Characterisation of nystagmus waveforms in eye-tracker signals

Rosengren, William

2020

Document Version:
Publisher's PDF, also known as Version of record

Characterisation of nystagmus waveforms in eye–tracker signals

William Rosengren

LUND
UNIVERSITY

Doctoral Dissertation
Biomedical Engineering
Lund, January 2021

Dissertation for the degree of Doctor of Philosophy in Biomedical Engineering.

Department of Biomedical Engineering
Lund University
P.O. Box 118, SE-221 00 LUND, SWEDEN

ISBN: 978-91-7895-676-0 (print)
ISBN: 978-91-7895-677-7(pdf)
ISRN: LUTEDX/TEEM-1122-SE
Report No. 1/20
Printed in December 2020 Sweden by *Tryckeriet i E-huset*, Lund.

Populärvetenskaplig Sammanfattning

Den största delen av de sinnesintryck som vi människor får från omvärlden kommer från ögonen. Synen är därför viktigt för oss, och det innebär att om vi får problem med vår syn så kan det i stor utsträckning påverka vardagsliv, sociala relationer och yrkesutövning. Att studera synen och dess funktion är därför ett viktigt forskningsområde.

I denna avhandling studeras en ögonrörelserubbning som heter nystagmus, som kännetecknas av permanenta, ofrivilliga ögonrörelser. Nystagmussymptomen leder till att de som drabbas inte har möjlighet att fixera blicken, vilket ofta leder till nedsatt syn, yrsel och en risk för ökad uttröttning. Det finns flera olika anledningar till varför nystagmus uppkommer och det är ofta svårt att ställa en korrekt diagnos vad orsaken är till att en patient har nystagmus. Exempel på sjukdomar som kan leda till nystagmus är multipel skleros och stroke. I de flesta fall är störningarna livslånga, även om själva ögonrörelsemönstret kan förändras över tiden.

För att kunna förbättra både diagnostik och behandling av patienter krävs en noggrann analys av de ögonrörelser som är associerade med nystagmus. Denna avhandling handlar om att utveckla analysmetoder för att göra det möjligt. För att kunna göra analysen, krävs att ögonrörelserna kan mätas noggrant och precist. För detta används en teknik som kallas för ögonrörelsemätning (eng. eye tracking). Den typ av ögonrörelsemätning som har använts i huvudsak i denna avhandling är baserad på en kamera som filmar ögonen, och som med hjälp av bildanalys identifierar pupillen. Pupillens position ger en indikation om var personen tittar, vilket sedan kan användas för att mäta hur ögonen rör sig.

När datan är insamlad är det viktigt att kunna analysera tillståndet hos varje patient på ett tillförlitligt sätt. Denna avhandling handlar främst om att utveckla metoder

för automatisk analys av nystagmusögonrörelser, det vill säga analys baserat på algoritmer. Arbetet har inneburit inspelning av ny nystagmusdata och utveckling av fyra olika metoder. I det första arbetet har en metod som avser att förbättra kalibreringen av mätuppställningen när man spelar in ögonrörelser från nystagmuspatienter tagits fram. Kalibrering är en process som är central för ögonrörelsemätning, då den används för att relatera pupillens position i bilderna till en blickvinkel. De kommersiella system som finns idag för videobaserad ögonrörelsemätning har ofta inte en kalibreringsprocedur som är avsedd för att användas på personer som har problem med att fixera blicken. Detta gör det svårt att använda dessa metoder för att studera nystagmuspatienter. Med den presenterade metoden blir kalibreringen mer robust jämfört med tidigare metoder.

I det andra arbetet har en metod för att modellera olika typer av nystagmusrörelsemönster utvecklats. Modellen är baserad på ett antagande om att ögonen vid nystagmus har ett oscillerande rörelsemönster. Denna modell används även i det tredje arbetet, där en metod för att klassificera olika ögonrörelsemönster introduceras. I detta arbete analyseras ögonrörelser som är inspelade både under fixeringar och följerörelser. Analys av följerörelser skulle kunna ge en fördjupad kunskap om patients tillstånd och därmed kunna förbättra diagnostiken.

I det sista arbetet utvecklas en metod för att detektera olika typer av rörelsemönster i nystagmusdata och klassificera dessa i enlighet med kliniskt etablerade nystagmusmönster annoterade av experter. Metoden använder koncept introducerade i arbete 2 och 3, och sammanför två olika analysmetoder som tidigare har använts separat. Syftet med denna metod är att underlätta vid kliniska undersökningar av nystagmus genom automatisk tolkning av inspelade ögonrörelser. Metoden ger tillgång till en detaljerad beskrivning av olika tillstånd som kan användas vid såväl diagnostik som vid utvärdering av olika behandlingar.

Abstract

This thesis deals with the analysis of eye–tracker signals recorded from nystagmus patients. Nystagmus is an eye movement disorder caused by an underlying condition, and patients who suffer from nystagmus express involuntary oscillating eye movements. The oscillatory patterns expressed by these patiens are typically linked to the underlying condition, but it is usually difficult to precisely diagnose each individual. The main focus of this thesis is to develop methods for automatic and robust analysis of nystagmus eye movements. These methods are developed with the purpose of providing diagnostic support for clinicians, or for evaluation of treatment effects.

This thesis comprises an introduction and four papers describing various aspects of nystagmus analysis. In all four papers, eye movement signals recorded using an eye tracker are used as input to the proposed methods. In the first paper, a method to robustly calibrate eye–tracker data recorded from nystagmus patients is proposed. Calibration of data from nystagmus patients using video–based systems is difficult since the calibration process relies on an ability to accurately and precisely fixate calibration targets, which is difficult for nystagmus patients. Due to the nystagmus oscillations, it is difficult to obtain calibration results that are acceptable in terms of accuracy. In this work, a novel approach to find outliers in the calibration data is implemented, and a linear Procrustes transformation is used as the calibration mapping function. The results show that the proposed approach leads to reduced gaze estimation variance, and a higher robustness against outliers in the calibration data.

In the second paper, a method to model different nystagmus waveform morphologies is presented. This model is used to characterise the nystagmus oscillations and to assert the quality of the analysed eye–tracker signals. The modelling approach is based on a stationary harmonic series, and the signals are modeled in short segments, allowing for tracking of local changes in signal characteristics. Each segment is assessed using a metric referred to as the *normalised segment error*, which is used to determine whether or not the segment contains measurement disturbances. The re-

sults show that the model is well suited to distinguish between nystagmus oscillations and disturbances in the signal.

The harmonic model from the second paper is used in the third paper in order to analyse data acquired during both smooth pursuit and fixation eye movements. Smooth pursuit eye movements may carry valuable clinical information, and reliable modelling of smooth pursuit eye movements is therefore of interest. The harmonic model is used to parametrise the different waveforms. Based on the parametrisation, a *waveform distance index* is defined, which is a metric used to measure similarity between waveforms, as well as for clustering of waveforms. Eleven different clusters are defined using known reference nystagmus waveforms, and all recorded fixation and smooth pursuit waveforms are assigned to one of the eleven cluster centers. The results show that the waveform clustering is robust, is able to distinguish between recordings from different individuals, and is suitable for analysis of smooth pursuit recordings.

In the fourth paper, a novel method to combine cycle analysis and morphological classification is proposed. The goal of this work is to provide a diagnostic tool to identify subtle differences between patients, and over time in longer or recurring recordings. The cycle analysis method uses adaptive thresholds in order to detect breaking saccades, fast phases, foveations and slow phases. Eighteen template waveforms are used to create a profile of identified morphologies for each recorded waveform. The method is evaluated against expert annotations from a public dataset. The results show that the method is capable of analysing nystagmus eye movement recordings from both video–based and magnetic scleral search coil techniques. The waveform classification is reliable for both recording techniques.

The methods presented in this thesis are used to improve the robustness and reliability for analysis of nystagmus eye movements recorded using an eye–tracker. In total, the four proposed methods constitute a complete framework showing how analysis of nystagmus eye–tracker signals may be used to improve diagnostics in nystagmus patients.

List of Publications

Included

The doctoral dissertation is based on four journal publications. The publications are included in the order listed below.

[1] William Rosengren, Marcus Nyström, Björn Hammar and Martin Stridh "A robust method for calibration of eye tracking data recorded during nystagmus," in *Behavior research methods*, pp. 1–15, 2019.

[2] William Rosengren, Marcus Nyström, Björn Hammar, Markus Rahne, Linnea Sjödahl and Martin Stridh, "Modeling and quality assessment of nystagmus eye movements recorded using an eye-tracker," in *Behavior research methods*, , pp. 1–15, 2020.

[3] William Rosengren, Marcus Nyström, Björn Hammar and Martin Stridh "Waveform characterisation and comparison of nystagmus eye-tracking signals," under review.

[4] William Rosengren, Marcus Nyström, Björn Hammar and Martin Stridh "Cycle–to-cycle nystagmus waveform characterisation," submitted for publication.

Contribution statement

For all papers, William Rosengren and his three supervisors (Marcus Nyström, Björn Hammar och Martin Stridh) are listed as authors. In all cases, William Rosengren is the main contributor and has worked on all parts from ideas, experiment design and data collection to data analysis and writing of the papers. Marcus Nyström and Martin Stridh have contributed with supervision related to eye tracking and signal processing, respectively, and with guidance when writing the papers. Björn Hammar has been

responsible for patient referrals, medical diagnostics of all patients and other clinical guidance. Markus Rahne and Linnea Sjödahl collected data and contributed to the drafting process for the second paper: *Modeling and quality assessment of nystagmus eye movements recorded using an eye-tracker*.

Related

Parts of the work have been presented at conferences. The related conference papers listed below are not included in the thesis.

[5] William Rosengren, Marcus Nyström, Björn Hammar and Martin Stridh "Suitability of calibration polynomials for eye-tracking data with simulated fixation inaccuracies," in Proceedings - ACM Symposium on Eye Tracking Research, Warsaw, Poland, p. 66, June 2018.

[6] William Rosengren, Marcus Nyström, Björn Hammar, Markus Rahne, Linnea Sjödahl and Martin Stridh, "Quality assessment of nystagmus data from eye tracker recordings," in Alicante, Spain, pp. 172, August, 2019.

Acknowledgments

After having studied at LTH for more than 9 years, there are many people that I would like to thank. I would like to start by expressing my gratitude to Martin Stridh who has been my main supervisor throughout my PhD studies. Martin has always been enthusiastic about my work, pushing me forward and helping me to get a better, deeper understanding of the concepts related to signal processing and research. I would also like to thank my two co-supervisors Marcus Nyström and Björn Hammar. You have helped me to understand eye movements, both from a technical and from a clinical perspective. Without your knowledge of eye movements and your dedication for this field, this PhD project would never have been feasible.

It has been a privilege to be a part of the signal processing group at the department of biomedical engineering. I would like to thank Leif, Martin, Frida, Hamid, Mikael, Mattias, Hesam, Mostafa, Johan, Ricardo, Mattias and Felix. I highly value the knowledge in this group within the field of biomedical signal processing, and you have helped me to develop as an engineer. A special thanks goes to Linnéa Larsson who helped me in the beginning of my Phd studies with both research and teaching related questions.

I would also like to thank all the members at the department of biomedical engineering for being great colleagues over the years. A special thanks goes to my friends at BMC. You are really great and it has been a lot of fun to hang out with you.

I would also like to thank Bengt-Arne Molin, who helped and inspired me when I was in high school, and who questioned me when I started to study business and economics instead of engineering. Without your entusiasm about science and engineering, I might never have ended up at LTH.

A big thank you goes to all participants in my studies, and a special thanks to my co-authors in the second paper Markus and Linnea. You did a great job with data collection, and without your effort, I might never have been able to conduct as

advanced studies as I have.

I would like to thank all of my family and friends for supporting me through out this long journey towards my doctoral degree. Finally I would like to thank the people who are closest to me: my sweet Tove, my sisters Karolina and Julia and my parents Kristina and Jonas, for all you love and support throughout this project and all other days. You mean the world to me!

William
Malmö, December 14, 2020

Contents

Part I

Introduction

Chapter 1
Thesis introduction

This thesis deals with the analysis of eye movements in patients with *nystagmus*, which is an eye movement disorder characterised by involuntary oscillating eye movements. The focus of this thesis has been to develop various methods for analysis of these eye movements with the hope to provide a better understanding of the symptoms expressed by the patients who suffer from nystagmus. Data in this thesis were acquired using an *eye tracker*, which is a tool for recording eye movements and estimating where a person is looking. Using eye tracker data, algorithms for analysis of nystagmus eye movements have been developed.

Our visual system is important for how we interpret and interact with the world around us, and a majority of our sensory receptors, about 70 %, are found in the eyes [1]. However, the field of vision where we can resolve objects in high resolution is limited to approximately 0.5° of the visual field [2]. In order to resolve objects in a high level of detail, we therefore need to look directly at them, and keep our gaze stable on the object. Because of this, eye movement control is an important part of our visual system. Since patients who suffer from nystagmus have a reduced ability to control their eye movements, their visual function is often impaired compared to healthy individuals.

There are two main categories of nystagmus eye movements: physiological and pathological [2]. There are several forms of physiological nystagmus, one is a reflex found in healthy subjects, where oscillations arise as a response to a moving environment. These eye movements may be elicited, e.g., when sitting on a train and looking at trees passing by. This type of nystagmus is often referred to as optokinetic nystagmus (OKN). Physiological nystagmus is not within the scope of this thesis, and nystagmus will therefore in this thesis refer to pathological nystagmus unless otherwise specified.

Pathological nystagmus can be caused by factors such as disease, trauma or sub-

stance abuse. In some cases there is no known pathology, and in such cases the nystagmus is called idiopathic [2, 3]. It has a reported prevalence of approximately 0.2 % in the general population [4]. Pathological nystagmus is often further divided into acquired and early onset nystagmus, where the former is developed from, for example, neurological disease and the latter is developed before or shortly after birth [5]. The main characteristic of nystagmus is the oscillating eye movement pattern observed in this patient group. These oscillations have traditionally been referred to as the *nystagmus waveform*, or just the waveform. The waveform describes important features of the eye movements such as the direction of the oscillations and whether it is a jerk, pendular, or torsional form of nystagmus. Other interesting characteristics of nystagmus are the amplitude and frequency of the oscillations. Their product, the *intensity*, has been used when evaluating the current condition of a patient as well as when evaluating the effect of various medical treatments [2]. The medical implications of nystagmus may be a decrease in visual acuity and *oscillopsia*, which is a sensation that the world is in constant motion. Interviews with patients diagnosed with nystagmus have revealed social problems linked to the condition, such as poor self-esteem, problems with their social life and a feeling that the medical community had abandoned them [6].

As described above, diagnostics and treatment evaluation in nystagmus patients require extraction of detailed characteristics in the individual patient. However, obtaining these characteristics, for example oscillation amplitude, is difficult using only the naked eye. Today, video–based eye trackers are preferred by many researchers due to their ease of use and non–invasiveness. A problem when recording patients with nystagmus using commercial eye–tracker setups, is that such systems often are not designed to record people who have impaired eye movement control. For example, calibration of nystagmus patients using video–based eye trackers is very difficult. During the calibration, it is assumed that the participant is able to gaze directly at a target while keeping their eyes still for an extended time period, often one or two seconds. This is typically not possible for nystagmus patients, which results in invalid calibrations for many patients. Eye–tracking data are often used to identify different types of eye movements, known as *events* [7]. Event detection refers to segmentation of recorded eye movement signals into different types of eye movement categories, such as fixations and saccades. Algorithms for event detection are, however, in general also not designed to handle nystagmus eye movements. Instead, specialised algorithms for the analysis of nystagmus data have been developed, often with the aim to characterise the various nystagmus waveforms that may be observed in a patient with nystagmus.

In this thesis, several algorithms aimed at analysing nystagmus eye movements have been developed. A total of four research papers have been included in this thesis and they deal with the following concepts:

I The development of a protocol for calibrating nystagmus data using video–based eye trackers.

II The construction of a harmonic sinusoidal model used both to parameterise nystagmus waveforms and to assess the quality of the recorded eye movement signals.

III A method for comparison between eye movement data when viewing static and dynamic stimuli, and clustering of waveforms based on model features proposed in Paper II.

IV The development of a nystagmus waveform classification system based on cycle–to–cycle analysis.

The introduction of this thesis continues as follows. In Chapter 2, a description of the human visual system, including the function of the eyes and the muscles controlling the eye movements, is provided. The eye tracking technologies used for data collection in this thesis are introduced in Chapter 3. A detailed description of nystagmus is provided in Chapter 4. Finally, each of the included papers are briefly summarised in Chapter 5.

Chapter 2
The human visual system

The main purpose of the visual system is to detect visible light and convert it to something that our brain can interpret. Before considering nystagmus eye movements, it is of importance to understand the healthy visual system. This is a complex system, containing many different parts, and some of these will be presented here. This chapter is divided into the following two sections. First, a description of the anatomy and physiology of the visual system is provided in Section 2.1. Eye movements and the muscles controlling them are discussed in Section 2.2.

2.1 The anatomy and physiology of the visual system

The eyes are an important part of the visual system, and they are responsible for converting incoming light to electrical impulses that are transmitted to and processed in the brain [8]. As is illustrated in Figure 2.1 [9], the eye itself consists of various tissues, all of which are important for the function of our vision. The most anterior part of the eye is a thin structure called the *cornea*, and when light passes through the cornea it is refracted. After the light has passed through the cornea it goes through the *pupil*. The pupil may be described as a hole, and the size of the pupil changes depending on the light conditions. During dark conditions it increases in size in order to let more light enter the eye. After the light has passed through the pupil, it goes through the lens, which has the purpose to focus the gaze.

The retina covers the inside of the eye globe and consists of different highly specialised neural cells organised in interconnected retinal layers. One cell type in the retina is the photoreceptor cells, which transform light into electrical implulses [8]. In the human eye, two types of photoreceptors are found: *cones* and *rods*. Cones are used for color and contrast vision whereas the rods are used to differentiate between light and dark conditions. In total, there are more rods than cones in the retina, but

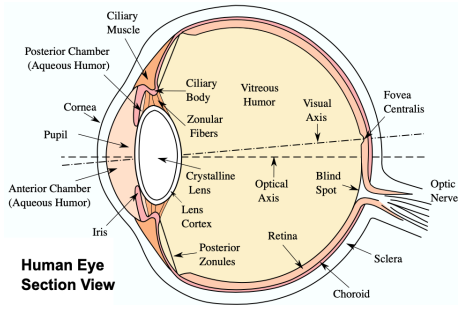


Figure 2.1: Illustration of the human eye. Figure reprinted from Wikimedia commons [9].

there is a region on the retina called the *fovea*, where the concentration of cones is higher than the concentration of rods, and there is one spot in the eye where there are no photoreceptors at all, called the *blind spot*, which is where the *optic nerve* enters the eye.

