Title
Stability of the Fixated Retinal Image During Monocular and Binocular Tasks

Permalink
https://escholarship.org/uc/item/8hs3d1wv

Author
Bowers, Norick Richard

Publication Date
2022

Peer reviewed|Thesis/dissertation

Stability of the Fixated Retinal Image During Monocular and Binocular Tasks

by

Norick R Bowers

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Vision Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Austin Roorda, Chair
Professor Michael Deweese
Professor Susana Chung

Spring 2022

Stability of the Fixated Retinal Image During Monocular and Binocular Tasks

Abstract

Stability of the Fixated Retinal Image During Monocular and Binocular Tasks

by

Norick R Bowers

Doctor of Philosophy in Vision Science

University of California, Berkeley

Professor Austin Roorda, Chair

The movements of the eye play an integral part in constructing a smooth perception of the world around us. The fact that our eyes are foveal, with a high resolution center and a radial decrease in resolution as we move away from this center, necessitates a constant shifting of our gaze around our environment. Although our eyes are in constant motion we do not perceive the world as a series of static snapshots and constant jitter, instead our visual system is able to effectively integrate information over time to render the world veridically; a complex array of color, depth, and motion. These eye movements are essential to vision for a wide variety of reasons. In fact, movement is such a key aspect of vision that if the projection of the world onto the retina is rendered perfectly still, our vision will slowly fade. Our brain relies on a constantly fluctuating signal in order to generate a visual percept, and without this temporal variation, vision is impossible. Even when fixating on a point, our eyes are still engaging in a series of small movements to examine and enhance the fine details of the foveated stimuli. Collectively, these small movements are referred to as fixational eye motion (FEM). In order to truly understand how these small temporal variations will contribute to vision, it is necessary to understand how these movements interact with the photoreceptors in the retina. To this end we utilized an Adaptive Optics Scanning Laser Ophthalmoscope to project stimuli onto the photoreceptor mosaic and simultaneously recover an unambiguous trace of the motion of these stimuli as the eye engages in fixational eye movements.

Dedication

This work is dedicated to my father Rick Bowers, who has been a constant source of support throughout my entire life. His patience and kindness instilled in me an ambition for learning and his guidance taught me the resilience to get through the difficult times.

Acknowledgments

First and foremost, I thank my advisor, Austin Roorda for all the advice and support he has offered throughout the years. I've learned a tremendous amount under his guidance and I thank him for his encouragement and willingness to push me to learn and understand things outside my comfort zone. Having the opportunity to work in this environment vastly expanded my knowledge beyond my psychological background, and I've developed skills and know-how beyond anything I've expected.

I would also like to thank the current and former members of the Roordalab for their support and advice throughout the years: notably Kavitha Ratnam for sharing her knowledge when I first began, Ethan Bensinger for his kindness and friendship, and Ally Boehm for her advice and technical help as a coauthor. I, of course, also extend my heartfelt thanks to the countless other members of the lab who have been a source of inspiration and support.

I also want to extend my thanks to other members of the Vision Science community here at Berkeley, specifically Martin Banks, Susana Chung, Mike Deweese, and my coauthor Josselin Gautier. Marty has been an invaluable source of advice and information, even going so far as to be an unofficial member of my qualifying exam committee as well as offering his advice and help for my dissertation. Susana Chung has offered so much kindness and practical advice, as well as being kind enough to let us utilize her system for our experiments. Mike was kind enough to sit on my committee and offer me his expertise. Joss has been such an incredible source of knowledge for me. He has been kind enough to freely share his expertise and exchange information even beyond the scope of our work together.

I would also like to acknowledge my friends who have taught me what's important in life and given me countless laughs. Julian Quinn, for his intelligence and wit and Erica (Paco) Luke for her support and kindness. I would also like to extend my thanks and love to the memory of Anthony Delgado. Our time together was priceless and altogether too short.

Lastly I want to thank my family who's been a source of stability and grounding for me through my entire life. My father Rick Bowers and my step-mother Raena Bowers, who've shown me incalculable kindness. My sister, Nikki Bowers, who's been one of my best friends. My brother Nathan McWilliams, who has always encouraged me to be better. My brother Nolan Bowers and his wife Kayla for welcoming me into their family, as well as my niece Violet, my nephew Xavier, and my as-yet-unnamed and soon-to-be-born nephew. You're going to be awesome kid.

Contents

List of Figures

Chapter 1

Introduction

1.1 Fixational Eye Motion

When fixating our gaze on an object, our eyes are never truly at rest. Even while staring at a small object, like the bottom row of a Snellen acuity chart, our eyes are constantly in motion. These movements can shift a stimulus over dozens of photoreceptors every second. Fixational eye motion has typically been categorized to fall into three main components: (a) Microsaccades, small ballistic movements similar to larger saccades, (b) Ocular drift, a slow Brownian-like movement shifting the gaze only a few arcminutes, and (c) Tremor, a high-frequency oscillatory jitter roughly the size of a foveal cone (Ditchburn and Ginsborg, 1953; Ezenman et al., 1985; Ko et al., 2016; Ratliff and Riggs, 1950; Rucci and Poletti, 2015). These fixational eye movements are a nuisance for many ophthalmic measurements such as imaging, microperimetry or retinal and refractive surgery. However, owing to the fact that FEM are the finest motor control system in the human body, they offer an opportunity for early detection and monitoring of neurological disorders (Hunfalvay et al., 2021; Montesano et al., 2018; Sheehy et al., 2020) and insight into the visuomotor system at large (Yarbus, 1968). There has been a resurgence of interest in studying fixational eye motion in the last few decades as new techniques have become available to examine their properties and influence on visual perception. Each of the three main components of fixational eye motion have been examined extensively and new discoveries have been made in recent years.
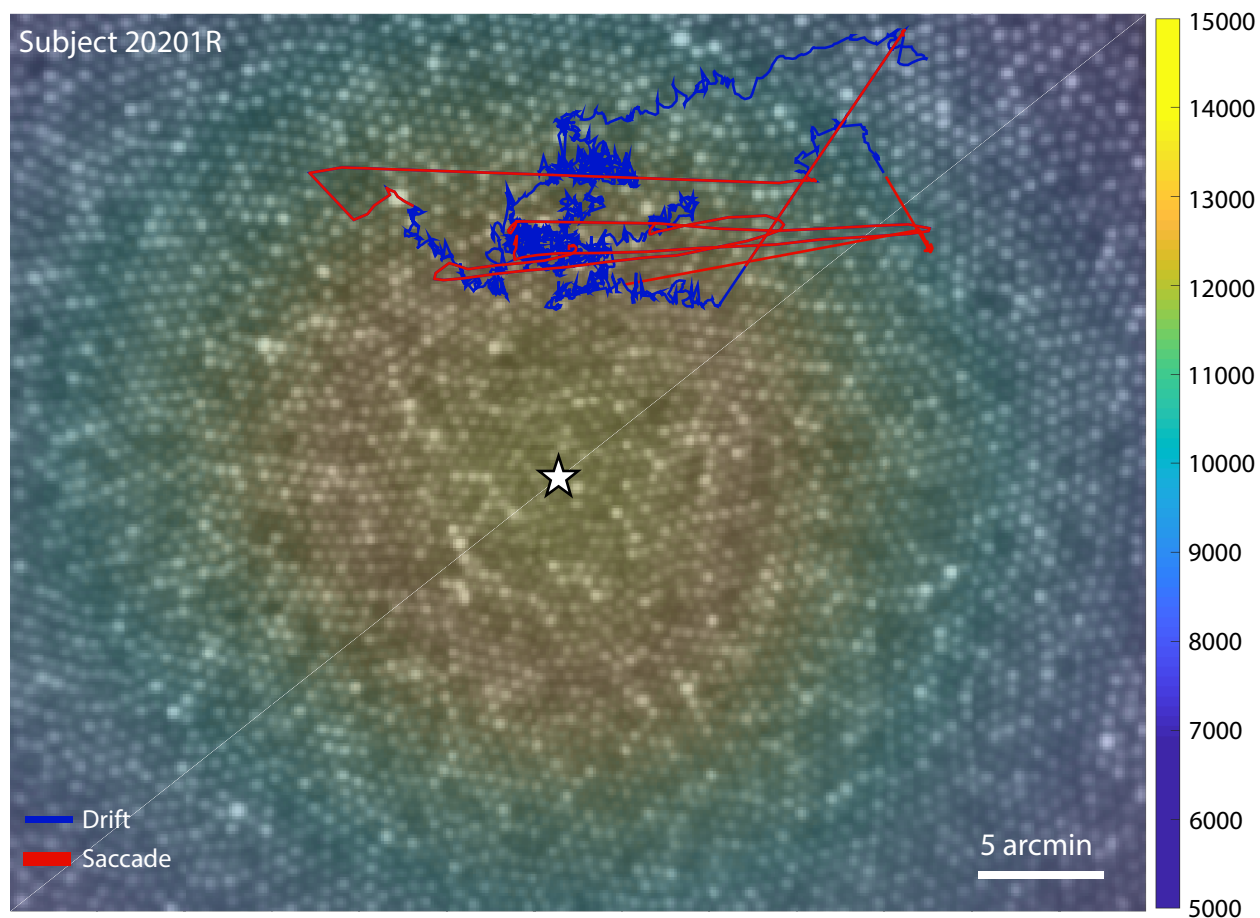
Figure 1.1: An example of the path of a small stimuli as it moves across the retinal mosaic with the fixational eye motion. Drift is highlighted in blue and microsaccades are highlighted in red. The underlying cone mosaic is shown with the peak cone density marked with a star. The underlying heatmap represents the thinning out of cones away from this center. The colorbar on the right denotes the cones per degree. Note that the subject is placing the stimuli slightly offset from the peak cone density. This phenomenom will be discussed in greater detail in Chapter 4

1.1.1 Microsaccades

Microsaccades are the most thoroughly examined component of fixational eye motion, likely owing to their relatively larger size compared to drift and tremor. Examination into the perceptual uses of microsaccades has seen a resurgence lately as new techniques for eye tracking, analysis, and modelling of the traces have been discovered. Microsaccades are an essential part of everyday visual behavior and have been found to improve and contribute to the visual percept in a wide variety of ways (Rucci and Victor, 2015; Rolfs, 2009). Microsaccades are used to precisely relocate the gaze during reading of small text to relevant portions of words

and sentence structure (Bowers and Poletti, 2017). They also play a large part in preventing visual fading of foveated stimuli (Martinez-Conde et al., 2006). They also play a vital role in repositioning fine stimuli within the fovea (Ko et al., 2010; Poletti et al., 2013; Intoy and Rucci, 2020) , which has traditionally been considered a relatively homogenous region of the retina in terms of acuity. In fact, performance of a high acuity visual task is greatly reduced when the stimulus is placed even slightly away from the optimal position within the foveola (Rossi and Roorda, 2010). Neurophysiological signatures have also been found to be associated with microsaccades (Kagan et al., 2008b), suggesting that there exists components of the visual system in the brain that are specifically tuned to detect and respond to the visual transients caused by microsaccades. Microsaccades have also been associated with shifts of covert attention (Hafed and Clark, 2002; Engbert and Kliegl, 2003), which strongly suggests a pairing between the oculomotor system and more attention-oriented cognitive mechanisms.

1.1.2 Drift

Although drift is less well studied when compared to microsaccades, it is an essential aspect of the oculomotor system that relays information from the retina to the brain. Oculomotor drift is the movement of the eyes between microsaccades. It is well modeled as a Brownian random-walk (Ko et al., 2016; Rucci et al., 2018), and demonstrates the expected $1/f^2$ that Brownian motion often shows in nature. The temporal effects of drift on the image falling on the retina is an often overlooked aspect of vision. This is surprising given that during drift is when most information is being relayed. In fact, recent research suggests that microsaccades show the effects of saccadic suppression in a similar manner to larger saccades (Mostofi et al., 2021), which means that during drifts is essentially the only time that information is being relayed to the brain. It's been shown that fixational drift has a synergistic relationship with the spatial properties of natural scenes. The natural world is dominated by low spatial frequencies; however these frequencies often carry redundant information and it is in fact the high spatial frequencies that transmit more useful information, such as edges of objects and fine details. Fixational drift has been shown to shift the power of natural scenes by transforming the spatial properties of these scenes into the temporal domain, essentially acting as a whitening mechanism and enhancing the relative power of higher spatial frequencies (Rucci et al., 2007; Kuang et al., 2012). This filtering of natural images is generally thought of as happening in the visual cortex but this recent work demonstrates the importance of the transients fixational eye motion imparts on our representation of the world. In fact, much like microsaccades, there are cells in the brain that selectively respond to the transients introduced by drift (Kagan et al., 2008b). Recent work has also proposed methods for how the motion itself plays a vital role in encoding information over time to heighten discriminability (Anderson et al., 2020; Ratnam et al., 2017).

1.1.3 Tremor

Tremor is the least studied component of fixational eye motion. Fixational tremor is a high frequency oscillatory jitter roughly the size of a foveal cone ($\sim$30arcsec). This type of movement is often difficult to see in the raw eye trace but it is frequently visible in the power spectra of ocular drift as a deviation in the expected $1/f^2$ pattern between 50-100Hz. There is however some disagreement about the exact amplitude of tremor, with some studies finding it only has an amplitude of $\sim$10arcssec (Ko et al., 2016), and some finding it has an amplitude as high as 1 arcmin (Ratliff and Riggs, 1950). Tremor was first discovered in the 1950's (Ratliff and Riggs, 1950) and measurements taken from that time were largely corroborated in the following decades as more advanced eye tracking techniques were developed (Ezenman et al., 1985; Kagan et al., 2008a; Rucci and Victor, 2015). The perceptual consequences of tremor were never fully understood. Some theories suggested that tremor could help to facilitate the synchronization of retinal ganglion cells (Greschner et al., 2002) or be involved in stochastic resonance of visual noise (Hennig et al., 2002). However, a recent study looking into tremor on the retinal image directly have found that the amplitude of tremor was many times smaller than traditionally reported and would be highly unlikely to contribute to the visual percept (Bowers et al., 2019). This study will be discussed in much greater detail in Chapter 3.

1.1.4 Eye Tracking Techniques

Early studies examining fixational eye motion relied on the optical lever (or plane-mirror) technique to record the eye movements. In this technique a contact lens is fitted onto the eye through suction with a reflective surface adhered to the side. A beam of light is shone onto the reflective surface attached to the contact lens and the light reflecting from it is shone onto photographic paper which can measure the movement of the beam. If the head is sufficiently stabilized then the trace of the beam on the photographic paper will be an accurate reflection of the movements of the eye. Early experiments studying fixational eye motion obtained surprisingly accurate records of the eye trace during fixation on a variety of targets (Ratliff and Riggs, 1950; Ditchburn and Ginsborg, 1953). These early studies were able to identify many components of fixational eye movements that have been corroborated with more advanced modern techniques, such as the microsaccade rate, amplitude, and overshoot, as well as some degree of binocular coordination. Although the technology these experimenters were working with was in many ways limited, it truly speaks to their ingenuity in designing techniques that were so accurate that they are comparable with those that were developed decades later.

Although there are a variety of eye tracking devices for clinical, research, and commercial use today, most of these trackers cannot render truly unambiguous records of a stimulus' motion across the photoreceptor mosaic with high spatial and temporal resolution. Many high-end eye trackers, such as the Dual-Purkinjie Image Eye Tracker or scleral search coils can record the direction of gaze with high temporal ($\sim$1Khz) and spatial ($<$1arcmin) resolution

(Cornsweet and Crane, 1973; Crane and Steele, 1985). However, they can be biased by their calibration. Many of these systems use a point calibration where the subject is instructed to fixate on a grid of points and the readout from the eye tracker is used to assess where the subject is looking on the screen used to display the stimuli. Systematic, noisy, or even accidental misalignment between the grid points and the true direction of the subject's gaze will cause offset in the trace that may not be readily evident to the experimenter. Although these systems are cleverly designed and lead to great discovery, they still cannot render with complete certainty the path of a stimuli's image across the retina.

In order to achieve an unambiguous trace of a stimulus' path on the retina, it is necessary to image the retinal surface directly while simultaneously tracking the gaze direction. This combination presents an extraordinary challenge. A significant difficulty in obtaining a clear image of the retinal surface arise from the optics of the eye itself, which are far from perfect. Light passing through the eye must pass through the cornea, lens, and aqueous humor, all of which have slightly different refractive indices and structural flaws which will contribute significantly to aberate the light before it reaches the retina. The performance of any optical system can be measured using the Strehl Ratio, which is the ratio of the peak luminance of a focused point from an abberated optical system over the peak luminance of a focused point of a diffraction-limited system. A perfect system limited only by diffraction will have a Strehl ratio of 1. However, the emmetropic human eye has a Strehl Ratio of ~0.1, although this can vary significantly due to changes in the pupil diameter as well as from person to person. The effects of this distortion is easily visible in the point-spread function (PSF) of the human eye. The PSF represents the effects an optical system has on a point source of light that passes through it. A perfect optical system will cause a point source to be focused into an Airy Disc due to the inescapable effects of diffraction. The Airy Disc is a bright central region with very faint concentric rings surrounding it. The distance seperating the rings is defined by the wavelength of light forming the pattern. An imperfect optical system will distort and smear the image, and will lead to loss of focus. If the system is perfect, it will produce a perfect Airy Disc. However, as the system becomes more and more aberrant, the smearing will worsen and the peak luminance will drop, leading to a smeared ill-defined image. A demonstration of this smearing can be seen in Figure 1.2 below.
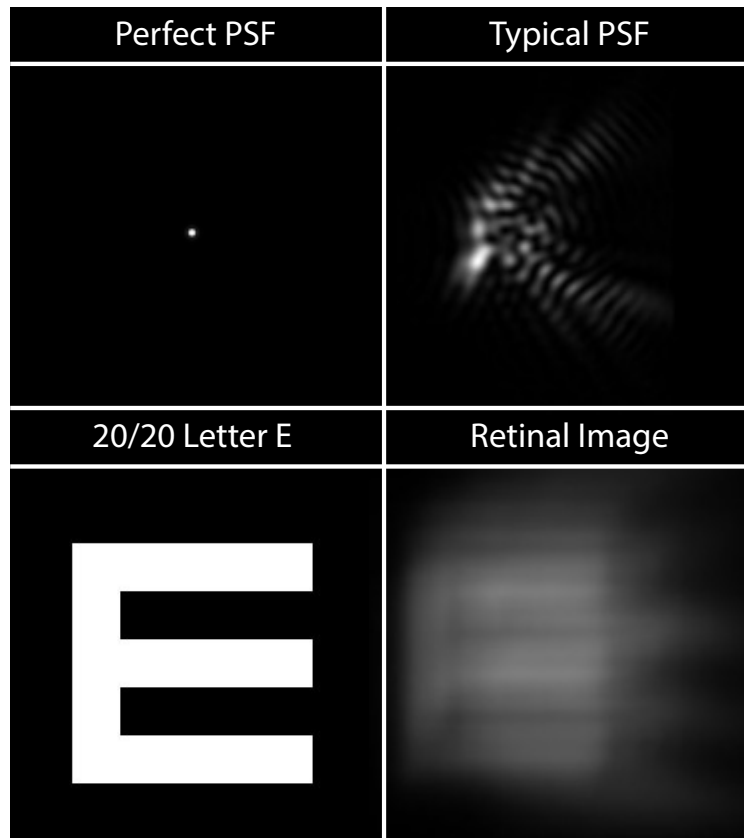
Figure 1.2: An example of a PSF for a "perfect" eye (top, left) and a typical eye (top right) for a pupil size of 5mm. In this instance, the concentric ring pattern of the Airy Disc is too faint to be visible. If the PSF of a typical eye is convolved with a 20/20 letter E from a Snellen acuity chart (bottom, left) the retinal image will be dramatically smeared (bottom, right). This smearing, while quite extreme, does not preclude normal vision in day-to-day life.

1.2   Scanning Laser Ophthalmoscopy

The first step in acquiring an unambiguous record of the gaze position is imaging the retina. The scanning laser ophthalmoscope (SLO) is a device that can provide a sharp, high-contrast image of the photoreceptor mosaic with high enough resolution to resolve foveal cone cells (Roorda et al., 2002). The SLO functions similarly to a scanning laser microscope. It uses single point source of light that is swept across the retina in a raster pattern, pixel by pixel, to construct a high resolution image of the photoreceptor mosaic. A major limitation in the acquisition of these images that is not present in more conventional laser microscopy is that the retinal surface is never still. Small eye movements constantly shift the retina as the image is acquired. These eye movements, although frequently fractions of a degree in scale, are sufficiently large enough to cause distortions in the retinal image acquired by the

sweeping laser due to the high spatial resolution of the SLO. For example, a movement of the eye by 1/2 degree can shift the image by roughly half the size of the frame. The exact magnitude of the motion in relation to the frame can be adjusted by increasing or decreasing the size of the raster, however a raster size of roughly 1° is common. More detail on these types of eye movements and how they're addressed in SLO imaging is discussed later.

1.2.1   System Layout

The AOSLO system is laid out on an optics table that helps to eliminate any movement of components from ambient vibrations in the environment. A supercontinuum laser provides a point source of light that is relayed through the optical path and scanned across the retina in a raster pattern utilizing two scanners, a 16 kHz fast horizontal scan and a 30 Hz slow vertical scan. The reflected light is descanned through the optical path and directed to a custom-built Shack–Hartmann wavefront sensor and through a confocal pinhole to a photomultiplier tube (Hamamatsu, Japan). The Shack–Hartmann wavefront sensor is used to measure the optical aberrations and send a correction to the deformable mirror in the optical path. This device and the necessity of utilizing adaptive optics in this system is discussed in further detail in Chapters 3 and 4. Light detected by the PMT and the positional information from the scanner are combined to construct videos of the retina with 512×512 pixel sampling resolution at a frame rate of 30 Hz (the speed of the slow vertical scanner). The subject's pupil is often kept in a fixed position relative to the AOSLO beam by restraining the subject's head movement through the use of a dental bite bar and temple mounts while the nonimaged eye is covered with a patch. The system has the capacity to use a variety of wavelengths. The current system uses an infrared (840nm), red (680nm) or green (543nm) beam for imaging and stimulus presentation and an infrared beam (940nm) for wavefront measurement. The field size appears as a bright red (for the 840nm or 680nm beam) or bright green (for the 543nm beam) square that flickers at a rate of 30Hz (the speed of the slow vertical scanner). Fixation targets can be presented to the subject within the field by turning off the scanning laser using an acousto-optic modulator (Brimrose Corp, MD) at the appropriate time points during the raster scan. To the subject, these targets appear as black decrements. The stimuli are very sharp and have high contrast owing to the use of adaptive optics on the input scanning beam. Importantly, these decrements are also encoded directly into the output video, which allows for an unambiguous measurement of the motion of the image of the fixation target over the retina. Extracting the motion of the stimulus on the retinal surface is discussed in further detail below. This system is capable of obtaining near diffraction-limited images of the photoreceptor mosaic and delivering stimuli with the precision of 1 pixel (approximately 6 arcseconds).

1.2.2   Adaptive Optics

Measurement of a wavefront is an essential step in adaptive optics, a technique for continuously measuring and correcting for optical aberrations to enhance the resolution of a noisy

image. Adaptive optics is most often used in the field of Astronomy to correct for atmospheric distortions while imaging celestial bodies. However, the same techniques (even the same technology) can be used to correct the aberrations of the eye for retinal imaging. A common technique to measure and correct for optical aberrations is through the use of a coupled Shack-Hartmann wavefront sensor and a deformable mirror.

The Shack-Hartmann wavefront sensor is used to measure the optical aberrations of the point source of light in the AOSLO system (Shack and Platt, 1971). Importantly, a small part of the same light used in imaging is used to measure the wavefront, ensuring the wavefront is being corrected on the same region as the image. The Shack-Hartmann wavefront sensor consists of a series of tiny lenses, called lenslets. These lenslets are arranged in an array and placed in the optic path of the beam. The lenslets all have equal focal lengths and measure a small cross-section of the beam. A photon sensor (often a CCD) is placed at the focal distance of the lenslet array. The grid of points the lenslets project onto the sensor can be used to assess the distortions of the wavefront. For a perfectly collimated wavefront the array of points projected onto the sensor will be equidistant and regular. Deviation from a perfect array at any given point will indicate that the subset of light at that particular point is distorted (lagging or leading) and is not perfectly collimated relative to the rest of the wavefront. The resolution of the wavefront measurement depends primarily on the number of lenslets used in the array. For light-starved systems, like astronomy telescopes, the array may only consist of as few as four lenslets. However, since the SLO has an abundance of light, the wavefront sensor for the SLO consists of over 500 lenslets, allowing for a much finer measurement of the wavefront. The measurement of the aberrated wavefront is sent to a deformable mirror to essentially counter-aberate the light so the resultant reflection is planar. The deformable mirror (7.2 mm diameter, 97 actuators membrane; AL-PAO, Montbonnot-Saint-Martin, France) consists of a silver film allows for great precision in wavefront correction. This design can achieve near diffraction-limited images of the retinal surface with a Strehl Ratio of ∼0.9. This allows for imaging of features of the retina of only a few microns in size (Roorda et al., 2002).
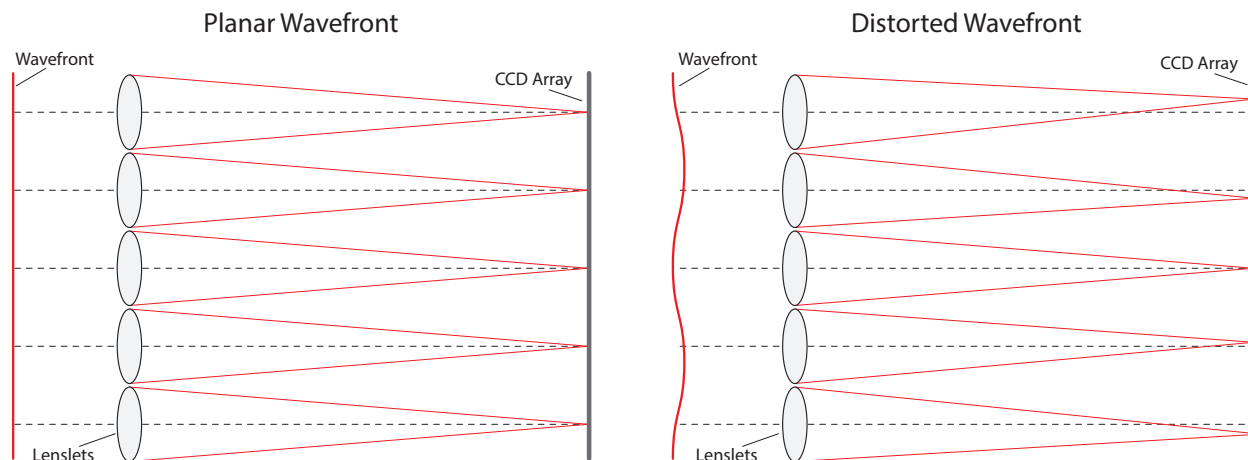
Figure 1.3: A one dimensional example of the effects a distorted wavefront will have on focus points of light from the lenslet array. Left: An example of a perfectly planar wavefront falling upon the lenslet array. Note that the focal points are regular and equidistant in this case. Right: An example of a distorted wavefront falling upon the lenslet array. Each lenslet effectively subsamples the wavefront and determines the local distortions. In this one-dimensional example the lenslets will measure the local slope of the distorted wavefront, however the same principle applies on the actual 2 dimensional lenslet array.

