Glaucoma and disability: which tasks are affected, and at what stage of disease? *Current Opinion in Ophthalmology*, 20(2):92–98, mar 2009.

[52] L. F. Schettino, S. V. Adamovich, and H. Poizner. Effects of object shape and visual feedback on hand configuration during grasping. *Experimental Brain Research*, 151(2):158–166, Jul 2003.

[53] L. M. Schmitt, E. H. Cook, J. a. Sweeney, and M. W. Mosconi. Saccadic eye movement abnormalities in autism spectrum disorder indicate dysfunctions in cerebellum and brainstem. *Molecular Autism*, 5(1):47, 2014.

[54] A. L. Shelton and J. D. E. Gabrieli. Neural correlates of encoding space from route and survey perspectives. *Journal of Neuroscience*, 22(7):2711–2717, 2002.

[55] N.-J. Shih, C.-Y. Lin, and C.-H. Yang. A virtual-reality-based feasibility study of evacuation time compared to the traditional calculation method. *Fire Safety Journal*, 34(4):377–391, jun 2000.

[56] R. W. Skelton, S. P. Ross, L. Nerad, and S. A. Livingstone. Human spatial navigation deficits after traumatic brain injury shown in the arena maze, a virtual morris water maze. *Brain Injury*, 20(2):189–203, jan 2006.

[57] SMI. Smi eye tracking glasses 2 wireless. https://www.smivision.com/wp-content/uploads/2017/05/smi_prod_ETG_120Hz_asgm.pdf, 2017 (accessed 2018-02-28). Published by SensoMotoric Instruments.

[58] SR Research. Eyelink 1000 plus technical specifications. https://www.sr-research.com/wp-content/uploads/2017/11/eyelink-1000-plus-specifications.pdf, 2017 (accessed 2018-02-28). Published by SR Research Ltd.

[59] Y.-C. Tham, X. Li, T. Y. Wong, H. A. Quigley, T. Aung, and C.-Y. Cheng. Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis. *Ophthalmology*, 121(11):2081 – 2090, 2014.

[60] E. C. Tolman. Cognitive maps in rats and men. *Psychological Review*, 55(4):189–208, 1948.

[61] K. A. Turano, G. S. Rubin, and H. A. Quigley. Mobility performance in glaucoma. *Investigative Opthalmology & Visual Science*, 40(12):2803–2809, nov 1999.

[62] I. J. M. van der Ham, A. M. E. Faber, M. Venselaar, M. J. van Kreveld, and M. Loffler. Ecological validity of virtual environments to assess human navigation ability. *Frontiers in Psychology*, 6, may 2015.

[63] J. M. Wiener, O. de Condappa, M. A. Harris, and T. Wolbers. Maladaptive bias for extrahippocampal navigation strategies in aging humans. *Journal of Neuroscience*, 33(14):6012–6017, apr 2013.

[64] L. R. Young and D. Sheena. Survey of eye movement recording methods. *Behavior Research Methods & Instrumentation*, 7(5):397–429, Sep 1975.

[65] T. Zaehle, K. Jordan, T. Wüstenberg, J. Baudewig, P. Dechent, and F. W. Mast. The neural basis of the egocentric and allocentric spatial frame of reference. *Brain Research*, 1137:92–103, mar 2007.