The distribution of the two photoreceptor cells on the retina as a function of the angle from the fovea is illustrated in Figure 2.2 [10]. In order to see an object as clearly as possible, it should fall within $0.5°$ from the foveal center [2]. This property of the retina makes it important for us to move our eyes in order to image new objects in high contrast. Once the conversion of light has occurred in the photoreceptor cells, the signal is propagated through the so called bipolar cells to the ganglion cells. The axons of the ganglion cells form the optic nerve, and the electric impulses are transmitted through the optic nerve to the brain, where the impulses are processed [8].

When the light has been converted by the neurons in the retina, the neural signal is sent through the visual pathways to the visual cortex, where the signals are processed by the brain. The visual pathways are illustrated in Figure 2.3 [11]. As is illustrated in the figure, each eye's visual field is divided by a vertical meridian through the fovea into a right and a left hemifield. The neurons from the right hemifield (green) of each eye converge at the optic chiasm [12], and then travel together in the left hemisphere to the left visual cortex. The left hemifields are organised in the same way and their neurons end up in the right visual cortex.

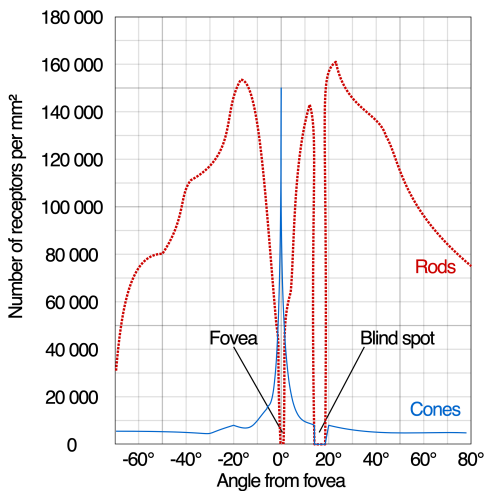


Figure 2.2: Illustration of the photoreceptor distribution on the retina. Figure reprinted from Wikimedia Commons [10].

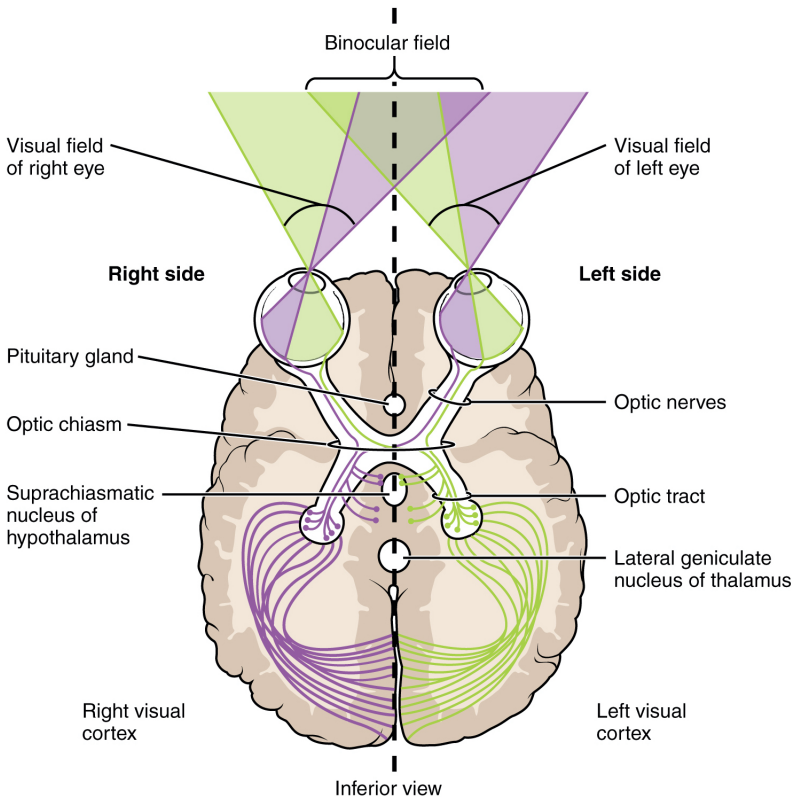


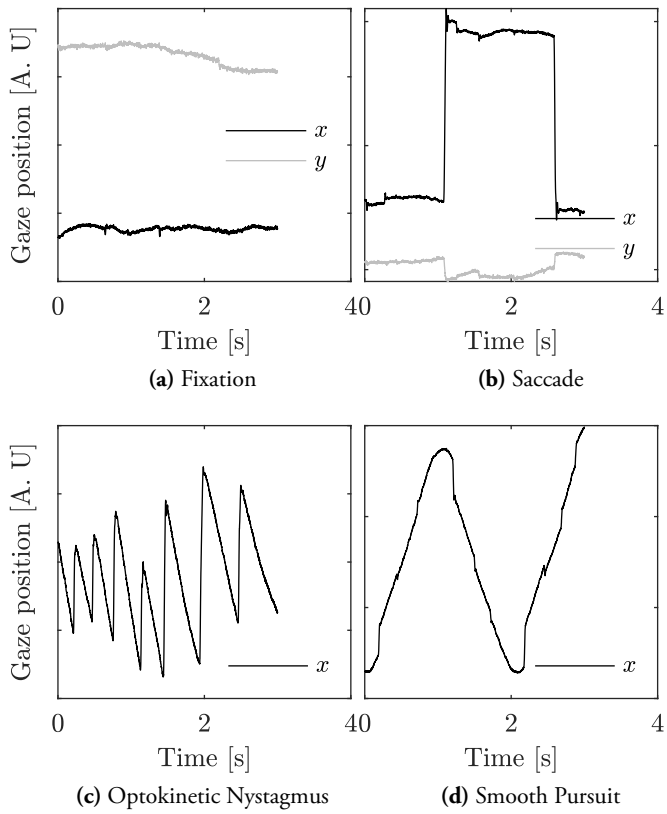
Figure 2.3: Illustration of the visual pathways to the visual cortex and the optical fields from the inferior view. Figure reprinted from Wikimedia Commons [11].

2.2 Eye muscles and eye movements

As described in Section 2.1, it is important to control our eye movements since high contrast vision is limited to only 0.5° of the visual field. There are three muscle pairs that allow us to move our eyes in three planes: the horizontal, the vertical and the torsional planes [1]. The muscle pairs are the *lateral* and *medial recti*, the *superior* and *inferior recti* and the *superior* and *inferior obliques*. These muscles pairs are used to generate different types of eye movements. Two main classes of eye movements based on their function are the *gaze stabilising* and *gaze shifting* eye movements [2]. Four different eye movements are considered in this section, *fixations*, *saccades*, *optokinetic nystagmus* and *smooth pursuit*, and eye tracker recordings illustrating these are presented in Figure 2.4. The fixation, optokinetic nystagmus and smooth pursuit are all in the gaze stabilising class of eye movements, and the saccades belong to the gaze shifting eye movements [2].

Gaze stabilising eye movements are used to keep the image of the intended fixation target as close to the fovea as possible. Fixations are used when viewing static objects, and during fixations the retinal drift is minimised [2]. An example of a fixation is when looking at a word in a book. The purpose of the smooth pursuit eye movement is to keep an object that is in motion on or close to the fovea [2]. Smooth pursuit eye movements may be observed when a person follows a flying bird with the gaze. It is not possible to generate smooth pursuit eye movements without a moving object, or the perception of a moving object to follow. Optokinetic nystagmus is an oscillatory eye movement caused by the optokinetic reflex [13], and it occurs when the surrounding environment is in motion, as opposed to smooth pursuit, where an object is moving. Optokinetic nystagmus is usually divided into two different phases: the *slow phase* (SP) and the *fast phase* (FP) [14]. The oscillations begin with a drift of the eye in the direction of the motion of the environment, the slow phase, followed by a corrective eye movement in the opposite direction to the environment motion, the fast phase. Optokinetic nystagmus is typically expressed in a repetitive pattern which is not the case with smooth pursuit.

The saccade is a gaze shifting eye movement, which is used to shift the gaze between different objects [2]. These are fast eye movements with short duration, typically less than 100 ms. It is important to note that no standardised definitions of different types of eye movements exist, and therefore different researchers may use different definitions [15]. In this thesis, the definitions from [2] are mainly followed, unless otherwise stated.



(a) Fixation **(b)** Saccade

(c) Optokinetic Nystagmus **(d)** Smooth Pursuit

Figure 2.4: Examples of four eye movements: fixation, saccade, optokinetic nystagmus and smooth pursuit are presented in panels (a) - (d), respectively. The x (black) and y (gray) signals represents the horizontal and vertical eye positions, respectively. The x-axis in each figure represents the time and the y-axis represents the eye position. The presented data were collected using an EyeLink 1000 Plus.

Chapter 3
Eye tracking

Eye tracking is a technique that has been used in research since at least the beginning of the twentieth century [16]. Since then, it has become a tool for conducting research in various areas such as reading, psychology, consumer behaviour and medicine [17]. The technique may either be used to determine where a person is looking, i.e., gaze direction, or as a tool to measure eye movements relative to head orientation. There are numerous approaches to measuring eye movements and some of the more well known techniques are: *Electrooculography*, *dual purkinje imaging*, *magnetic scleral search coils* and *video–based* eye trackers. The first two techniques are not used in this thesis, and readers are referred to other sources for a more in–depth description of the methods [2, 17].

Today the most common eye tracker technology used in research is the video–based eye tracker [17], and this type is therefore the main focus of this chapter. This chapter is structured as follows: A description of the general principles of video–based eye trackers is provided in Section 3.1. Some of the data used in Paper IV were taken from a publicly available dataset, which was recorded using a magnetic scleral search coil system. A brief description of this technique is given in Section 3.2.

3.1 Video-based eye tracking

Video–based eye tracking is based on capturing images of the eye(s) using a video camera system [17], and then extract features from the captured images, which are used to estimate the orientation of the eye. The video–based eye tracking system used in this thesis is the EyeLink 1000 Plus (SMI Research, Ontario, Canada). This system is a specific type of video–based eye tracker that uses the pupil and the *corneal reflection* to estimate the gaze, and will be referred to as PCR eye tracking in thesis. In Section 3.1.1, the feature detection of the PCR based eye tracker used in this

thesis is discussed. Coordinate systems that are relevant for the experiments and gaze estimation are presented in Section 3.1.2. The calibration and gaze estimation are described in Section 3.1.3. The experiment setup used to record data are presented in Section 3.1.4. Finally, the concept of data quality is discussed in Section 3.1.5. In Figure 3.1, a block diagram depicting the various steps used in PCR eye tracking to estimate gaze is shown. The calibration step is used to convert measurements within

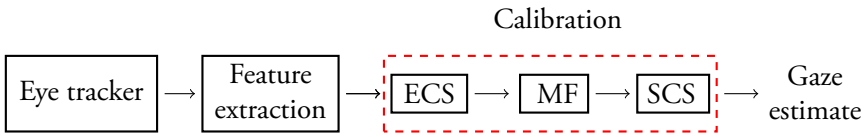


Figure 3.1: Gaze estimation process. Gaze estimation process. Block diagram of the gaze estimation process in video–based eye tracking. The blocks inside the red rectangle represent the calibration process, i.e., the transformation between the eye tracker coordinate system (ECS) and the stimuli coordinate system (SCS) using some mapping function (MF). The input data represented in the eye tracker coordinate system are generated by extraction of features in the images captured by the eye tracker. The result of the calibration process is a gaze direction estimate.

the eye tracker to eye movements related to the stimulus screen.

3.1.1 Feature detection

Pupil-corneal reflection eye tracking systems use the pupil center as the primary feature for gaze estimation. A second feature, called the corneal reflection, is used as additional input to the estimation of the gaze position. The corneal reflection is sometimes referred to as the first Purkinje image [17], and originates from the reflection of an external light source on the corneal surface. The problem that is adressed by adding the corneal reflection as a feature is to make the eye tracker less sensitive to head movement relative to the camera. [18].

The detection of these two features is often implemented in the software of commercial eye tracking systems, and is therefore in general not accessible for researchers. There are various ways to detect the pupil and corneal reflections from the captured images. One approach is to assume that the darkest area in the eye image is the pupil, and then threshold the image in order to get an estimate of the pupil center [19]. Similar thresholding techniques may be used to estimate the corneal reflection [17]. Another approach, that may be paired with the thresholding technique, is to use a shape model of the pupil to detect the pupil area in the eye image [19].

The corneal reflection may be viewed as a point of reference, and if using both the pupil center and the corneal reflection as features for gaze estimation, it is common to create a two dimensional coordinate. In this work it is called the *pupil - corneal reflection vector* (PCRV), and it is created by subtracting the corneal reflection coordinate from the pupil center position [17]. An example of these features are presented in Figure 3.2 [20]. The red cross illustrates the pupil center position, and the green cross

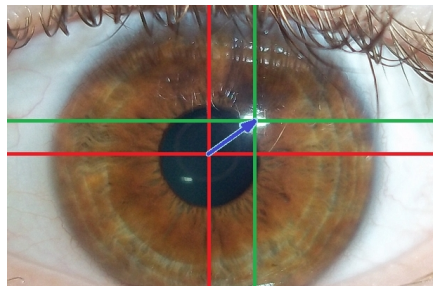


Figure 3.2: Example of features in the eye image. The pupil center and the corneal reflection are represented by red and green crosses, respectively. The vector between the respective centers is referred to as the pupil-corneal reflection vector, shown as a blue arrow. Figure reprinted from Wikimedia Commons [20].

shows the corneal reflection center. The vertical and horizontal difference between the pupil center and the corneal reflection creates the PCRV.

3.1.2 Coordinate systems

When using eye trackers similar to the one used in this thesis, there are three coordinate systems that are important for the gaze estimation: the *stimulus coordinate system*, the *eye tracker coordinate* system, and the *participant coordinate system*. In an experiment, the participant is asked to view some stimuli on a computer screen, which is defined by some coordinate system. Here, this coordinate system is called the stimulus coordinate system, since the stimulus presented during the experiments are defined using these coordinates. The eye movements, however, are recorded by an eye tracker, which generally has a coordinate system that is independent of the stimulus coordinate system. For most researchers it is of interest to know where the person looked on the screen and not what the eye tracker coordinates are. To complicate things further, each participant has different characteristics such as eye shape and distance between the eyes, which need to be accounted for when performing gaze estimation. Below, an explanation of each of these coordinate systems is provided.

An illustration of the stimulus coordinate system is presented in Figure 3.3, and the coordinate system is typically expressed in degrees instead of meters. For each sample, the current gaze position is represented as a two dimensional metric where the x and y values correspond to the horizontal and vertical position, respectively. The coordinate system is defined such that movements to the right and upwards are positive, with the screen center being defined as coordinate $(0°, 0°)$. In Figure 3.3, the eye movement from point A to point B would result in a difference of $(-\alpha°, 0°)$ in the horizontal and vertical positions respectively.

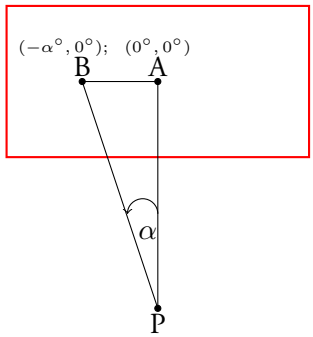


Figure 3.3: Illustration of the stimulus coordinate system. The distance from point A to point B is $-\alpha°$ in the horizontal plane and $0°$ in the vertical plane. The point P represents the position of the participant. The screen coordinate system is bounded by the red rectangle.

The PCR eye trackers provide information about the locations of the pupil and corneal reflection on the camera sensor, which typically define the eye tracker coordinate system. The eye tracker coordinate system, however, is not always known to the users, and the measurements obtained from the eye tracker are in arbitrary units, which are related to the sensor of the camera. In order to establish a mathematical relationship between the eye tracker coordinate system and the stimulus coordinate system, a calibration is required, which is further discussed in Section 3.1.3.

When performing the gaze estimation, individual differences between participants need to be taken into account. For example, the shape and location of the pupil influences the gaze estimation, and varies between participants [21]. Individual characteristics are typically accounted for by a calibration, and it is in general not suitable to apply the calibration model of one person to another person's data.

3.1.3 Calibration and Gaze Estimation

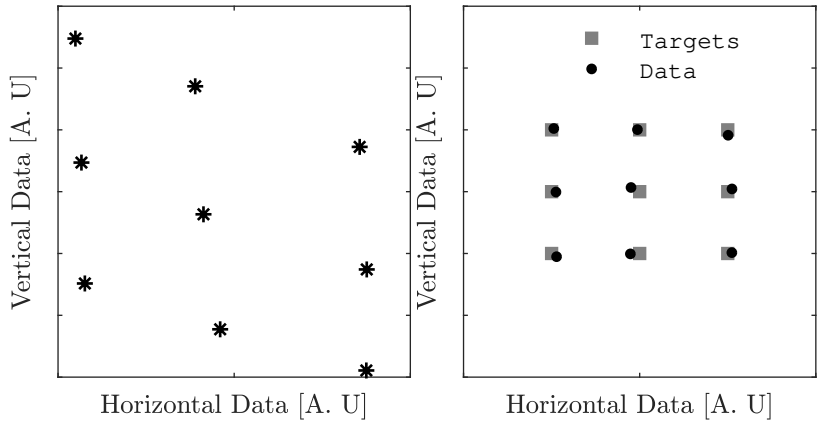
As described in Section 3.1.2, data recorded in the eye tracker coordinate system cannot directly be used to determine the direction of gaze. In order to retrieve this information, the eye tracker coordinate system must be linked to the stimulus coordinate system, as well as to account for the individual differences of each participant. This process is referred to as calibration. If using only pupil and corneal reflection data without calibration, it is possible to retrieve information about temporal changes in the eye movement signal, such as the frequency of nystagmus oscillations. It is, however, not possible to assess the magnitude of oscillations nor the screen position of the gaze. The calibration process is therefore an initial step of most experiments. Commercial software packages typically have calibration protocols built into the system and the calibration protocol is automatically run when an experiment is initiated. These protocols are in general not open to the user of the system, which means that it is not possible to know exactly how the systems handle participants who are hard to calibrate.