1.3   Strip-Based Eye Tracking with SLO

Although the SLO system can acquire high spatial resolution videos of the retinal surface, the videos themselves are at a 30Hz framerate. The small eye movements that occur during acquisition of the SLO movies operate at a much higher temporal resolution than 30Hz, necessitating a finer scale tracking of the gaze direction in time than the framerate of the movies would allow. However, because this system uses a raster scanning technique (i.e., each frame is acquired over time), eye motion information is available beyond the 30Hz frame rate. Any eye motion that occurs during the acquisition of a frame appears as a distortion, either a shearing (or horizontal eye motion) or a compression or expansion (for vertical eye motion). This information can be extracted to achieve eye traces at temporal resolution many times greater than the frame rate of the movies (Stevenson and Roorda, 2005; Vogel et al., 2006). First the reference frame is generated from a subset of frames within the movie. One frame is used as the initial reference and then subsequent frames are cross correlated onto this first frame. Those frames that correlate well (typically those that do not contain saccades or blinks) are overlaid and averaged onto this frame in order to generate a coarse reference that is a composite of many frames. By averaging many frames together the signal-to-noise ratio of these images is greatly improved. This coarse reference frame is then refined by taking the same frames used to generate it and dividing those frames into horizontal strips. These horizontal strips are the cross-correlated onto this coarse reference and overlaid together to

generate a fine reference. By dividing the frames into strips intraframe distortion due to fixational eye motion can be ameliorated. Once the fine reference is generated, every frame of the movie is then divided into strips and cross-correlated against this fine reference to generate a final eye trace, as well as a stabilized movie and a stabilized frame. The stabilized movie shows the motion of the eye but with the frame moving as features of the retina remain in place. The stability of features within the stabilized movie gives an unambiguous record for the quality of the eye trace. The stabilized frame is simply the frames of the stabilized move summed and averaged together, which gives an high quality static image of the retinal lattice. This analysis is done offline using custom software written in Matlab (The Mathworks, Natick, MA (Stevenson et al., 2010; Agaoglu et al., 2018). This technique allows collection of eye traces at high spatial ($<$1 arcmin) and temporal ($\sim$1KHz) resolution.
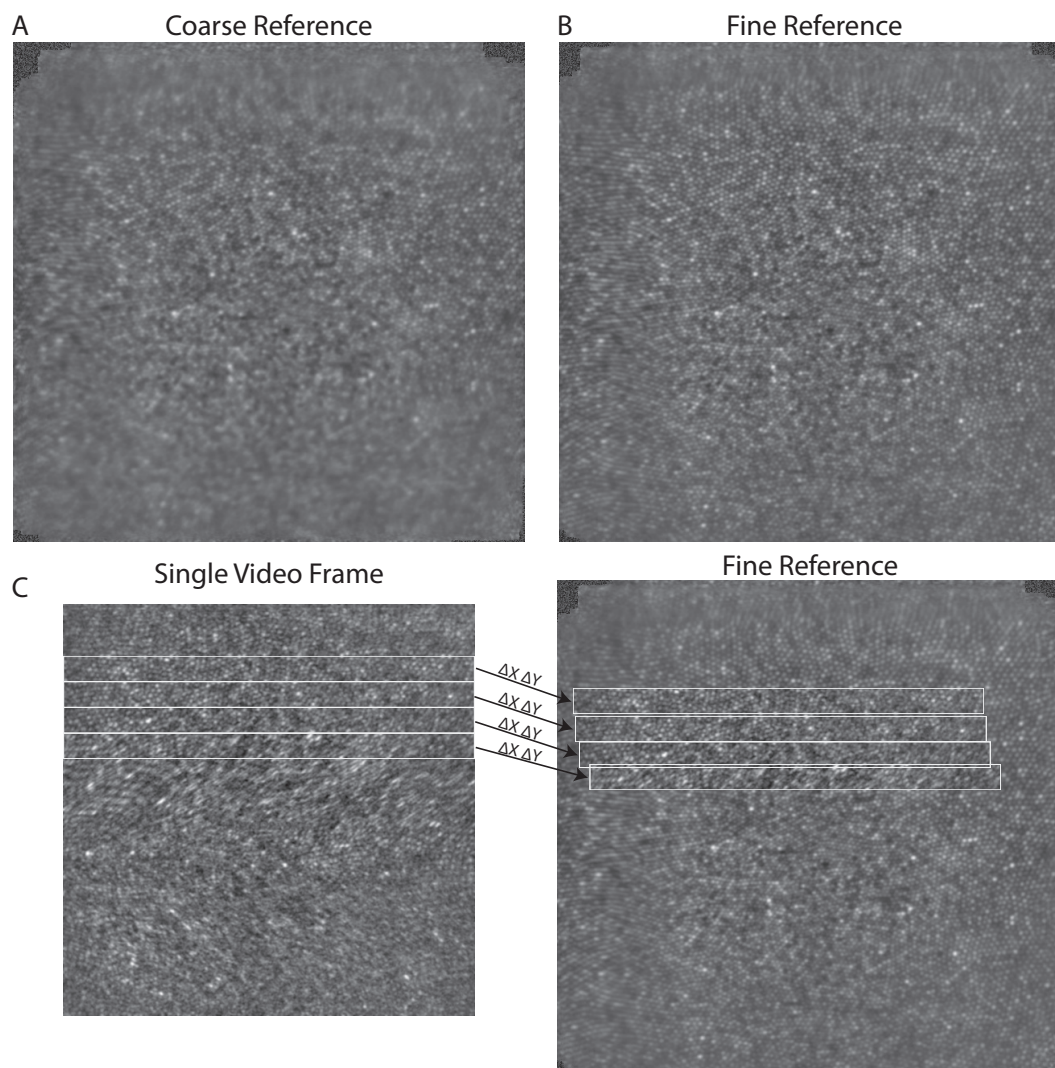
Figure 1.4: An example of the stabilization procedure used to extract the eye motion from an AOSLO movie. In this example 500 frames were used for generating the references. A: The coarse reference frame is first generated by aligning a subset of frames from a single AOSLO movie onto one another. Note the blurring at the edges due to small movements of the eye. B: The fine reference frame is then generated by using the same subset of frames from the coarse reference and dividing them into a number of strips and aligning the strips individually. Note the increased signal-to-noise ratio between the coarse and fine reference. This particular subject imaged extremely well and their underlying cone mosaic is clearly visible even in the coarse reference. C: A representation of the final strip analysis used to extract the eye motion. Each individual video frame is broken into a number of strips and cross-correlated against the fine reference. The ΔX and ΔY values needed to align these strips to the reference give a record of the eye motion over the acquisition of that particular frame.

1.3.1 Artifacts from Torsion

Through the use of high quality reference frames and the strip-based eye tracking technique described above, a high resolution measure of the stimuli's path across the retinal surface can be reconstructed. However the AOSLO system has a major drawback that is not present in more conventional eye trackers. Torsion, which represents a rotation of the eye along the roll axis (as opposed to pitch and yaw which leads to vertical and horizontal motion respectively) causes a rotation of an entire frame relative to the reference that the current strip-based technique is unable to extract. Instead, when using strip-based eye tracking the best correction to account for a torsional rotation of the eye is to shift the strips in the top part of the frame in one direction and the strips in the bottom part of the frame in the opposite. This leads to a sawtooth pattern in the eye motion trace that occurs at the framerate of the system. This is readily evident in the power spectrum of the drift eye motion, which shows clear and distinct peaks at the framerate of the system and all higher harmonics. An example of this sawtooth artifact can be seen in Figure 1.5
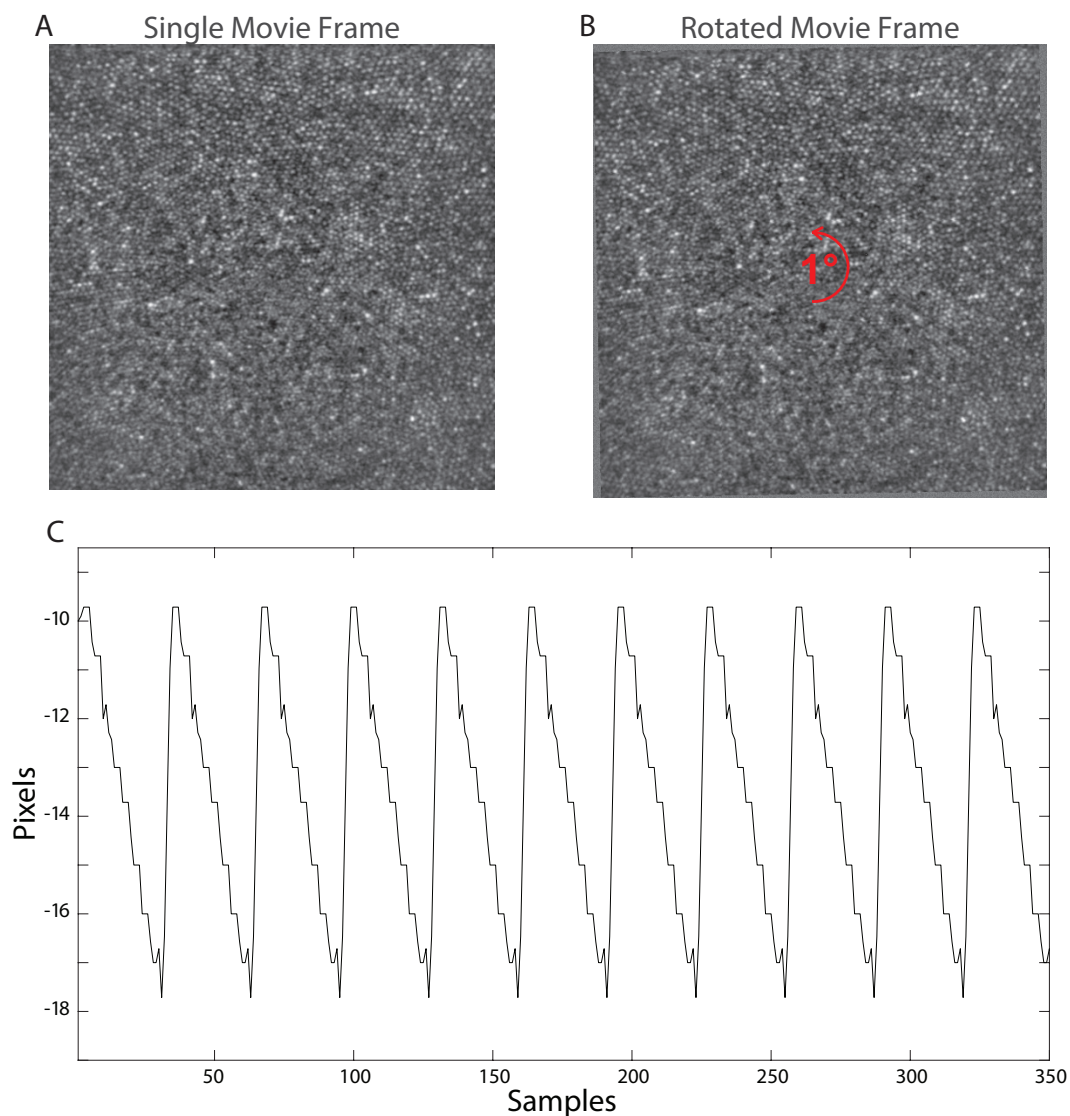
Figure 1.5: Effects of torsion on the strip-based stabilization technique. In this example a single movie frame was taken and rotated counterclockwise 1 degree. Then a movie was constructed using this rotated frame. This movie simply consisted of this rotated frame repeated 10 times, with no motion or variation. After this movie was created it was stabilized using the strip-based technique described above with the single original movie frame as the reference. The temporal resolution was 960Hz (32 strips per frame) A: A single AOSLO movie frame. B: This same frame rotated counterclockwise by 1 degree. C: The output horizontal eye position. Notice the clear and distinct sawtooth pattern from what is otherwise a static image.

1.4 Binocular Coordination of FEM

The binocular coordination across the two eyes is a vital aspect of depth perception and stereopsis, and without a tight coupling between the movements of the two eyes diplopia (double-vision) can occur as well as a loss of stereovision. However, in normal day-today life we rarely suffer from diplopia, indicating that these small eye movements do not preclude normal binocular vision and depth perception. The binocular coordination between the two eyes is of particular interest to those looking into certain oculomotor pathologies, such as amblyopia or strabmismus. However, little research has been conducted examining the coordination of fixational eye motion across the two eyes, largely owing to the difficulty of obtaining reliable measurements. It is remarkably difficult to study the coordination of fixational eye movements. Most systems lack the requisite resolution to accurately and precisely track the line of sight on such a small scale, and the necessity of doing so in a perfectly temporally synced way for both eyes simultaneously only heightens the difficulty. For a thorough review on binocular eye movements and steropsis see Howard and Rogers (2008) and Otero-Millan et al. (2014).

1.4.1 Conjugacy of Fixational Eye Motion

Most studies examining the characteristics of fixational eye movements under binocular conditions find that the eyes have a strong tendency to move in a conjugate manner, that is the direction and amplitude of eye movements tend to be near equal across the two eyes. This is particularly true for microsaccades, likely owing to their relatively large amplitudes compared to drift and tremor. Many studies looking to examine the conjugacy of microsaccades found that microsaccades were very well coupled between the two eyes (Krauskopf et al., 1960; Møller et al., 2002; Ditchburn and Ginsborg, 1953). Each eye has a strong tendency to make microsaccades in the same direction and with the same amplitude, thereby changing the version position without changing the vergence position. These studies have found a small subset of microsaccades would be orthogonal to one another, generally to correct for vergence errors during fixation. Fixational drift has not been as thoroughly examined as microsaccades, owing to the high spatial and temporal resolution needed to obtain an accurate measure of this tiny motion. Those few studies that have been able to obtain a reliable measure of drift during binocular vision have found that drift is also often conjugate across the two eyes, but not nearly as strongly as microsaccades (Ditchburn and Ginsborg, 1953; Simon et al., 1984). Disconjugate drift motion will change the vergence position of the eyes and can be a source of vergence error. However, across a long enough time scale drift has been found to be largely corrective for vergence errors on average (Engbert and Kliegl, 2004; Stevenson et al., 2016). Interestingly, Ditchburn and Ginsborg (1953) reported decades ago that drift would often exhibit a wave-like behavior, where the two eyes would move together but orthogonally while frequently reversing direction.

Chapter 2

Software Development

Although the scanning laser ophthalmoscope can obtain high resolution traces of a stimuli's motion over the photoreceptor mosaic, eye tracking was not the original use of this technology. This technology developed to image the cells on the retinal surface and study retinal diseases. Using the SLO as an eye tracker presents it's own set of unique challenges compared to more conventional eye tracking technology. In order to properly utilize the traces extracted from stabilizing SLO movies, we've developed a toolbox to process and analyze these data, as well as correct some of the more unique artifacts that are present in SLO traces. This toolbox is publicly available at `https://github.com/AbsenceOfRick/AOSLO-TSLO-Analysis`.

2.1 Analysis Software

This analysis pipeline was created to handle to raw eye traces that the stabilization software (Stevenson et al., 2010; Agaoglu et al., 2018) generates in order to create a more functional eye trace for analysis. This software was designed to be universally applicable to any needs for analyzing eye motion from SLO movies. It is made of a single master function and a number of modular subfunctions that makes it easy to add new analyses into the pipeline. The stabilized eye traces, while being an accurate reflection of the eye motion during the acquisition of a movie, need further processing to be usable for an in-depth analysis of fixational eye motion. This software creates a more coherent picture of the eye's motion by accounting for a number of inconsistencies in the raw traces, such as unlabeled events, missing samples, noisy or aberrant eye positions, and artifacts. More detail on how this is done will be discussed below.

2.1.1 Trace Labelling

The eye position from the SLO movies are a raw representation of the path a stimulus moves over the retinal mosaic. However, as discussed earlier, fixational eye motion is a patterned discrete behavior with different components. The eye traces from the SLO movies

do not differentiate between drifts, saccades, and tremor. This toolbox was written, first and foremost, to correct and label these eye traces.

In order to properly label these eye traces this toolbox uses straightforward tried-and-true techniques to identify different components. Similar techniques have been used to label fixational eye motion from other high-resolution eye trackers, and even more conventional eye trackers, before. The most obvious component of the eye traces is microsaccades. Microsaccades can be identified by a sudden ballistic increase in speed, eventually landing on a new point in space some distance away. These events are readily evident to a trained observer examining the raw eye traces. In order to identify microsaccades, a simple speed threshold is used. When the eye moves above a certain threshold (default 1.5deg/sec) this is labeled as the beginning of a saccade, with the end of the saccade being the point when the eye falls back below this threshold. Consecutive microsaccades that occur too closely in time to one another (default 15ms) are merged into one. This is to eliminate post-saccadic overshoot from being flagged as a separate event, as well as handle cases where the tracking may fail or mislabel some part of the saccades, causing the instantaneous speed to briefly fall below the threshold. Blinks are identified in the output traces when the luminance of a movie falls below a threshold (default 10 in the 8-bit frames). However this means that blinks are labeled on a frame by frame basis, instead of on the scale of the sampling rate. This is a unique circumstance due to the nature of the SLO movies, but it does not pose any significant complication in the processing of the eye trace. One major issue that occurs when attempting to label blinks is that, depending on how the movies were stabilized, sometimes the blinks are not labeled at all. Instead the frames where the blinks occurred are simply removed and the trace corresponding to the frames on either side are stitched together. In this instance, the blinks are identified as a discontinuity in the time axis and a series of NaN values are inserted into the trace to make the time axis continuous. These NaN values are labeled as blinks. After blinks and saccades are identified, all remaining samples are labeled as drifts.

2.1.2   Manual Validation

Although the automatic algorithm labels the eye traces fairly accurately, it is still helpful to manually verify the labels. It is possible, due to intersubject variability or misalignment of the strips during registration, that the eye trace is either mislabeled or noisy. In these cases having the ability for the experimenter to view and verify the traces is paramount. To this end we've developed a manual validation module for this toolbox that allows the experimenter to view the eye trace and correct for a number of common mistakes the automatic algorithm tends to make. An example of this interface can be seen in Figure 2.1 below. Given the large amount of intersubject variability in fixational eye motion, it is possible that a specific subject might have lower or higher oculomotor speed and acceleration. This can cause the automatic saccade detection software to either miss obvious saccades (due to insufficient sensitivity) or mislabel particularly fast segments of obvious drift as a saccade (due to oversensitivity). This manual validation software allows the researcher to manually identify missed saccades or

shift the beginning/end points of saccades in case the automatic algorithm mislabels regions of the trace. This module also allows the user to manually select regions of the trace that were obviously misaligned and reject them from the analysis. This often happens when the cross correlation software will incorrectly place the strips in nonsensical locations. Although this is a relatively rare occurrence, the presence of these noisy sections will contaminate analysis.
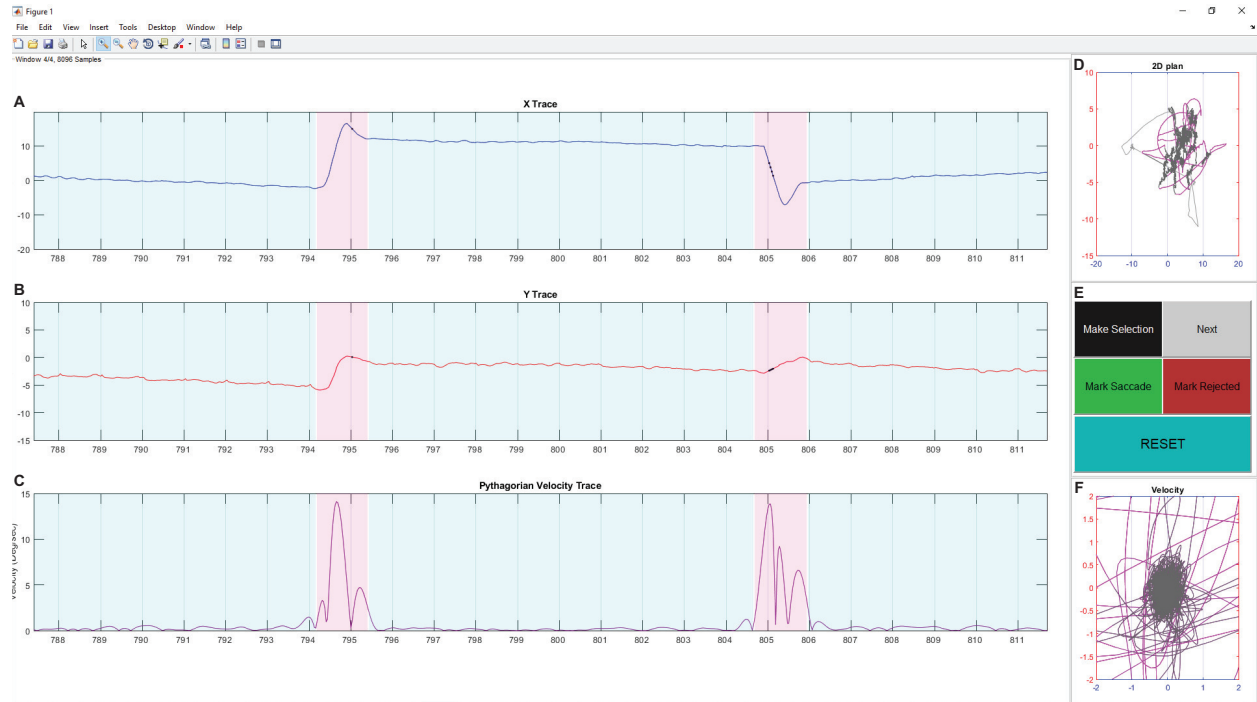
Figure 2.1: An example of the GUI used to manually verify the eye motion. For the main plots the stems denote frame demarcations and the colors denote different events, with saccades in pink and drift in light blue. A: Horizontal motion (arcmin). B: Vertical motion (arcmin). C: Instantaneous 2 dimensional speed (arcmin/sec). D: 2 dimensional position trace (arcmin) E: GUI interface. The black button (Make Selection) is for manually selecting a segment of the trace. The green button (Mark Saccade) marks the selected portion as a saccade. This is used to identify a missed saccade or to adjust the beginning/end of a mislabeled saccade. The red button (Mark Rejected) marks the selected portion as rejected in case of noisy tracking (not shown in this example). The blue button (Reset) resets the GUI in case of accidental selection. The grey button (Next) continues to the next portion of the trace. F: 2 dimensional velocity trace (arcmin/sec)

2.1.3 Torsion Correction

The effects of torsion on the AOSLO eye trace outlined in Chapter 1 are serious artifacts that are not present in more conventional eye trackers. Although this artifact does not

preclude an accurate tracking of a stimulus's position over the retina, severe torsion can cause a systematic offset in this estimate. Luckily, since the top of the frame is shifted in one direction and the bottom is shifted in the opposite, the net motion from torsion on a frame-by-frame basis is 0. This can still cause spurious measurements of the finest parts of the eye trace, such as drift. The periodic nature of this motion means it will not contribute any significant DC offset in the drift eye motion. However, if a researcher wished to examine the drift in detail, this artifact will be a serious hindrance. We have developed a technique that can measure and correct the torsion on a frame-by-frame basis. This technique relies on measuring the mean motion of the frames (taking advantage of the net 0 impact torsion has on the trace on a frame-by-frame basis) over some period of time to get a rough estimate of the real trajectory of the eye. This is accomplished by taking the slope of a line fitted to the mean position of a specific frame and the frame immediately preceding it as well as the frame immediately following it. Then, this mean motion is subtracted from a single frame's samples and any remaining slope within each frame is attributed to torsion. An example of how this is done can be seen in Figure 2.2 below. This technique has been verified by examining the power spectra of fixational drift. The power spectra of a perfect sawtooth consist of a spike at the frequency of the sawtooth and gradually degrading spikes at all higher harmonics. These spikes are visible overlaid onto the $1/f^2$ power spectra of drift. This technique was used thoroughly in Chapter 3 and the effects of the sawtooth's removal on drift power spectra will be discussed in much greater detail in Chapter 3.
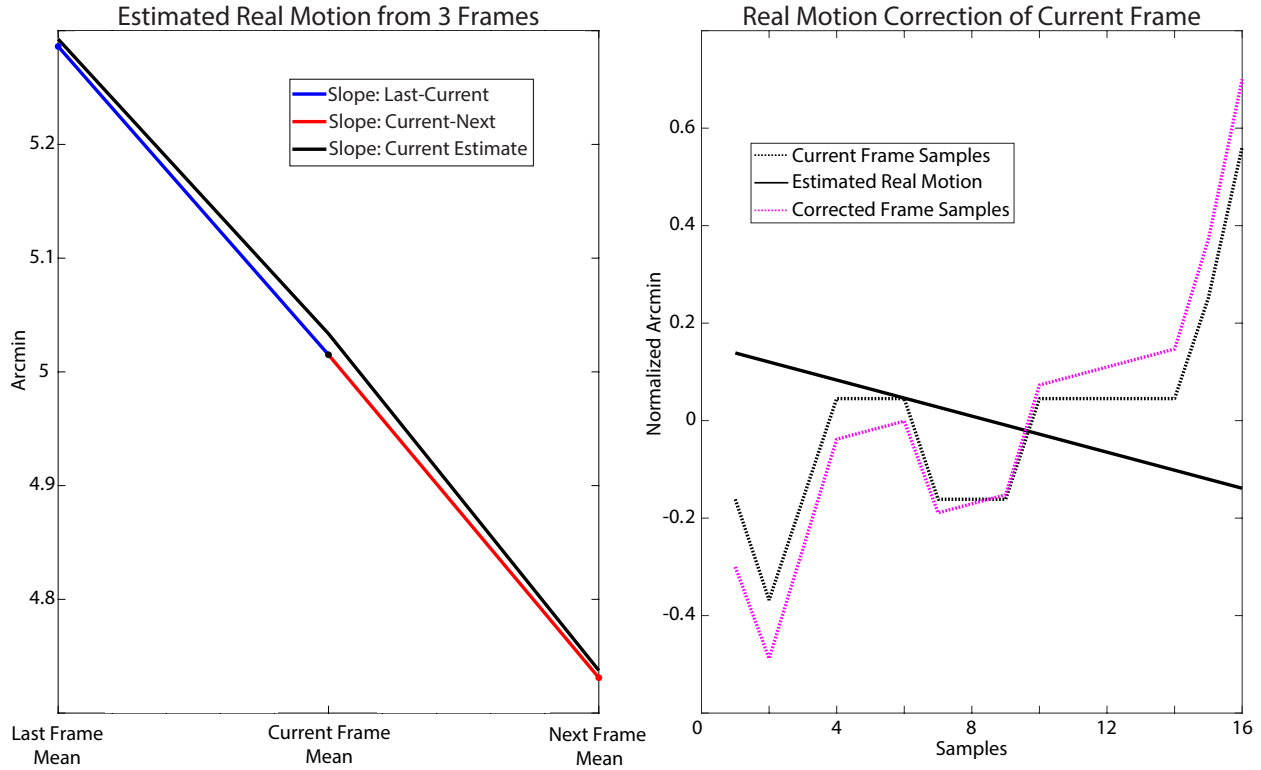
Figure 2.2: An example of how the sawtooth artifact arising from torsion is removed from an AOSLO trace. Left: Since torsion has a net zero motion within a single frame, the average motion of a frame is used to acquire an estimate of the real underlying eye motion trajectory. The slope of a line connecting the mean position of the previous frame to the mean position of the current frame (blue slope) is averaged with the slope of a line connecting the mean position of the current frame to the mean position of the next frame (red slope) in order to obtain a rough trajectory of the real underlying eye motion (black slope). Right: The slope of this motion (black slope) is removed from a single frame's worth of samples (black, dotted). Any remaining slope in the samples after the real motion correction (magenta, dotted) is attributed to torsion and removed from the current frame's samples to get a corrected estimate of the motion of the eye during acquisition of the current frame without the sawtooth artifact from torsion.

Chapter 3

The Effects of Fixational Tremor on the Retinal Image

3.1 Abstract

The study of fixational eye motion has implications for the neural and computational underpinnings of vision. One component of fixational eye motion is tremor, a high-frequency oscillatory jitter reported to be anywhere from ∼11–60 arcseconds in amplitude. In order to isolate the effects of tremor on the retinal image directly and in the absence of optical blur, high-frequency, high-resolution eye traces were collected in six subjects from videos recorded with an adaptive optics scanning laser ophthalmoscope. Videos were acquired while subjects engaged in an active fixation task where they fixated on a tumbling E stimulus and reported changes in its orientation. Spectral analysis was conducted on periods of ocular drift, with all drifts being concatenated together after removal of saccades from the trace. The resultant amplitude spectra showed a slight deviation from the traditional 1/f nature of optical drift in the frequency range of 50–100 Hz, which is indicative of tremor. However, this deviation rarely exceeded 1 arcsecond and the consequent standard deviation of retinal image motion over the tremor band (50–100 Hz) was just over 5 arcseconds. Given such a small amplitude, it is unlikely tremor will contribute in any meaningful way to the visual percept.

3.2 Introduction

3.2.1 Fixational Eye Motion

Even during intersaccadic periods of fixation the eye is never still; small eye movements constantly shift the retinal image over the photoreceptor mosaic. These movements can shift a stimulus over dozens of photoreceptors every second. Fixational eye motion has typically been categorized to fall into three main components: (a) Microsaccades, small ballistic movements similar to larger saccades, (b) Ocular drift, a slow Brownian-like movement shifting

the gaze only a few arcminutes, and (c) Tremor, a high-frequency oscillatory jitter roughly the size of a foveal cone (Ditchburn and Ginsborg, 1953; Ezenman et al., 1985; Ko et al., 2016; Ratliff and Riggs, 1950; Rucci and Poletti, 2015). Extensive research has been conducted on the functional and perceptual consequences of microsaccades and drift (Bowers and Poletti, 2017; Burak et al., 2010; Engbert, 2006; Kagan et al., 2008a; Ko et al., 2010; Kuang et al., 2012; Martinez-Conde et al., 2006; Ratnam et al., 2017; Rolfs, 2009; Rucci et al., 2007), but the perceptual consequences of tremor and its possible functional role are still not fully understood.

3.2.2 Tremor

Reports of tremor vary widely on the statistical nature and magnitude of this motion. Tremor is generally defined as an increase in eye motion amplitude at high frequencies. The bandwidth of tremor is often reported as falling between 50 Hz and 100 Hz, whereas the magnitude of motion has been found to be as small as 11.1 arcseconds (average standard deviation of eye movements within the tremor band reported by Ko et al., 2016) and as large as 1 arcminute (visual observation of eye traces by Ratliff & Riggs, 1950) and some reports question the existence of tremor at all (Stevenson et al., 2010). Few studies have set out to examine the implications of tremor for vision, largely due to the technical difficulties of accurately measuring tremor with conventional eye trackers. There is some evidence that tremor could contribute to perception by synchronizing retinal ganglion cells (Greschner et al., 2002) or through stochastic resonance of visual noise (Hennig et al., 2002).

3.2.3 Eye Tracking

Most reports of tremor stem from the use of high-resolution eye tracking techniques, such as dual-Purkinje image (DPI) eye tracking (Crane and Steele, 1985; Ko et al., 2016), scleral search coils (Houben et al., 2006), reflections from small mirrors placed on contact lenses (Ditchburn and Ginsborg, 1953; Ratliff and Riggs, 1950; Riggs and L, 1968; Steinman et al., 1973; Yarbus, 1968), and reflections from the cornea directly (Ezenman et al., 1985). Each of these trackers has the potential spatial and temporal resolution to measure tremor, but each relies on tracking some part of the anterior segment or lens of the eye and inferring the motion on the retina. The current study looks to examine the effects of tremor on the retinal image directly using an adaptive optics scanning laser ophthalmoscope (AOSLO), a relatively novel method of tracking the eye that relies on imaging the retinal surface directly (Stevenson and Roorda, 2005; Vogel et al., 2006).

3.3 Methods

3.3.1 Adaptive optics system

Movies of the retina are obtained through the use of an adaptive optics scanning laser ophthalmoscope (AOSLO; Roorda et al. (2002)). In the AOSLO system, a focused point of light is scanned across the retina in a raster pattern to obtain high-resolution movies of the photoreceptor mosaic during fixation. In the most recent version of the system, a supercontinuum light source provides a point source of light that is relayed through the optical path to a fast horizontal resonant scanner (16 kHz) and a slow vertical scanner (30 Hz), which together sweep the point across the retina in a raster pattern, and the deformable mirror (7.2 mm diameter 97 actuator membrane; ALPAO, Montbonnot-Saint-Martin, France), which compensates for the aberrations of the eye. The light reflecting from the retina is relayed and descanned back through the optical path to a custom Shack-Hartmann wavefront sensor, which measures the aberrations, through a confocal pinhole (conjugate to the retinal plane of focus), and then to a photomultiplier tube, which is used to record the scattered light, pixel-by-pixel, to reconstruct an image of the retinal surface. The AOSLO system used is equipped with four wavelength channels: 840, 680, and 543 nm channels are used for imaging and stimulus delivery, and a 940 nm wavelength channel is used for wavefront sensing. In this particular experiment, 840 nm (40–60 microWatts average power) was used for imaging and 543 nm was used to provide a stimulus for fixation (see Experimental design). The vergence of all wavelengths were adjusted in the light delivery arm to compensate for longitudinal chromatic aberration so that all wavelengths were in simultaneous focus on the retina (Grieve et al., 2006; Harmening et al., 2012). Custom software was used to operate the entire AO control system. Measurement and correction were performed over the entire pupil diameter up to a maximum of 7.2 mm. This system generally obtains near diffraction-limited images of the retina with high enough spatial resolution to resolve foveal cones.