The transformation of uncalibrated data to gaze data may be stated using the following equation

$$\boldsymbol{S} = f(\boldsymbol{E}), \tag{3.1}$$

where the matrices $\boldsymbol{S}^{n\times 2}$ and $\boldsymbol{E}^{n\times m}$ represents the extracted features in the stimuli coordinate system and eye tracker coordinate system, respectively, and f represents the mapping function. The variable n is the number of coordinate pairs that are used to estimate the mapping function parameters, and the variable m describes the number of features used in each transformation. In order to obtain the n calibration data points, the participant is asked to view n targets, displayed in the stimuli coordinate system. The data recorded during the calibration are defined in the eye tracker coordinate system, and are recorded using the eye tracker. Once data for all calibration targets have been recorded, the parameters of the mapping function are estimated.

There are various options for how to model the transformation between the eye tracker coordinate system and the stimuli coordinate system, and some common choices include regression modes, neural networks and geometric modelling [19]. An example of the transformation of uncalibrated data, i.e., data in the eye–tracker camera domain, to calibration targets is presented in Figure 3.4. In Figure 3.4a, uncalibrated data, i.e., the PCR-vector, collected during the calibration process are shown. These data are defined in the eye tracker coordinate system. The transformation of the uncalibrated data $\boldsymbol{E}$ through the mapping function f, results in an estimation of the target positions $\boldsymbol{S}$, which are displayed in Figure 3.4b. The calibration is often validated by asking the participant to look at another set of targets after the calibration, and calculate the difference between these target locations and the estimated gaze locations. This difference is known as the *accuracy* of the gaze estimation and



(a) Uncalibrated data **(b)** Calibration targets and calibration target estimation

Figure 3.4: (a) uncalibrated pupil - corneal reflection coordinates in the eye tracker coordinates system. (b) Visual targets that participants are asked to fixated during the calibration process along with calibrated data in the stimulus coordinate system.

this metric is further discussed in Section 3.1.5.

3.1.4 Experiment setup

All eye tracking data in this thesis, besides the data taken from [22], have been collected using the EyeLink 1000 Plus, which is manufactured by SR Research (Ontario, Canada), using the host software v. 5.09 and the DevKit 1.11.571. An illustration of the recording setup is presented in Figure 3.5. The setup consists of four components: the eye tracker camera, an infra red (IR) light source, the participant, and the computer screen used to present the stimulus. The geometry of this setup is in accordance with the manufacturer recommendations [23] and it has been used throughout this thesis. The participants have been seated approximately 80 cm away from the computer monitor displaying the various stimuli. The data have been recorded binocularly at 1000 Hz. All stimuli have been presented using the PsychoPy toolbox version 1.83 for Python 2.7 [24].

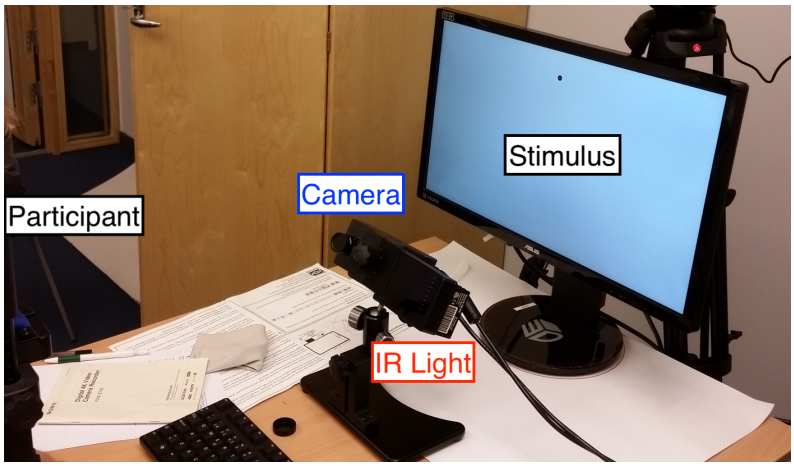


Figure 3.5: Illustration of the experiment setup. The participant position, the infra red (IR) light source, camera and stimulus screen are marked in the image.

3.1.5 Data quality

For data recorded using video–based eye trackers, several factors may interfere with the recordings, which may lead to decreased data quality. The most common data performance metrics are accuracy and *precision* [17]. Accuracy measures the difference between the true and the estimated gaze location, and is often operationalized as the distance between a visual target and the data recorded from a participant looking at the target. This is a measurement of the systematic error of the eye–tracker data. Precision measures the sample–to–sample variation of the data, and represents random errors. They are estimated through a validation of the calibration model. In this section, one type of disturbance that is commonly found when using video–based eye trackers and that may influence data quality is blinks. Data loss here means missing samples in the eye movement signal.

Blinks are movements of the eyelid, during which the eyelids meet and cover the eyeball. The main function of the blink is to prevent the eye from drying out, by remoistening the eye. The blink introduces two different problems for video–based eye trackers [17]. First, when the eyelid occludes the pupil, it is not possible for the video camera to find the pupil, and thus no eye tracker data are available. Second, when the eye is partially occluded, the pupil feature is different compared to when

the eye is not occluded. The movement of the pupil may appear as a saccade–like eye movement in the eye tracker signal [17]. In Figure 3.6, an example of a saccade–like eye movement caused by a blink in shown. As may be seen in the figure, there are three

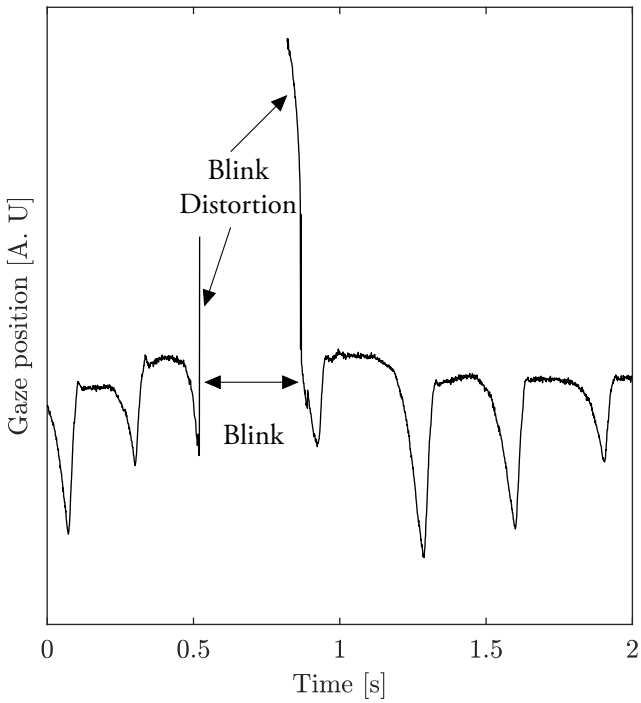


Figure 3.6: Blink distortions. An example of the eye tracker data right before, during and right after a blink. Right before and right after, there are high frequency components, which do not reflect actual eye movements, but rather artefacts associated with inaccurate pupil center detection.

components to the blink which has an effect on the resulting eye movement recording. Right before the blink, there is a high frequency component, that is not a reflection of the true eye movement, but rather an inaccurate description of the current eye position, due to the occlusion of the pupil by the eyelid. When the eyelid completely covers the pupil, there is no data available. When the eyelid begins to open, there is another high frequency component, which again is an inaccurate description of the current eye position. Due to this, blinks are considered as disturbances in the signals and excessive blinks may corrupt an entire recording.

3.2 Magnetic scleral search coil eye tracking

Some of the data used in paper IV are publicly available and collected using the magnetic scleral search coil technique [22] `http://www.omlab.org/OMLAB.html`. This technique is based on placing contact lens(es) on the eye(s) of the participant, where either one or two metal wires are embedded in the lenses [25]. The participant sits in a weak magnetic field, and when the eyes move, the wires give rise to an electric potential that is proportional to the eye movement. This technique has several advantages compared to the video–based eye tracking system. First, the signal–to–noise ratio is superior using this technique. It is also possible to calibrate the device using another individual, e.g., a healthy individual, which is advantageous for nystagmus recordings. When using two wires inside the lens, torsional eye movements may also be recorded [25]. The magnetic scleral search coil system is generally considered as the gold standard for eye tracking recordings, since the data quality is superior compared to all other methods [25].

The main drawbacks of this technique is that it is invasive and usually requires anesthetics before the lens(es) are placed on the eye(s) [2]. It is typically not possible to collect longer recordings, since the participants experience discomfort after a while.

Chapter 4
Nystagmus

In this Chapter, a thorough characterisation of various aspects of nystagmus is provided. As discussed previously, pathological nystagmus is usually divided into two subcategories: early onset (infantile) and acquired nystagmus [26, 27]. Early onset nystagmus is developed early in life, either before or up to a few months after the child is born. Some pathologies linked to early onset nystagmus include strabismus, congenital cataracts and albinism [28]. If the cause of early onset nystagmus is unknown, it is referred to as idiopathic nystagmus [2]. Acquired nystagmus, on the other hand, usually arises later in life [29], and may develop after head trauma, substance abuse or neurological disease [2, 3]. The cause of nystagmus differs between individuals, and it is therefore of clinical interest to investigate and find the specific cause for each individual.

Diagnostics of nystagmus is difficult, and the symptoms requires detailed examination in order to establish the underlying cause [2]. As stated in the thesis introduction, the purpose of this work is to contribute with methods for analysis of nystagmus eye movements, and to provide tools that can improve the medical analysis of patients with the condition. In order to do so, it is important to have a basic understanding of the causes for nystagmus, the current tools for analysing the symptoms and the treatments that exist.

In Section 4.1, a general description of nystagmus is provided, where terminology for important concepts is introduced. This is followed by Section 4.2, where the causes and different types of nystagmus are discussed in detail. Eye movement measurements and calibration of nystagmus patients are discussed in Section 4.3 and methods to characterise the overall condition of nystagmus patients are described in Section 4.4. In Section, 4.5, a brief overview of different ways to treat some causes of nystagmus is presented. An overview of methods to analyse nystagmus waveforms is presented in Section 4.6.

4.1 Nystagmus oscillation analysis

Nystagmus may occur in three different planes: horizontal, vertical or torsional [2]. The oscillations can also occur in a combination of these planes. The two main categories of nystagmus eye movement patterns are jerk and pendular nystagmus [2]. For jerk nystagmus, the oscillations have traditionally been divided into cycles, where each cycle begins with a movement away from the currently fixated object. This is called the slow phase (SP) of the oscillation. During the fast phase (FP), the slow phase movement is corrected by an eye movement in the opposite direction, towards the starting point of the cycle. The nystagmus cycles always begin with the slow phase, in contrast to the saccadic intrusions, where the cycle starts with a fast eye movement away from the intended fixation target [2]. The division into a slow and fast phase is illustrated in Figure 4.1 where the slow phases are shown as dash-dotted lines and the fast phases are shown as solid lines. This decomposition into slow and fast phases was created using the algorithm described in [30]. To divide the signal into oscillation cycles serves as the bases for various algorithms aimed at analysis nystagmus eye movements [29, 31]. For pendular nystagmus, there is no clear difference between

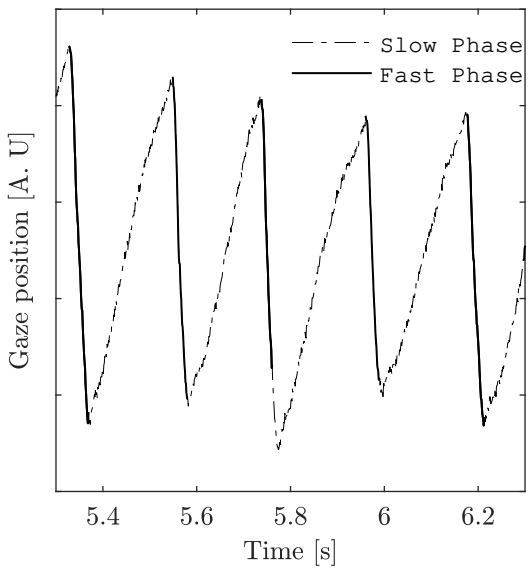


Figure 4.1: Slow and Fast Phases in Nystagmus Oscillation. The slow phases are plotted as dash-dotted lines and the fast phases are plotted as solid lines. The horizontal eye tracker signal is shown.

the slow and fast phases of the eye movement, and the oscillations are similar to a

sinusoidal signal [2].

When describing nystagmus eye movements patterns, four properties are often used: amplitude, frequency, *intensity*, and *waveform morphology* [27, 32]. Intensity is the product of amplitude and frequency, and is usually expressed in degrees per second. The nystagmus properties are illustrated in Figure 4.2. Since nystagmus is defined as involuntary oscillating eye movements, the definitions of the four properties are in this thesis constructed using an *oscillation cycle*. The oscillation cycle begins with the start of a slow phase and ends with the end of a fast phase. The frequency is the inverse of the time to complete one cycle, and the amplitude is the difference between the minimum and maximum eye positions during one oscillation cycle. The waveform morphology is defined as the eye movement pattern during one oscillation cycle. In some nystagmus waveform morphologies, a *foveation period* is present. A foveation period is defined as a segment of the oscillation cycle where the fovea is aligned with the intended fixation target, and the eye velocity is sufficiently low to allow for high visual acuity [33]. The presence of a foveation period is an indication of that the patient has a higher visual function compared to a patient without a foveation period [2]. The degree of binocular synchronisation is also sometimes used

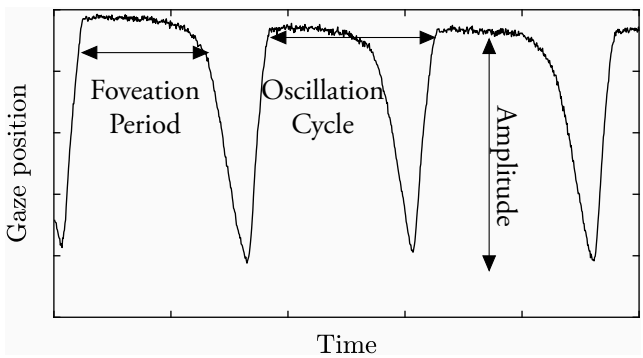


Figure 4.2: Waveform Features. Illustration of peak–to–peak amplitude, oscillation cycle and foveation period in a recorded nystagmus signal.

for diagnostic purposes. If the eyes move in synchrony, the nystagmus is said to be *conjugate*, whereas if the eyes do not move in synchrony, the nystagmus is said to be *disconjugate* [2].

The *null zone* is defined as the gaze angle for which the nystagmus intensity is at a minimum [32], and this does not necessarily coincide with the *primary position*. The primary position is usually defined as the gaze position where the head is aligned along the sagittal plane, and the eyes are directed straight ahead.

The beat direction, i.e., the direction of the fast phase has also been used to describe the oscillations in patients. However, the slow and fast phases alone do not provide a sufficiently detailed description of all possible waveforms in order to distinguish them [34]. The waveform morphology is used to describe the overall shape of the oscillation cycle, and is thus a more in-depth description compared to the beat direction. There are eighteen waveform morphologies commonly described in the literature [34]. The waveform morphology is more difficult to describe compared to the amplitude or frequency. From the waveforms described in [34], six are described as having one left and one right beating version. The twelve waveforms, not including the right beating waveforms, are illustrated in Figure 4.3. The waveforms have been reconstructed using the method provided in [35]. The labels on the right hand side of each waveform are the same as in the original work [34]. The presented waveforms from the bottom to the top are: asymmetric pendular (AP), bidirectional jerk - left (BDJ-L), dual jerk - left (DJ-L), jerk with extended foveation - left (EF-L), jerk - left (J-L), pendular with foveating saccades (P_{FS}), pendular (P), pseudo cycloid - left (PC-L), pseudo jerk - left (PJ-L), pseudo pendular with foveating saccades (PP_{FS},) pseudo pendular (PP), and triangular (T).

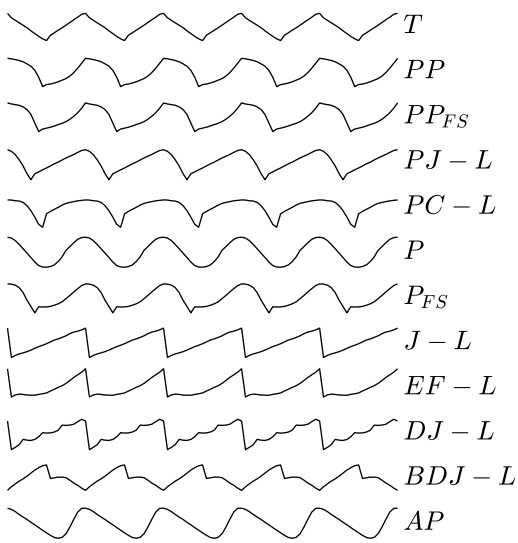


Figure 4.3: Waveform Examples. Illustrations of twelve different waveforms described in [34]. The waveforms have been reconstructed from the original work using the method described in [35].

Once the waveform morphology has been determined there are even deeper lay-

ers of characterisation of nystagmus oscillations that have been described in research papers and used clinically. The directional component, i.e., the oscillations as a function of the gaze angle, is interesting in order to determine the null zone [2]. The short–term component, i.e., the beat-to-beat changes in waveform characteristics, is also clinically relevant [34]. A clinical guide for nystagmus diagnostics based on the eye movement pattern found 27 different causes of nystagmus [36]. The steps in the guide included a detection of the main axis of eye movement (horizontal or vertical), the conjugacy, the temporal variability and other characteristics describing, for example, periodic alternating eye movements.

4.2 Cause and different types of nystagmus

Nystagmus symptoms vary between patients, and different oscillatory patterns are linked to different underlying causes [25]. It is important to understand from where nystagmus may originate and why the oscillations occur in order to give the best possible treatment. In this section, a brief overview of some of the conditions leading to nystagmus are presented. Note that the beat direction is defined by the direction of the fast phase [2].

4.2.1 Early onset nystagmus

Early onset nystagmus may be divided into three subcategories: *infantile nystagmus*, *latent* nystagmus and *spasmus nutans* [37]. A brief description of these different categories follows.

Infantile nystagmus syndrome (INS)

Infantile nystagmus syndrome, previously called *congenital nystagmus*, is a form of nystagmus that develops in infancy [38]. About 30 % of patients diagnosed with infantile nystagmus are classified as idiopathic [37], meaning that there is not an established pathology to explain the oscillatory eye movements. Infantile nystagmus syndrome is in a majority of cases horizontal and conjugate and the slow phases are accelerating [38]. Infantile nystagmus syndrome symptoms often increase during attempted fixations, or increased levels of attention or arousal [2]. In some patients with infantile nystagmus syndrome, there is a gaze angle for which the nystagmus intensity is the lowest, which is referred to as the null zone. In some patients with a null zone, the head is turned so that the nystagmus intensity is reduced [39].

Some waveforms found in infantile nystagmus syndrome patients have a signature foveation period, discussed in Section 4.1, and it is likely a main contributor to

why some nystagmus patients have a near normal visual function [2]. If there is no foveation period in the cycle, it is likely that the person suffers from decreased visual function.

Other forms of early onset nystagmus

Latent nystagmus is a form of nystagmus that appears when an eye is occluded. There are two forms of latent nystagmus: latent–latent and manifest–latent nystagmus. In the former case, nystagmus oscillations are observed only when one eye is closed, and the other eye oscillates as a jerk waveform beating away from the closed eye [2]. In manifest–latent nystagmus the oscillations are always present, but the intensity increases when one eye is occluded [37]. For these patients, the intensity increases when the person's gaze is in the direction of the fast phase, which is known as Alexander's law [37]. Latent nystagmus is present in patients since infancy and linked to strabismus [2].

Spasmus nutans is a disconjugate nystagmus with a small amplitude and high frequency (3 Hz – 11 Hz) [2, 37]. In most cases the eye movements are pendular in the horizontal plane, but there have been reports of both vertical and torsional components [2]. The eye movements are developed early in life and usually disappears within a few years after onset [2]. The pathology of these eye movements has not been established [37], although it has been observed in monozygotic twins [2].

4.2.2 Acquired nystagmus

Unlike early onset nystagmus, acquired nystagmus is developed later in life. As described in section 4.1, there are various reasons to why a person develops acquired nystagmus. It has been linked to diseases such as multiple sclerosis and stroke as well as visual loss, hypoxic encephalopathy and oculopalatal tremor [2, 37]. Different characteristics have been linked to different symptoms. For example, the frequency of the oscillations, conjugacy of the eye movements and the direction (horizontal, vertical or torsional) may indicate whether the nystagmus is due to demyelinating diseases or if it is a symptom of oculopalatal tremor [2].

Foveation periods in acquired nystagmus are rarer compared to INS patients, which could explain the higher percentage of reported cases of osciollopsia in acquired nystagmus. Foveation periods are thought to be the reason why many patients with early onset nystagmus are able to perform well on sight tests, and would thus indicate that patients with acquired nystagmus have poorer vision.

4.2.3 Clinical assessment of nystagmus

Historically, nystagmus has been diagnosed and evaluated by medical professionals, such as neuro–ophthalmologists. The diagnostics of nystagmus is complex and often requires long clinical experience. In this section, a brief review of the clinical assessment of nystagmus is presented. Although eye movement recording methods exist, they are not necessarily used in clinical practice [37].

The first step of nystagmus diagnostics is to determine whether the patient experiences blurred or decreased vision [2]. A visual acuity test for both near and far vision is usually performed. If the patient turns their head, and this is observed from old photographs, it indicates a null zone. In order to establish the null zone, observations of the eyes are conducted for various gaze angles. The visual fixation capabilities both in near and far view are established. Latent forms of nystagmus may be diagnosed by observing the eye movements when one eye is occluded. With the use of *Frenzel googles*, the nystagmus oscillations may be observed without sustained fixations [2]. These googles consist of glass with large magnification, inhibiting the wearer to fixate the gaze. This may reveal nystagmus that is not observable if the patient is fixating. The nystagmus eye movements during saccades and smooth pursuit, are also clinically relevant to analyse [2]. For example, gaze evoked nystagmus has been linked to impaired smooth pursuit [2].

4.3 Nystagmus and eye movement recordings

Eye tracking technology has allowed more accurate recordings of eye movements from nystagmus patients compared to human observation. This in turn has enabled more detailed research of the various conditions causing the oscillations compared to the previously used manual inspection of nystagmus eye movements. When using eye tracking to record nystagmus eye movements, there are some important methodological issues to consider, which are discussed in this section.

4.3.1 Calibration

The main issue when using video–based eye trackers for recording nystagmus eye movements is the calibration step. The reason for this is that most calibration protocols assume stable and accurate fixations, which nystagmus patients in general are unable to produce. As discussed in Section 3.1.3, the purpose of the calibration step is to transform uncalibrated data from the eye tracker coordinate system to the stimuli coordinate system, typically a computer screen. There are a number of practical issues that need to be addressed when using video–based eye trackers to record nystagmus eye movements. First of all, nystagmus amplitudes are often greater than 1°, which is

a common accuracy threshold for accepting the calibration results [19]. In some eye tracking software, a calibration protocol is provided, and the software may not accept unstable fixations. This may result in a rejection of all nystagmus recordings, since stable fixations are not available. If present, it is usually desired to detect and use the foveation period for calibration, since this is the segment of the oscillation cycle where the fovea is typically best aligned with the intended fixation target [2]. Strabismus is common among nystagmus patients, and it is recommended to perform calibration for each eye separately, by covering one eye and calibrating the other [40].

There are a few different approaches that have been used to detect the foveation periods in nystagmus data prior to calibration. A manual selection method has been proposed [40], where an operator selects foveation periods and does a post experiment calibration of the data. This method is suitable to use when working with small datasets where it is crucial that the selection method is controlled by the person conducting the experiment. A benefit is that it is not necessary to write an algorithm to handle all types of waveform morphologies. The problem with manual selection of foveation periods is that it is time consuming, inconsistent and requires training of the operator performing the data selection. Manual annotation makes it impractical to use real time experiment paradigms, i.e., change the stimuli based on the participants eye movements, and makes it difficult to compare results between different studies, since the foveation period selection may be different.

Another approach is to use an algorithm to detect the foveation periods. The advantages of this are that the method is consistent, and that it may be programmed in advanced, allowing for online calibration. The method may be distributed to researchers in other labs, which makes it easier to compare results between different studies. The main issue with algorithm–based methods is that there are many different forms of nystagmus eye movement patterns. If it is of interest to use an algorithm for all nystagmus types, the variance in waveform morphology when constructing the algorithm needs to be considered.

Various methods have been constructed for automatic detection of foveation periods. Most algorithms use position, velocity, or acceleration thresholds or some combination of all three to detect foveation periods. For example, the extended nystagmus acuity function (NAFX) [41] uses both position and velocity for the detection. A similar approach is called the nystagmus optimal fixation function (NOFF), that was developed to handle cases where the NAFX was not reliable, e.g., when recording children [42]. Both of these approaches rely on calibrated eye velocity and acceleration data. Since the purpose of the calibration process is to get calibrated data, they are not suitable to be used in a calibration algorithm.

In order to obtain foveations periods in uncalibrated data, a method was developed using adaptive thresholds instead of absolute thresholds [29]. This method is based on the division of the oscillation cycles into slow and fast phases, and then to

identify the segment of the slow phase that has the lowest velocity. This serves as a proxy for the foveation period. For each calibration target, the median of each of these foveation periods is computed, which results in a coordinate pair for each calibration target.

An additional problem, other than the eye movements themselves, when calibrating nystagmus patients, is head movements. As described in Section 4.1, some patients have a null zone, where the nystagmus intensity is at a minimum. For these patients, it is common to turn the head so that the intensity is at a minimum when looking at an object. If multiple fixation targets are presented during an experiment, there is a risk of head movements, which may distort the eye movement data. It is therefore common to use a chin and head–rest during recordings, so that the head is stabilised as much as possible.

4.4 Characterisation of Nystagmus

As has been discussed before in this thesis, there are various ways that nystagmus may express itself. Nystagmus characterisation is an important topic in this thesis and some fundamental concepts are discussed in this section. For example, the amplitude, frequency and waveform morphologies of each person with nystagmus may depend on the gaze angle, various treatments for the condition or how far the symptoms have progressed. One case study reported monthly waveform morphology changes in an infant from no apparent nystagmus oscillations, to square wave jerks, to different forms of nystagmus oscillations [43]. For the intended purpose of this work, characterisation refer to a description of the current eye movements of a nystagmus patient based on *spatial* and/or *temporal* properties of the oscillations. The spatial characteristics refers to the oscillation properties as a function of the gaze angle. The temporal characteristics may be divided into two different subcategories: short–term and long–term. Short term characterisation refers to how the different nystagmus properties change over a short timeframe, i.e., during the recording of one fixation target. This is interesting when studying nystagmus conditions that have a time dependent component, such as periodic alternating nystagmus. Long–term temporal characteristics refers to the nystagmus symptoms between recordings, i.e., when performing the same task separated by some time period, e.g., a week, a month or a year. This may be used when studying the development of nystagmus in infancy or acquired forms of nystagmus as well as during treatment or postoperative evaluation of a surgical procedure.

The characterisation may be important for many different reasons. First, characterisation allows monitoring of the development of nystagmus for a specific patient. In order to study the effects of a drug treatment, it is crucial to know what the eye

movements looked like before the treatment. Without characterising or quantifying the eye movements before starting the treatment, it is not possible to measure the effects of the treatment. Second, characterisation is also of interest since it allows a comparison between patients. For example, if two patients have similar symptoms and similar history, it is possible that a treatment that was effective for the first patient would also be effective for treating the second patient. By studying the responses from both patients when exposed to the same treatment, it may be possible to improve existing treatment procedures or find new ones.

In order to illustrate the variation in spatial and short–term temporal characteristics, example signals are presented in Figure 4.4. The figure illustrates data taken from two different participants, both diagnosed with early onset nystagmus. The gray signals (two upper signals) represent data recorded from one person looking at one fixation target, 8° above the primary position in the stimuli coordinate system. These two segments are separated by approximately ten seconds. The gray signals illustrate where a reversal of the beat direction can be observed. The black signals (two lower signals) are recorded from a different participant looking at two different fixation targets. The bottom black signal is taken from a patient looking left and the signal above is recorded from the primary position. As can be seen, these signals do not only have different beat directions, but the waveforms are completely different.

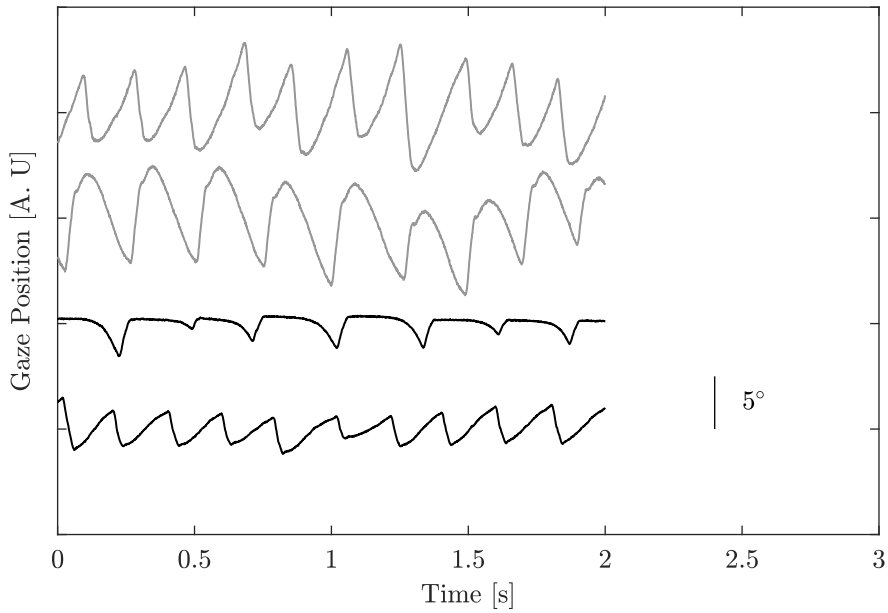


Figure 4.4: Illustration of the spatial and temporal variance of nystagmus waveform morphologies from horizontal eye movements. The two gray waveforms are taken from the same participant looking at the same fixation target 10 seconds apart. The two black waveforms are taken from another participant looking at different fixation targets. In both cases, it is obvious that the waveforms are different, i.e., the eye movement pattern has changed. The black bar to the right of the traces spans 5°.

4.5 Treatment of Nystagmus

Thus far, this chapter has focused on presenting methods for analysing nystagmus eye movements. The application of this analysis is intended to improve diagnostics and treatment of nystagmus symptoms. In this section, a brief review of different methods used to treat nystagmus symptoms is provided.

In order to analyse the outcome of a treatment, it is essential to measure improvements of the nystagmus oscillations. Some common measures include improvement in the visual function of patients [27], nystagmus amplitude, frequency and intensity analysis [27] and the range and angle of the null zone [44]. The Anderson-Kestenbaum procedure is aimed at changing the angle of and broaden the null zone, and the method was independently developed by Anderson and Kestenbaum in 1953 and 1954 [39], respectively. The surgical procedure is performed by reattaching the

extra ocular muscles in order to shift the null zone to the primary position. The null zone and visual acuity is measured preoperative for varying gaze angles in order to determine the condition of the patient [2]. A similar surgery is called *tenotomy and reattachment procedure*, and it has been performed in order to broaden the null zone of patients with both acquired and infantile nystagmus [45, 46].

For a number of nystagmus conditions, pharmacological treatment of nystagmus is a good option. The choice of drugs is dependent on the type of nystagmus [45], and it is therefore crucial to correctly diagnose each individual patient. Both Memantine and Gabapentin have resulted in improved visual acuity and decreased nystagmus intensity for patients with both infantile and acquired nystagmus [27, 47, 48]. An intensity reduction in acquired forms of nystagmus has been observed after administration of Baclofen [2]. In patiens with down- or upbeat nystagmus, 4-aminopyridine has been proven to surpress the nystagmus symptoms [45].

4.6 Modelling of Nystagmus waveforms

As has been illustrated so far in this thesis, there are a plethora of nystagmus waveforms, and it is important to identify each specific eye movement pattern in each individual in order to offer the best diagnostics and treatment. This is not a trivial task, but it may be aided by the use of waveform modelling methods. Below follows a brief review of various modelling methods for nystagmus waveforms found in the literature.

The purpose of the modelling methods is to parametrise the observed eye movement patterns. This could then be used to assert the signal quality [35], comparing and clustering data [49], or to reconstruct various nystagmus waveforms, which is useful for simulations. The modelling methods are divided based on their function into two categories: *system*–based and *signal*–based [35]. System–based methods aim to understand the underlying neurological process that is responsible for generating the nystagmus eye movement patterns. Some forms of early onset nystagmus have been simulated using a model where a deficit in the saccadic system was introduced [50, 51], and non–linear dynamic models were used to simulate the *neural integrator*. This is a function of the central nervous system that is related to the gaze stability in eccentric eye positions [2]. A model based on control theory was used to analyse the origin of INS waveforms, shifts in the null zone and the presence of oscillopsia [52].

The signal–based methods deal with characterisation and classification of measured eye movement signals. These methods include, e.g., amplitude and frequency estimation [29], morphological classification [49] and determination of the visual function of a patient [41]. Some examples of signal–based modelling methods include linear combination of sine and saw–tooth waves [49], harmonic series analysis

[53] or NAFX [41]. In the first two examples, the goal of the modelling is to estimate a set of unknown parameters, which may be used to classify, e.g., waveform morphology. The NAFX metric may be used more directly as a measure of the development of symptoms, or as an evaluation tool for a treatment.

Chapter 5
Summary of the included papers

In this chapter, each of the four included papers is summarised. In the first paper, a protocol for calibrating nystagmus data is proposed. This method provides users with the possibility to calibrate nystagmus data, which has proven to be a challenge when using eye trackers for nystagmus research. The second paper introduces a model for nystagmus waveforms, which is used to describe different nystagmus waveforms and to measure the quality of recorded data. The method presented in this paper serves as a basis for the analysis in paper III and IV. A waveform clustering method applied to fixation and smooth pursuit data is developed in Paper III. A distance metric referred to as the *waveform distance index* is defined and used in the clustering process. All waveforms in both fixation and smooth pursuit data are labelled, and the clustering performance is evaluated. In the fourth and final paper, a method aimed at characterising nystagmus waveforms is evaluated against manual expert annotations. Together the proposed methods constitute a complete framework for analysis of nystagmus waveforms from calibration to cycle–to–cycle characterisation and comparison between recorded waveforms during different tasks. A brief summary of each paper follows below.

Paper I: A Robust Method for Calibration of Eye Tracking Data Recorded During Nystagmus

A novel protocol for calibrating eye tracker data for nystagmus patients is proposed and evaluated. The method consists of an outlier correction algorithm (OCA) constructed to address the uncertainty during the calibration caused by the nystagmus oscillations. Different calibration mapping functions are differently sensitive to outliers, which may affect the resulting waveform after calibration. An example of different gaze estimations obtained using different mapping functions is shown in Figure 5.1. The $\mathcal{G}$ polynomial is quadratic, $\mathcal{A}_1$ is linear and $\mathcal{A}_4$ is a fourth order polynomial. As may be observed in the figure, they lead to different gaze estimations. This is an

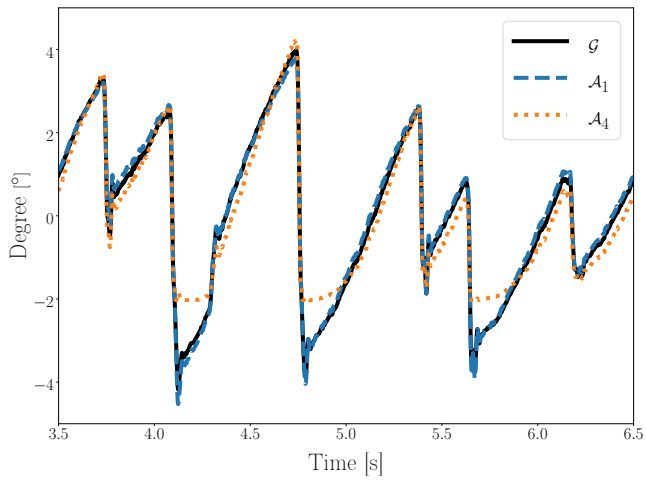


Figure 5.1: PoR Estimation Examples. Illustration of three different PoR polynomial estimations.

issue, since it is not possible to know which one is most correct. The OCA utilises a predetermined model of the distribution of the calibration data in order to detect and adjust data points which are likely to be outliers. A Procrustes transformation model for calibration of data from the eye tracker coordinate system to the stimuli coordinate system is evaluated and compared to four different mapping functions, which previously have been used for calibration of nystagmus data.