3.3.2 Strip-based eye tracking

The images obtained by the AOSLO system were compiled together in a continuous sequence to create movies of the retina. The movies were acquired at 30 Hz (the frequency of the slow vertical scanner) and were composed of 512×512 pixels. The size of the raster on the retina was computed by imaging a calibration grid on a model eye to be $0.9°$ so that each arcminute is subtended by $\sim$10 pixels. Eye movement traces were acquired from the movies using an offline algorithm that utilized a strip-based cross correlation technique to obtain eye traces at higher temporal resolution than the frame rate of the system (Stevenson and Roorda, 2005; Vogel et al., 2006). Since each frame was acquired over time, additional temporal information on eye movements, which manifest as unique distortions within each frame, is available beyond the 30 Hz frame rate. The top of each frame occurs earlier in time than the bottom, and by dividing each frame into strips and analyzing the movement in a

strip-wise manner, eye motion traces can be acquired with a much higher temporal sampling rate than the frame rate of the movie (30 Hz). The eye motion sampling rate is the frame rate multiplied by the number of strips per frame, so the temporal resolution of the eye motion can be adjusted by increasing or decreasing the number of strips used in the cross-correlation. For the current study, 64 strips were used per frame, giving an eye-motion sampling rate of 1920 Hz. The eye motion correction has been done in real time (Arathorn et al., 2007) and offline (Stevenson et al., 2010). For all analyses in this paper, eye motion computations were done offline after the videos were acquired. Eye motion traces were converted from pixels to units of arcminutes using the scaling described above. For offline analysis, an oversized composite reference frame is generated for each movie by averaging together and roughly aligning selected frames of the movie. The size of the composite reference is dependent on the extent of eye motion during recording. Each frame of the movie is then divided into 64 strips that are 8 pixels in height and run the entire 512 pixels of the frame width. Each strip is cross-correlated against the reference frame in order to obtain vertical and horizontal offsets of the eye position compared to the reference at that instance. Each strip represents one sample of the eye trace and the strips from each frame are strung together into a continuous sequence to obtain eye traces at a rate of 1920 Hz from the 30 Hz AOSLO movies. This technique can detect eye motion with amplitudes smaller than one arcsecond (Stevenson et al., 2010).

3.3.3 Eye movement parsing

Once the raw eye traces were acquired, they were parsed using an automatic algorithm. First, erroneous or noisy eye motion traces recorded during blinks or during periods in which the image quality was very poor were identified by labeling frames in the movie in which the average luminance of the total frame fell below a threshold. Second, saccades were identified using a speed threshold, wherein saccade onset was considered the point in which the eye moved above 1.5°/s and saccade offset was considered the point in which the eye fell below 1.5 °/s. Saccades falling below an amplitude threshold (3 arcminutes) were not considered for analysis. Any consecutive saccadic events that occurred within 15 ms of each other were merged into one event in order to automatically eliminate saccade overshoot. This technique is similar to the one employed by Ko et al. (2016). Drift was identified as all intersaccadic periods of eye motion. Saccade detection was verified by human observers manually to identify any saccades the automatic algorithm may have missed. In the data sets collected here on young healthy eyes with normal fixation (see Experimental design), the automatic algorithm captured most saccades and only a small number had to be manually identified. Whenever a frame contained poor data or a saccade, the entire frame was flagged and not considered in the analysis. The first sample of each drift frame was repositioned to align with the last sample of the previous drift segment to eliminate discontinuous jumps from saccades and blinks. This technique was used because intersaccadic periods of drift are generally small (see statistics in the Results section), which poses a constraint on the resolution of the Fourier analysis. Drifts stitched together in this method were cropped

together using only full frames worth of samples ($\sim$3 ms periods), that is, any portion of a drift that began or ended in the middle of a frame was not included. On average, 40 ms were cut from the beginning and end of each drift before stitching them together, although the exact amount varied from drift to drift. This was done in order to better eliminate periodic artifacts at the frame rate of the system arising from torsion or reference frame distortions (see section under heading Reference frame and torsion correction). This method assumes stationarity of intersaccadic drift segments (disregarding any polar bias in drift direction). All of the eye motion traces are available for download in the resources section of this website: roorda.vision.berkeley.edu. Each data file contains two traces (a) a complete eye motion trace with each sample identified as a drift, saccade, blink, or bad data, and (b) the concatenated drift segment that was used for the spectral analysis.

3.3.4   Reference frame and torsion correction

Although the eye motion traces after parsing produce continuous segments of isolated drift at high sampling rates, the traces still contain motion artifacts caused by distortions in the reference frame as well as torsional eye movements. Torsional eye movements produce a sawtooth waveform that repeats at the frame rate of the system, whereas reference frame distortions present as a short random walk overlaid onto each frame's motion. In the case of the current study, a reference frame distortion will be a random walk constructed of 64 samples (the number of strips/samples per frame) and overlaid onto every set of 64 samples throughout the trace. Fortunately, both of these artifacts are periodic and introduce peaks in the amplitude spectra that are isolated to the frame rate of the system and higher harmonics only; they do not affect the underlying amplitude spectrum anywhere else. In the eye motion traces from the offline-processed videos, reference frame artifacts are largely removed by using multiple frames to generate a composite reference frame (Stevenson et al., 2010). By combining a series of frames, the distortions of the individual frames are averaged out (Bedggood and Metha, 2012). Torsion, however, may change over the course of a video and the periodic sawtooth must be removed from each frame individually. (for a full description of the algorithms to measure and remove torsional artifacts see Chapter 2) For the purposes of this paper, it is sufficient to state that the sawtooth artifact was measured and removed from the eye motion trace prior to further analysis.

3.3.5   Eye motion analysis

The amplitude spectra of these drift segments were then analyzed by doing a multitaper spectral analysis on the X and Y eye motion traces expressed in units of amplitude in arcminutes versus time in seconds, originally proposed by (Thomson, 1982) and more recently reviewed by (Babadi and Brown, 2014). The analysis was done using the command pmtm in the MATLAB Signal Processing Toolbox (MathWorks, Natick, MA). Briefly, the spectral analysis method involves running an fast Fourier transform (FFT) on a motion trace using overlapping and mutually orthogonal tapers (discrete prolate spheroidal sequences). For

the current manuscript, the analysis was run on overlapping segments (50% overlap) of 1 s, which comprised 1,920 samples each, allowing for analysis of frequencies sampled in 1-Hz steps up to 960 Hz. In the MATLAB function, the time half-bandwidth product was fixed at 2.5 and the last taper was dropped to maximize spectral concentration ratios in the Slepian sequences. In general, we found that variations in the bin size, amount of overlap, and time half-bandwidth product had a minimal effect on the final output in the frequency range of tremor (50–100 Hz). The square root of the output spectra were taken to convert from power to amplitude in arcmin/Hz in order to better capture the motion on the retina and to facilitate a more direct comparison with previously published results. To compute the actual motion of the retinal image caused by tremor, we employed similar methods used by Ko et al. (2016) wherein we bandpass-filtered the eye motion traces to contain only the motion within the tremor band (50–100 Hz) and computed the standard deviation of the filtered eye motion trace.

3.3.6 Validations

Two tests were done in order to validate the ability of the AOSLO system to track eye motion. The first validation aimed to test the ability of the entire AOSLO system to record motion. A model eye was attached to a galvanometer and oscillated in a diagonal sinusoidal pattern at four frequencies (4 Hz, 16 Hz, 64 Hz, and 256 Hz) while being recorded with the AOSLO system. To enable analysis of the noise floor of the system, a video was also recorded of the nonmoving model eye. Thirty-second movies were collected at each frequency with a fixed amplitude of 0.45 arcminutes at an angle of $\sim$26.6° to give horizontal and vertical amplitudes of $\sim$0.4 and $\sim$0.2 arcminutes, respectively. The motion traces from these movies were analyzed by averaging together the amplitude spectra of a series of nonoverlapping 1-s segments of the trace. Unlike the eye motion traces, a simple FFT computation was used here since it is more suitable for single-frequency motion traces.

The second validation was a simulation aimed to test the ability of the strip-based cross correlation technique to recover motion from actual AOSLO movies. First, a real AOSLO movie was manipulated digitally to add a distortion that would result from tremor. Specifically, we added motion from a parabolic-weighted band of frequencies ranging from 50–100 Hz with a peak in the amplitude spectrum of 2 arcseconds. Following the manipulation, the movie was analyzed using the strip-based cross correlation technique described above to obtain traces of eye motion at 1920 Hz. The modified AOSLO movie was analyzed using the same multitaper spectral analysis methods described above. The amplitude spectra of the motion trace from the manipulated movie were compared against the amplitude spectra of the motion trace from the original movie. We felt that validation with an actual AOSLO video was important because in a model eye, the luminance and contrast of the image is static and the retina moves in only the direction of the galvanometers. The AOSLO movie, in contrast, contains actual eye motion and has more variable luminance owing to changes in the adaptive optics correction as well as actual changes in reflected intensity from the retina (Pallikaris et al., 2003). If our eye motion analysis can recover the frequency and amplitude

of high frequency motion that has been added to an actual movie, then we can be confident that the strip-based eye-tracking algorithm would be able to detect real eye motion at these frequencies.

3.3.7 Experiment design

Six healthy subjects with normal vision were recruited for the study. Informed consent was obtained from each subject and all experimental procedures were reviewed and approved by the UC Berkeley Institutional Review Board and adhered to the tenets of the Declaration of Helsinki. To prepare subjects for AOSLO imaging, one drop of tropicamide (1%) and phenylephrine (2.5%) solution were administered topically 15 min prior to imaging to temporarily dilate the pupil and paralyze accommodation. For measurements of the motion amplitude spectrum in human eyes, AOSLO videos were recorded while subjects fixated a letter E optotype (Figure 3.1). The E was projected directly onto the retina at the center of the raster scan using the green channel (543 nm) in the AOSLO system while imaging was done with 840 nm near infrared (NIR) light. The combination of NIR light and very weak green background light (caused by light leaking through the acousto-optic modulator (AOM) that was used to project the letter E), formed a dim, reddish background over the extent of the raster scan. The E was 5 arcminutes in height, corresponding to a 20/20 letter. To keep the subject engaged in the fixation task over the course of each video, the E changed orientation to random positions at random time points ranging from 0.5 to 1.5 seconds. The subject was instructed to report its orientation every time it changed via the use of arrow keys on a keypad. Each subject completed five 30-s trials, for a total of 2.5 min of fixational eye motion per subject (with the exception of one subject, who only had four 30-s trials).
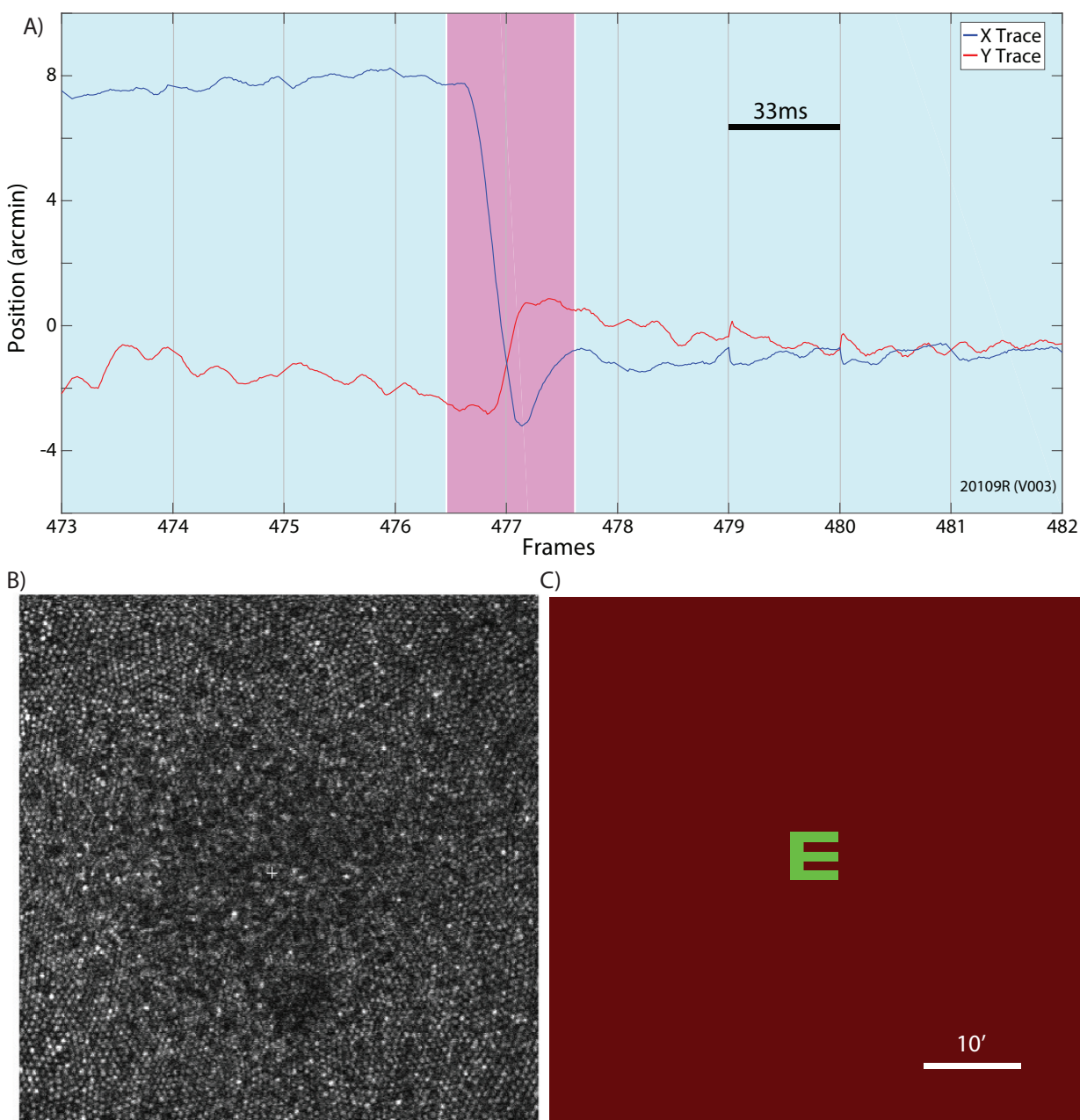
Figure 3.1: (A) An example of an eye trace taken from an AOSLO movie. A microsaccade (magenta background) is clearly distinguishable from the ocular drift (blue background). Gray vertical gridlines demarcate frame boundaries from the AOSLO movie. Each frame is acquired over 33 ms as indicated by the scale bar. (B) An example of an image/frame from an AOSLO movie. The cone mosaic can be resolved even at the fovea. (C) An example of the AOSLO raster with a green letter E as it would appear to the subject. The small discontinuities in the eye trace at the boundaries between frames 478–479 and 480–481 are likely the result of tracking errors that occur at the edges of the frame. They are infrequent and an example is included here for full disclosure. Errors like this contribute to the peaks in the amplitude spectrum at the frame rate and higher harmonics. All original eye motion traces are available for download at http://roorda.vision.berkeley.edu/

3.4   Results

The purpose of the current study was to examine ocular tremor using the AOSLO system. The AOSLO system has advantages over other types of high-resolution eye tracking techniques due to its ability to image the retina directly, instead of having to infer retinal image motion from measurements taken from the anterior segment of the eye. Since the AOSLO system is a relatively novel eye tracker, the capabilities of the AOSLO system to measure small movements was validated in two ways. The first validation aimed to test the entire AOSLO system's ability to detect sinusoidal oscillations from a moving model eye. The resultant amplitude spectra plotted in Figure 3.2 showed a clear peak at the frequency of the sinusoidal oscillation for all input frequencies. Even though the motion of the sinusoidal oscillation was just a fraction of an arcminute, the AOSLO system was able to reliably recover the amplitude of the input motion. We suspect that the slight reduction in measured amplitude at the higher frequencies resulted from small relative shifts in the frequency between the galvanometer scanner or the AOSLO frame rate over the course of the video. Such shifts will cause slight distributions in amplitude to nearby frequencies in the spectrum. Human eye motion, which is broadband in frequency, will not be seriously affected by this limitation.
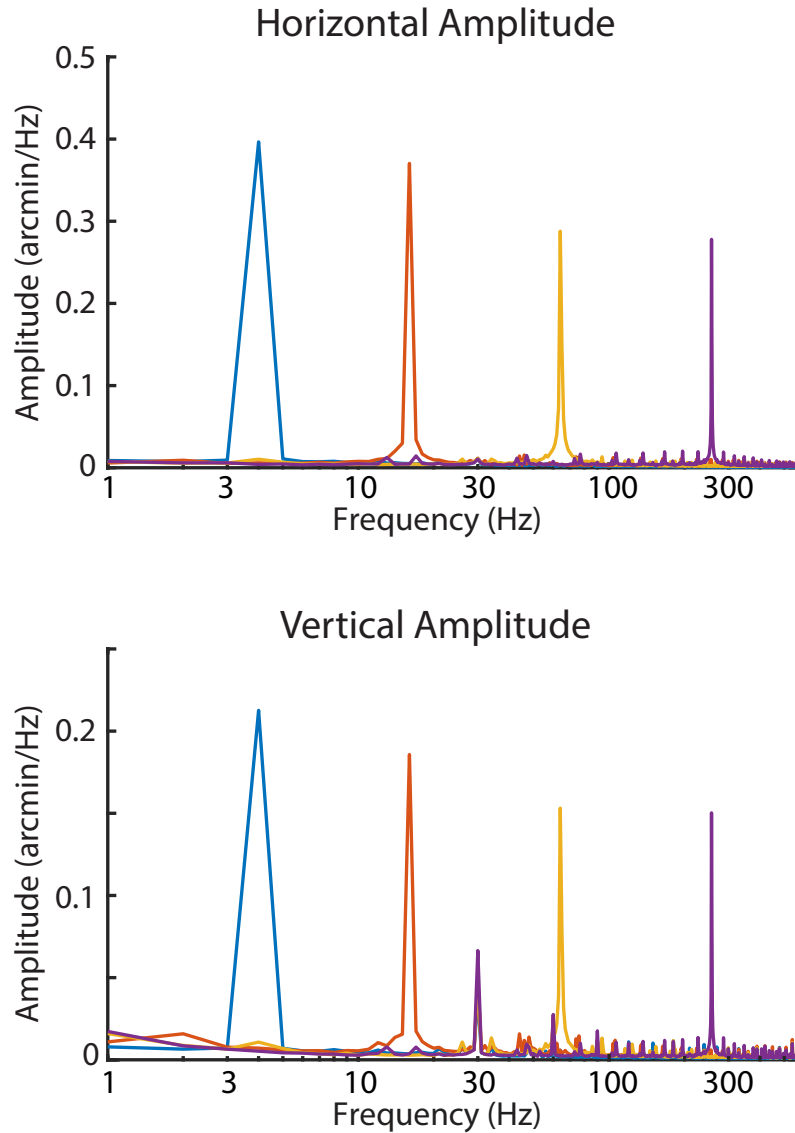
Figure 3.2: Amplitude spectra of motion from a moving model eye. Four videos were recorded from a model eye that was oscillated at four frequencies (indicated by the colors) with a fixed amplitude. For each of the input frequencies, the resultant amplitude spectra showed a peak of ~0.4 arcminutes (horizontal) and just above ~0.2 arcminutes (vertical) for each corresponding frequency. This simulation shows that we can recover high frequency, low amplitude motion with high fidelity from AOSLO videos.

The second validation was a simulation aimed to test the capabilities of the strip-based cross-correlation technique to recover motion from AOSLO movies. A real AOSLO movie was manipulated to add distortions consistent with a bandwidth of tremor-like motion between 50–100 Hz with a peak in the amplitude spectrum of ~2 arcseconds. The distorted movie was then processed using the same strip-based cross correlation technique used on the human

eye motion. The resultant amplitude spectrum of the manipulated movie compared to the original movie plotted in Figure 3.3 showed strong correlation except for a large bump in the amplitude spectrum between 50–100 Hz consistent with the input motion. Note that the amplitude spectra should theoretically match perfectly at all other frequencies, however the inclusion of the tremor-like signal subtly changed the samples which were flagged as saccades, so some small discrepancy between the two is to be expected. Regardless of this discrepancy, the strip-based cross-correlation technique was able to recover the tremor-like motion added to the modified movie. We compared the standard deviation of the retinal motion within the tremor band (50–100 Hz) with and without the artificially added tremor and found that the magnitude of motion added as expected. The standard deviation of the bandpass-filtered traces from the original movie was 4.8 arcseconds. The standard deviation of the tremor-like signal inserted into the movie was 10.9 arcseconds. The vector sum of the two components (square root of the sum of the squares) was 11.9 arcseconds, which matches up closely with the standard deviation of the modified movie (11.7 arcseconds).
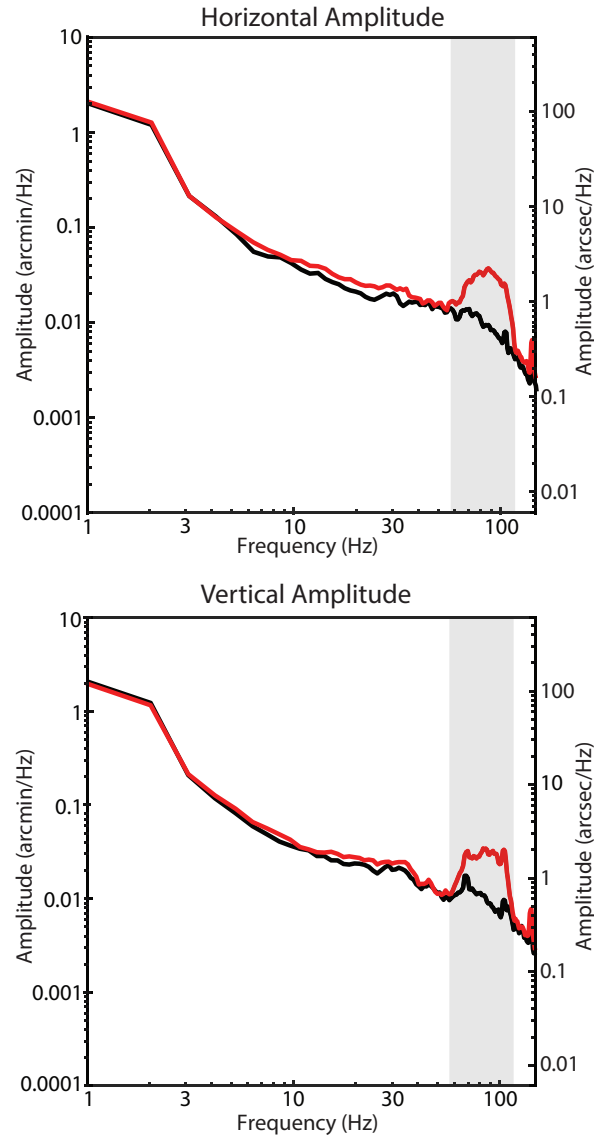
Figure 3.3: Amplitude spectra of an original AOSLO movie (black) and an AOSLO movie (red) that was digitally modified to include distortions from simulated tremor. The bump in the amplitude spectra from the modified AOSLO movie is clearly visible and matches the spectrum of the tremor that was digitally added.

Prior to performing the spectral analysis, basic metrics to describe the fixational eye motion were computed from the traces for each of the six subjects. Heat maps for saccade landing positions and drift segments are shown on Figure 3.4. All subjects showed normal fixational eye motion. Subjects made microsaccades roughly once per second (1.10 ± 0.57 microsaccades/s) with a normal amplitude (7.5 ± 1.5 arcminutes), speed (375 ± 49 arcminutes/s), and duration (36.9 ± 6.9 ms). Intersaccadic periods of drift also showed relatively normal statistics. Amplitude (3.8 ± 0.9 arcminutes); span, defined as the maximum dis-

tance from the mean location (3.2 ± 0.7 arcminutes); speed, defined as the mean of the instantaneous speed between each pair of samples within each drift segment (79.6 ± 15.6 arcminutes/second); and duration (620 ± 245 ms, min ~33 ms, max 5.8 s) were all within normal parameters and were consistent with previous findings (Martinez-Conde et al., 2004; Rucci and Poletti, 2015). All values are ±1 SD. Overall subjects did show slightly better fixational stability than average. This is likely due to a combination of the task (subjects had to attend to the orientation of a tumbling E) and the fact that all subjects were experienced in psychophysics experiments using the AOSLO system.
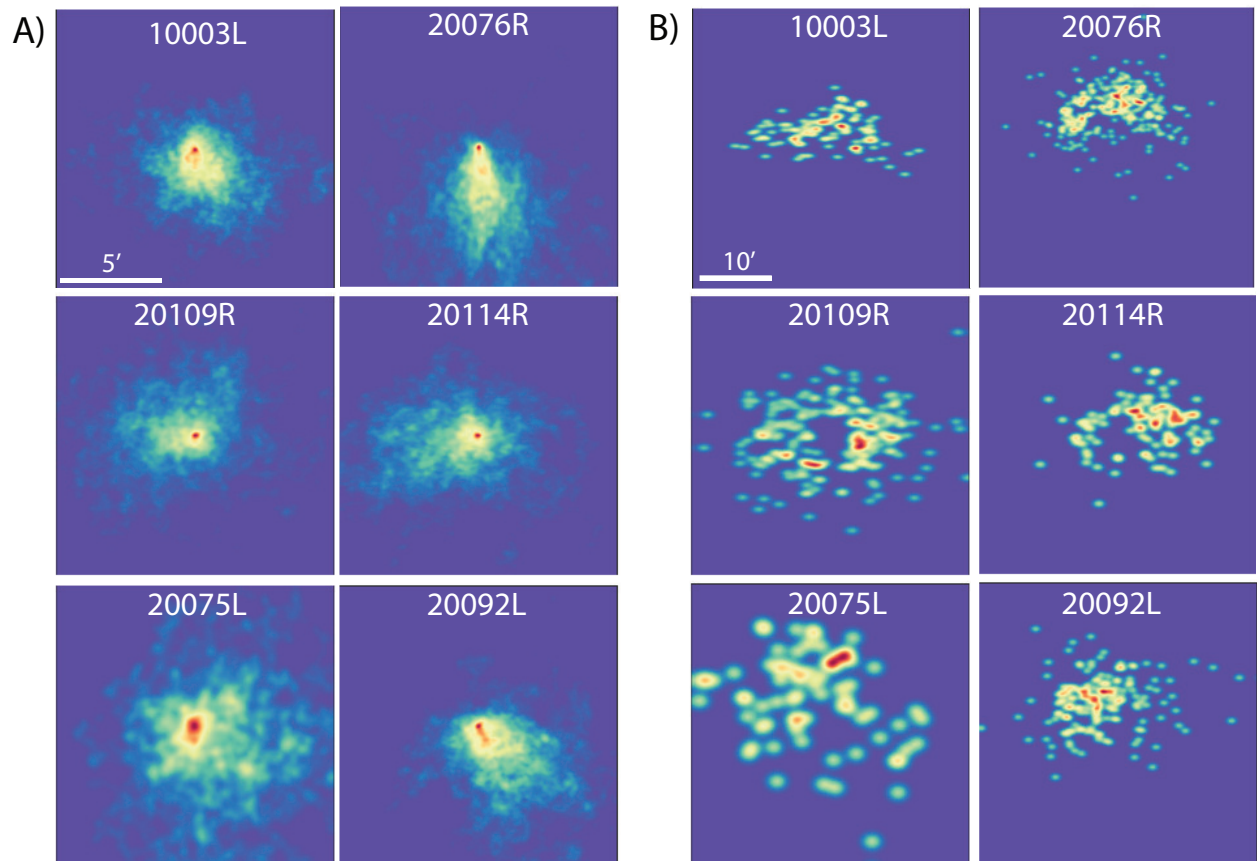
Figure 3.4: Heat maps of horizontal and vertical position for (A) drift and (B) saccades in space. All drift segments were offset to begin at 0,0 for display purposes, and the saccade heat maps show end points of each saccade relative to an origin of 0,0. Note that tendencies for some eyes to drift in specific directions are balanced by saccades in the opposite direction (e.g., 20076R). Scale bars of 5 and 10 arcminutes are shown on the top left panels of A and B, respectively. The color scale for each panel is normalized based on the amount of data available for that subject, with red indicating the highest frequency of occurrence.