Three different metrics were used for evaluation: accuracy, *calibration plane distortion* and *waveform robustness*. The calibration plane distortion is a measurement of

the suitability to use different mapping functions for data transformation from the eye tracker coordinate system to the stimulus coordinate system. The waveform robustness metric is related to distortions caused by these five mapping functions on the nystagmus waveform itself. The evaluation data for this method consisted of two different datasets where the first, nystagmus data experiment (NDE), consisted of six recordings from five nystagmus patients. The second dataset, control data experiment (CDE), consisted of seven healthy controls recorded once each. The outlier correction algorithm was evaluated using the NDE dataset in order to study improvements in accuracy. The calibration plane distortion and the waveform robustness were tested using the CDE dataset. The calibration protocol for the nystagmus participants consisted of nine calibration targets and four validation targets, where the four validation targets were used to assess the accuracy of the calibration. For the CDE dataset, the calibration protocol consisted of 81 targets, where eight out of nine targets were introduced with either a horizontal or vertical offset, simulating the uncertainty when recording a participant with nystagmus. Optokinetic nystagmus data were also recorded for these participants. The optokinetic data were estimated using either a mapping function fitted to calibration data with no offset or to calibration data with a random horizontal offset. The calibration plane distortion and the waveform robustness were evaluated using the difference between the OKN data estimated with no offset and the OKN data estimated with the random horizontal offsets.

Looking at the results, the accuracy of the nystagmus recordings is improved by the OCA, and four out of the five mapping functions generates accuracy results that are similar to one another. However, when examining the the calibration plane distortion and waveform robustness, the proposed Procrustes calibration method outperforms the other mapping functions in terms of waveform robustness and achieves similar results when comparing the calibration plane distortion.

To conclude, the outlier correction algorithm improves the accuracy of the nystagmus recordings, and it is preferable to use the linear Procrustes mapping function in favour of the previously suggested mapping functions.

Paper II: Modelling and Quality Assessment of Nystagmus Eye Movements Recorded Using an Eye-Tracker

In this work, a sinusoidal model is developed for modelling of nystagmus waveform morphologies and for signal quality assessment. Each recording is divided into shorter segments, and each segment is modeled using a harmonic sinusoidal series. By splitting the signal into shorter segments, it is possible for the model to detect local changes in amplitude, frequency and waveform morphology. For the evaluation, both simulated signals reconstructed from a set of template nystagmus waveforms, and patient recordings are used. A metric to assess the signal quality, the *normalised segment error* (NSE), is constructed in order to separate recording artefacts and other disturbances from nystagmus oscillations. The NSE is defined using the mean normalised energy in the error between the observed signal and the modelled reconstructed signal. The evaluation of the NSE was performed using the simulated template signals. Noise, frequency changes, amplitude modulation and waveform changes were introduced into these signals, and the aim of the optimisation of the NSE was to detect the segments where no nystagmus oscillations are present. The nystagmus participant dataset, PD, was constructed using two subsets: one where five different participants were recorded, and one consisting of multiple recordings (n=6) from the same participant. All participants were asked to view a sequence of five fixation targets with a duration between 15 and 35 seconds for each target.

In order to separate the nystagmus oscillations from noise, a threshold NSE value, ϵ, was estimated by utilising receiver operating characteristics (ROC). The ROC curve is optimised by computing the shortest distance to the (0, 1) position in the False Positive - True Positive coordinate system, where the optimisation resulted in $\epsilon = 0.18$. All segments with a normalised segment error above 0.18 are considered noise and all other segments are considered as nystagmus oscillations.

The ROC analysis of the simulated template signals resulted in a true positive and false positive rate of 0.97 and 0.07 respectively. These results are shown in Figure 5.2. The model parameters from the signal segments classified as noise did not match the expected reference values of the template signals, which indicates that the signals in this class do not represent the corresponding template waveforms. The model parameters of the segments with a normalised segment error ≤ 0.18, however, were in good agreement with the expected template waveform values. When analysing the repeated participant recordings, repeated waveform parameterisation is observed, which is according to expectations. From these results, it was concluded that the presented method may be used to analyse and model nystagmus waveforms, as well as detect segments in the data that are unlikely to be nystagmus oscillations.

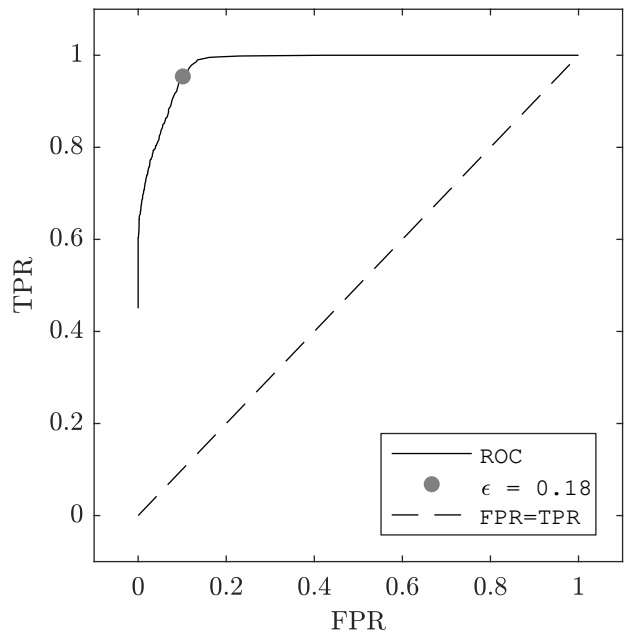


Figure 5.2: Receiver Operating Characteristics. The false positive rate (FPR) is plotted against the true positive rate (TRP) for various values of the error threshold. Setting ϵ to 0.18 results in a false positive rate of 0.07 and a true positive rate of 0.97 for the simulated signals.

Paper III: Analysis of Fixation and Smooth Pursuit Eye Movements from Nystagmus Participants using Waveform Clustering

In this paper, a method for comparison of waveforms within and between fixation and smooth pursuit data is proposed. Each waveform is modeled using the normalised waveform model developed in Paper II. Signal features are derived from this model describing time-varying properties in the nystagmus waveform. A new model is estimated for every 0.5 s. A clustering procedure is employed to identify frequently occurring predefined waveform morphologies within each recording.

Before clustering, the model parameters are normalised in amplitude, frequency and phase using the method described in Paper II. This is performed in order to make waveform models in different segments and signals comparable. The waveform in each signal segment is represented by a vector $\boldsymbol{\theta}$, consisting of eight features. The proposed method classifies each observed waveform into one of eleven possible waveform classes, which have been derived from a subset of the template waveforms described in [34]. The classification is based on a metric called the *waveform distance index*, which is introduced in order to allow comparison in terms of similarity between two different $\boldsymbol{\theta}$ vectors. This analysis results in a morphological profile showing the prevalence of different classes and a morphological trend showing when different classes are detected.

A method called *signal segment matching* is introduced in order to compare morphological trends between different signals. In this work, this method is used to compare smooth pursuit eye movement signals to fixation signals at different gaze angles. A total of six participants diagnosed with nystagmus and nine healthy controls were recorded. Some participants were recorded multiple times, and in total 14 nystagmus and 17 control recordings were used in the evaluation. One participant diagnosed with nystagmus was repeatedly recorded over the course of several months. Both fixation and smooth pursuit eye movements were recorded for all nystagmus patients and all controls. In order to validate the smooth pursuit analysis, simulated nystagmus smooth pursuit data were generated by superimposing fixation data from nystagmus patients onto smooth pursuit data from healthy controls.

The proposed method was evaluated by studying recordings from different nystagmus patients with different eye movement patterns, as well as by studying repeated recordings from the same participant. The results show that the proposed method is able to discriminate between different eye movement patterns for different patients during fixations. Repeated recordings from the same participant were classified into the same template class.

When studying the signal segment matching between repeated recordings of smooth

pursuit data, the results show that segments recorded when performing the same movement task are matched to each other. When comparing smooth pursuit data to fixation data, segments with similar waveforms are matched to each other. An example of smooth pursuit data matched to different fixation signals is presented in Figure 5.3 The blue sections of the smooth pursuit signal (top) is matched to the

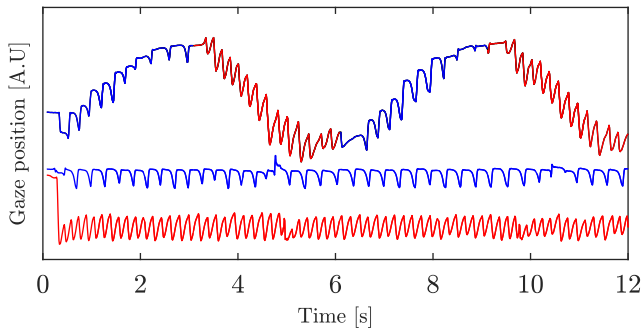


Figure 5.3: Signal segment matching between a smooth pursuit signal (top) and two different fixation signals (bottom). The blue segments of the smooth pursuit signal are matched to fixation signal in the middle, and the red segments are matched to the fixation signal at the bottom.

blue fixation signal (middle), and the red sections of the smooth pursuit signal are matched to the red fixation signal (bottom). Overall, the proposed method is a useful explorative tool for the investigation of eye movement patterns during smooth pursuit movements.

Paper IV: Cycle-to-Cycle Nystagmus Waveform Characterisation

In this work, a method for cycle–to–cycle waveform classification of nystagmus eye–tracker signals is proposed, where each nystagmus cycle is analysed independently. The purpose of the method is to perform computer guided waveform classification of nystagmus signals, which can be used as a complement to manual annotation of nystagmus waveforms and as decision support for diagnostics of nystagmus.

The cycle–to–cycle classification consists of two parts: The first part is referred to as *morphological classification*, and is used to analyse the overall shape of each cycle. The morphological classification is based on principle component analysis of the 18 template waveforms presented in [34]. Ten principle components are used, as they describe more than 99 % of the variability of the nystagmus waveform morphologies. The second part is called *cycle feature analysis*, and is used to detect characteristic nystagmus features such as foveation periods and breaking saccades. The analysis of cycle features and the morphological classification are combined into a waveform class, which is used to describe the features in each cycle.

Two datasets were used in the evaluation of the proposed method. The first dataset was recorded using magnetic scleral search coils, where each recording was manually annotated by the authors of the study [22]. It should be noted that the annotations apply to the entire recordings, and not to individual cycles. The second dataset was recorded using a video–based eye tracker, and for this dataset, no annotations exist. A total of 16 patients, 8 patients from the first dataset and 8 patients from the second dataset, were included in the study. All included data were recorded in the primary gaze position.

The manual waveform annotations were used to evaluate the performance of the proposed method. Here, true positives are defined as waveforms where the waveform class from the proposed method matches the manual waveform annotation, and false positives are defined as waveforms for which detected waveform classes are not listed in the annotation for that signal. In order to compensate for that annotations are applied to the entire recording and not to individual cycles, unannotated and false positive waveforms are compared in terms of similarity to true positive detections of the same waveform class in other signals. In this way, it can be shown that false positive or unannotated detections match annotated nystagmus waveforms on a cycle level. Here, the waveform distance index proposed in Paper III is used to measure similarity between waveforms. If the waveform distance index between the average waveform of the unannotated or false positive waveforms and the corresponding true positive average waveform(s) in another signal is sufficiently small, the unannotated or false positive waveform is considered to be correctly detected.

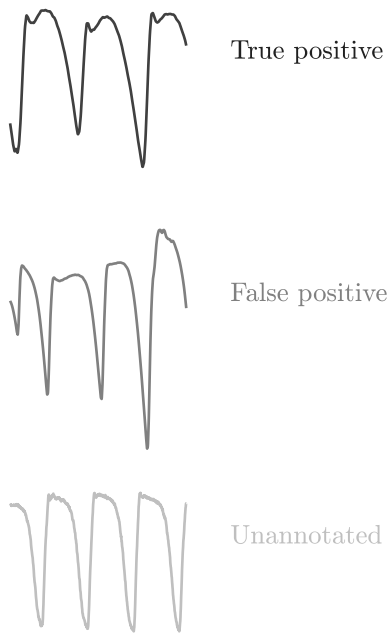


Figure 5.4: Example of classification for the extended foveation right waveform. The black waveform (top signal) is a true positive, the gray waveform (middle signal) is a false positive and the light gray waveform (bottom signal) is unannotated. These cycles come from three different signals, and they have all been assigned to the extended foveation right (EF-R) waveform class.

These results show that there is a high similarity between the detected waveforms classes in different recordings and patients. This is illustrated in Figure 5.4, where true positives, false positives and unannotated extended foveation right (EF-R) waveforms are shown. Overall, the results show that the proposed method can be used as a complement to manual annotations when studying nystagmus eye movements.

References

[1] E. N. Marieb and K. Hoehn, *Human anatomy & physiology*. Pearson Education, 2007.

[2] R. J. Leigh and D. S. Zee, *The neurology of eye movements*. Oxford University Press, USA, 2015.

[3] N. Karbach, N. Kobrenko, M. Myers, and A. S. Gurwood, "How drug abuse affects the eye," *Review of Optometry*, vol. 155, no. 9, pp. 80–89, 2018.

[4] N. Sarvananthan, M. Surendran, E. O. Roberts, S. Jain, S. Thomas, N. Shah, F. A. Proudlock, J. R. Thompson, R. J. McLean, C. Degg, *et al.*, "The prevalence of nystagmus: the leicestershire nystagmus survey," *Investigative ophthalmology & visual science*, vol. 50, no. 11, pp. 5201–5206, 2009.

[5] I. Gottlob and F. A. Proudlock, "Aetiology of infantile nystagmus," *Current opinion in neurology*, vol. 27, no. 1, pp. 83–91, 2014.

[6] R. J. McLean, K. C. Windridge, and I. Gottlob, "Living with nystagmus: a qualitative study," *British Journal of Ophthalmology*, vol. 96, no. 7, pp. 981–986, 2012.

[7] L. Larsson, *Event Detection in Eye-Tracking Data for Use in Applications with Dynamic Stimuli*. PhD thesis, Lund University, 2016.

[8] A. J. Vander, J. H. Sherman, and D. S. Luciano, *Human physiology: the mechanisms of body function*. New York, US: McGraw-Hill, Sixth Edition, 1994.

[9] Wikimedia-Commons, "*The Human Eye.*," 2007.

[10] Wikimedia-Commons, "*Photoreceptor Distribution.*," 2019.

[11] Wikimedia-Commons, "*Optical Fields*.", 2013.

[12] A. C. Ireland and I. B. Carter, "Neuroanatomy, optic chiasm," in *StatPearls [Internet]*, StatPearls Publishing, 2019.

[13] J. R. Naegele and R. Held, "The postnatal development of monocular optokinetic nystagmus in infants," *Vision Research*, vol. 22, no. 3, pp. 341–346, 1982.

[14] U. Büttner and O. Kremmyda, "Smooth pursuit eye movements and optokinetic nystagmus," in *Neuro-Ophthalmology*, vol. 40, pp. 76–89, Karger Publishers, 2007.

[15] R. S. Hessels, D. C. Niehorster, M. Nyström, R. Andersson, and I. T. Hooge, "Is the eye-movement field confused about fixations and saccades? a survey among 124 researchers," *Royal Society open science*, vol. 5, no. 8, p. 180502, 2018.

[16] N. Wade, B. W. Tatler, *et al.*, *The moving tablet of the eye: The origins of modern eye movement research*. Oxford University Press, USA, 2005.

[17] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer, *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford, 2011.

[18] L. R. Young and D. Sheena, "Survey of eye movement recording methods," *Behavior research methods & instrumentation*, vol. 7, no. 5, pp. 397–429, 1975.

[19] D. W. Hansen and Q. Ji, "In the eye of the beholder: A survey of models for eyes and gaze," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 478–500, 2010.

[20] Wikimedia-Commons, "*Visible light eye-tracking algorithm*.," 2015.

[21] H. J. Wyatt, "The form of the human pupil," *Vision research*, vol. 35, no. 14, pp. 2021–2036, 1995.

[22] L. Dell'Osso, J. Jacobs, and A. Serra, "The sub-clinical see-saw nystagmus embedded in infantile nystagmus," *Vision research*, vol. 47, no. 3, pp. 393–401, 2007.

[23] SR-Research, "*EyeLink 1000 User Manual*.," 2010.

[24] J. W. Peirce, "Psychopy psychophysics software in python," *Journal of Neuroscience Methods*, vol. 162, no. 1-2, pp. 8–13, 2007.

[25] R. B. Daroff and M. J. Aminoff, *Encyclopedia of the neurological sciences*. Academic press, 2014.

[26] N. Hussain, "Diagnosis, assessment and management of nystagmus in childhood," *Paediatrics and Child Health*, vol. 26, no. 1, pp. 31–36, 2016.

[27] R. McLean, F. Proudlock, S. Thomas, C. Degg, and I. Gottlob, "Congenital nystagmus: Randomized, controlled, double-masked trial of memantine/gabapentin," *Annals of Neurology*, vol. 61, no. 2, pp. 130–138, 2007.

[28] R. V. Abadi, "Mechanisms underlying nystagmus," *Journal of the royal society of medicine*, vol. 95, 2002.

[29] M. Dunn, *Quantifying perception and oculomotor instability in infantile nystagmus*. PhD thesis, Cardiff University, 2014.

[30] W. Rosengren, M. Nyström, B. Hammar, and M. Stridh, "A robust method for calibration of eye tracking data recorded during nystagmus," *Behavior research methods*, pp. 1–15, 2019.

[31] G. Pasquariello, M. Cesarelli, M. Romano, A. La Gatta, P. Bifulco, and A. Fratini, "Waveform type evaluation in congenital nystagmus," *Computer Methods and Programs in Biomedicine*, vol. 100, no. 1, pp. 49–58, 2010.