Using the technique outlined in the methods, the amplitude spectra of fixational drift for each of the six subjects were calculated. The average horizontal and vertical amplitude

spectra for all six subjects are shown in Figure 3.5. Similar to that reported in Ko et al. (2016), the spectra show a steeper than 1/f fall-off in amplitude but becomes 1/f after 10 Hz. However, unlike Ko et al. (2016) and others who used different tracking methods, the characteristic deviation from 1/f in the amplitude spectra indicative of tremor was very small. Although there was a slight elevation within the band of 50–100 Hz, the average amplitude of this deviation never exceeded 1 arcsecond and only one subject had an amplitude greater than 1 arcsecond within that range, much smaller than previous reports of tremor. To estimate the magnitude of motion on the retina caused by tremor, we bandpass-filtered all eye traces between 50 and 100 Hz and computed the standard deviation of the resulting traces. The distributions of the bandpass-filtered traces are shown in Figure 3.6. The standard deviation of the bandpass-filtered traces was 5.10 ± 0.66 arcseconds horizontally and 5.51 ± 0.57 arcseconds vertically. The standard deviation of the bandpass-filtered noise floor was 0.13 arcseconds horizontally and 0.55 arcseconds vertically, well below the measurements of the actual eye traces.
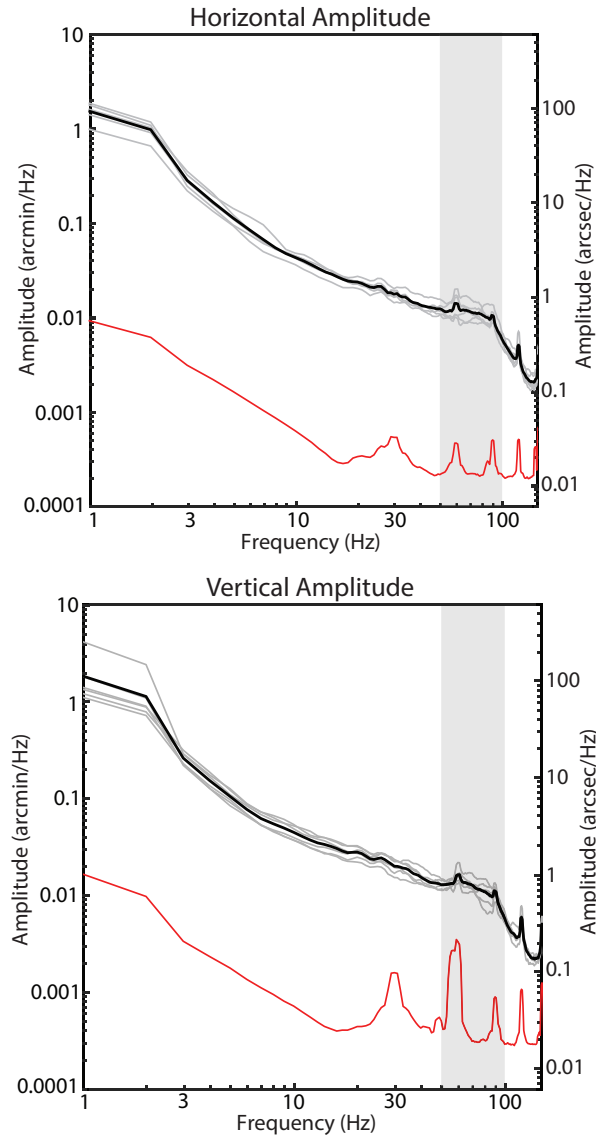
Figure 3.5: Amplitude spectra for six subjects and noise floor measured from a nonmoving model eye (red). Individual subjects' amplitude spectra are plotted in gray and the mean is plotted in black. The slight deviation from a linear 1/f falloff between 50–100 Hz (highlighted by the vertical gray bar) indicative of tremor is, on average, just over 1 arcsecond, much smaller than all previous reports of tremor. The spikes in the spectra are from periodic artifacts in the traces caused by torsion and residual reference frame distortions. The peaks are slightly broadened due to the multitaper spectral analysis method (Babadi and Brown, 2014). Note that the peaks appear larger in the noise floor, but this is due to the logarithmic scaling.

Bandpass-Filtered Horizontal Traces

10003L, Std: 5.06

20076R, Std: 5.31

20109R, Std: 4.19

20114R, Std: 6.07

20075L, Std: 4.56

20092L, Std: 5.39

Distribution of eye motion (arcseconds)
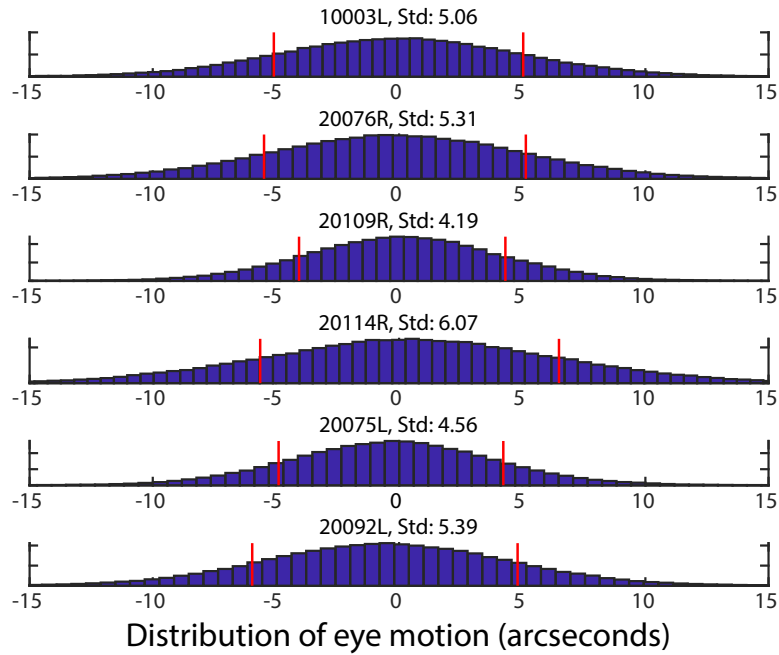
Bandpass-Filtered Vertical Traces

10003L, Std: 5.29

20076R, Std: 5.40

20109R, Std: 5.67

20114R, Std: 6.57

20075L, Std: 5.11

20092L, Std: 5.00

Distribution of eye motion (arcseconds)

Figure 3.6: Histogram distribution of the bandpass-filtered drift traces. Each plot is a combination of all trials for each subject. The red stems indicate ±1 SD from the mean. All values are in seconds of arc.

3.5 Discussion

To assess the functional role that the tremor component of eye motion might have for human vision, its characteristics (amplitudes and frequencies) on the retinal image must be known. However, reports of the tremor in the literature v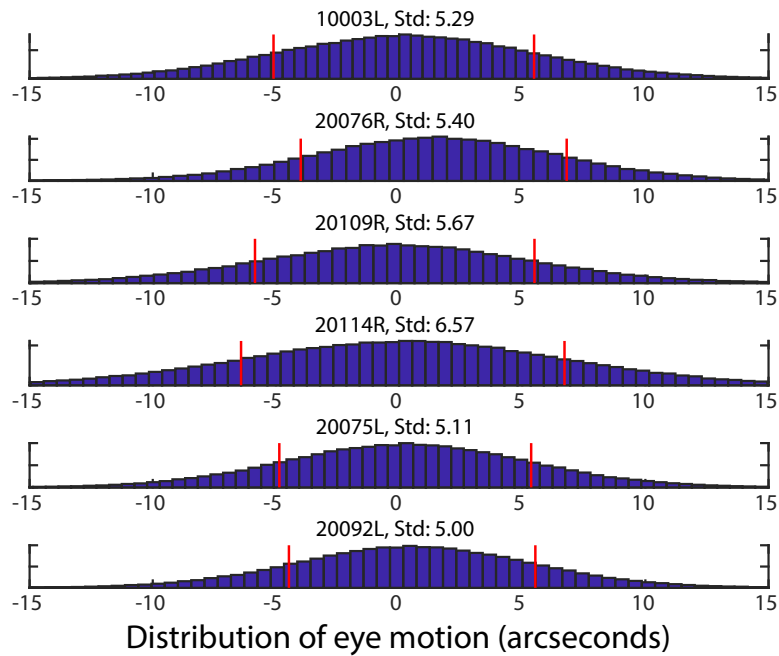ary widely and all published measurements are based on measurements made from the anterior segment, either from corneal reflections or a combination of reflections from the cornea and lens. In this paper, we present the first measurements of tremor on the retinal image directly employing the AOSLO as a retinal-image–based eye tracker.

As it is a new technology for eye tracking, we first validated that the AOSLO is capable of recording microscopic eye movements with high fidelity up to very high frequencies. The first validation used a moving model eye to test the AOSLO system's ability to precisely record eye motion from movies of high frequency and low amplitude. The second validation was a simulation where we used a digitally modified AOSLO movie to examine the capabilities of the offline analysis software to extract eye motion traces from movies of a real human retina as the subject engaged in an active fixation task. We then used the AOSLO to record fixational eye movements in six normal subjects to measure tremor on the retinal image directly while they engaged in an active fixation task, in this case reporting the orientation of a small rotating letter E optotype.

The temporal sampling rate of eye traces from the AOSLO system (1920 Hz) allows for analysis of frequencies up to 960 Hz, well beyond the 50–100 Hz bandwidth of tremor. The noise floor in the 50–100 Hz range measured from a non-moving model eye is <0.03 arcseconds in the amplitude spectrum, which is well below the amplitude of any eye motion previously reported, including tremor. The eye tracking capabilities of the AOSLO system, combined with direct retinal imaging, is therefore uniquely capable of analyzing the effects of small eye movements on the retinal image.

The measurements of retinal image motion from the six normal subjects during active fixation showed some evidence of tremor in the frequency range of 50–100 Hz, but it was very small with only one subject having an amplitude greater than 1 arcsecond at any point across that range. Even when there was evidence of tremor, the amplitude spectrum was monotonic (continuously declining) with increasing frequencies and was little more than a slight deviation from a 1/f curve. The resultant motion of the image on the retina caused by tremor was just over 5 arcseconds.

3.5.1 Comparison with previously published results

We compared our measurements of tremor to three other reports in the literature wherein a similar spectral analysis of fixational eye motion was performed. With regard to the peak height of the tremor component in the amplitude spectrum, Ko et al. (2016) reported peak amplitudes of 4.8 and 3 arcseconds for the horizontal and vertical directions, respectively; Ezenman et al. (1985) reported a peak of 6 arcseconds and Findlay (1971) plots show peaks ranging from 3 to 4 arcseconds. By comparison, the average height of the peak amplitude

within the tremor band in our study was less than 1 arcsecond with only one subject out of the six having a slightly higher peak (Figure 5). It is important to note that the peak amplitude cannot be used directly to compute the magnitude of motion on the retina caused by tremor. To estimate the actual motion on the retina, we bandpass-filtered the eye motion traces to contain the tremor band only (50–100 Hz), computed the standard deviation of the resultant motion, and found the standard deviation of motion to be 5.10 ± 0.66 arcseconds horizontally and 5.51 ± 0.57 arcseconds vertically. Ko et al. (2016) did a similar analysis (their bandpass filter was between 40 and 80 Hz) and they found motions with standard deviations of 13.2 and 9 arcseconds for vertical and horizontal motion respectively. Neither Eizenman et al. (1985) nor Findlay (1971) did a bandpass-filtered eye motion analysis but, given the similarity between the spectra for those studies to Ko et al. (2016), we expect the motion caused by tremor to be similarly higher than ours.

Although there are some suggestions tremor could contribute to the visual percept through synchronization of retinal ganglion cells or through influencing the behavior of visual neurons in the brain (Greschner et al., 2002; Hennig et al., 2002), these studies generally assume the amplitude of tremor is around the scale of a foveal cone. Given that tremor on the retina is much lower than this, the possibility of this movement influencing the visual percept will need to be reexamined.

3.5.2   Effects of cycloplegia

In the current study, it was necessary to dilate and cycloplege subjects' eyes in order to achieve the best image quality for image-based eye tracking. Cycloplegia relaxes the ciliary muscle but, being that it is a sphincter muscle, it actually leads to an increase in the tension on the lens. Decreasing the tension on the lens is known to increase lens wobble (He et al., 2010). How this intervention might affect the magnitude and or presence of tremor is not well known. To address this question, we performed similar eye-tracking measurements in a tracking scanning laser ophthalmoscope (for details on that system, see Sheehy et al. (2012)) for subjects that had not been cyclopleged. One 1-min video was collected for six subjects (four of whom also participated in the previous experiment) and the amplitude spectra were assessed using the same method described above. The results are shown in Figure 3.7. Compared to AOSLO, the spatial sampling resolution was ~5 times lower (0.5 arcminutes per pixel) and the optical image resolution was worse. Nevertheless, the frequency bandwidth was the same and the noise floor was sufficient to detect tremor. Amplitude spectra measures without cycloplegia showed a similar small amplitude (>2 arcseconds) of tremor.

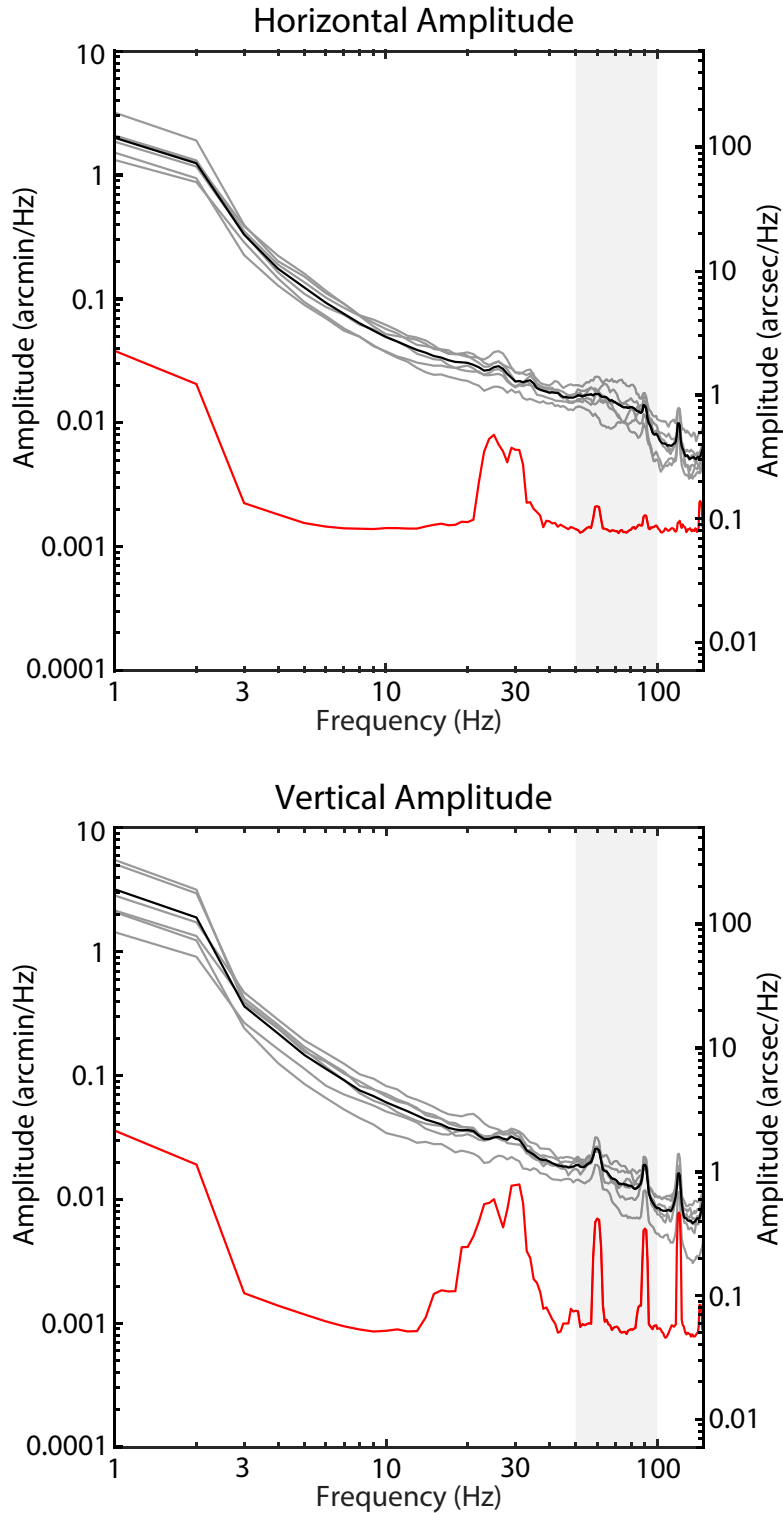Figure 3.7: Amplitude spectra from fixational drift measured using the tracking scanning laser ophthalmoscope from six subjects without cycloplegia and noise floor measured from a nonmoving model eye (red). Individual subjects' amplitude spectra are plotted in gray and the mean is plotted in black. Eye motion with or without cycloplegia show a similar lack of any significant tremor on the retinal image.

3.5.3 Why are the current measurements lower than all previous reports?

At first glance, the negligible increase in the amplitude spectrum indicative of tremor on the retinal image is a surprising finding considering previous reports. Ultimately, tremor is not a well-defined concept and there appears to be large differences across subjects, eye tracking techniques, and quantitative analyses. Nevertheless, tremor has been consistently observed in eye motion traces obtained from a number of high-resolution eye tracking systems. It is important to note however, that all reports of tremor to date have relied on tracking eye motion from the anterior segment of the eye, and retinal image motion has only ever been inferred. The current study is based on unambiguous, high-resolution measurements of the retinal image motion directly.

In the following sections, we describe a combination of optical and biophysical factors that may serve to reduce the amplitude of tremor of the retinal image relative to entire eyeball. Some evidence indicating the presence of such reduction can be found in a report where AOSLO and DPI traces were recorded simultaneously (Stevenson and Roorda, 2005). In that report, eye motion traces from the two modalities were very similar except after microsaccades. However, no effort to compare estimates of tremor between the two tracking modalities was attempted in that paper. They speculated that overshoots caused by lens wobble that were detected in the DPI trace resulted in much smaller overshoots in movement of the retinal image as recorded in the AOSLO traces. Their experimental finding confirmed predictions made by (Deubel and Bridgeman, 1995).

There are two stages to modeling the differences between eye motion measured from the anterior segment and eye motion measured from retinal images. First is optical modeling and second is an analysis of the temporal dynamics of the lens.

3.5.4 Optical modeling

We used optical design software (Zemax, LLC, Kirkland WA) to model the effects of lens tilt and decentration in the schematic eye model available from the Zemax website (`http://customers.zemax.com/os/resources/learn/knowledgebase/zemax-models-of-the-human-eye`). Based on the optical model, we found that lens displacements and tilts both give rise to retinal image motion Figure 3.8. The relationships between lens displacements and retinal image movement are:

$$R = 0.95 \times d \tag{3.1}$$

and

$$R = 0.042 \times \theta \tag{3.2}$$

where R is the retinal image displacement in degrees, d is the lens translation in mm, and θ is the lens tilt in degrees. Lens translations have the most profound effect: Tilting the lens also displaces the retinal image: the retinal image moves in the same direction as

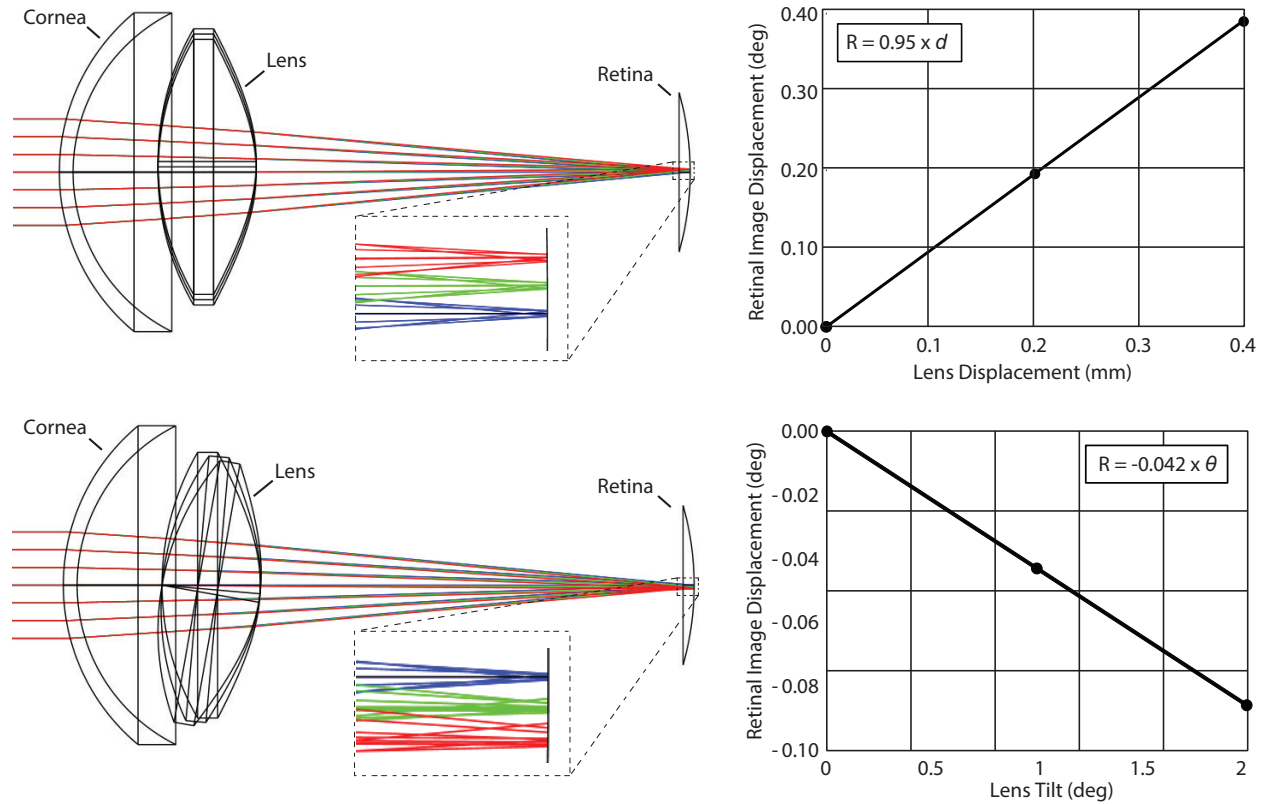the movement of the optical axis of the lens, but the magnitude is negligible compared to displacement.

Figure 3.8: Zemax simulations. Top: Lens displacements of 0, 0.2, and 0.4 mm with their respective ray traces in blue, green, and red are illustrated in the drawing. A magnified inset is indicated by the dashed box. The focused spot is displaced in the same direction as lens displacement; 1 mm of lens displacement gives rise to just under 0.95° of displacement of the retinal image (assuming that 1° of visual angle corresponds to 300 microns on the retina). Bottom: Lens tilts of 0°, 1°, and 2° were tested (amplified tilts of 0°, 5°, and 10° with their respective ray traces in blue, green, and red are shown in the schematic and magnified inset to help to visualize the effect). In this case, the image displaces in the same angular direction as the tilt, but the effect is very small: 1° of tilt gives rise to -0.042° of retinal image displacement.

3.5.5   Temporal dynamics of the lens

We considered how the lens might move within the tremoring eye, due to the fact that it is elastically supported by the zonules and ciliary body. The manner in which the crystalline lens moves in the eyeball following a saccade is a classic example of damped harmonic motion. (He et al., 2010) confirmed this to be the case when they measured the motion of the lens

following a saccade. A frequency analysis of one of the subjects in the study revealed the resonant frequency of lens oscillation to be about 20 Hz. The oscillation of the lens over time, x(t), dampens quickly, after about 1–2 cycles, following the equation:

$$x(t) = ae^{-vt}/2cos(\omega_1 t - \phi) \tag{3.3}$$

where ω_1 is the oscillating frequency, given by:

$$\omega_1 = (\omega_0^2 - \frac{v^2}{4})^{\frac{1}{2}} \tag{3.4}$$

ω_0 is the resonant frequency, a is the amplitude, ϕ is the phase offset (not very important), and v is a constant with dimensions of angular frequency indicating the strength of damping. Based on visual observation of the lens wobble artifact in He et al. (2010), we estimated the constant v to be about half of the resonant frequency. A damped oscillation with a constant $v = \omega_0/2$ is plotted on Figure 3.9A and shows that the oscillation relaxes after one or two cycles.
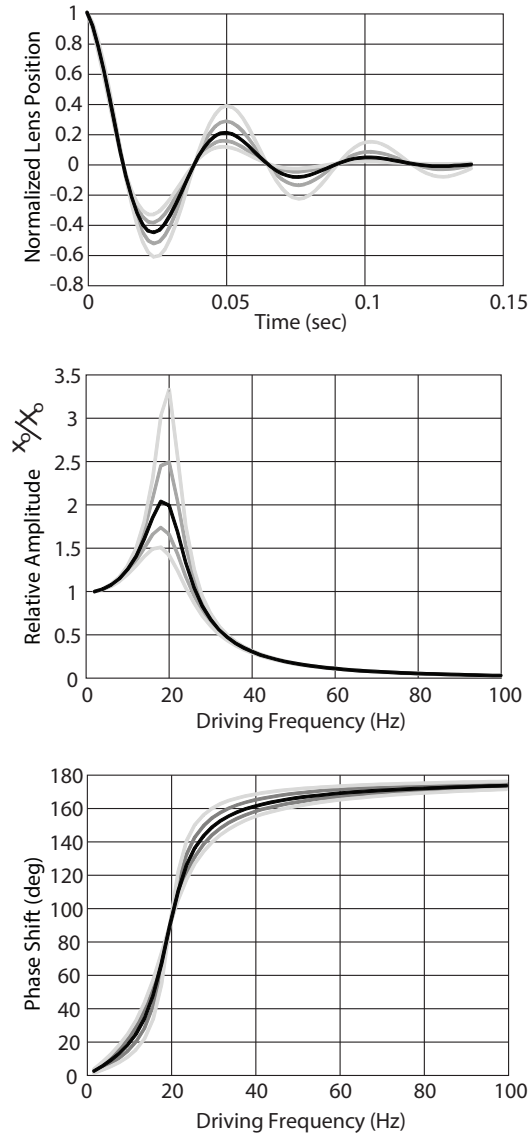
Figure 3.9: (A) Model of the position of a lens behaving as damped harmonic oscillator. In this case, the damping coefficient is 20π (10 Hz), or half of the resonant frequency of 20 Hz, which leads to about two cycles of oscillation before relaxing to zero. Gray shaded lines in all three plots show calculations with $\pm 20\%$ and $\pm 40\%$ changes in the damping coefficient. (B) Model of the amplitude of a lens behaving as a driven damped harmonic oscillator with resonant frequency of 20 Hz and range of damping coefficients. Driving frequencies near the resonant frequency gives rise to amplified motion of the lens. Driving frequencies within the range of tremor give rise to lens oscillations that are, on average, about 0.1 of the driving amplitude. Amplitudes of lens motion at the resonant frequency are highly dependent on the damping coefficient, but outside of that the trends are relatively similar. (C) Model of the phase shift of the lens motion. When driving frequencies are slow, the lens moves along with the eyeball as expected. For driving frequencies that are in the range of tremor (50–100 Hz) the lens moves in counter-phase with the eyeball. Changes in the damping coefficient have little effect on this trend.

The manner in which tremor affects the motion of the lens is a classic example of driven damped harmonic oscillation. The eyeball rotates about its center of rotation, and how the lens reacts to this motion depends on the damping constant υ and the resonant frequency ωo in the following way:

$$x_0 = \frac{\omega_0^2 X_0}{[(\omega_0^2 - \omega^2)^2 + \upsilon^2 \omega^2]^{\frac{1}{2}}} \tag{3.5}$$

and

$$\phi = tan^{-1} \left(\frac{\upsilon \omega}{\omega_0^2 - \omega^2} \right) \tag{3.6}$$

where χ_0 is the amplitude of the oscillating mass (the lens), Xo is the amplitude of the driver (eye tremor), ω_0 is the resonant frequency, ω is the frequency of the oscillating driving force (eye tremor), and ϕ is the phase shift between the driving oscillation and the oscillating mass.

Figure 3.9B and 3.9C plot the amplitude and phase of the lens oscillation that has a resonant frequency of $\omega_0 = 40\pi$ (20 Hz) and a damping coefficient of $\upsilon = 0.5\omega_0$ over a range of driving frequencies. Note that for low frequencies, we see the expected behavior where the lens moves in phase with the eye with the same amplitude. When the driving frequency reaches the resonant frequency of the lens, the amplitude oscillation increases to about 2 times. At the same time, the phase of the lens movement transitions to counterphase motion. When the driving frequency increases further, the motion is in counterphase and the amplitude approaches zero. In the limit, the lens remains perfectly fixed in place relative to the tremoring eye. The modeled behavior remains qualitatively similar with variations in the damping constant υ, which were deduced from plots in He et al. (2010). Changes in the damping coefficient by up to 40% do not change the general trends in the plots. Similarly, changes in the resonant frequency will shift the curves but not in a manner that would alter the main conclusion, which is that during tremor, the lens moves with lower amplitude than the eyeball and it oscillates in counterphase to the eyeball.

A counterphase motion of the lens relative to the eyeball would, in effect, amplify the motion that is estimated by the DPI eye tracker. The situation is illustrated schematically on Figure 3.9. In the DPI, the eye rotation is assumed to be proportional to the separation between the reflection from the cornea, P1, and the reflection from the back surface of the lens, P4, which are situated approximately at the center of curvatures of the surfaces that generate them (Cornsweet and Crane, 1973). If the lens and eyeball rotate together, as they would with slow rotations, then the separation between the two reflections (assuming small angles) is $\sim 7\phi$, where ϕ is the rotation angle (in radians; Cornsweet and Crane (1973)). If the lens moves in counterphase to the eyeball rotation with 1/10 of the amplitude (as per estimations on Figure 3.8C), then the separation is $\sim -7.3\phi$. The overall magnitude of the separation between the two reflections is 4.2% larger than the actual motion, and the sign is opposite.
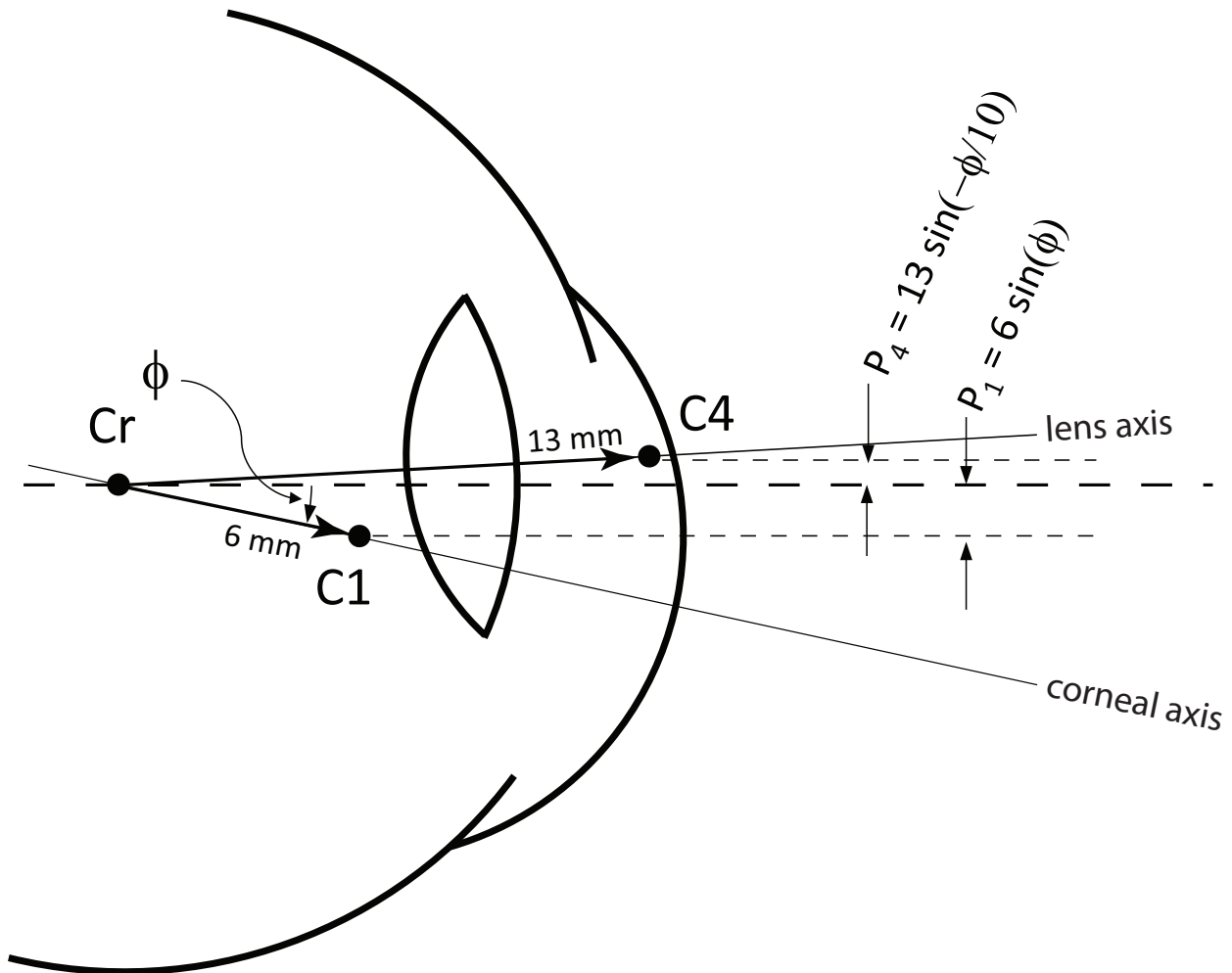
Figure 3.10: Schematic drawing showing how Purkinje reflections get displaced with rotations of the eye and lens. In the schematic, the eyeball has rotated downward by an angle ϕ and the lens has undergone a counterphase movement upward with 1/10 of the amplitude. For collimated incident light, the Purkinje reflections P1 and P4 are approximately positioned at the centers of curvature of their respective reflecting surfaces. The two reflections are displaced according to the equations on the figure. The total separation between P4 and P1 is $13_{sin}(-\phi/10) - 6_{sin}(\phi)$, which, for small angles, is -7.3ϕ.

The same lens translations that give rise to an overestimation of eye motion in the DPI would reduce the motion of the retinal image due to the prismatic effect illustrated in Figure 3.8. Consider an eyeball undergoing a tremor rotational motion with an amplitude of 1 arcminute. The amplitude of pupil displacement associated with the rotational eye movement is about 0.0029 mm (since the pupil is displaced about 10 mm from the center of rotation of the eye, giving rise to displacement of $SP = 10sin(\phi)$).

The displacement of the lens relative to the pupil will move 0.0029 mm plus an additional

10% due to the counterphase oscillation (as indicated by the charts in Figure 3.10) for a total displacement of 0.0032 mm. According to Equation 1, these lens displacements will cause the amplitude of the retinal image motion to be smaller by 0.003° (0.18 arcminutes), or 82% of the eyeball rotation. In total, AOSLO measures of tremor amplitude should be 82% - 4.2% ≈ 78% of DPI measures.

However, tremor is not only detected in DPI traces, but has been seen in eye motion traces from search coil measurements as well as from traces of reflections from the cornea, neither of which involve a measurement of the lens. In these cases, assuming there is no artifact from the search coil or corneal reflection measurement, the only effect that will diminish the movement of the image on the retina from the overall eyeball motion is the translation of the lens which, according to the Zemax model, will give rise to a retinal motion amplitude that is 82% of the eyeball rotation.

Considering the arguments made above, the reported peak tremor amplitudes of 4.8 and 6 arcseconds from Ko et al. (2016) and Eizenman et al. (1985), respectively, would reduce to 3.75 and 4.92 arcseconds. Ko et al.'s (2016) average standard deviation of eye motion within the tremor band of 11.1 arseconds would reduce to 8.7 arcseconds. Our results remain lower than previous reports, even after correction, but they are in the same order of magnitude. The remaining differences could be due to the actual fixation task: our subjects were actively engaged in a fixation/acuity task, but Ko et al. (2016) had their subjects simply fixate the center of a blank screen and Eizenman et al. (1985). had their subjects fixate on a small source at 50 cm. Additionally, other biophysical factors, such as damping of lateral motion due to the elasticity of the retinal surface, could also be present. More experiments would be required to assess the effects of these possible causes for the differences.

3.6   Conclusion

In this study, we measured the tremor component of eye motion by directly measuring the motion of the retina using the AOSLO eye tracker. We started by validating the AOSLO as a high-resolution, retinal-image–based eye-tracking technique that is uniquely able to resolve small movements of the retinal image caused by fixational eye movements. We first validated the capabilities of the AOSLO system by reliably recovering both the frequency and amplitude from recorded movies of an oscillating model eye. We were also able to recover a band of tremor that was artificially inserted into a real AOSLO movie. The AOSLO system was able to reliably measure the fixational eye motion of six human subjects as they participated in a tumbling E task. The statistics of both microsaccades and fixational drift were generally found to be within normal parameters of previous reports using high-resolution eye tracking. However, the amplitude of fixational tremor was smaller than all previous reports. Some, but not all, of the discrepancy is accounted for by optical and biomechanical factors that amplify measures of tremor-like eye motion based on measures from the anterior segment while damping the motion of the retinal image. Regardless of the cause, this paper shows that the amplitude of tremor in the eye during fixation is likely too

small to have meaningful visual consequences.

3.7   Acknowledgments

Chapter 4

Fixational Eye Movements in Passive Versus Active Sustained Fixation Tasks

4.1 Abstract

Human fixational eye movements are so small and precise that high-speed, accurate tools are needed to fully reveal their properties and functional roles. Where the fixated image lands on the retina and how it moves for different levels of visually demanding tasks is the subject of the current study. An Adaptive Optics Scanning Laser Ophthalmoscope (AOSLO) was used to image, track and present a variety of fixation targets (Maltese cross, disk, concentric circles, Vernier and tumbling-E letter) to healthy subjects. During these different passive (static) or active (discriminating) tasks under natural eye motion, the landing position of the target on the retina was tracked in space and time over the retinal image directly with high spatial (<1 arcmin) and temporal (960Hz) resolution. We computed both the eye motion and the exact trajectory of the fixated target's motion over the retina. We confirmed that compared to passive tasks, active tasks elicited a partial inhibition of microsaccades, leading to longer drift periods compensated by larger corrective saccades. Consequently, the overall fixation stability during active tasks was on average 57% larger than during passive tasks. The preferred retinal locus of fixation was the same for each task and did not coincide with the location of the peak cone density.

4.2 Introduction

When fixating our gaze on an object, our eyes are never truly at rest. Even while staring at a small object, like the bottom row of a Snellen acuity chart, our eyes are constantly in motion. Small, fast microsaccades and slow drifts constantly shift the image of the fixation target

over the photoreceptor lattice. These fixational eye movements (FEM) are a nuisance for many ophthalmic measurements such as imaging, microperimetry or retinal and refractive surgery. However, owing to the fact that FEM are the finest motor control system in the human body, they offer an opportunity for early detection and monitoring of neurological disorders (Hunfalvay et al., 2021; Montesano et al., 2018; Sheehy et al., 2020). It is also increasingly clear that these small movements are not simply noise in the oculomotor system (Rucci and Victor, 2015), but that they serve a number of important functions in the visual system, such as preventing fading (Martinez-Conde et al., 2006). Microsaccades have been found to be associated with shifts in attention (Hafed and Clark, 2002; Engbert and Kliegl, 2003) and fine scale repositioning stimuli within the foveola (Ko et al., 2010; Poletti et al., 2013; Intoy and Rucci, 2020). Whereas drift has been shown to enhance the discrimination of fine spatial details through a combination of spatiotemporal enhancement at high frequencies (i.e. whitening of the power spectrum) (Rucci et al., 2007; Rucci, 2008) and enrichment of the information relayed from the retina to the brain through dynamic sampling (Anderson et al., 2020; Ratnam et al., 2017; Burak et al., 2010; Ahissar and Arieli, 2012).

A major limitation in studying the smallest FEM is the instrument used to measure the gaze itself (Poletti and Rucci, 2016). Modern video-based eye trackers are convenient but often lack the resolution of earlier systems. Scleral search coils, which were among the first high-resolution eye trackers, have high temporal and spatial resolution (Robinson, 1963; Collewijn et al., 1975) but are invasive and difficult to use. Dual Purkinje Image (DPI) eye trackers came to use shortly after search coils and offered a noninvasive way to attain high accuracy for tracking the gaze (Crane and Steele, 1985) and still represent a reliable eye tracker today (Fourward Technologies, Gallatin, MO). However, the DPI system has been shown to have its own drawbacks. The crystalline lens, which is the source of the 4th Purkinje image that is tracked in a DPI system, can move independently from the rest of the eyeball giving rise to spurious measurements of gaze direction especially for tremor or post-saccadic overshoot. (Tabernero and Artal, 2014; Nyström et al., 2015; Bowers et al., 2019; He et al., 2010). On the other hand, modern video eyetrackers, today's most commonly used instruments in research and industry, suffer from pupil size changes. (Choe et al., 2016; Nyström et al., 2016; Hooge et al., 2016). They also simply lack the requisite spatial resolution to accurately estimate the gaze position produced by these fixational eye movements (Kimmel et al., 2012; Holmqvist and Blignaut, 2020).

FEM have drawn renewed interest in research lately as the field progresses toward higher and higher resolution structural and functional measurements of the retina and as the role of FEM in fine-scale vision continues to be examined. The effects of fixation target and task have not been thoroughly examined using high resolution retinal-image-based tracking techniques. Published studies, which primarily relied on video eyetrackers, generally showed that smaller targets elicit modestly less overall FEM compared with larger targets (McCamy et al., 2013; Kazunori et al., 2016). Other target properties, such as shape, color, contrast, blur and luminance in eliciting improved fixation stability have been more scarce (Thaler et al., 2013; Steinman, 1965; Bhattarai et al., 2019; Ukwade and Bedell, 1993) with few trends emerging except that a bull's eye and cross targets (or a combination of both) elicit

the least FEM (Thaler et al., 2013). Most experiments that require fixation use a simple static fixation target (see Thaler, 2013 for a comprehensive overview of the variety of targets commonly used).

The implication of FEM in encoding visual information (Otero-Millan et al., 2008; Mc-Camy et al., 2014), as well as their modulation during fine discrimination tasks (Rucci and Victor, 2015) is well documented. The variations in subjects' stability over different kinds of controlled tasks is less understood. A more comprehensive characterization of FEM and fixation target is therefore important for several reasons. A high resolution set of unambiguous oculomotor data reporting intra- and inter-individual variability during various tasks offers an important baseline for improved interpretation of FEM in health and disease as well as to better understand their functional roles. Finally, a more practical reason is simply to learn what target and/or fixation task might minimize overall FEM in clinical settings where motion and its consequent blurred or distorted retinal images can be detrimental, such as with fundus photography, OCT scans or fundus-guided microperimetry.

The current study aims to compare and contrast FEM during *active* tasks - those that contain temporal variation and require subject input - and *passive* tasks, where the subject is simply instructed to maintain fixation on a target. An Adaptive Optics Scanning Laser Ophthalmoscope (AOSLO) is used as an eye tracker to acquire high spatial (<1 arcmin) and temporal (960 Hz) resolution eye traces. Since the AOSLO can also obtain an unambiguous record of the motion of the target that is projected onto the retinal surface, we compare how the preferred retinal locus for fixation (PRL) relates to the location of peak cone density (PCD) for each type of fixation target.

4.3 Methods

Eight healthy subjects (self-reported), 3 male and 5 female, with normal or corrected-to-normal vision participated in the experiment. Subject ages ranged from 23-53 years old. All experimental procedures adhered to the conditions set by the institutional review board of the University of California, Berkeley and followed the tenets of the Declaration of Helsinki. Each subject read and signed a written informed consent document. Prior to imaging, the subjects' eyes were dilated and cyclopleged using 1 drop each of 1% tropicamide and 2.5% phenylephrine. The drops were used to provide maximum dilation for imaging as well as to paralyze accommodation, both of which help to ensure high quality images in the AOSLO. No detectable difference has been found between eye traces measured in an SLO system with or without dilation (see Chapter 3, (Bowers et al., 2019).

4.3.1 AOSLO System

Data were recorded using the Adaptive Optics Scanning Laser Ophthalmoscope (AOSLO) (Roorda et al., 2002), which is used to image and track the retina as well as to provide the fixation targets used in this experiment. For imaging, a point source of light is relayed

through the optical path and scanned across the retina in a raster pattern utilizing two scanners, a 16kHz fast horizontal scan and a 30Hz slow vertical scan. The reflected light is descanned through the optical path and directed through a confocal pinhole to a custom-built Shack-Hartmann wavefront sensor and a photomultiplier tube (Hamamatsu, Japan). The Shack-Hartmann wavefront sensor is used to measure the optical aberrations and send a correction to the deformable mirror (7.2mm diameter, 97 actuators membrane; ALPAO, Montbonnot-Saint-Martin, France) in the optical path. Light detected by the PMT and the positional information from the scanner are combined to construct videos of the retina with 512x512 pixel sampling resolution at a frame rate of 30Hz (the speed of the slow vertical scanner). In this experiment, the eye's pupil was kept in a fixed position relative to the AOSLO beam by restraining the subject's head movement through the use of a dental bite bar and temple mounts. The non-imaged eye was covered with an eye-patch. The imaging wavelength was 680nm, with 940nm used for wavefront sensing. The field size of the video was 0.9 x 0.9°. Using an average power of 50-70 μW, the raster scan field appeared as a bright red square flickering at a rate of 30Hz to the subject. Fixation targets were presented to the subject within the red field by turning off the scanning laser using an acousto-optic modulator (Brimrose Corp, MD) at the appropriate time points during the raster scan. To the subject, these targets appeared as black-on-red decrements. The stimuli were very sharp and had high contrast owing to the use of adaptive optics on the input scanning beam. Importantly, these decrements are also encoded directly into the video, which allows for an unambiguous measurement of the motion of the image of the fixation target over the retina. This system is capable of obtaining near diffraction-limited images of the photoreceptor mosaic and delivering stimuli with the precision of 1 pixel ($\sim$6 arcseconds). An example video from one of the concentric circle trials in this experiment is shown in the Supplementary Materials. This system has been explained in greater detail in previous manuscripts from our group. (Poonja et al., 2005; Rossi and Roorda, 2010).

4.3.2   Experiment Design

The experiment consisted of 5 different conditions: Maltese cross, disk, concentric circles, Vernier acuity, and a tumbling E (M, D, C, V, and E respectively). The Maltese cross condition (M) was chosen as it has been suggested to provide a better fixation target than the simple dot that is commonly used in fixation tasks. The disk condition (D) consisted of an annulus within the center of the raster that the subjects were instructed to fixate. Both of these conditions were simple fixation tasks where subjects were instructed to hold their gaze on the target. The concentric circles condition (C) consisted of concentric rings moving in a constricting radial motion. There were 6 rings ranging in size from 10 to 1 arcmin that were presented over the course 18 frames (3 frames per ring size) and replayed every 30 frames for a frequency of 1 Hz. The aim of the concentric rings was to provide a fixation task that was similar to the passive task in that it required no subject response, but was also similar to the active task in that it was dynamically changing. As such, it provided a control to distinguish whether any FEM differences could be attributed to the active task

or due to the fact that the stimulus was dynamic. The Vernier hyperacuity condition (V) required subjects to judge the relative displacement of two horizontal bars which appeared at random intervals (seven 6-arcsecond steps). The tumbling E condition (E) consisted of a tumbling E task where the subjects were asked to report the orientation of a letter E as it rotated randomly. The size of the E varied in seven steps, from 20/6 to 20/20 Snellen acuity. For both V and E tasks the stimulus was presented for 0.5 sec (15 consecutive frames) and there were random time intervals between presentations - evenly spread over 0.5 to 1.5 sec - where nothing was presented. The random time intervals were used so that subjects could not anticipate the next trial and were therefore compelled to maintain fixation the entire time. The V and E condition can be differentiated from the others as they both required subject judgment and response, as well as providing temporal variation. These conditions were further categorized into passive tasks (M, D) and active tasks (E, V) for further analysis depending on whether they required subject response and varied in time. The different fixation targets were presented in a pseudo-random order to eliminate any training or fatigue effects. Furthermore, subjects were given consistent instructions from a script to avoid known changes in behavior due to instruction (Steinman et al., 1967). The full script is provided in the Supplementary Materials but the primary emphasis in the instruction was for the subject to maintain their gaze throughout the entire duration of each 36-second trial task. There were five 36-second trials for each condition in total.

	Maltese	Disk	Circles	Vernier	Letter

Task (Distinction)	Passive (Static)		Attention Grabbing	Active (Responsive)	
Motion	No			Yes	
			Inward Radial Motion	Presented every 0.5s	
Size	Fixed		Variable		
	6′	11′	10′ to 1′ steps	Seven 6″ steps	20/20 to 20/6*
Response	No			Yes	

Figure 4.1: Table 1: Illustration of the 5 experimental conditions and their respective parameters. The different colors indicate distinctions between passive (green), active (purple), and mixed (orange) tasks. This color scheme will be used throughout to differentiate the 5 conditions. *Snellen fraction

4.3.3 Eye Tracking and Video Processing

Since this system utilizes a raster scanning technique (i.e. each frame is acquired over time), eye motion information is available beyond the 30Hz frame rate. This information can be extracted to achieve eye traces at temporal resolution many times greater than the frame rate of the movies (Stevenson and Roorda, 2005; Vogel et al., 2006). In order to acquire eye traces at higher temporal resolution than the 30Hz frame rate, each frame of the AOSLO movie is broken into horizontal strips and cross-correlated against a reference frame. This analysis is done offline using custom software written in Matlab (The Mathworks, Natick, MA) (Stevenson et al., 2010). This technique allows collection of eye traces at high spatial (<1 arcmin) and temporal (960Hz) resolution. Eye traces were separated into drifts and saccades using a semi-automatic software and the output was manually verified by the authors. Saccade onset was defined as the point when instantaneous speed exceeded 1.5 deg/sec and offset was defined as the point when the trace fell back below this threshold. Blinks were defined as frames of the AOSLO movie when the mean luminance fell below a threshold that was defined on a per-subject basis dependent on the average brightness of the respective movies.

The AOSLO records high-resolution videos of the retina for each trial and the fixation target is directly encoded into the video, thereby making it possible to plot the exact path of the fixation target over the photoreceptor mosaic directly. This is done in the following way. First, the eye motion traces extracted from AOSLO videos indicate how the entire retina moves, but do not directly indicate where the fixation target lands on the retina. Computing the actual retinal trajectory requires computing the ΔX and ΔY offsets that need to be applied to each eye trace to anchor it to the exact position of the fixation target on the retina. To accomplish this step, we first generate a high quality master retinal image chosen from one of the best videos recorded in the experimental session for each subject. Then we use the same cross correlation methods to align strips containing the encoded stimulus from each video with that master retinal image and determine the position of the stimulus on the master retinal image. The X-Y position corresponding to the strip that contains the stimulus is then aligned to that exact position on the master retinal image using these ΔX and ΔY offsets. In theory, the offset only needs to be computed once for a single strip, but the match between a single strip and the master retinal image can have small errors due to noise in the strip or torsion in the retinal image. So, to improve accuracy, we compute the average offsets from at least 20 unique strips, ensuring that the standard deviation of the offsets is less than 2 pixels ($\sim$0.2 arcmin). These processing steps yield accurate trajectories in retinal coordinates for every trial and every condition, all referenced to a single master retinal image.

For all of our subjects the master retinal image was of sufficient quality to label all cones across the image. Cones were labeled across the entire foveal region using a combination of automatic cone-finding (Li and Roorda, 2007) with manual intervention when necessary. Cone density was computed within a 10-arcmin diameter circular window while it traversed, pixel by pixel, across the mosaic (using a convolution process). The 10-arcmin averaging

window was chosen since it has been shown to strike an optimal balance between minimizing noise and maximizing resolution (Wang et al., 2019). The point of maximum cone density was expressed as the pixel location with the highest density value. This analysis allows us to determine how the location on the retina the subject used to examine the stimulus (the Preferred Retinal Locus of fixation, or PRL) differed from the peak cone density (PCD) on the retinal lattice. Figure 4.2 shows an example of a master retinal image from one subject with selected structural and functional measures overlaid onto it.

4.3.4   Eye Movements, ISOA and PRL Analysis

The Isoline Area (ISOA) method was used to measure fixation stability as it was proposed as a better alternative to Bivariate Contour Ellipse Area in the presence of multiple loci of fixation positions (Castet and Crossland, 2012). It does not make any assumption on the nature of the random variables underlying the distribution of data points, which is specifically appropriate for people with eccentric fixation (Whittaker et al., 1988), but also normal subjects whose fixational eye positions have been shown to not be randomly distributed (Cherici et al., 2012). The ISOA and PRL are computed through Kernel Density Estimation (KDE) of the 2D probability density function (PDF) of eye positions. The ISOA is the area within the non-uniform contour that encompasses 68% of the entire eye trace. The PRL is computed as the corresponding peak of the 2D PDF. In other words, the isoline contour encloses all the eye positions that lie within 1 SD from the PRL, if we could assume normality and a unique PRL (which we observed). In order to assert both non-normality of the 2D distribution of eye positions and the non-separability between pairwise distributions obtained during different conditions, one-way and two-way 2D Kolmogorov-Smirnov tests were used, respectively. This implementation relies on Fasano & Franceschini's generalization (Fasano and Franceschini, 1987) for two dimensions.
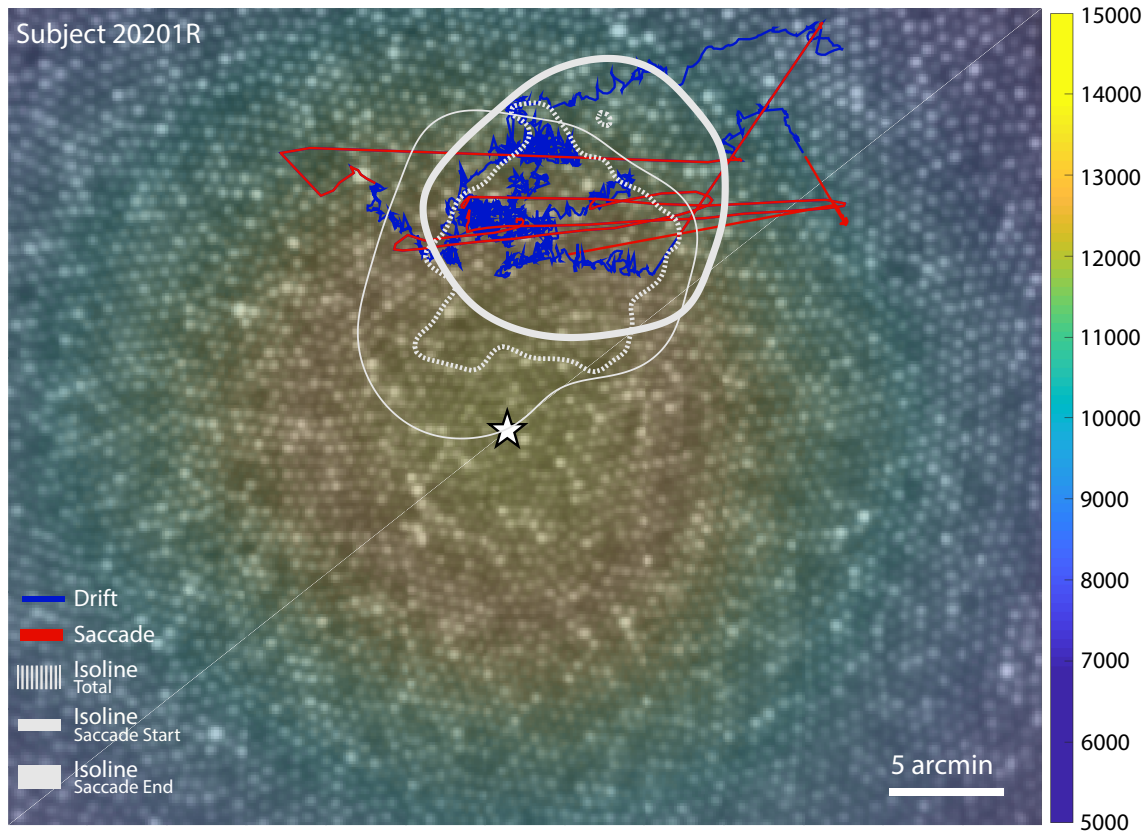
Figure 4.2: Master retinal image for subject 20201R with functional and structural measures overlaid. The star indicates the point of maximum cone density (PCD) and the underlying colormap represents the cone density in cones per degree^2. 5 seconds of eye movement is plotted on the retinal image representing the stimulus motion on the retina with saccades (red) and drifts (blue) highlighted. The isoline contours for all the eye positions obtained during the Vernier condition (dotted), as well as saccade start (thin) and end points (thick) for this condition are shown in gray (see Figure 4.4 and 4.5 for contours for all conditions and all subjects).

4.4 Results

4.5 Global eye movement statistics

Selected FEM measurements for all subjects and all conditions are plotted in Figure 4.3. The figure reveals expected extensive differences in FEM between subjects (Cherici et al., 2012), which will be discussed later. Comparing between conditions, we found that subjects had a lower saccade rate (repeated measures ANOVA $F(4, 28) = 9.91$, $p < 0.001$, Post-Hoc Tukey $p < 0.05$ in all passive vs active comparisons) and a higher saccade amplitude

(repeated measures ANOVA $F(4, 28) = 11.4$ with Greenhouse-Geisser correction, $p < 0.001$, Post-Hoc Tukey $p < 0.05$ in all passive vs active comparisons) during the two active tasks compared to the two passive tasks. Correspondingly, the drift amplitude (repeated-measures ANOVA with Greenhouse-Geisser correction, $F(4, 28) = 26.46$, $p < 0.001$, Post-Hoc Tukey $p < 0.01$ in all passive vs active comparisons) as well as the drift duration (repeated-measures ANOVA with Greenhouse-Geisser correction, $F(4, 28) = 21.67$, $p < 0.001$, Post-Hoc Tukey $p < 0.05$ in all passive vs active comparisons) were smaller in the passive tasks compared to the active tasks. Although the saccade rate was lower in the two active tasks, the overall area encompassed by the FEM measured by ISOA was larger in the active tasks compared to the passive tasks (repeated-measures ANOVA with, $F(4, 28) = 13.57$, $p < 0.001$, Post-Hoc Tukey $p < 0.001$ in all passive vs active comparisons).