[32] R. V. Abadi and C. M. Dickinson, "Waveform characteristics in congenital nystagmus," *Documenta Ophthalmologica*, vol. 64, no. 2, pp. 153–167, 1987.

[33] N. Sheth, L. Dell'Osso, R. Leigh, C. Van Doren, and H. Peckman, "The effects of afferent stimulation on congenital nystagmus foveation periods," *Vision research*, vol. 35, no. 16, pp. 2371–2382, 1995.

[34] L. Dell'Osso and R. Daroff, "Congenital nystagmus waveforms and foveation strategy," *Documenta Ophthalmologica*, vol. 39, no. 1, pp. 155–182, 1975.

[35] W. Rosengren, M. Nyström, B. Hammar, M. Rahne, L. Sjödahl, and M. Stridh, "Modeling and quality assessment of nystagmus eye movements recorded using an eye-tracker," *Behavior Research Methods*, pp. 1–15, 2020.

[36] L. Frisén, "Nystagmus lathund," *Läkartidningen*, vol. 78, no. 22, pp. 2229–2232, 1981.

[37] M. Dunn, "Clinical assessment of nystagmus," *Optometry Today*, vol. 56, no. 7, pp. 80–85, 2016.

[38] R. W. Hertle and S.-O. M.-B. Vision, "A classification of eye movement abnormalities and strabismus (cemas)," in *Report of a National Eye Institute Sponsored Workshop from the committee for the Classification of Eye Movement Abnormalities and Strabismus (CEMAS) Workshop*, pp. 1–56, 2001.

[39] J. Lee, "Surgical management of nystagmus," *Journal of the Royal Society of Medicine*, vol. 95, no. 5, pp. 238–241, 2002.

[40] L. Dell'Osso, "*Recording and Calibrating the Eye Movements of Nystagmus Subjects*. omlab report 011105, 1-4, 2005."

[41] L. Dell'Osso and J. Jacobs, "An expanded nystagmus acuity function: intra-and intersubject prediction of best-corrected visual acuity," *Documenta Ophthalmologica*, vol. 104, no. 3, pp. 249–276, 2002.

[42] J. Felius, V. L. Fu, E. E. Birch, R. W. Hertle, R. M. Jost, and V. Subramanian, "Quantifying nystagmus in infants and young children: relation between foveation and visual acuity deficit," *Investigative ophthalmology & visual science*, vol. 52, no. 12, pp. 8724–8731, 2011.

[43] I. Gottlob, "Infantile nystagmus. development documented by eye movement recordings.," *Investigative ophthalmology & visual science*, vol. 38, no. 3, pp. 767–773, 1997.

[44] L. Dell'Osso, "Extraocular muscle tenotomy, dissection, and suture: a hypothetical therapy for congenital nystagmus," *J Pediatr Ophthalmol Strabismus*, vol. 35, no. 4, pp. 232–233, 1998.

[45] M. J. Thurtell and R. J. Leigh, "Treatment of nystagmus," *Current treatment options in neurology*, vol. 14, no. 1, pp. 60–72, 2012.

[46] R. W. Hertle, L. F. Dell'Osso, E. J. FitzGibbon, D. Yang, and S. D. Mellow, "Horizontal rectus muscle tenotomy in children with infantile nystagmus syndrome: a pilot study," *Journal of American Association for Pediatric Ophthalmology and Strabismus*, vol. 8, no. 6, pp. 539–548, 2004.

[47] T. Shery, F. A. Proudlock, N. Sarvananthan, R. J. McLean, and I. Gottlob, "The effects of gabapentin and memantine in acquired and congenital nystagmus: a retrospective study," *British journal of ophthalmology*, vol. 90, no. 7, pp. 839–843, 2006.

[48] M. J. Thurtell, A. C. Joshi, A. C. Leone, R. L. Tomsak, G. S. Kosmorsky, J. S. Stahl, and R. J. Leigh, "Crossover trial of gabapentin and memantine as treatment for acquired nystagmus," *Annals of neurology*, vol. 67, no. 5, pp. 676–680, 2010.

[49] M. Theodorou and R. Clement, "Classification of infantile nystagmus waveforms," *Vision Research*, vol. 123, pp. 20–25, 2016.

[50] O. Akman, D. Broomhead, R. Abadi, and R. Clement, "Eye movement instabilities and nystagmus can be predicted by a nonlinear dynamics model of the saccadic system," *Journal of mathematical biology*, vol. 51, no. 6, pp. 661–694, 2005.

[51] D. S. Broomhead, R. A. Clement, M. R. Muldoon, J. P. Whittle, C. Scallan, and R. V. Abadi, "Modelling of congenital nystagmus waveforms produced by saccadic system abnormalities," *Biological cybernetics*, vol. 82, no. 5, pp. 391–399, 2000.

[52] C. M. Harris, "Problems in modelling congenital nystagmus: towards a new model," in *Studies in Visual Information Processing*, vol. 6, pp. 239–253, Elsevier, 1995.

[53] R. Abadi and R. Worfolk, "Harmonic analysis of congenital nystagmus waveforms," *Clinical vision sciences*, vol. 6, no. 5, pp. 385–388, 1991.

Part II

Included Papers

Paper I

A Robust Method for Calibration of Eye Tracking Data Recorded during Nystagmus

Abstract

Eye tracking is a useful tool when studying the oscillatory eye movements associated with nystagmus. However, this oscillatory nature of nystagmus is problematic during calibration since it introduces uncertainty about where the person is actually looking. This renders comparisons between separate recordings unreliable. Still, the influence of the calibration protocol on eye movement data from people with nystagmus has not been thoroughly investigated.

In this work, we propose a calibration method using Procrustes analysis in combination with an outlier correction algorithm, which is based on a model of the calibration data and on the geometry of the experimental setup. The proposed method is compared to previously used calibration polynomials in terms of accuracy, calibration plane distortion and waveform robustness.

Six recordings of calibration data, validation data and optokinetic nystagmus data from people with nystagmus and seven recordings from a control group were included in the study. Fixation errors during the recording of calibration data from the healthy participants were introduced, simulating fixation errors caused by the oscillatory movements found in nystagmus data.

The outlier correction algorithm improved the accuracy for all tested calibration methods. The accuracy and calibration plane distortion performance of the Procrustes analysis calibration method were similar to the top performing mapping functions for the simulated fixation errors. The performance in terms of waveform robustness was superior for the Procrustes analysis calibration compared to the other calibration methods. The overall performance of the Procrustes calibration methods was best for the datasets containing errors during the calibration.

William Rosengren, Marcus Nyström, Björn Hammar, and Martin Stridh,
"A Robust Method for Calibration of Eye Tracking Data Recorded during Nystagmus," in. *Behavior research methods*, pp. 1 - 15, 2019

1 Introduction

Eye tracking is a useful tool to record and study eye movements. However, the nystagmus eye movements disturb the calibration procedure for individual recordings, causing comparisons of waveforms between recordings unreliable. For example, the calibration protocol assumes an ability to fixate the gaze, which is limited in people with nystagmus. Using the default calibration protocol may lead to unreliable eye tracker data, which in turn may misrepresent or even invalidate data analysis. In this paper, we explore the problems associated with calibration and propose a method that secures a repeatable and reliable gaze estimation, referred to as *point-of-regard* (PoR), which is crucial for detailed computer based nystagmus diagnostics and objective evaluation of treatment effects between recordings.

1.1 Description of nystagmus

Nystagmus could be a symptom of an underlying oculomotor disorder, which causes involuntary movements of the eye(s) and the condition may lead to decreased visual acuity [1, 2]. There are two broad types of nystagmus: *early-onset nystagmus* and *acquired nystagmus* [2, 3], where the former condition is developed in the months after birth and the latter is developed later in life [4]. The eye movement pattern, sometimes referred to as a *waveform*, can be classified into different categories and there are at least 12 different types of nystagmus waveforms according to a classification study [2, 5, 6].

Different treatments strategies, for instance drug treatment [3] and surgery [7], have been suggested to improve the visual acuity in people with nystagmus. In order to evaluate the different strategies, eye movements before and after the treatment can be studied. Treatment effects are difficult to asses in detail without an objective evaluation of the eye movements, since people with nystagmus are often considered to be hard to diagnose by clinicians [2].

Nystagmus can also be found in visually healthy subjects. *Optokinetic nystagmus* (OKN) is a reflex found in humans [8], which causes oscillatory eye movements similar to the oscillations found in some forms of nystagmus such as pure jerk nystagmus. It can easily be elicited by keeping the head still in a moving environment [8].

1.2 Calibration of a camera based eye tracker

Nystagmus eye movements can be studied in detail with the use of an *eye tracker*. The video-based eye tracker, referred to as video-oculography (VOG) [9], records eye movements using eye images captured by an infrared camera. The data from the VOG system are in this work obtained by finding the pupil center (PC) and the reflection

off the cornea caused by an infrared illuminator, called the *corneal reflection* (CR). The vector between the PC and CR positions is a measure called the *pupil-corneal reflection vector* (PCRV). This measure is unique for each eye orientation and can therefore be used to estimate the PoR. In order to do this estimation from the PCRV, a relationship between the PCRV data and the corresponding PoR data is needed. The process to identify this relationship is referred to as *calibration*, which is dependent on the geometry of the experiment as well as the individual eye anatomy of each participant [9].

The goal of the calibration is to find a *mapping function* (MF), e.g. a polynomial, which describes the relationship between the PCRV data and the PoR data. By presenting targets at known positions during an experiment, referred to as *calibration targets*, and simultaneously recording the corresponding PCRV data, it is possible to estimate the mapping function parameters. The number of calibration targets can vary, but common choices are 2, 5, 9, 13 and 16 targets [9].

The structure of the mapping function needs to be determined before its parameters can be estimated. The selection of the structure is difficult and the choice may significantly affect the resulting PoR estimation. This is illustrated in Figure 1, where three different polynomial structures are used to estimate the same eye movement. The PoR estimations are not the same, which means that one would have to decide which of these is most likely to represent the actual eye movement.

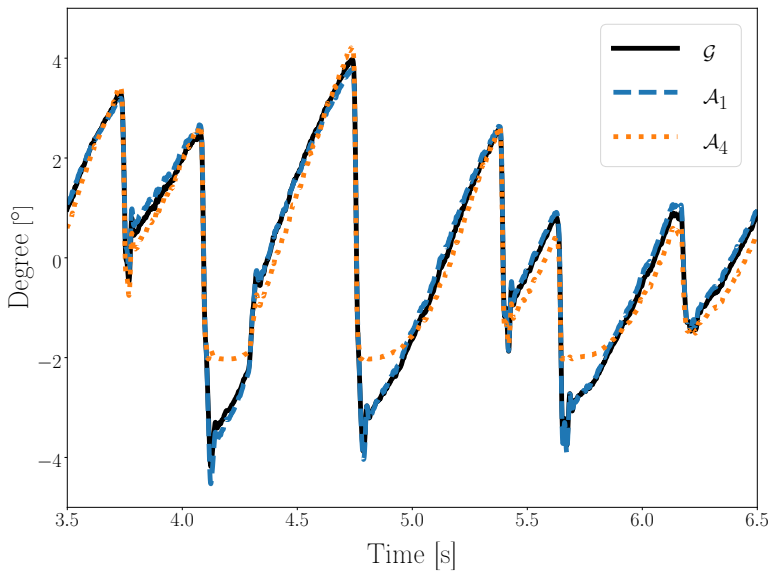


Figure 1: PoR Estimation Examples. Illustration of three different PoR polynomial estimations (see [10, 3, 11] and equations (5), (8) and (7)), of the same recorded PCRV data. The data is obtained from a healthy participant viewing an OKN-stimulus, generating oscillatory eye movements. When comparing the three PoR estimations, it is apparent that $\mathcal{A}_4$ is different from the other two.

1.3 Previous work

Several calibration polynomials for video-based eye tracking have previously been studied. One study investigated more than 400,000 polynomials and evaluated their performance based on the *average error* (accuracy), *maximum error*, *standard deviation* of the estimated PoR, *number of polynomial parameters* and *head movement tolerance* [12]. Another study tested polynomial structures based on accuracy and the number of calibration targets [13]. The two studies were using simulated data or data from participants with no visual impairments. In both [12] and [13], *accuracy* was used to evaluate the calibration MFs. As is pointed out perfect accuracy, or goodness of fit, can be achieved by using the same model order as the number of calibration targets [13]. The calibration polynomial is, however, used also for other gaze positions and should be tested also for these [13].

Table 1: Summary of nine different studies, their calibration and validation protocols, the calibration methods and the calibration method performance. All studies used an EyeLink system (except [3] which used the SMI EyeLink) and were concerned with technical, analytical or clinical applications of eye tracking and nystagmus. The columns represent from left to right: the reference to the study (with number of citations as of December 7, 2017), the eye tracker used in the study, the calibration target positions, the calibration data selection strategy, the calibration polynomial structure (note that no explicit equations are written), the validation protocol and the reported data quality. The calibration data strategy includes manually selection (M), automatic selection (A) or not explicitly stated (U).

Study	System	Calibration Positions	Data Selection Method	Calibration Polynomial	Validation	Reported Data Quality
McLean et. al [3] (101)	SMI Eye Link 250 Hz	1: 3X3 grid, 0° and ±20° Horizontal , ±15° Vertical; 2: 3° steps, −24° to 24°. Start (−24°, −24°), Stop (24°, 24°)	1:Information Missing (U); 2:Fixation (U)	1: Information Missing; 2: Fourth Order Polynomial	Information Missing	Information Missing
Tai et. al [14] (6)	EyeLink 1000 500 Hz	0° and ±10° Horizontal and Vertical	Not explicitly specified (U)	Information Missing	Information Missing	Information Missing
Abel et.al [15] (11)	EyeLink II	Information Missing	Foveation Periods (U)	Information Missing	Information Missing	0.5° - 1.0° Manufacturer Numbers
Barot et. al [10] (16)	EyeLink II	30° Left to 30° Right in steps of 3°	Foveation Periods (A)	Best line of fit	Information Missing	Information Missing
Dell'Osso et. al [16] (24)	EyeLink II 500 Hz	Information Missing	Foveation Periods (U)	Information Missing	Information Missing	0.5° - 1.0° Manufacturer Numbers
Hertle et. al [17] (19)	Ober 2 or EyeLink, 500 Hz or 1000 Hz	1° targets or 3° pictures	End of fast phase (U)	Information Missing	Information Missing	Information Missing
Taibbi et. al [18] (28)	EyeLink II 500 Hz	Information Missing	Foveation Periods (U)	Information Missing	Information Missing	Information Missing
Thomas et. al [19] (56)	EyeLink 250 Hz	0° and ±15° Horizontal and Vertical	Foveation Periods (U)	Information Missing	Information Missing	Information Missing
Dunn [4] (1)	EyeLink 1000	±5° Horizontally, ±3° and (0°, 0°)	Automatic Foveation Algorithm [4] (A)	Regression with cross term. Degree unspecified.	Self Validation	Mean and standard deviation for horizontal and vertical values

Previous work on nystagmus calibration

Different approaches for *calibration data selection* for nystagmus applications have previously been published. This is an important part of the calibration since the selected calibration data should represent that the participant looked at the displayed calibration target. If the selected calibration data do not represent the "correct" fixation, there is a risk of misrepresenting eye movement data.

A method to find saccades in eye movement data based on adaptive acceleration thresholds was presented in [20]. The intent of the method was not calibration of nystagmus data, but it served as the basis for the development of a method designed for the nystagmus case. The nystagmus specific version identified the slowest eye movement velocities, referred to as *foveation periods* [4]. The method is based on an

algorithm for saccade detection in uncalibrated data, which is used to divide the waveform into fast and slow eye movements. The foveations are found in the slow phase of the data. Another approach to find foveations was presented in [21], where manual annotation to mark the start and end times of the foveations, was used. While there has been some work on how calibration data are selected, literature on the suitability of various polynomials for nystagmus recording purposes is sparse.

Many papers concerning nystagmus and eye tracking do not report how calibration was performed and evaluated. In Table 1, nine different studies are summarised. As can be seen from the table, only three of the nine studies ([3], [4], [10]) report any sort of calibration model structure, although the calibration MF details are not explicitly presented in any of the papers. Only three of the studies ([4], [15],[16]) report any type of data quality measure or accuracy. In the two first, the reported accuracies are taken from the manufacturer's specification sheet and therefore reveals no information about the accuracy for participants in these particular studies.

Calibration polynomials

Table 2: Summary of the calibration polynomials found in eye tracking and nystagmus related studies. *:The polynomial suggested in [22] has been slightly changed compared to the original proposal. The corner correction terms $m[q]$ and $n[q]$ used in [22] are not estimated for each quadrant but rather for the entire plane. There are not any direct references in nystagmus research to this method in the literature presented in this work. Since it is a common calibration polynomial it was included.**: Polynomials which were not explicitly stated. Instead they have been interpreted from the context.

Study	Polynomial [P]	Eye tracking Data Vector $[\boldsymbol{u}_{PC}]$	Property
Barot et. al **[10]	$\mathcal{A}_1$	$\begin{bmatrix}1 & x_{PC} & y_{PC}\end{bmatrix}^T$	Linear mapping (Linear)
Dunn** [4]	$\mathcal{B}$	$\begin{bmatrix}1 & x_{PC} & y_{PC} & x_{PC}y_{PC}\end{bmatrix}^T$	Linear mapping + Rotation (Non-linear)
Stampe* [22]	$\mathcal{G}$	$\begin{bmatrix}1 & x_{PC} & y_{PC} & x_{PC}^2 & y_{PC}^2 & x_{PC}y_{PC}\end{bmatrix}^T$	Quadratic Mapping + Rotation (Non-linear)
McLean et. al **[3]	$\mathcal{A}_4$	$\begin{bmatrix}1 & x_{PC} & x_{PC}^2 & x_{PC}^3 & x_{PC}^4 & y_{PC} & y_{PC}^2 & y_{PC}^3 & y_{PC}^4\end{bmatrix}^T$	Fourth order (Non-Linear)

There are various references to calibration polynomials used in nystagmus eye tracking research. Four of these polynomials are evaluated in this paper. As described above, calibration data are used to estimate the polynomial coefficients where the input to the calibration polynomial is PCRV, denoted PC in the equations, data and the output is PoR data. Table 2 summarises the characteristics of the four selected polynomials previously used in the nystagmus eye tracking literature.