As can be seen in Figure 4.3, the FEM between the two passive tasks (Maltese cross and disk) were statistically similar as well as between the two active tasks (Vernier and tumbling E) (post-hoc Tukey, $p > 0.05$ in all cases, with a single exception in Drift Duration between the two active tasks). So, to highlight the comparisons between passive and active fixation tasks, the tasks in each category were combined. The concentric circles condition did not show consistently significant differences from either task and so it was ignored for this analysis. Once combined, the difference between active and passive tasks were clear for all FEM metrics (paired-samples t-test, $p < 0.001$). Figure 4.4 plots the combined and averaged data. The differences between tasks are more readily evident when the active and passive tasks are pooled together.
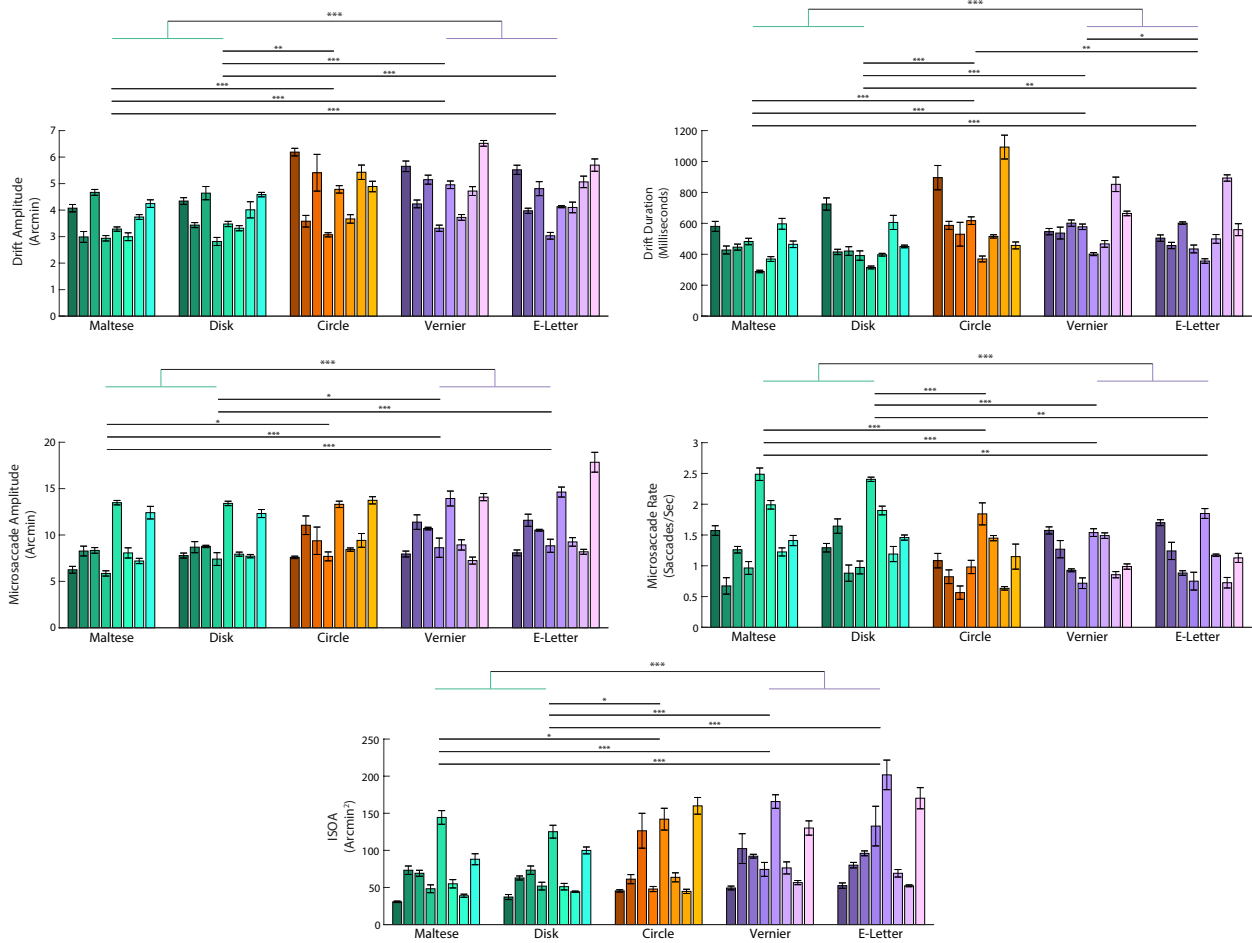
Figure 4.3: Individual subject performance across the 5 conditions. The passive tasks are represented by the green bars and the active tasks are represented by the purple bars. The mixed task (circles) is represented by the orange bars. Differently shaded bars indicate the mean and S.E.M. of each subject's performance across their 5 respective trials for each condition. The horizontal lines above the bar plot with asterisks represent levels of significance ($p<0.05$, $p<0.01$, and $p<0.001$ respectively) from post-hoc Tukey test ran on a repeated-measures ANOVA across the five conditions for each variable. The uppermost horizontal bar with green and purple stems represents the same levels of significance from a paired-samples t-test but compares the two active tasks pooled together vs the two passive tasks pooled together.
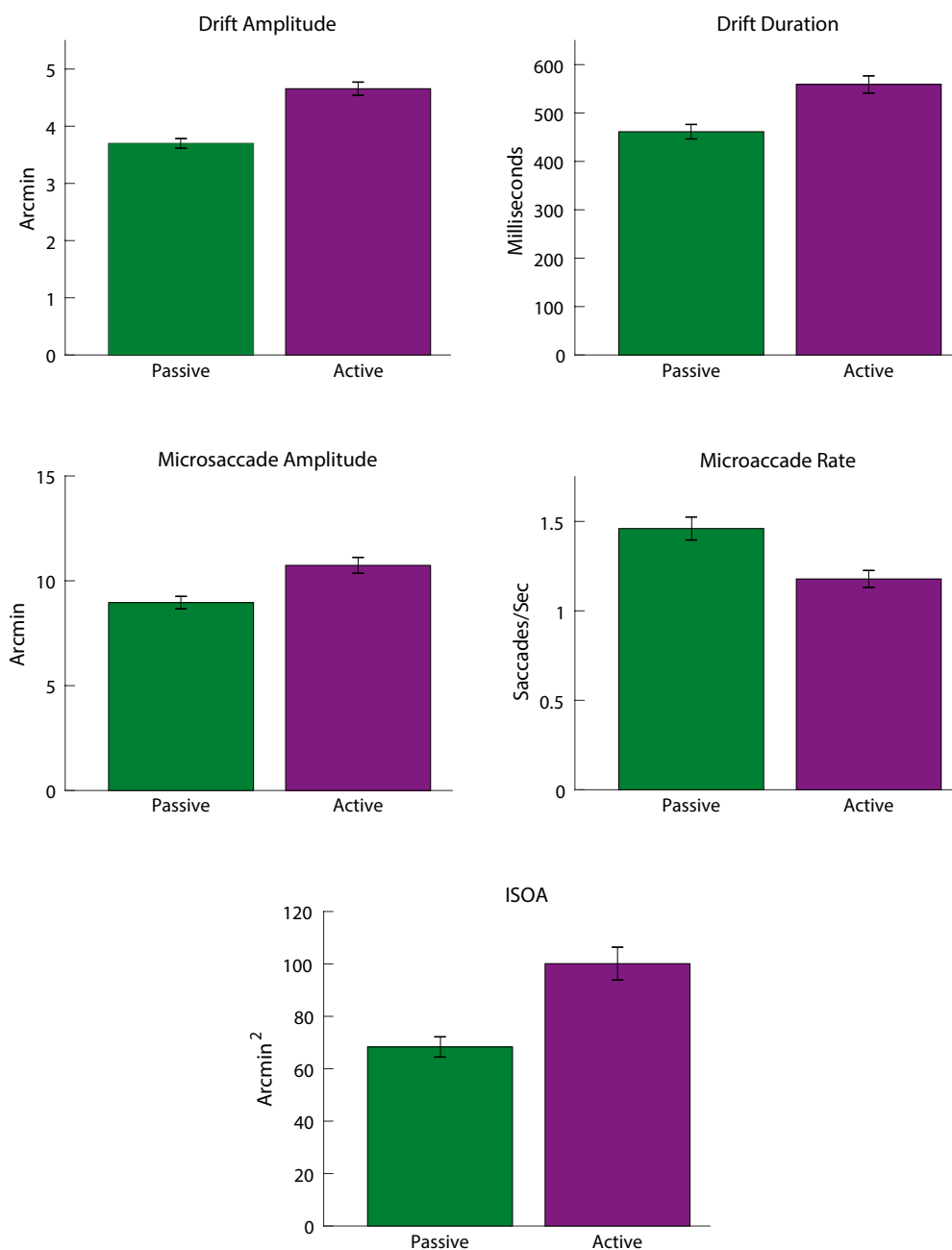
Figure 4.4: Comparison of FEM with the passive and active tasks pooled together. Error bars represent s.e.m. across all subjects for the combined conditions. Overall there were fewer, but larger microsaccades in the active tasks compared to the passive. There were also larger and longer drifts in the active tasks. This led to an overall higher fixation area in active tasks as measured by the ISOA.

Figure 4.4 shows the 68% isoline contours for each subject for each condition centered on the PRL, which is defined at the peak of the fixation positions' PDF. The stimulus for

each condition is drawn in the center of each graph for reference, but the stimulus will sweep across the retina based on the extent of the eye movement. The overall fixation area was larger for the active tasks compared to the passive tasks (paired-samples t-test, p<0.001). Extensive intersubject variability is readily apparent in Figure 4.3. The average standard deviation of the ISOA between subjects for each condition (columns in Figure 4.4) was 43.53 arcmin^2, whereas the average standard deviation between conditions for each subject (rows in Figure 4.4) was 21.26 arcmin^2. Although intersubject variability is extensive (over double the size of the difference between conditions), there is still a significant difference in fixation behavior between the five conditions. Large differences in fixation behavior between subjects is expected in measurements of fixational eye movement, especially when psychophysical expertise is taken into account (Cherici et al., 2012).

Figure 4.5 shows the same data as shown in Figure 4.4 but in this case the ISOAs from all subjects and conditions are overlaid on a single plot with all subjects' respective PCDs at (0,0). This figure reveals several phenomena. First, the PRL rarely coincides with the PCD and is displaced, on average by 5.20 arcmin (SD = 2.54 across tasks and between subjects, SD = 0.23 across subjects and between tasks). This is largely in line with other reports that the PRL does not perfectly correspond with the PCD (Putnam et al., 2005; Li et al., 2010; Wilk et al., 2017; Wang et al., 2019). Second, the PRL tends to be displaced above the PCD in fundus coordinates. This is consistent with recently published reports (Reiniger et al., 2021). Finally, subjects adopt a consistent PRL regardless of the task and its visual demand. The Euclidean distance between the PRL and the peak cone density did not significantly differ from one another (repeated-measures ANOVA with Greenhouse-Geisser correction ($F_{(2.155, 15.087)} = 0.313$, p = 0.751). A Kolmogorov-Smirnov test for two-dimensional distributions was used to determine whether eye position distributions, and therefore the PRL, differed between the different conditions. There was no difference in the PRL location between conditions for any subject (p-value <0.001 in each of the 5 conditions and 8 subjects). This is in agreement and extends on the finding that the PRL for a static Maltese cross target remains stable between days (Kilpeläinen et al., 2020).

Figure 4.5 shows two analyses of microsaccades. The left matrix of plots are 68% isoline contours for saccade start and end points, represented by thin and thick contours respectively. The data for each subject are further distilled into the rose plots on the right which aggregate the microsaccade data from all conditions. Some clear and distinct patterns emerge here. First, the ISOA for saccade start and end points cover a larger area than the conventional ISOAs of Figure 4.4 which encompass both saccade and drift periods. Second, the distribution of microsaccades tends to be more frequent and of larger amplitude in the horizontal direction, which is largely in agreement with other research (Cherici et al., 2012; Sheehy et al., 2020; Thaler et al., 2013). Finally, despite the more extensive horizontal spread, every subject shows a tendency to make saccades, on average, with an upward component (binomial test, looking for a proportion of 50%, p < 0.001 for all subjects). The saccades move the image upward on the fundus, straddling either side of the PRL. If the image moves up during a saccade, this means that the fovea moves down relative to it. In gaze coordinates this corresponds to a saccade that redirects the gaze upward as the coordinates between

fundus view and gaze coordinates are inverted. This movement could be classified as a form of spontaneous upbeat micro-nystagmus (Eggers et al., 2019) although in these instances, the upbeat nystagmus clearly does not indicate a pathological condition. A similar behavior is reported in other papers (Mestre et al., 2021; Stevenson et al., 2016), but not observed universally.
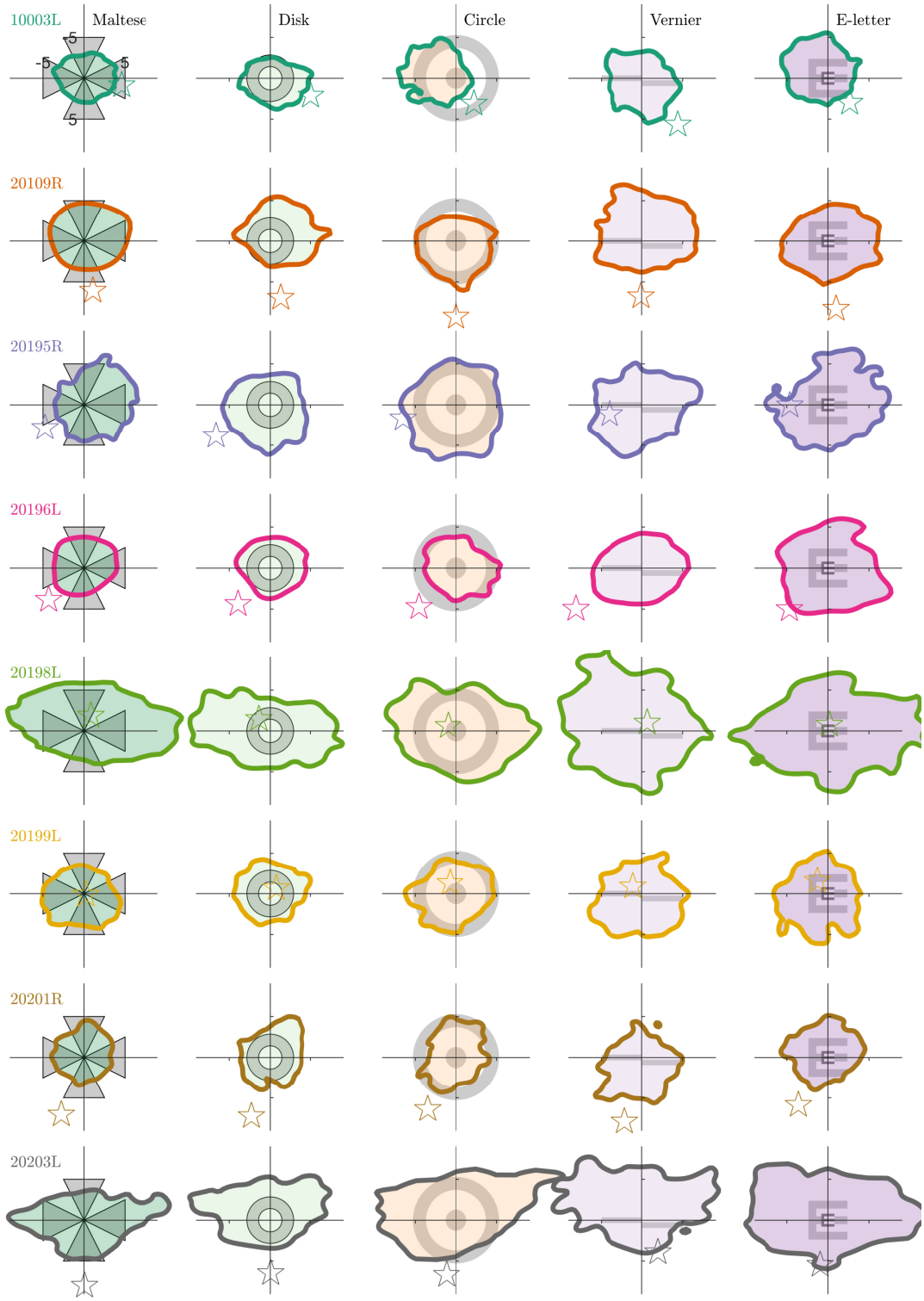
Figure 4.4: 68% isoline contours (ISOA) of the entire eye movement trace. Each row is a subject and each column is a fixation condition, the stimulus for which is drawn centered on each plot for reference. Position (0,0) on the plot corresponds to the PRL, or the peak of the fixation position distribution. Note that idiosyncrasies in the eye movement can cause this peak to appear displaced from the center of the isoline contour. The position of these distributions represent the location of the fixated image and are plotted in fundus-view coordinates (same as Figure 4.2). The star in each plot indicates the relative location of the PCD. Axis units in the upper left plot are in minutes of arc.
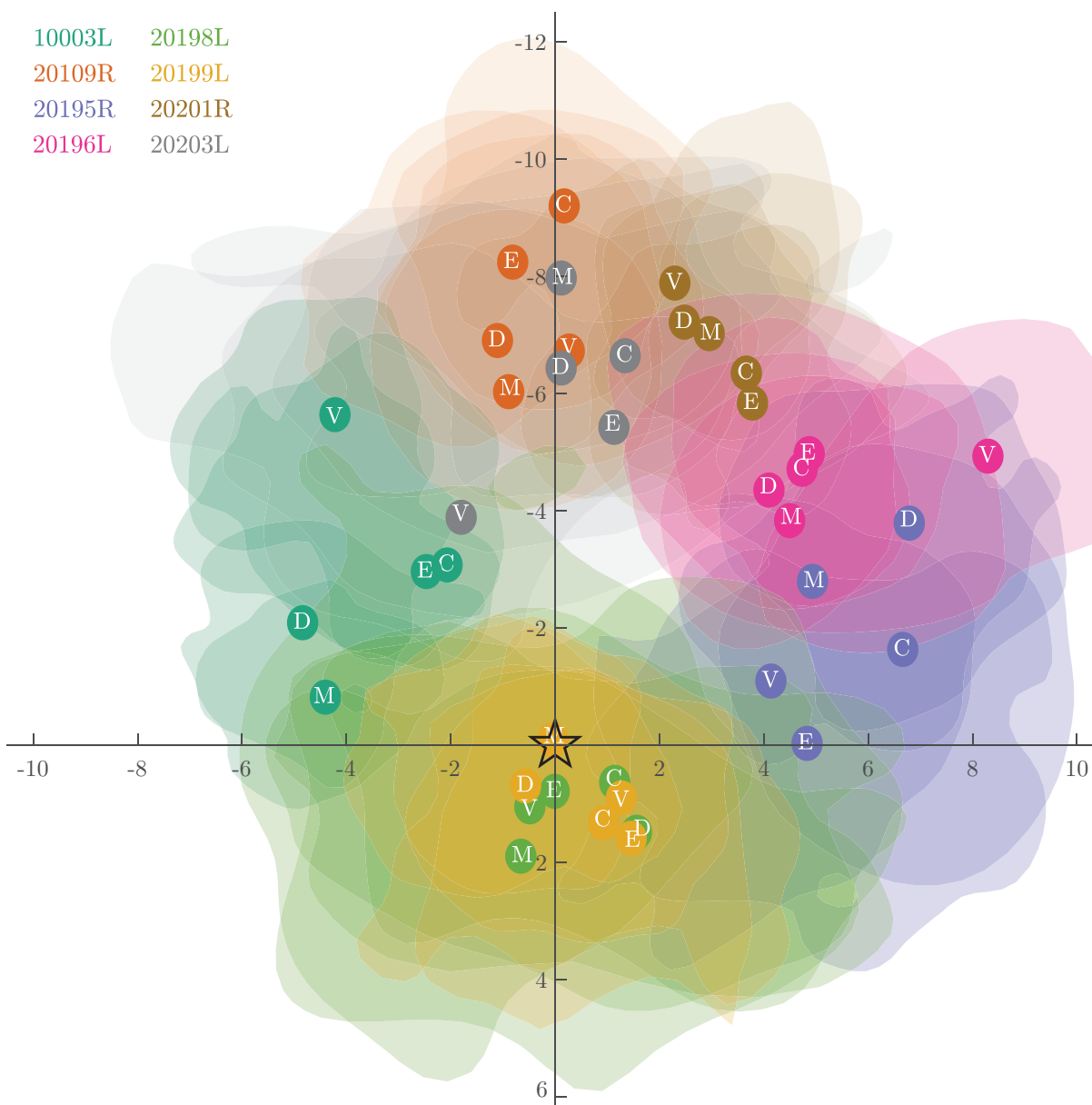
Figure 4.5: Individuals' PRL locations plotted relative to their PCD centered to (0,0). Location of the PRL's are shown as opaque circles overlaid onto the transparent isoline contours with the condition defined in white text within the center. This figure highlights that each subject's PRL tends to fall off their respective PCDs. However, the various PRLs defined in each condition all group together closely. To enhance visibility, the isoline contours here encompass only 38% (0.5 S.D.) of the fixation trace instead of the 68% (1 S.D.). Axis units are in minutes of arc.

63

Figure 4.5: 68% isoline contours for the location of the fixated image at the start and end of saccades, drawn as thin and thick contours respectively. Each row is a subject and each column is a fixation condition, the stimulus for which is also drawn on the figure. Position (0,0) on the plot corresponds to the PRL location from Figure 4.4. The position of these distributions represent the location of the fixated image and are plotted in fundus-view coordinates (same as Figure 4.2). The star indicates the location of the P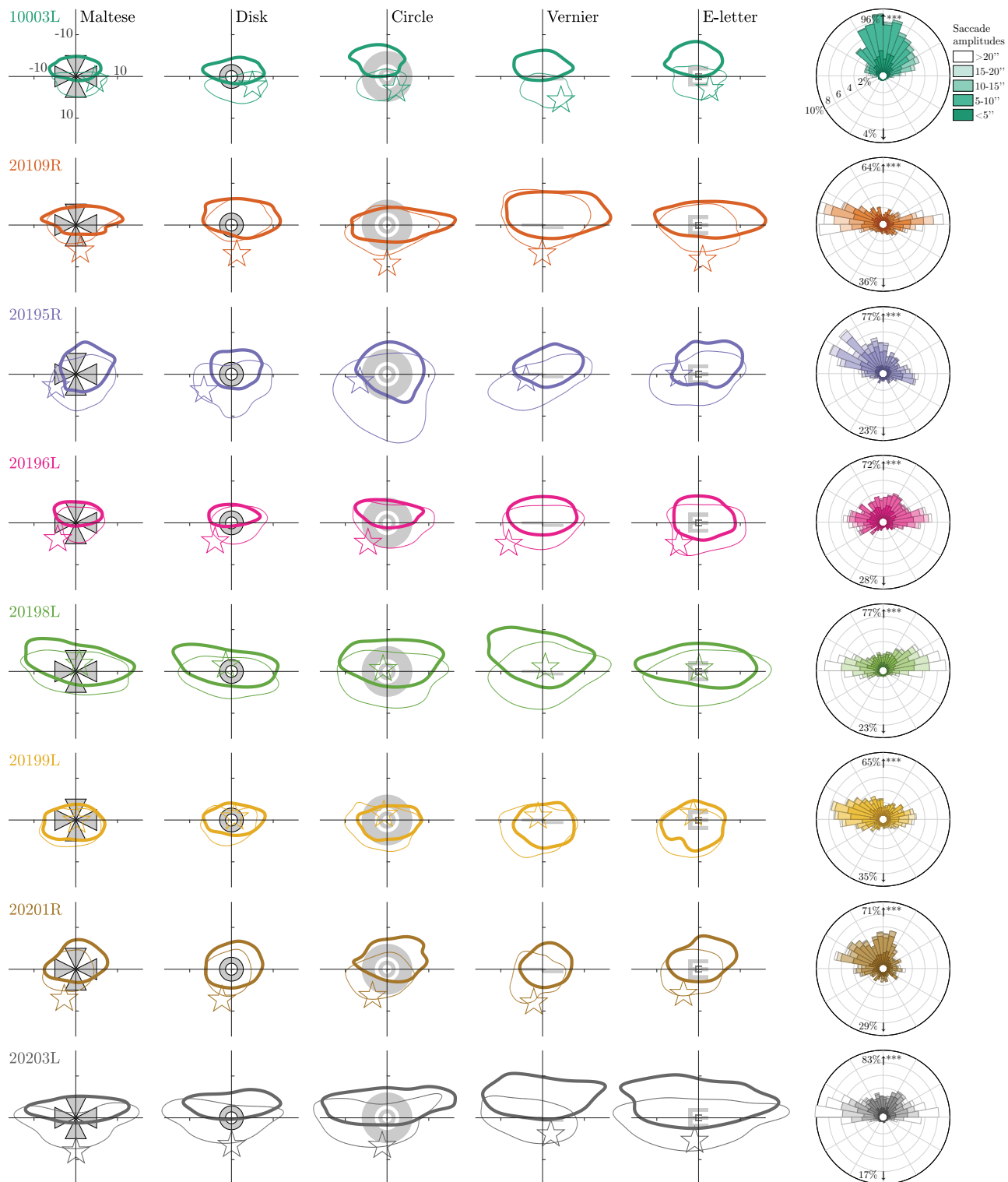CD. Axis units in the upper left plot are in minutes of arc. The right column contains rose plots indicating the percentage of microsaccades as a function of direction for each subject. Each petal in the plot is further broken down by saccade length. The percentage values written in the upper and lower fields indicate the total proportions of microsaccades with an upward vs a downward component, respectively. All subjects have a significant tendency for an upward component (see text).

4.6   Discussion

Our study shows large and significant variations of microsaccade and drift kinematics between subjects and between different tasks. We have confirmed that an individual's FEM behavior depends on the task involved. Specifically, we have confirmed that active tasks result in less frequent microsaccades (Bridgeman and Palca, 1980; Martinez-Conde et al., 2004) giving rise to correspondingly longer and larger drift epochs which, in turn, cause the microsaccades to be larger. This pattern of behavior leads to overall larger ISOAs during active tasks. Passive tasks, by comparison, are marked by shorter and more frequent microsaccades and smaller, briefer drift epochs, all leading to a smaller ISOA. The increase in ISOA from passive to active tasks was 57% (+/- 23% S.D.) on average. These results are consistent with the notion that FEM constitute a behavior that is subconsciously mediated to serve different functions depending on the task at hand. Indeed, evidence of reduced microsaccade rate and a seemingly optimal control of drift eye movements was recently reported for subjects asked to read a line on a Snellen acuity chart (Intoy and Rucci, 2020). The exact extent that FEM may be modulated to enhance vision is still unclear and a matter of ongoing experimentation (Rucci et al., 2018; Kagan et al., 2008b). During the active tasks subjects tended to suppress their microsaccades because the rapid transients from these movements can be detrimental to fine-scale discrimination due to microsaccadic suppression (Mostofi et al., 2021; Bridgeman and Palca, 1980), either from blurring of the retinal image or central suppression.

The consistency of the PRL location between tasks was tested using a 2-dimensional Kolmogorov-Smirnov test and was found to remain the same regardless of the fixation target. This extends on a recent report that the PRL for a Maltese cross target does not change between hours and across days (Kilpeläinen et al., 2020). However, the possible shifts in the location of the PRL have not been investigated for binocular viewing conditions or for more complex viewing experiences, such as during smooth pursuit or fixation within extended

scenes.

Although the peak cone density on the retina offers the best location for photoreceptor spatial sampling (according to the Nyquist sampling limit), the PRL rarely aligns with it exactly. Previous research has consistently shown the same (Putnam et al., 2005; Wang et al., 2019; Wilk et al., 2017). We found the average separation between the PCD and the PRL to be 5.20 arcmin. In the majority of cases the PRL is positioned superior to the PCD (in fundus-view coordinates). This reflects a similar tendency for the PRL in individuals with a central scotoma to adopt an eccentric PRL in the superior retina (Messias et al., 2007; Verdina et al., 2017). This means that the retinal location with peak cone density is sampling a part of the visual field just above the direction of gaze. This tendency has been reported previously by another group (Reiniger et al., 2021) who also used an AOSLO. These displacements are very small and in our opinion, as discussed in one of our previous papers (Wang et al., 2019), seem unlikely to have any functional importance.

A subclinical form of upbeat nystagmus was present to varying extents in all of our subjects (see Figure 4.5 right column). Similar behavior was reported for some, but not all of the subjects in two other studies (Stevenson et al., 2016; Mestre et al., 2021) and is further evident in a slight upward tendency (although not commented on) in the saccade distribution plots of other papers (Cherici et al., 2012; Thaler et al., 2013). Interestingly, these saccades had the tendency to direct the gaze above the PRL, with the following drift generally bringing the gaze downward toward the PRL. This pattern was present in all subjects and suggests that, when classifying the PRL, it is important to consider the complimentary relationship between drifts and saccades. More work is necessary to assess the PRL overshoot behavior and complementary behavior of the following drift segment. In any case, the minutiae of FEM reveals that a PRL that is identified by any of the current methods, including the ISOA approach used here, may be ill-defined. This is a topic of an ongoing investigation.

While we measured significant and informative differences in FEM between conditions, we found that differences in FEM between individuals are even greater. The standard deviation of the ISOA between individuals, for example, was roughly twice that between conditions. These differences can be partly explained by experience (Cherici et al., 2012). All of our subjects were recruited from within the UC Berkeley School of Optometry community and therefore had some experience sitting for visual psychophysics experiments and/or for clinical examinations. But subjects 10003, 20109 and 20196, who all had ISOAs that were lower than the mean, have logged dozens of hours in psychophysics experiments related to eye tracking, including AOSLO psychophysics experiments.

When considering fixational eye movements it is prudent to consider the goals for maintaining a subject's steady fixation, since all fixations are not equal. If the goal of the fixation target is to minimize the overall movement of the eyes (as is the case in many clinical situations), then one must consider which types of FEM are most likely to be an impediment. If the rapid transients from saccades are most likely to have a deleterious effect then it is preferred to rely on an active task so the subject will suppress their microsaccades in order to perform the task. If the goal is to minimize the total area covered by the fixation, then choosing a more passive fixation task is likely to be most effective. Of course, given the effect

of intersubject variability seen in these data, as well as other studies, it is also prudent to keep in mind that subject instructions and recruitment play a large role in the stability of fixation as well.

4.7 Conclusion

This study examined the influence of different fixation targets and tasks on FEM and the location of the PRL in healthy eyes. Using an AOSLO, we developed a new method to locate and follow the target projected on the retina over time relative to the PCD. We confirmed the non-normality of the eye motion distribution, hence the necessity to rely on better descriptors of fixation stability indices such as ISOA and its accuracy to estimate each individual's PRL. The different fixation tasks consisted of active tasks, which had temporal variation and required subject responses, and passive tasks, where the subjects were instructed to simply hold their gaze on the target. The active tasks elicited larger but fewer microsaccades. Consequently, the amplitude and duration of intersaccadic drifts were significantly larger. Larger and longer drifts combined with larger microsaccades led to larger overall fixation instability, as quantified by the ISOA. Our result suggests that subjects suppress their microsaccades during active tasks, and the subsequent longer drift epochs would cause the object to move away from the PRL, thereby requiring a relatively larger microsaccade to reorient. Finally, although the FEM were significantly modulated by the task, the intersubject variability was expectantly substantial. The two-to-four times larger effect on fixation stability across individuals compared to task suggest that experimenters might, when aiming to better control the user's eye position, put a greater emphasis on instructions, training, and subject recruitment rather than on the fixation stimulus itself.

4.8 Acknowledgments

4.9 Supplemental

All supplementary materials can be found under the original article in the Journal of Vision (`https://doi.org/10.1167/jov.21.11.16`).

4.9.1 Example Video

The supplementary movie shows on the left how the AOSLO raster appears to the subject for the concentric circles task. The right side shown the simultaneously recorded video of the retina of subject 20201R. 3 seconds out of a full 36-second video is shown. The movement of the retina arising from fixational eye motion and its amplitude relative to the size of the stimulus is readily evident. This segment also highlights how the decrement stimulus is inscribed into the video file directly, providing a completely unambiguous record of the stimulus' position across the photoreceptor mosaic. Note that this segment has been compressed to keep the file size small and does not represent the raw videos used in stabilization and analysis.