The PoR estimation, $\boldsymbol{p}_{PoR} = \begin{bmatrix}x_{PoR} & y_{PoR}\end{bmatrix}^T$, is computed using a polynomial, $\boldsymbol{P}$, and eye tracker data, $\boldsymbol{u}_{PC}$, as,

$$\boldsymbol{p}_{PoR} = \boldsymbol{P}\boldsymbol{u}_{PC}. \tag{1}$$

where $\boldsymbol{u}_{PC} = \begin{bmatrix} x_{PC} & y_{PC} \end{bmatrix}^T$. The selected structure of $\boldsymbol{P}$ determines the structure of $\boldsymbol{u_{PC}}$ (see Table 2). The purpose of the calibration is to estimate the coefficients of the polynomial

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{p_h} \\ \boldsymbol{p_v} \end{bmatrix}, \tag{2}$$

where $\boldsymbol{p_h}$ and $\boldsymbol{p_v}$ are the horizontal and vertical polynomials respectively. The coefficients are estimated using a least squares solution according to

$$\boldsymbol{p_d} = (\boldsymbol{U}_{PC}^T \boldsymbol{U}_{PC})^{-1} \boldsymbol{U}_{PC}^T \boldsymbol{t}_d, \tag{3}$$

where d is either the horizontal or the vertical direction, $\boldsymbol{U}_{PC}$ is a matrix containing the calibration data vectors for each calibration target,

$$\boldsymbol{U}_{PC} = \begin{bmatrix} \boldsymbol{u}_{PC}(1) \\ \vdots \\ \boldsymbol{u}_{PC}(n) \end{bmatrix}, \tag{4}$$

$\boldsymbol{t}_d$ is a vector with calibration targets of direction d, and n is the number of calibration targets. The different polynomials evaluated in this work are given in the equations below:

$$\mathcal{A}_1 = \begin{bmatrix} a_{0,c} & a_{0,x} & 0 \\ a_{1,c} & 0 & a_{1,y} \end{bmatrix}, \tag{5}$$

$$\mathcal{B} = \begin{bmatrix} b_{0,c} & b_{0,x} & b_{0,y} & b_{0,xy} \\ b_{1,c} & b_{1,y} & b_{1,y} & b_{1,xy} \end{bmatrix}. \tag{6}$$

$$\mathcal{G} = \begin{bmatrix} g_{0,c} & g_{0,x} & g_{0,y} & g_{0,x^2} & g_{0,y^2} & g_{0,xy} \\ g_{1,c} & g_{1,x} & g_{1,y} & g_{1,x^2} & g_{1,y^2} & g_{1,xy} \end{bmatrix}, \tag{7}$$

$$\mathcal{A}_4 = \begin{bmatrix} a_{0,c} & a_{1,c} \\ a_{0,x} & 0 \\ a_{0,x^2} & 0 \\ a_{0,x^3} & 0 \\ a_{0,x^4} & 0 \\ 0 & a_{1,y} \\ 0 & a_{1,y^2} \\ 0 & a_{1,y^3} \\ 0 & a_{1,y^4} \end{bmatrix}^T. \tag{8}$$

1.4 Aim of this paper

The aims of this paper are to propose and evaluate a new calibration MF generating consistent PoR estimations across recording sessions and participants and compare it to other calibration mapping functions previously used in nystagmus research. The main objective is to find an MF which reliably can be used to evaluate the effects of different nystagmus treatments, even when the participant fails to accurately fixate the calibration target.

2 Proposed method

In this section a new calibration method is proposed. It is developed for video-based eye trackers using a nine-point calibration and a geometrical setup similar to that of an EyeLink 1000 Plus in desktop mode. The method consists of two parts: First, an outlier correction algorithm aimed at correcting inaccuracies in the recorded calibration data. Second, a linear mapping function based on *Procrustes analysis* is proposed. The method is based on 5 seconds of data recorded for each calibration target, as will be presented in more detail in Section 3.

2.1 The outlier correction algorithm

For the recommended setup of the eye-tracker used in this work, the horizontal data typically have the following structure; the horizontal PoR data are dependent only on the horizontal PCRV data, and not on the vertical PCRV data. Thus, horizontal PCRV for a horizontal gaze position is approximately the same, regardless of the vertical gaze position. This characteristic is used to create an algorithm to reduce errors in the calibration dataset. The algorithm is based on nine calibration targets distributed in a 3×3 grid where the calibration data for each calibration target are mapped to one coordinate pair. In this case there are 9 two-dimensional coordinates; one for each two-dimensional calibration target. The outlier correction algorithm consists of two stages.

Stage I

(a) Divide the data into six groups with three adjacent data points in each. Half of the groups share a horizontal calibration target value (see Figure 2a) and the other half share the vertical calibration target value (see Figure 2b).

(b) Fit a line to the three data points in each of the six groups.

(c) Compute the angle between each of the vertically fitted lines and each of the horizontally fitted lines (3×3 computations).

(d) If the angle deviates more than 25° from the expected 90°, the vertical line is considered to contain an outlier. The value of 25° was chosen empirically.

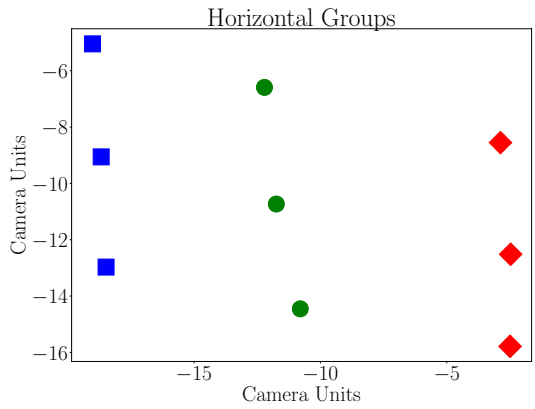
If one or more outliers were found during Stage I, Stage II is initiated.

Stage II

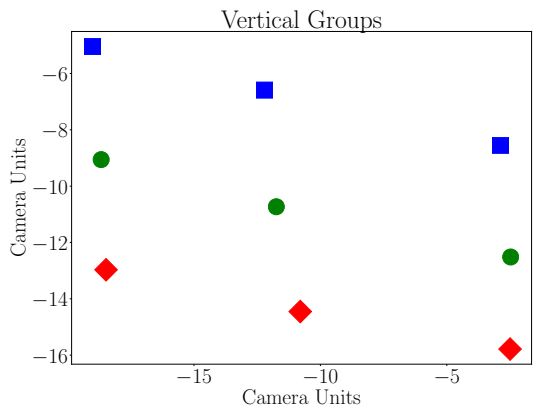
(a) An outlier is detected by finding the datapoint with the largest horizontal deviation from the vertical line.

(b) Corrected coordinates of the outlier are computed as the average of the other data points on each of the intersecting horizontal and vertical lines, i.e., the new horizontal data point value is computed as the average of the corresponding horizontal data points of the vertical line, and the new vertical data point value is computed as the average of the corresponding the vertical data points of the horizontal line.

An example of calibration data points before and after outlier correction is shown in Figure 3.



(a)



(b)

Figure 2: Group Division. The data points divided after the horizontal target values (2a) and vertical target values (2b). All data points of the same shape and colour belong to the same horizontal group (2a) or same vertical group (2b).

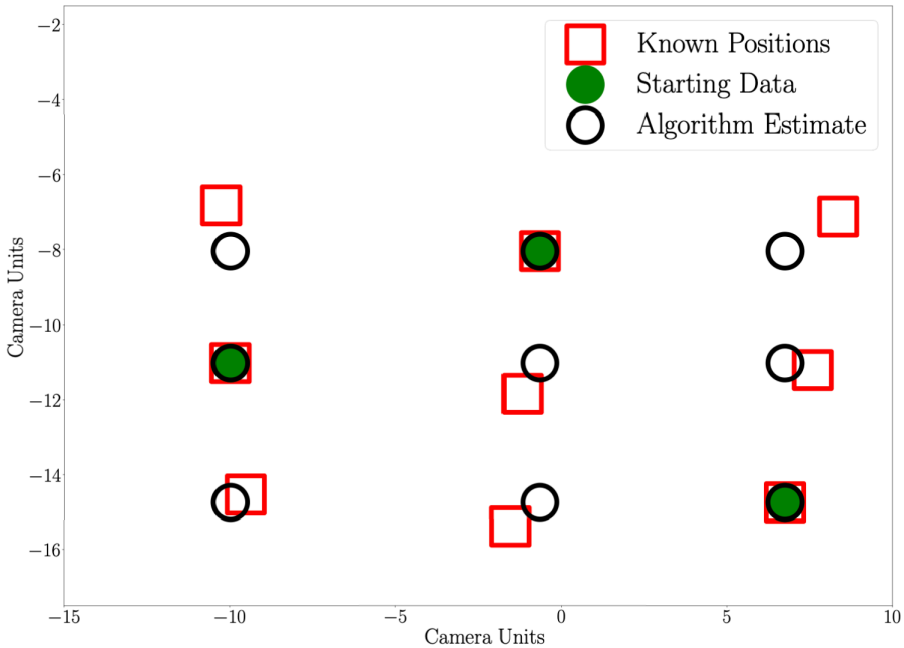


Figure 3: Illustration of the outlier correction algorithm where 6 out of 9 calibration data positions have been removed from the data set. The red squares represent the known calibration data positions, the green filled circles represent the known data before the algorithm estimation, and the black circles represent the estimated calibration data positions.

2.2 Procrustes calibration

In the calibration process, a set of n (here $n = 9$) two-dimensional data points (calibration data) are fitted to another set of n two-dimensional data points (calibration targets). Both of these data sets can be viewed as two-dimensional shapes, and the objective of the calibration is to identify the best transformation from the calibration data shape to the calibration target shape. In this work, *Procrustes analysis* [23] is used to compare and align the two datasets. Three steps are involved in the Procrustes analysis: translation, scaling and rotation. Once they have been estimated, they can be used to compute the gaze positions from PCRV data.

The three transformations have been implemented in the following way:

(a) Construct the calibration data matrix $\boldsymbol{D} = \begin{bmatrix}\boldsymbol{x}_d & \mathbf{y}_d\end{bmatrix}^T$ as a $2 \times n$ matrix where n is the number of calibration targets, and the calibration target matrix $\boldsymbol{T} = \begin{bmatrix}\boldsymbol{x}_t & \boldsymbol{y}_t\end{bmatrix}^T$ contains the corresponding calibration targets.

(b) Center both the calibration data and calibration target datasets by subtracting their respective horizontal and vertical averages from each data set to create $\boldsymbol{D_\mu}$ and $\boldsymbol{T_\mu}$.

$$\boldsymbol{D}_\mu = \begin{bmatrix}\boldsymbol{x}_d - \bar{x}_d \\ \boldsymbol{y}_d - \bar{y}_d\end{bmatrix} = \begin{bmatrix}\boldsymbol{x}_{d,c} \\ \boldsymbol{y}_{d,c}\end{bmatrix}, \tag{9}$$

$$\boldsymbol{T}_\mu = \begin{bmatrix}\boldsymbol{x}_t - \bar{x}_t \\ \boldsymbol{y}_t - \bar{y}_t\end{bmatrix} = \begin{bmatrix}\boldsymbol{x}_{t,c} \\ \boldsymbol{y}_{t,c}\end{bmatrix}, \tag{10}$$

where $\bar{x}_d$ is the average of $\boldsymbol{x_d}$, $\bar{y}_d$ is the average of $\boldsymbol{y_d}$, $\bar{x}_t$ is the average of $\boldsymbol{x_t}$ and $\bar{y}_t$ is the average of $\boldsymbol{y_t}$.

(c) Compute the norms, N_D and N_T, using

$$N_D = \sqrt{\sum_{i=1}^{n} x_{d,c}^2(i) + \sum_{i=1}^{n} y_{d,c}^2(i)} \tag{11}$$

where $x_{d,c}(i) \in \boldsymbol{x}_{d,c}$ and $y_{d,c}(i) \in \boldsymbol{y}_{d,c}$,

$$N_T = \sqrt{\sum_{i=1}^{n} x_{t,c}^2(i) + \sum_{i=1}^{n} y_{t,c}^2(i)} \tag{12}$$

and $x_{t,c}(i) \in \boldsymbol{x}_{t,c}$ and $y_{t,c}(i) \in \boldsymbol{y}_{t,c}$. The datasets are scaled according to:

$$\boldsymbol{D}_N = \frac{\boldsymbol{D}_\mu}{N_D} \tag{13}$$

$$\boldsymbol{T}_N = \frac{\boldsymbol{T}_\mu}{N_T} \tag{14}$$

(d) The rotation, $\boldsymbol{R}$, is computed using singular value decomposition (SVD). In general, the SVD decomposes a matrix $\boldsymbol{M}$ into two orthonormal matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ and a diagonal matrix $\boldsymbol{S}$ that contains the singular values $\sigma_l, l \in [1, k]$. In Procrustes analysis, $\boldsymbol{M} = \boldsymbol{D}_N^T \boldsymbol{T}_N$.

$$\boldsymbol{D}_N^T \boldsymbol{T}_N = \boldsymbol{U}\boldsymbol{S}\boldsymbol{V}^H, \tag{15}$$

where

$$\boldsymbol{R} = \boldsymbol{U}^H \boldsymbol{V}. \tag{16}$$

and

$$\boldsymbol{S} = diag(\sigma_1, \ldots, \sigma_k). \tag{17}$$

(e) Once the translation, scaling and rotation parameters have been estimated, the PoR estimation, $\boldsymbol{p}_{PoR}$, is computed as follows:

$$\boldsymbol{p}_{PoR} = \kappa \boldsymbol{R} \boldsymbol{p}_{PC} - \boldsymbol{L} \tag{18}$$

where

$$\kappa = \frac{N_T}{N_D} \sum_{i=1}^{k} \sigma_i, \tag{19}$$

$$\boldsymbol{L} = \kappa \begin{bmatrix} \bar{x}_d \\ \bar{y}_d \end{bmatrix} \boldsymbol{R} - \begin{bmatrix} \bar{x}_t \\ \bar{y}_t \end{bmatrix}, \tag{20}$$

and

$$\boldsymbol{p}_{PC} = \begin{bmatrix} x_{PC} \\ y_{PC} \end{bmatrix}. \tag{21}$$

This method is denoted as $\mathcal{P}$.

3 Calibration method evaluation

In this Section the evaluation strategy of the proposed method is presented. The Section consists of three main parts; the recording of *nystagmus data* (Sec 3.2), the recording of *control data* (Sec 3.3) and the performance evaluation measures (Sec 3.4).

3.1 Hardware and software

Binocular, raw pupil and CR data were recorded with an EyeLink 1000 Plus (desktop mode) with a sampling frequency of 1000 Hz using the host software v. 5.09 and the DevKit 1.11.571. The center of mass tracking mode was used. The eye tracker camera was placed in accordance with the recommendations of the manufacturer [24]. PsychoPy (version 1.83) [25] was used to present all stimuli. The stimulus was presented on an ASUS VG248QE monitor with a resolution of 1920×1080 pixels, with dimensions $53cm \times 30cm$. The participant to monitor distance was 80 cm. A chin and forehead rest was used for all participants. The analysis software was written in Python (version 2.7).

3.2 The nystagmus data experiment (NDE)

Participants

The nystagmus data experiment was performed with patients diagnosed with nystagmus. The diagnosis was performed by Björn Hammar (MD), senior consultant at the neuro-ophthalmology unit at Skåne University Hospital in Lund, Sweden. This dataset is denoted **NDE data**. A total of eight patients with nystagmus were recorded, two of which were recorded twice totalling ten separate recordings. Two of the participants were female and six were male. Out of the ten recordings, four were excluded from the data set; one due to lack of validation data, two due to loss of calibration data (too many blinks during the recording of calibration data) and one due to too small oscillations. For this participant, only the data from one out of the nine calibration targets consisted of oscillations with an amplitude larger than 1° and a frequency higher than 2 Hz. Out of the six remaining recordings, from five different participants, all were diagnosed with infantile nystagmus (M = 35.3 [year], SD = 15.9[year]).

Data recording

The experiment included calibration and validation data recordings. Both calibration and validation data were recorded monocularly for both eyes by covering one eye and recording the other eye. Nine calibration targets were presented to each patient in a randomised order. The calibration targets were placed in a 3×3 grid. The horizontal target positions were 0° and $\pm 18^\circ$ and the vertical target positions were 0° and $\pm 10^\circ$. The validation targets were placed in a 2×2 grid where the horizontal and vertical validation target positions were $(\pm 5^\circ, \pm 5^\circ)$ respectively. The calibration target was a black circle with radius of 0.6° with a red circle of radius 0.15° in the center. The targets were presented on a grey background. The presentation duration of each calibration target and validation target was decided manually. The goal duration

for each target was 5 seconds (M = 5.02 [s], SD = 1.24 [s]). The experiment also included fixation, smooth pursuit, saccade and optokinetic nystagmus tasks which were not included in this work.

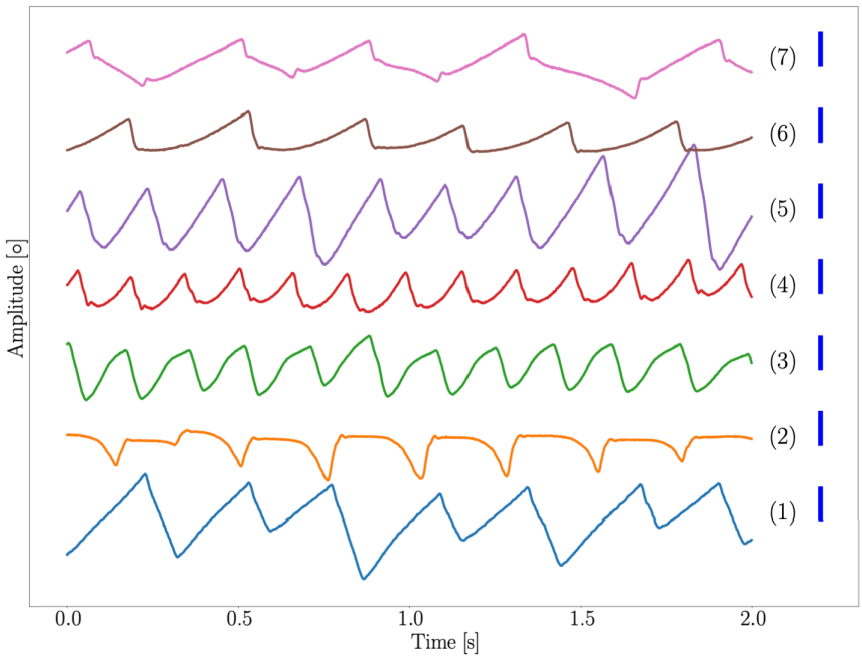


Figure 4: Various waveform recordings from the **NDE data** set for different participants. These are representative of the waveforms found in the dataset from the six participants. The length of blue scale bar at the right side of each signal is 4°. The calibration was preformed using the Procrustes calibration method.