4.9.2 Instructions

The instructions given to each subject were identical from one subject to the next. The below script was read out to each subject for each condition and any clarifying questions were answered before the experiment began. Note that experimenters asked the subject to blink when their tear film degraded to the point of interfering with the video quality.

Maltese: "In this experiment you will be required to look towards the center of the cross for the entire duration of the 36 second video. There will be a sound played when the video starts and stops recording. You are allowed to blink as needed and I will ask you to blink if necessary. Press the start button when you are ready to go."

Disc: "In this experiment you will be required to look towards the center of the disc for the entire duration of the 36 second video. There will be a sound played when the video starts and stops recording. You are allowed to blink as needed and I will ask you to blink if necessary. Press the start button when you are ready to go."

Concentric circles: "In this experiment you will be required to look towards the center of the target for the entire duration of the 36 second video. You will see a series of rings that become smaller. There will be a sound played when the video starts and stops recording. You are allowed to blink as needed and I will ask you to blink if necessary. Press the start button when you are ready to go."

Vernier: "In this experiment you will be required to look at the two lines for the entire duration of the 36 second video. The two lines will vary in position at random intervals which will be announced by an audible cue. Using the up and down buttons, report if the right line is higher or lower than the left line. There will be a sound played when the video starts and stops recording. You are allowed to blink as needed and I will ask you to blink if necessary. Press the start button when you are ready to go."

Tumbling E: "In this experiment you will be required to look at the "E" for the entire duration of the 36 second video. The letter "E" will change size and direction at random intervals which will be announced by an audible cue. Using the buttons, report the direction

the "E" is facing: up, down, left, or right. There will be a sound played when the video starts and stops recording. You are allowed to blink as needed and I will ask you to blink if necessary. Press the start button when you are ready to go."

4.9.3 Extra tables

The supplementary .csv files contain the raw data for each of the 5 eye motion parameters shown in Figures 4.3 and 4.4. Rows indicate different subjects and columns indicate different conditions. The variable names and related units are shown in the top, left of each table.

Chapter 5

Accuracy of the PRL Across Vergence Demand as a Measure of Objective Fixation Disparity

5.0.1 Abstract

The preferred retinal locus of fixation (PRL) is the region of the retina we direct stimuli toward when fixating. In terms of binocular vision, the PRL can be thought of as the corresponding point that represents the gaze direction at the distance of vergence (i.e. the fixation point). However, it is commonly reported that the exact center of fixation changes with different vergence demand, a phenomenon known as fixation disparity. These offsets in fixation can derive either from placing the center of the gaze in front of (eso fixation disparity) or behind (exo fixation disparity) the target in depth, leading to an over- or under-convergence respectively. This would imply that the PRL will shift with increasing vergence demand. However, many of these studies rely on an assumption about the exact placement of a stimuli's position on the retinal lattice from eye tracking measurements taken from the anterior segment of the eye (often through the use of video eye trackers). The current study aims to measure the stability of the PRL for both eyes simultaneously under different vergence demands, and consequently the objective fixation disparity, on the retinal image directly through the use of a binocular tracking scanning laser ophthalmoscope.

5.0.2 Introduction

Studies have often found that subjects may not properly fixate a target when the vergence demand of the target is increased, a phenomenon known as fixation disparity. Subjects tend to fixate either in front of (eso) or behind (exo) the target. An example of eso and exo fixation disparity, as well as it's effects on the retinal projection in fundus view, can be seen in Figure 5.1. Many studies looking to examine fixation disparity rely on subjective methods, such as measuring the alignment of dichoptic nonious lines as subjects maintain fixation on

a point in depth (Fogt and Jones, 1998; Duwaer and Van Den Brink, 1981; Jaschinski et al., 2005). However, objective measures of gaze direction taken during this task (through the use of eye tracking) tend to find fixation disparity is much greater than the subjective measures would suggest. The subjective alignment of the nonious lines would often suggest fixation disparities of 6-20arcsec, while the objective measures report a larger offset, sometimes as large as 1° (Fogt and Jone, 1998; Fogt and Jones, 1998; Duwaer and Van Den Brink, 1981; Jaschinski et al., 2005, 2010). The difference in fixation disparity measured objectively versus subjectively is often attributed to a shifting of the mapping between neurons in the brain and receptive fields of cells in the retina used for stereopsis. However, a study by Hillis and Banks found no evidence of shifting of retinal corresponding points with increasing vergence demand (Hillis and Banks, 2001). The implications of this shifting of receptive fields would suggest that the center of gaze will not always be consistent with the anatomical fovea, and instead subjects could "fixate" an object using their peripheral retina, as far as 1 degree away from the foveal center when their vergence demand changes.
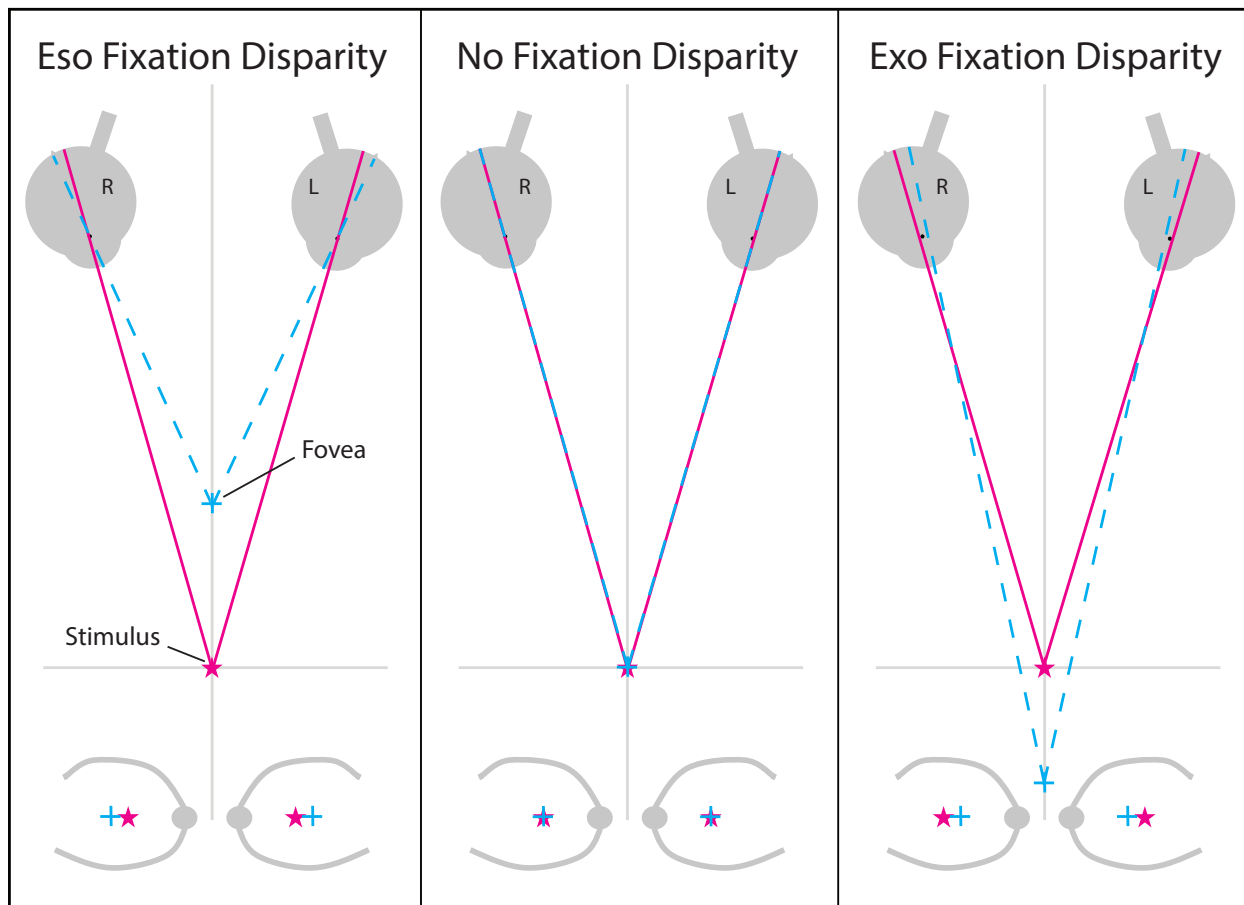
Figure 5.1: An example of eso and exo fixation disparity. The magenta star indicates the stimulus position while the cyan cross indicates where the foveas (or monocular PRLs) are pointing. The bottom plots indicate schematic retinal images in fundus view coordinates with the projection of the stimulus and the foveal locations highlighted in each case. During eso fixation disparity (left), the eyes over converge. This causes the stimulus to be shifted nasally (to the left in the left eye and to the right in the right eye). During exo fixation disparity (right), the eyes under converge. This causes the stimulus to be shifted temporally (to the left in the right eye and to the right in the left eye).

Recent research has been conducted examining how accurately and reliably subjects actually place the image of a fixated object onto the same region of their retina. When fixating on an object subjects do not place a stimulus onto their exact foveal center, deviating by an average of ~5 arcmin away from the peak cone density of the retina (Reiniger et al., 2021; Wang et al., 2019; Wilk et al., 2017; Bowers et al., 2019). This region of the retina is known as the preferred retinal locus of fixation (PRL). The PRL has been found to be remarkably stable within a subject. It does not change if the subjects are given different visual tasks (Bowers et al., 2019) nor across different days (Kilpeläinen et al., 2020; Reiniger et al., 2021). However, it is unknown if the PRL will change with increasing vergence

demand. The reports of over- and under- convergence found by researchers looking into binocular fixation stability would suggest that the PRL will shift with vergence demand. The majority of measurements of objective fixation disparity, however, are taken using video eye trackers, which may lack the requisite spatial resolution to accurately track the gaze on a small scale (Kimmel et al., 2012; Holmqvist and Blignaut, 2020; Niehorster et al., 2021). These video eyetrackers are also known to suffer from artifacts caused by changes in pupil size (Choe et al., 2016; Nyström et al., 2016; Hooge et al., 2016). Furthermore, no previous study has been able to tie a stimuli's position under different vergence demands to the retinal image directly. The current study aims to examine the stability of the PRL for both eyes simultaneously under different vergence demands by imaging the fixated stimulus on the retinal surface directly.

5.0.3   Methods

5.0.3.1   Experiment Design

Data were collected through the use of a binocular tracking scanning laser ophthalmoscope (TSLO). The binocular TSLO consists of two separate TSLO systems that are temporally synced. The TSLO system has the capability to image and track the retina, as well as display stimuli onto the retina directly. For imaging, an infrared point source (840nm) is swept across the retina pixel by pixel in a raster pattern. The system utilizes two galvonometer scanners, a 16Khz fast horizontal scan and a 30Hz slow vertical scan. In this instance, the size of the raster was 6° with a 512x512 pixel sampling resolution, giving a pixel resolution of ∼42". Light reflected from the retina was descanned through the optical path and images of the retina were acquired through the use of a photomultiplier tube. The images obtained by the binocular TSLO were compiled together to create two temporally synced movies of the retina at a framerate of 30Hz (the frequency of the slow vertical scanner). For more detail on the TSLO system see Sheehy et al. (2012). Eye traces from the system were computed offline using a strip-based cross correlation method. This method has been described in detail in previous studies (Bowers et al., 2019; Stevenson et al., 2010) (also see Chapter 1 and 2). The size of the strips were set at 32 pixels (16 samples per frame), giving a temporal sampling resolution of 480Hz.

Five subjects participated in the experiment. Subjects were only involved if they had minimal refractive error (<1 diopter) and suffered no visual or neurological disorders. All subjects gave informed consent and all experimental protocols adhered to the conditions set by the institutional review board of the University of California, Berkeley and followed the tenets of the Declaration of Helsinki. All subjects were right-eye dominant. Subjects were given a very simple fixation task to keep them engaged. The stimuli were either an X or a cross shape within a circle and the stimuli would randomly switch every 2 to 4 seconds. Stimuli appeared to the subjects as black-on-red decrements. Subjects were instructed to fixate the target throughout the trial and indicate when the stimuli switched with a key press. An example of the stimuli in the subjects and the experimenters view can be seen

in Figure 5.2. The stimuli were $\sim$ 15arcmin in size. There were 4 separate experimental conditions: Monocular, 0° vergence, 4° vergence and 8° vergence. The monocular condition consisted of recording movies of a single eye with the fellow eye patched while the system was set to 0° vergence. The 0° vergence consisted of recording simultaneously from both eyes at a parallel gaze direction. The 4° and 8° vergence conditions consisted of recording simultaneously from both eyes with a 4° and 8° vergence demand respectively. The vergence demand was created by rotating the entire left TSLO system clockwise. An example of the rotated binocular TSLO can be seen in Figure 5.3. The vergence of the two rasters was measured using a model eye each time the system was rotated. There were three 15-second movies recorded from each condition. Movies were stabilized using ReVAS (Agaoglu et al., 2018) with a unique reference for each condition.
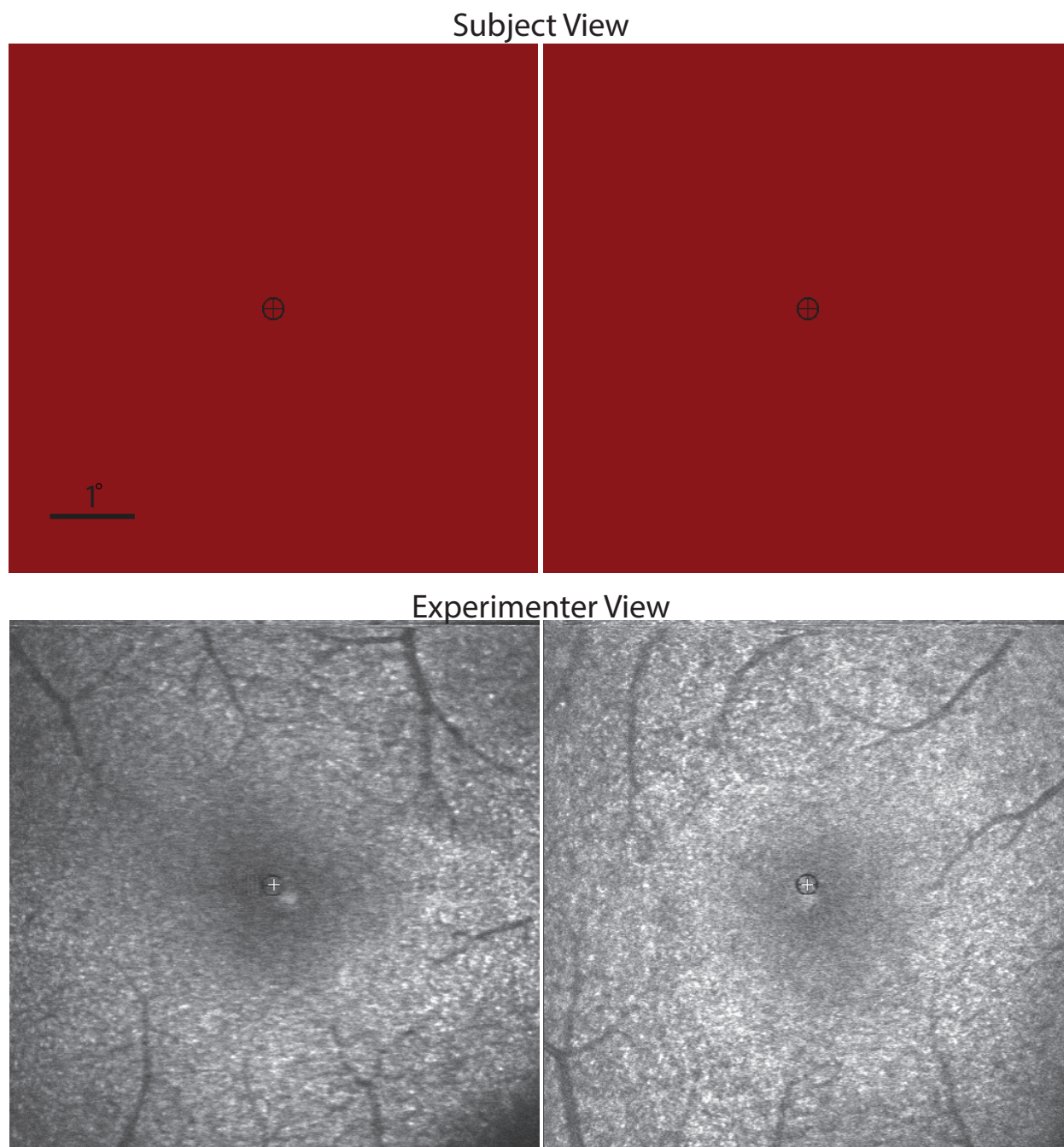
Figure 5.2: A comparison of the subject's view of the stimulus compared to the TSLO videos. Note that during binocular viewing the left and the right views were fused (ie seen as a single display). All images of the retina are in fundus view. Top: How the stimuli appeared to the subject. Stimuli were black decrements on a dim red background. Bottom: Images of the binocular stimulus for subject 20231 in the 0° vergence condition. Left and right images represent the left and right eyes respectively.

5.0.3.2 Eye Movement Alignment & PRL Identification

Eye traces acquired from the binocular TSLO system were processed using custom software. This software is identical to that used in previous studies (Bowers and Poletti, 2017; Bowers et al., 2019) (see Chapters 3 and 4 for more detail). The traces were manually verified by the experimenter to remove any noise or aberrant data using the manual validation software described in Chapter 2. In order to examine any possible movement of the center of fixation with increasing vergence angle, it's necessary to align all traces onto a single master reference image for each eye. Since the traces from each condition were relative to their own respective reference, this can be accomplished by finding the offset needed to align one reference to a master reference. For all subjects the 0° vergence reference frame was chosen as the master reference and the reference frames from each other condition was aligned to this master using cross-correlation. Any instances where the cross-correlation could not be used (often due to changes in retinal reflectance in different conditions), the references were aligned manually. This was only done for a subset of the total alignments (5 out of 30 total alignments). Once the X/Y offsets needed to align one condition's reference to the master were found, each reference frame was rotated and cross correlated against the master reference again in a series of 0.1° degree steps and the highest peak correlation was used to obtain a measure of torsional differences between a reference and the master. These alignments were manually verified by the author by ensuring that the position of the stimuli on the retina from the original movie matched both the reference frame for each condition as well as the master. Once all traces were aligned to a master reference frame, the center of fixation is determined by plotting an Isoline contour for each condition onto the master reference. The Isoline contour is a nonuniform contour that encompasses the central 68% of the data points consistent with the stimuli's path across the retina. The PRL is labeled as the peak of the kernel density function used to generate the Isoline contour. The Isoline contours for each condition on the master reference image can be seen in Figure 5.4. This technique is identical to the one described in (Bowers et al., 2019) (Chapter 4).
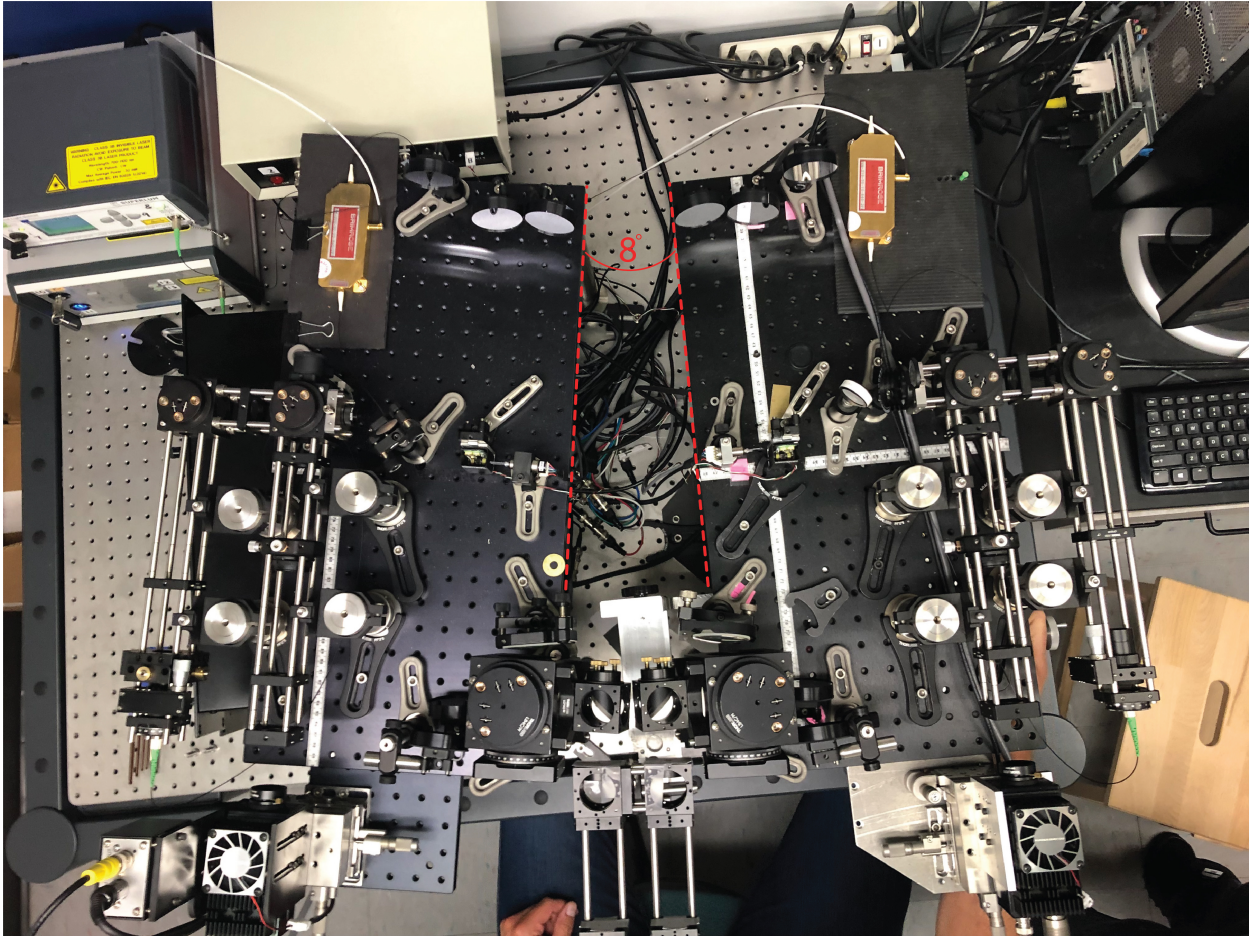
Figure 5.3: Image of the Binocular TSLO image, with the left TSLO rotated to generate an 8° vergence demand. Dashed lines along the edges of the breadboard are present to illustrate the angle of rotation.

5.0.4 Results

The PRL as identified using the peak of the kernel density function varied slightly from one condition to the next. This variance was largely consistent with previous findings (Reiniger et al., 2019; Putnam et al., 2005; Bowers et al., 2021) and likely represents normal fixation instability during the trial. Importantly, if the PRL was changing with vergence demand, we would expect to see the 4° and 8° conditions (red and green contours in Figure 5.4 respectively) shifting in an orthogonal direction along the horizontal axis compared to the 0° condition. For eso convergence we would expect the PRL in the left eye to be shifting to the left and the PRL in the right eye to be shifting toward the right. For exo convergence we would expect to see the opposite results. It is clear from figure 5.4 that there is no consistent shift in the PRL location with increasing vergence demand.
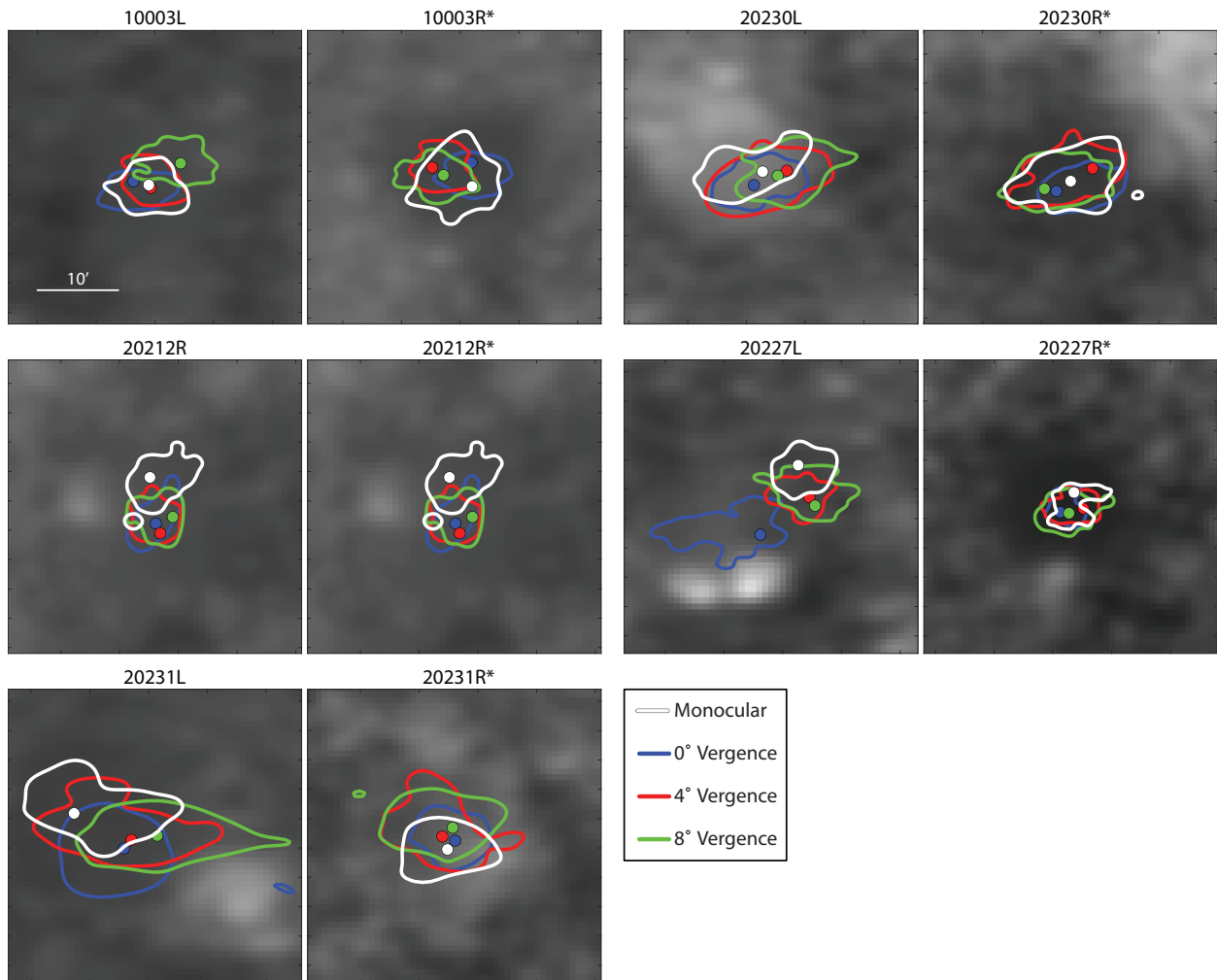
Figure 5.4: 68% Isoline contours for each condition plotted onto the master reference figure for each subject. The Isoline contour represents the central 68% of samples of the stimuli's path on the retina. White represents the Isoline areas from the monocular task while the fellow eye is patched; while blue, red and green are from the 0°, 4°, and 8° vergence task respectively. The central dots represent the PRL locations for each condition identified as the peak of the kernel density function. Because this is a TSLO with a 6° raster, foveal cones are not clearly resolved in these images. Asterisks indicate eye dominance. In this case, all subjects were right-eye dominant.

In order to examine the change in fixation disparity over time, the difference between the left and right eyes traces were plotted in Figure 5.6. First, the difference between the left and right traces were taken for each trial ($Trace_L$ - $Trace_R$). Trials were only included if there was an accompanying trace for both eyes in each trial. In instances where a single movie in one eye was not able to be stabilized, the accompanying trace for the fellow eye was not included in this analysis. This was only the case in 5 out of 45 pairs of

videos (not counting the monocular videos) and there was at least two usable trials for each condition for each subject. Abrupt brief variations in fixation disparity are generally due to the presence of microsaccades. Some subjects, particularly 20231, exhibited large variations in fixation disparity over the course of the trial that did not have the characteristics of microsaccades. Instead, one eye would engage in a fast drift-like behavior and lose fixation before realigning itself. This subject had, on average, the largest fixation disparity and would lose fixation of the target in depth by as much as 80 arcmin. However, this was generally a brief and uncommon occurrence and the subject would quickly reorient themselves. A 5 second example of the binocular eye trace and fixation disparity for this subject can be seen in Figure 5.5. Although all subjects had, on average, near 0 fixation disparity, there was considerable change in fixation disparity throughout the course of the trial for some subjects, as well as significant intersubject variability. The interquartile range for fixation disparity throughout the course of a trial for all subjects was on average 6.09 ± 4.11 arcmin (range: 2.28 - 14.57 arcmin). This means that while subjects may briefly experience, sometimes significant, fixation disparity, they were still fixating the target in depth correctly on average over the course of the entire trial. The two eyes showed the expected strong correlation to one another (for a review on the binocular coordination of fixational eye movements, see Otero-Millan et al. (2014)). Averaged across all subjects in all conditions the correlation coefficient was $R = 0.64 \pm 0.19$ (range: 0.34 - 0.92) between the left and the right traces. Lowered correlation was generally due to drift, which exhibits the occasional orthogonal "wave-like" behavior between the two eyes (Ditchburn and Ginsborg, 1953; Simon et al., 1984).
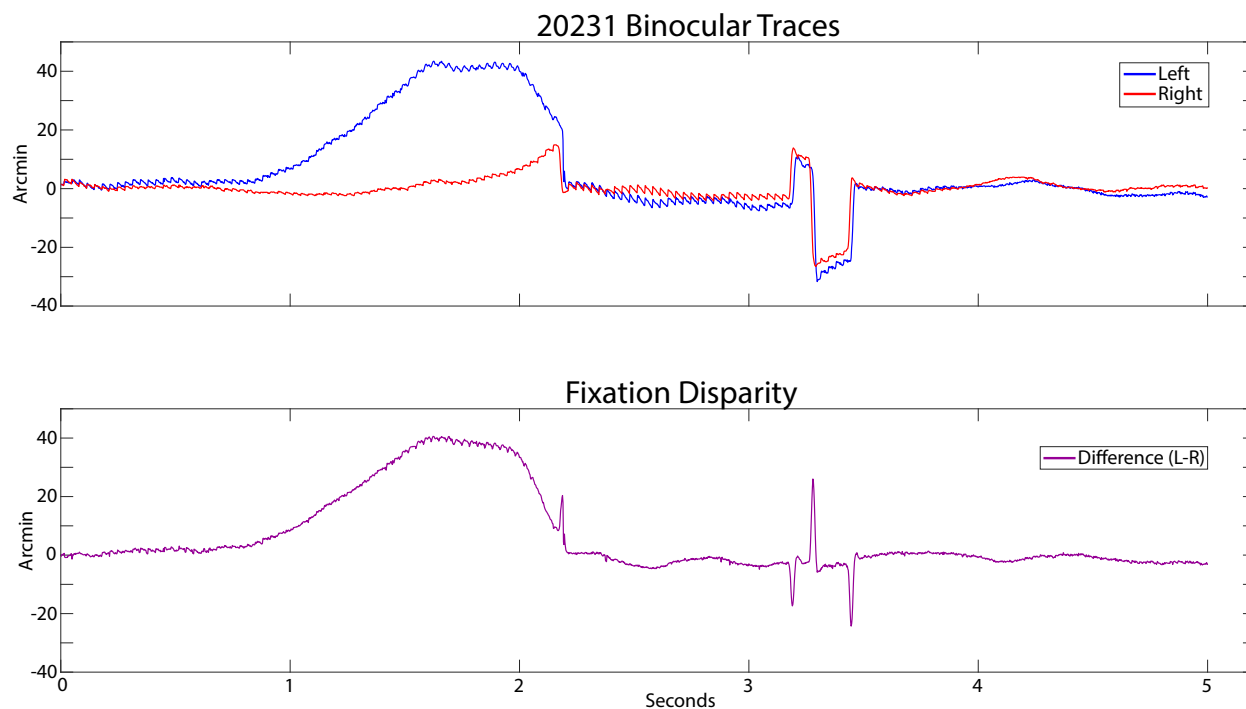
Figure 5.5: Top: Five seconds of eye motion for the left and right eyes for subject 20231. Between seconds 1 and 2 there is a large ($\sim$40arcmin) increase in fixation disparity primarily driven by the motion of the left eye. This motion does not show the characteristics of microsaccades even though it has a large amplitude. Interestingly, only one eye would lose fixation before reorienting itself. Bottom: fixation disparity measured over time. Fixation disparity is simply a measure of the difference between the left and right trace. Besides the large increase between seconds 1 and 2, large spikes in fixation disparity are visible during microsaccades, generally due to a small difference in timing between the left and the right eye. The source of this difference could be biophysical or a technical artifact.

Figure 5.6: Average change in fixation disparity for each vergence condition over the course of the trial. Each plot shows the fixation disparity over time for each subject averaged across their three trials. The histograms on the right represent the probability distributions of fixation disparity f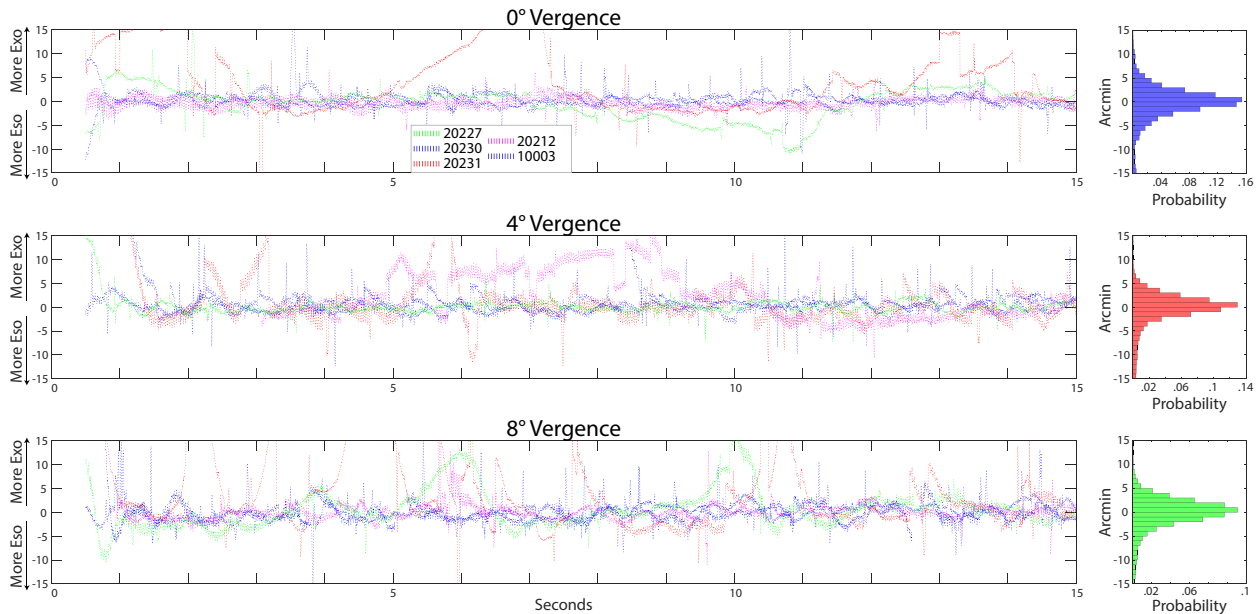or all samples and all subjects. Fixation disparity is simply calculated by taking the difference between the two traces (Trace_L - Trace_R). Data were only included if there were samples present in both traces at each time point, otherwise the data were rendered NaN and not included in the analysis. Positive values indicate an exo fixation disparity while negative values indicate an eso fixation disparity. Note that the first 500ms were dropped for each trial to allow subjects to converge onto the stimuli when they appeared. Units are in minutes of arc. These figures are constrained to $\pm$ 15 minutes of arc for clarity but some subjects would occasionally show brief but significant variations in fixation disparity over the course of the trial.

5.0.5 Discussion

No clear trend of eso or exo vergence was found with increasing vergence demand between $0°$, $4°$, and $8°$. While subjects did show variation in their fixation disparity over the course of the trial, on average this variation was near 0 for all subjects. Table 5.1 shows the mean fixation disparity over the course of a trial for each subject. This runs counter to previous findings on objective fixation disparity that tend to show either an eso or exo vergence of anywhere between 15' to upwards of $1°$. This can be attributed to a few causes. Although previous research has found that the PRL is remarkably stable across days (Reiniger et al., 2021) and tasks Bowers et al. (2019), and the current study shows a similar stability across vergence demand, it is possible that some combination of vergence and task could cause a shift in the PRL. Most studies examining the difference between subjective and objective

fixation disparity have the subjects aligning dichoptic nonious lines while maintaining fixation. It's possible that when performing this task the subjects will systematically fail to fixate as expected in depth, possibly due to the effort of trying to fuse peripheral stimuli while maintaining fixation. It is also possible that these measurements of objective fixation disparity can be attributed to a misalignment in the video eye trackers often used in these experiments. Recent research has shown that video eye trackers have insufficient spatial resolution to accurately and precisely measure minute changes in gaze direction (Kimmel et al., 2012; Holmqvist and Blignaut, 2020; Niehorster et al., 2021). Changes in vergence on the scale generally reported could be due to random noise in the video eye trackers, or worse, a systematic bias that occurs with increasing vergence demand. It is known that changes in pupil size can affect many video eye trackers (Choe et al., 2016; Nyström et al., 2016; Hooge et al., 2016), and there is a mutually intertwined system of vergence, accommodation, and pupil constriction associated with depth perception (Eadie and Carlin, 1995). In fact, Brautaset and Jennings (2006) found measures of objective fixation disparity of only 3.5arcmin when utilizing scleral search coils, which do not suffer from the same problems as video eye trackers.

Average Fixation Disparity			
Subjects	0° Vergence	4° Vergence	8° Vergence
10003	0.55	0.42	0.02
20212	-0.31	2.56	0.05
20227	-0.44	0.07	0.70
20230	-0.03	2.41	0.37
20231	4.47	1.81	4.86

Table 5.1: Mean change in fixation disparity over the course of the 15 second trial for each subject. First, the average fixation disparity for each time point across the three trials was taken for every subject. Samples were only included if there were usable data from each eye at each time point. Values represent the mean of this average across trials collapsed for all time points. Units are in minutes of arc. Positive values represent an exo fixation disparity and negative values represent an eso fixation disparity.

5.0.6   Conclusion

In this study we aimed to examine the stability of the PRL, and by extension the objective fixation disparity, on the retinal image of both eyes directly as subjects performed a simple discrimination task under different vergence demands. Three fifteen-second trials were recorded as subjects performed this task under monocular, 0°, 4°, and 8° vergence conditions. Movies of the retinal image were obtained through the use of a binocular tracking scanning laser ophthalmoscope and high resolution eye traces were extracted from these movies through a strip-based cross correlation technique. The PRL was identified in each condition as the peak of the kernel density function needed to fit an 68% isoline contour

over the 2 dimensional eye trace on the retinal image. There was no clear pattern of eso (over convergence) or exo (under convergence) fixation disparity between the 4 experimental conditions. Although subjects would show change in their fixation disparity over the course of the trial, these changes were often times transient, and most subjects had a mean fixation disparity of 0 over the entire 15 second trial. More research is needed to identify why and when subjects experience objective fixation disparity, and more importantly, whether this will play a role in normal day to day stereovision.

Bibliography

Agaoglu, M. N., Sit, M., Wan, D., and Chung, S. T. L. (2018). Revas: An open-source tool for eye motion extraction from retinal videos obtained with scanning laser ophthalmoscopy. *Investigative Ophthalmology Visual Science*, 59(9):2161.

Ahissar, E. and Arieli, A. (2012). Seeing via miniature eye movements: a dynamic hypothesis for vision. *Frontiers in Computational Neuroscience*, 6:89.

Anderson, A. G., Ratnam, K., Roorda, A., and Olshausen, B. A. (2020). High-acuity vision from retinal image motion. *Journal of Vision*, 20(7):34.

Arathorn, D. W., Yang, Q., Vogel, C. R., Zhang, Y., Tiruveedhula, P., and Roorda, A. (2007). Retinally stabilized cone-targeted stimulus delivery. *Optics Express*, 15:13731–13744.

Babadi, B. and Brown, E. N. (2014). A review of multitaper spectral analysis. *IEEE Transactions on Biomedical Engineering*, 61:1555–1564.

Bedggood, P. and Metha, A. (2012). De-warping of images and improved eye tracking for the scanning laser ophthalmoscope. *PLoS One*, 12:e0174617.

Bhattarai, D., Suheimat, M., Lambert, A. J., and Atchison, D. A. (2019). Fixation stability with bessel beams. *Optometry and Vision Science*, 96(2):95–102.

Bowers, N. R., Boehm, A. E., and Roorda, A. (2019). The effects of fixational tremor on the retinal image. *Journal of Vision*, 19(11):8.

Bowers, N. R., J, G., S, L., and A, R. (2021). Fixational eye movements in passive versus active sustained fixation tasks. *Journal of Vision*, 21(11):16.

Bowers, N. R. and Poletti, M. (2017). Microsaccades during reading. *PLOS ONE*, 12(9):1–14.

Brautaset, R. L. and Jennings, J. A. (2006). Measurements of objective and subjective fixation disparity with and without a central fusion stimulus. *International Medical Journal of Experimental and Clinical Research*, 12:MT1–MT4.

Bridgeman, B. and Palca, J. (1980). The role of microsaccades in high acuity observational tasks. *Vision Research*, 20(9):813–817.

Burak, Y., Rokni, U., Meister, M., and Sompolinsky, H. (2010). Bayesian model of dynamic image stabilization in the visual system. *Proceedings of the National Academy of Sciences*, 107(45):19525–19530.

Castet, E. and Crossland, M. (2012). Quantifying eye stability during a fixation task: a review of definitions and methods. *Seeing and Perceiving*, 25(5):449–469.

Cherici, C., Kuang, X., Poletti, M., and Rucci, M. (2012). Precision of sustained fixation in trained and untrained observers. *Journal of Vision*, 12(6):31.

Choe, K. W., Blake, R., and Lee, S. (2016). Pupil size dynamics during fixation impact the accuracy and precision of video-based gaze estimation. *Vision Research*, 118:48–59.

Collewijn, H., Van Der Mark, F., and Jansen, T. C. (1975). Precise recording of human eye movements. *Vision Research*, 15(3):447–450.

Cornsweet, T. N. and Crane, H. D. (1973). Accurate two-dimensional eye tracker using first and fourth purkinje images. *Journal of the Optical Society of America*, 63:921–928.

Crane, H. D. and Steele, C. M. (1985). Generation-V dual-Purkinje-image eyetracker. *Applied Optics*, 24(4):527–537.

Deubel, H. and Bridgeman, B. (1995). Fourth purkinje image signals reveal eye-lens deviations and retinal image distortions during saccades. *Vision Research*, 35:529–538.

Ditchburn, R. W. and Ginsborg, B. L. (1953). Involuntary eye movements during fixation. *The Journal of Physiology*, 119(1):1–17.

Duwaer, A. L. and Van Den Brink, G. (1981). Foveal diplopia thresholds and fixation disparities. *Perception and Psychophysics*, 30:321–329.

Eadie, A. S. and Carlin, P. J. (1995). Evolution of control system models of ocular accommodation, vergence and their interaction. *Medical Biological Engineering and Computing*, 33:517–524.

Eggers, S. D., Bisdorff, A., Von Brevern, M., Zee, D. S., Kim, J., Perez-Fernandez, N., Welgampola, M. S., Della Santina, C. C., and Newman-Toker, D. E. (2019). Classification of vestibular signs and examination techniques: nystagmus and nystagmus-like movements. *Journal of Vestibular Research*, 29(2-3):57–87.

Engbert, R. (2006). Microsaccades: A microcosm for research on oculomotor control, attention, and visual perception. *Progress in Brain Research*, 154:177–192.

Engbert, R. and Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research*, 43(9):1035–1045.

Engbert, R. and Kliegl, R. (2004). Microsaccades keep the eyes' balance during fixation. *Psychological Science*, 15:431–436.

Ezenman, M., Hallett, P., and Frecker, R. (1985). Power spectra for ocular drift and tremor. *Vision Research*, 25(11):1635–1640.

Fasano, G. and Franceschini, A. (1987). A multidimensional version of the Kolmogorov-Smirnov test. *Monthly Notices of the Royal Astronomical Society*, 225(1):155–170.

Findlay, J. M. (1971). Frequency analysis of human involuntary eye movement. *Kybernetik*, 8:207–214.

Fogt, N. and Jone, R. (1998). Comparison of fixation disparities obtained by objective and subjective methods. *Vision Research*, 38:411–421.

Fogt, N. and Jones, R. (1998). The effect of forced vergence on retinal correspondence. *Vision Research*, 38:2711–2719.

Greschner, M., Bongard, M., Rujan, P., and Ammermüller, J. (2002). Retinal ganglion cell synchronization by fixational eye movements improves feature estimation. *Nature neuroscience*, 5:341–7.

Grieve, K., Tiruveedhula, P., Zhang, Y., and Roorda, A. (2006). Multi-wavelength imaging with the adaptive optics scanning laser ophthalmoscope. *Optics Express*, 14:12230–12242.

Hafed, Z. M. and Clark, J. J. (2002). Microsaccades as an overt measure of covert attention shifts. *Vision Research*, 42(22):2533–2545.

Harmening, W. M., Tiruveedhula, P., Roorda, A., and Sincich, L. W. (2012). Measurement and correction of transverse chromatic offsets for multi-wavelength retinal microscopy in the living eye. *Biomedical Optics Express*, 3:2066–2077.

He, L., Donnelly, W. J., Stevenson, S. B., and Glasser, A. (2010). Saccadic lens instability increases with accommodative stimulus in presbyopes. *Journal of Vision*, 10(4):14.

Hennig, M., Kerscher, N., Funke, K., and Wörgötter, F. (2002). Stochastic resonance in visual cortical neurons: Does the eye-tremor actually improve visual acuity? *Neurocomputing*, 44-46:115–120.

Hillis, J. M. and Banks, M. (2001). Are corresponding points fixed? *Vision Research*, 41:2457–2473.

Holmqvist, K. and Blignaut, P. (2020). Small eye movements cannot be reliably measured by video-based P-CR eye-trackers. *Behavior Research Methods*, 52:2098–2121.

Hooge, I., Holmqvist, K., and Nyström, M. (2016). The pupil is faster than the corneal reflection (CR): Are video based pupil-CR eye trackers suitable for studying detailed dynamics of eye movements? *Vision Research*, 128:6–18.

Houben, M. M. J., Goumans, J., and Van Der Steen, J. (2006). Recording three-dimensional eye movements: Scleral search coils versus video oculography. *Investigative Ophthalmology and Visual Science*, 47:179–182.

Howard, I. P. and Rogers, B. J. (2008). *Seeing in Depth*. Oxford Scholarship Online.

Hunfalvay, M., Murray, N. P., and Carrick, F. R. (2021). Fixation stability as a biomarker for differentiating mild traumatic brain injury from age matched controls in pediatrics. *Brain Injury*, 35(2):209–214.

Intoy, J. and Rucci, M. (2020). Finely tuned eye movements enhance visual acuity. *Nature Communications*, 11(795):1–11.

Jaschinski, W., Jainta, S., and Kloke, W. B. (2010). Objective vs subjective measures of fixation disparity for short and long fixation periods. *Ophthalmic physiological optics : the journal of the British College of Ophthalmic Opticians (Optometrists)*, 30:379–390.

Jaschinski, W., Kloke, W. B., Jainta, S., and Buchholz, J. (2005). Horizontal fixation disparity measures with nonius lines. *Optometry and vision science*, 82:988–999.

Kagan, I., Gur, M., and Snodderly, D. M. (2008a). Saccades and drifts differentially modulate neuronal activity in v1: Effects of retinal image motion, position, and extraretinal influences. *Journal of Vision*, 8(13):1–25.

Kagan, I., Gur, M., and Snodderly, M. D. (2008b). Saccades and drifts differentially modulate neuronal activity in v1: Effects of retinal image motion, position, and extraretinal influences. *Journal of Vision*, 8(14).

Kazunori, H., Kana, O., Risako, K., Wakana, F., and Nobuyuki, S. (2016). Smaller fixation target size is associated with more stable fixation and less variance in threshold sensitivity. *PLoS One*, 11(11):e0165046.

Kilpeläinen, M., Putnam, N. M., Ratnam, K., and Roorda, A. (2020). The anatomical, functional and perceived location of the fovea in the human visual system. *Cell Press Sneak Peak*.

Kimmel, D., Mammo, D., and Newsome, W. T. (2012). Tracking the eye non-invasively: simultaneous comparison of the scleral search coil and optical tracking techniques in the macaque monkey. *Frontiers in Behavioral Neuroscience*, 6(49).

Ko, H., Snodderly, D. M., and Poletti, M. (2016). Eye movements between saccades: Measuring ocular drift and tremor. *Vision Research*, 122:93–104.

Ko, H. K., Poletti, M., and Rucci, M. (2010). Microsaccades precisely relocate gaze in a high visual acuity task. *Nature Neuroscience*, 13(12):1549–1553.

Krauskopf, J., Cornsweet, T. N., and Riggs, L. A. (1960). Analysis of eye movements during monocular and binocular fixation. *Journal of the Optical Society of America*, 50:572–578.

Kuang, X., Poletti, M., Victor, J., and Rucci, M. (2012). Temporal encoding of spatial information during active visual fixation. *Current Biology*, 22:510–4.

Li, K. Y. and Roorda, A. (2007). Automated identification of cone photoreceptors in adaptive optics retinal images. *Journal of the Optical Society of America*, 24(5):1358–1363.

Li, K. Y., Tiruveedhula, P., and Roorda, A. (2010). Intersubject variability of foveal cone photoreceptor density in relation to eye length. *Investigative Ophthalmology and Visual Science*, 51(12):6858–6867.

Martinez-Conde, S., Macknik, S. L., and Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, 5(3):229–240.

Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., and Dyar, T. A. (2006). Microsaccades counteract visual fading during fixation. *Neuron*, 49(2):297–305.

McCamy, M. B., Jazi, A. N., Otero-Millan, J., Macknik, S. L., and Martinez-Conde, S. (2013). The effects of fixation target size and luminance on microsaccades and square-wave jerks. *PeerJ*, 1(23).

McCamy, M. B., Otero-Millan, J., Di Stasi, L. L., Macknik, S. L., and Martinez-Conde, S. (2014). Highly informative natural scene regions increase microsaccade production during visual scanning. *Journal of Neuroscience*, 34(8):2956–2966.

Messias, A., Reinhard, J., Velasco e Cruz, A. A., Dietz, K., MacKeben, M., and Trauzettel-Klosinski, S. (2007). Eccentric fixation in Stargardt's disease assessed by Tubingen perimetry. *Investigative Ophthalmology and Visual Science*, 48(12):5815–5822.

Mestre, C., Bedell, H. E., Díaz-Doutón, F., Pujol, J., and Gautier, J. (2021). Characteristics of saccades during the near point of convergence test. *Vision Research*, 187:27–40.

Montesano, G., Crabb, D. P., Jones, P. R., Fogagnolo, P., Digiuni, M., and Rossetti, L. M. (2018). Evidence for alterations in fixational eye movements in glaucoma. *BMC Ophthalmology*, 18(191).

Mostofi, N., Intoy, J., and Rucci, M. (2021). Foveal vision at the time of microsaccades. *BioRxiv*.

Møller, F., Laursen, L. M., Tygesen, J., and Sjølie, A. K. (2002). Binocular quantification and characterization of microsaccades. *Graefes Archive for Clinical and Experimental Ophthalmology*, 240:765–770.

Niehorster, D. C., Zemblys, R., and Holmqvist, K. (2021). Is apparent fixational drift in eye-tracking data due to filters or eyeball rotation? *Behavior Research Methods*, 32:311–324.

Nyström, M., Andersson, R., Magnusson, M., Pansell, T., and Hooge, I. (2015). The influence of crystalline lens accommodation on post-saccadic oscillations in pupil-based eye trackers. *Vision Research*, 107:1–14.

Nyström, M., Hooge, I., and Andersson, R. (2016). Pupil size influences the eye-tracker signal during saccades. *Vision Research*, 121:95–103.

Otero-Millan, J., Macknik, S. L., and Martinez-Conde, S. (2014). Fixational eye movements and binocular vision. *Frontiers in Integrative Neuroscience*, 8.

Otero-Millan, J., Troncoso, X. G., Macknik, S. L., Serrano-Pedraza, I., and Martinez-Conde, S. (2008). Saccades and microsaccades during visual fixation, exploration, and search: foundations for a common saccadic generator. *Journal of Vision*, 8(14):21.

Pallikaris, A., Williams, D. R., and Hofer, H. (2003). The reflectance of single cones in the living human eye. *Investigative Ophthalmology and Visual Science*, 44:4580–4592.

Poletti, M., Listorti, C., and Rucci, M. (2013). Microscopic eye movements compensate for nonhomogeneous vision within the fovea. *Current Biology*, 23(17):1691–1695.

Poletti, M. and Rucci, M. (2016). A compact field guide to the study of microsaccades: Challenges and functions. *Vision Research*, 118:83–97.

Poonja, S., Patel, S., Henry, L., and Roorda, A. (2005). Dynamic visual stimulus presentation in an adaptive optics scanning laser ophthalmoscope. *Journal of Refractive Surgery*, 21(5):S575–S580.

Putnam, N. M., Hofer, H. J., Doble, N., Chen, L., Carroll, J., and Williams, D. R. (2005). The locus of fixation and the foveal cone mosaic. *Journal of Vision*, 5(7):632–639.

Ratliff, F. and Riggs, L. A. (1950). Involuntary motions of the eye during monocular fixation. *Journal of Experimental Psychology*, 40(6):687–701.

Ratnam, K., Domdei, N., Harmening, W. M., and Roorda, A. (2017). Benefits of retinal image motion at the limits of spatial vision. *Journal of Vision*, 17(1):30.

Reiniger, J. L., Domdei, N., Holz, F. G., and Harmening, W. M. (2021). Human gaze is systematically offset from the center of cone topography. *bioRxiv*.

Reiniger, J. L., Domdei, N., Linden, M., Holz, F. G., and Harmening, W. M. (2019). Relationship between the foveal photoreceptor mosaic and adaptive optics corrected visual acuity. *Investigative Ophthalmology and Visual Science*, 60(9):1777.

Riggs, L. A. and L, S. A. M. (1968). Accuracy of retinal image stabilization achieved with a plane mirror on a tightly fitting contact lens. *Vision Research*, 8:159–169.

Robinson, D. A. (1963). A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Transactions on Bio-medical Electronics*, 10(4):137–145.

Rolfs, M. (2009). Microsaccades: Small steps on a long way. *Vision Research*, 49(20):2415–2441.

Roorda, A., Romero-Borja, F., Donnelly, W. J., Queener, H., Herbert, T., and Campbell, M. C. W. (2002). Adaptive optics scanning laser ophthalmoscopy. *Optics Express*, 10(9):405–412.

Rossi, E. A. and Roorda, A. (2010). The relationship between visual resolution and cone spacing in the human fovea. *Nature Neuroscience*, 13(2):156–157.

Rucci, M. (2008). Fixational eye movements, natural image statistics, and fine spatial vision. *Network: Computation in Neural Systems*, 19(4):253–285.

Rucci, M., Ahissar, E., and Burr, D. (2018). Temporal coding of visual space. *Trends in Cognitive Sciences*, 22(10):883–895.

Rucci, M., Iovin, R., Poletti, M., and Santini, F. (2007). Miniature eye movements enhance fine spatial detail. *Nature*, 447(7146):852–855.

Rucci, M. and Poletti, M. (2015). Control and functions of fixational eye movements. *Annual review of vision science*, 1:499–518.

Rucci, M. and Victor, J. D. (2015). The unsteady eye: an information processing stage, not a bug. *Trends in Neurosciences*, 38(4):195–206.

Shack, R. and Platt, B. (1971). Production and use of a lenticular hartmann screen. *Journal of the Optical Society of America*, 61:656.

Sheehy, C. K., Bensinger, E. S., Romeo, A., Rani, L., Stepien-Bernabe, N., Shi, B., Helft, Z., Putnam, N., Cordano, C., Gelfand, J. M., et al. (2020). Fixational microsaccades: A quantitative and objective measure of disability in multiple sclerosis. *Multiple Sclerosis Journal*, 26(3):343–353.

Sheehy, C. K., Yang, Q., W, A. ., Tiruveedhula, P., de Boer, J. F., and Roorda, A. (2012). High-speed, image-based eye tracking with a scanning laser ophthalmoscope. *Biomedical Optics Express*, 3:2611–2622.

Simon, F., Schulz, E., Rassow, B., and Haase, W. (1984). Binocular micromovement recording of human eyes: - methods. *Graefes Archive for Clinical and Experimental Ophthalmology*, 221:293–298.

Steinman, . M., Haddad, G. M., Skavenski, A. A., and Wyman, D. (1973). Miniature eye movement. *Science*, 181:810–819.

Steinman, R. M. (1965). Effect of target size, luminance, and color on monocular fixation. *Journal of the Optical Society of America*, 55(9):1158–1164.

Steinman, R. M., Cunitz, R. J., Timberlake, G. T., and Herman, M. (1967). Voluntary control of microsaccades during maintained monocular fixation. *Science*, 155(3769):1577–1579.

Stevenson, S. B. and Roorda, A. (2005). Correcting for miniature eye movements in high resolution scanning laser ophthalmoscopy. *Proceedings of SPIE Vol. 5688A: Ophthalmic Technologies XV*, 5688:145–151.

Stevenson, S. B., Roorda, A., and Kumar, G. (2010). Eye tracking with the adaptive optics scanning laser ophthalmoscope. *Proceedings of the 2010 Symposium on Eye-Tracking Research*, pages 195–198.

Stevenson, S. B., Sheehy, C., and Roorda, A. (2016). Binocular eye tracking with the tracking scanning laser ophthalmoscope. *Vision Research*, 118:98–104.

Tabernero, J. and Artal, P. (2014). Lens oscillations in the human eye. implications for post-saccadic suppression of vision. *PloS One*, 9(4):e95764.

Thaler, L., Schütz, A. C., Goodale, M. A., and Gegenfurtner, K. R. (2013). What is the best fixation target? the effect of target shape on stability of fixational eye movements. *Vision Research*, 76:31–42.

Thomson, D. (1982). Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70:1055–1096.

Ukwade, M. T. and Bedell, H. E. (1993). Stability of oculomotor fixation as a function of target contrast and blur. *Optometry and Vision Science*, 70(2):123–126.

Verdina, T., Greenstein, V. C., Sodi, A., Tsang, S. H., Burke, T. R., Passerini, I., Allikmets, R., Virgili, G., Cavallini, G. M., and Rizzo, S. (2017). Multimodal analysis of the preferred retinal location and the transition zone in patients with stargardt disease. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 255(7):1307–1317.

Vogel, C. R., Arathorn, D. W., Roorda, A., and Parker, A. (2006). Retinal motion estimation in adaptive optics scanning laser ophthalmoscopy. *Optics Express*, 14(2):487–497.

Wang, Y., Bensaid, N., Tiruveedhula, P., Ma, J., Ravikumar, S., and Roorda, A. (2019). Human foveal cone photoreceptor topography and its dependence on eye length. *Elife*, 8:e47148.

Whittaker, S. G., Budd, J., and Cummings, R. W. (1988). Eccentric fixation with macular scotoma. *Investigative Ophthalmology and Visual Science*, 29(2):268–278.

Wilk, Melissa, A., Dubis, A. M., Cooper, R. F., Summerfelt, P., Dubra, A., and Carroll, J. (2017). Assessing the spatial relationship between fixation and foveal specializations. *Vision Research*, 132:53–61.

Yarbus, A. L. (1968). *Eye movements and vision*. New York: Plenum Press.