Calibration data selection

The calibration data selection algorithm presented in [4] was implemented. Some adjustments were made to the original algorithm:

(a) Instead of computing saccade velocity thresholds for the entire calibration data set, the thresholds were computed for each calibration target.

(b) The saccade acceleration threshold was not implemented, due to too heavy saccade rejection.

(c) The adaptive filter to find foveations was not implemented. Instead, each slow phase longer than 50 ms was considered as a potential foveation. The first 50 ms directly after the onset of the slow phase were considered to be the most likely foveation candidate.

3.3 The control data experiment (CDE)

The second experiment was designed to collect reference data to be compared to the **NDE data**.

Participants

A total of eight participants were included in the dataset, one female and seven male (M = 37.0 [year], SD = 7.7 [year]). This data set is denoted **CDE data** and was divided into two subsets, see Section 3.3. Data from one participant was excluded due to data loss (too many blinks during the recording of calibration data).

Data recording

The calibration protocol consisted of 81 calibration targets using a standard 3×3 grid with nine possible positions for each calibration target, one reference position and eight offset positions. The participants were recorded binocularly. The distribution of the targets is shown in Figure 5. The vertical and horizontal offset amplitudes were $\pm 0.5^\circ$ and $\pm 2.0^\circ$. Each target was shown for 1.5 seconds. The positions of the calibration targets were evenly distributed between -10° and 10° in the horizontal direction and -5° and 5° in the vertical direction, not counting the offsets. The calibration target was a white circle with radius of 0.6° with a black circle of radius 0.15° in the center. The targets were presented on a grey background. The calibration targets were presented in a randomised order and the offset magnitude at each calibration target was also randomised. Since no nystagmus is present in the **CDE data** a different method for calibration data selection was needed, see Section 3.3.

Two OKN tasks were recorded for this dataset. A black and white striped pattern, see Figure 6, was used to elicit OKN. In the first OKN task, the pattern was moving horizontally. In the second task, the pattern was rotated 90° and moved vertically. In both cases, the temporal frequency of the pattern was 8 cycles / second (in the moving direction), the spatial frequency was 0.5 cycles / degree and the duration was 15 seconds. Before the pattern was set in motion the participants were asked to look in the center of the screen and keep looking there as long as the pattern was moving. The experiment also included fixation, smooth pursuit and saccade tasks, which were

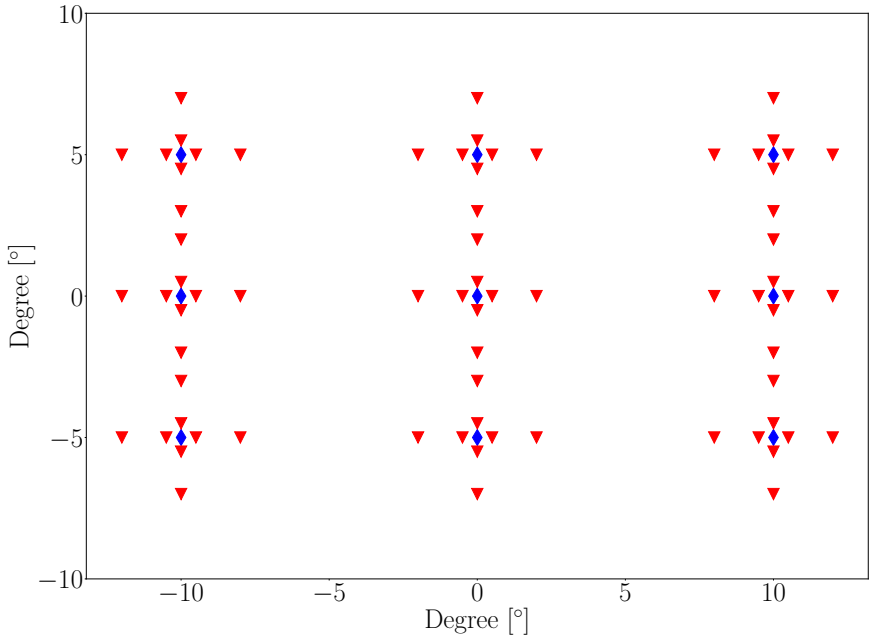


Figure 5: The 81 calibration targets used for the **CDE** calibration are shown. The blue diamonds represent calibration targets and the red triangles represents targets simulating fixation inaccuracies. The targets were presented in random order. Each target was shown for 1.5 seconds.

not included in this work. The OKN datasets were used to compute the *calibration plane distortion* and the *waveform robustness*, described in Section 3.4.

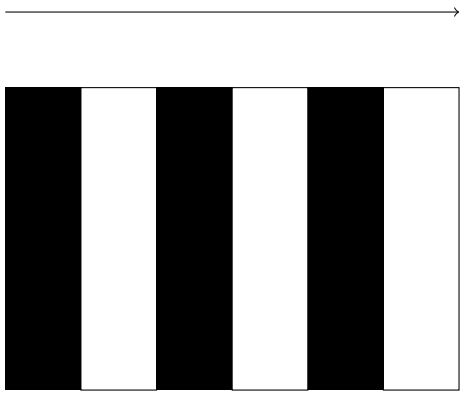


Figure 6: Illustration of the black and white striped pattern used to elicit nystagmus eye movement in healthy participants. The participant was asked to fixate in the center of the stripped screen when the pattern was in motion. The motion of the stimulus in the configuration illustrated above was horizontal. If the pattern is rotated 90° the stimuli moves vertically.

Two CDE subsets

The **CDE data** datasets were divided into two subsets: one which contains only calibration targets with no offset, **CDE - NO**, and one which contains calibration targets with a random offset for each calibration target, **CDE - O**. The notations NO and O represent datasets with no introduced offsets and with introduced offsets, respectively. While the **CDE - NO** data correspond to data from participants without any visual impairment, the **CDE - O** data simulate potential fixation inaccuracies caused by the nystagmus oscillations for different angles during the calibration.

The **CDE - O** dataset was created by repeating the calibration data selection process 50 times, each time assigning a horizontal random error (including 0°) to each calibration target. Each repetition was independent of previous repetitions.

Calibration data selection

The rationale for calibration data selection at each calibration target, is that the PCRV segment with the least variance best represents a fixation. The calibration data selection method is described below:

(a) First, in order to avoid influence of the time it takes to change positions after a new calibration target has appeared, the first 500 ms of the recorded data for each calibration target are removed.

(b) Second, the 200 ms window with the smallest variance of the following PCRV data are computed. The total variance, s_{tot}^2, is computed according to (22), where s_x^2 and s_y^2 are the horizontal and vertical variance respectively.

$$s_{tot}^2 = s_x^2 + s_y^2 \tag{22}$$

(c) Finally, the horizontal and vertical calibration data position estimates are computed as the averages of the 200 ms window found in step 2.

3.4 Comparing calibration methods

In this work, three different measures are used to compare the characteristics of the different mapping functions. These are accuracy, α, *calibration plane distortion*, μ_d, and *waveform robustness*, ξ. Accuracy is tested on a limited number of validation targets, which in this work is equal to four targets per participant. The calibration plane distortion is the distance between two PoR estimations from the same MF. Finally, the waveform robustness is computed as the difference between two PoR estimations after adjusting for the linear properties translation, rotation and scaling between the two PoR estimations.

Accuracy

The accuracy for validation target point k, α_k, is computed according to (23) where $x_{PoR}(k)$ and $y_{PoR}(k)$ are the mapping function estimates of the horizontal and vertical validation target positions, respectively, and $x_s(k)$ and $y_s(k)$ are their corresponding known validation target positions. The accuracy computation in (23) results in one single value for each validation target. A small accuracy value means good performance, while a large value means poor performance.

$$\alpha_k = \sqrt{(x_{PoR}(k) - x_s(k))^2 + (y_{PoR}(k) - y_s(k))^2} \tag{23}$$

The accuracy is presented in the following way. For each mapping function, the average accuracy of each eye of all validation data for one dataset is computed. This means for example that the **CDE O** dataset contain: 7 participants $\times$ 4 validation targets $\times$ 50 iterations = 1400 accuracy samples.

The accuracy is calculated separately for all three datasets. In order to evaluate the performance of the outlier correction algorithm (see Section 2.1), the accuracy results for the **NDE** dataset without the outlier correction algorithm are also calculated.

Calibration plane distortion & waveform robustness

The calibration plane distortion computations were implemented in the following way. If $\boldsymbol{P}_{PoR1} = \begin{bmatrix} \boldsymbol{v}_x & \boldsymbol{v}_y \end{bmatrix}^T$ and $\boldsymbol{P}_{PoR2} = \begin{bmatrix} \boldsymbol{w}_x & \boldsymbol{w}_y \end{bmatrix}^T$ are two matrices of dimension $L \times 2$ containing gaze estimations, the calibration plane distortion, μ, is defined as:

$$\mu(\boldsymbol{P}_{PoR1}, \boldsymbol{P}_{PoR2}) = \frac{1}{L} \sum_{l=1}^{L} \sqrt{\delta_x^2(i) + \delta_y^2(i)}. \tag{24}$$

where

$$\delta_x(i) = v_x(i) - w_x(i); v_x(i) \in \boldsymbol{v}_x, w_x(i) \in \boldsymbol{w}_x, \tag{25}$$

and

$$\delta_y(i) = v_y(i) - w_y(i); v_y(i) \in \boldsymbol{v}_y, w_y(i) \in \boldsymbol{w}_y. \tag{26}$$

The calibration plane distortion is used to compute how close two PoR estimations are in absolute terms, i.e., the distance on the stimuli screen. Even though it includes changes in the waveform, it is likely that translation effects are a dominant part of the calibration plane distortion value. In order to study the effects on the waveform itself, the waveform robustness measure was computed using the Procrustes Distance, D_P. It is defined as:

$$D_P(\boldsymbol{P}_{PoR1}, \boldsymbol{P}_{PoR2}) = 1 - (\sum_{l=1}^{k} \sigma_l)^2 \tag{27}$$

where $\boldsymbol{S} = diag(\sigma_1, \ldots, \sigma_k)$, is computed according to (15) and $D_P \in [0, 1]$. The $\boldsymbol{P}_{PoR1}$ and $\boldsymbol{P}_{PoR2}$ matrices correspond to the $\boldsymbol{D}$ and $\boldsymbol{T}$ matrices described in Section 2.2.

If $\boldsymbol{P}_{f\{k\},NO}$ is a gaze estimation from mapping function $f\{k\}$ from the **CDE - NO** dataset and $\boldsymbol{P}_{f\{k\},O}$ is a gaze estimation from mapping function $f\{k\}$ from the **CDE - O** dataset, where $f = \{\mathcal{A}_1, \mathcal{B}, \mathcal{G}, \mathcal{A}_4, \mathcal{P}\}$ and $k \in [0, 4]$, the calibration plane distortion, μ_k, and the waveform robustness, ξ_k, for mapping function k are defined in (28) and (29) respectively.

$$\mu_k = \mu(\boldsymbol{P}_{f\{k\},NO}, \boldsymbol{P}_{f\{k\},O}) \tag{28}$$

$$\xi_k = D_P(\boldsymbol{P}_{f\{k\},NO}, \boldsymbol{P}_{f\{k\},O}). \tag{29}$$

In order to reduce the influence of blinks and other artefacts in the OKN datasets, the blink removal algorithm used for the calibration data was applied to the OKN data before the computation of both the calibration plane distortion and the waveform robustness. The method is described in full detail in [4].

The results for calibration plane distortion and waveform robustness are presented as empirical *cumulative distribution functions* (CDFs), as well as the area under each CDF curve, A_{CDF}. The area computations for the calibration plane distortion were bounded to 1° as this is considered a good calibration accuracy [26]. The area computation for the waveform robustness was bounded to 0.2 as the results from Section 4.4 showed that $\text{Prob}(D_P > 0.2) \approx 0.01$ for the $\mathcal{G}$ MF. The A_{CDF} was adjusted such that $A_{CDF} \in [0, 1]$ by dividing the computed area with the maximum CDF-value for the area computation. Using this definition of the waveform robustness, the A_{CDF} for the Procrustes calibration method will be 1.0 be definition.

4 Results

4.1 Accuracy

The accuracies in all datasets are presented in Table 3. When comparing the accuracies for the **NDE** dataset with and without OA, it can be seen that the OA improves the accuracy at least for one of the eyes for all **mapping functions**. The most prominent improvements are seen for the **mapping functions** with a higher degree of freedom, i.e., $\mathcal{G}$ and $\mathcal{A}_4$. As expected, the $\mathcal{G}$ MF achieved the best accuracy for the **CDE - NO** dataset.

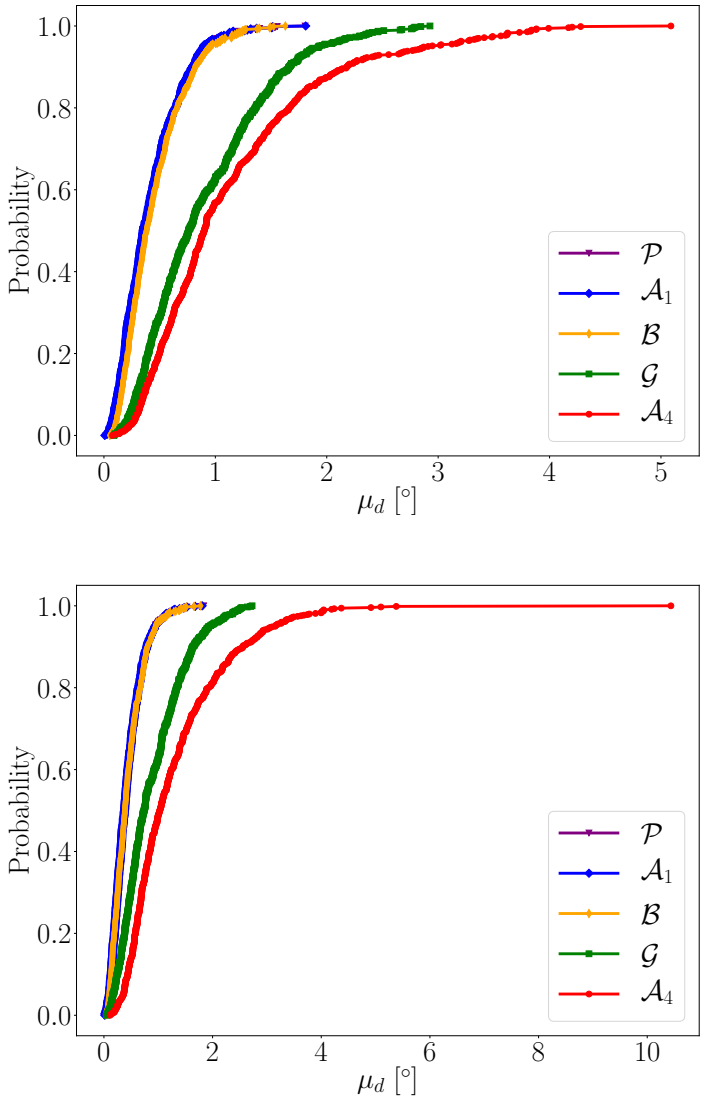
Table 3: The average and standard deviation of accuracy for all datasets. The format is *mean* $\pm$ *standard deviation*. The accuracies are divided and color-coded into three categories; *good*: $0^\circ - 0.5^\circ$, fair: $0.51^\circ - 1.5^\circ$ and **poor**: $1.51^\circ - \text{inf}^\circ$. The (OA) indicates that the outlier correction algorithm has been used during the calibration.

Dataset	$\mathcal{A}_1$	$\mathcal{B}$	$\mathcal{G}$	$\mathcal{A}_4$	$\mathcal{P}$
NDE					
L	**2.1**$^\circ \pm 1.6^\circ$	**2.5**$^\circ \pm 3.8^\circ$	**10.8**$^\circ \pm 22.6^\circ$	**9.8**$^\circ \pm 11.3^\circ$	**2.2**$^\circ \pm 1.7^\circ$
R	**1.7**$^\circ \pm 0.9^\circ$	**1.8**$^\circ \pm 0.9^\circ$	**1.7**$^\circ \pm 1.0^\circ$	**4.9**$^\circ \pm 5.7^\circ$	**1.8**$^\circ \pm 1.0^\circ$
NDE (OA)					
L	**1.8**$^\circ \pm 1.3^\circ$	**1.7**$^\circ \pm 1.4^\circ$	**1.9**$^\circ \pm 2.8^\circ$	**4.9**$^\circ \pm 6.4^\circ$	**1.6**$^\circ \pm 1.1^\circ$
R	**1.7**$^\circ \pm 0.9^\circ$	**1.7**$^\circ \pm 0.9^\circ$	1.5$^\circ \pm 1.0^\circ$	**4.9**$^\circ \pm 5.7^\circ$	**1.7**$^\circ \pm 0.9^\circ$
CDE - NO					
L	0.9$^\circ \pm 0.5^\circ$	0.8$^\circ \pm 0.4^\circ$	*0.5*$^\circ \pm 0.4^\circ$	1.4$^\circ \pm 0.9^\circ$	0.9$^\circ \pm 0.6^\circ$
R	0.9$^\circ \pm 0.4^\circ$	0.9$^\circ \pm 0.4^\circ$	*0.5*$^\circ \pm 0.3^\circ$	1.6$^\circ \pm 0.7^\circ$	1.0$^\circ \pm 0.6^\circ$
CDE - O (OA)					
L	1.2$^\circ \pm 0.7^\circ$	1.1$^\circ \pm 0.7^\circ$	1.1$^\circ \pm 0.9^\circ$	**1.9**$^\circ \pm 1.1^\circ$	1.2$^\circ \pm 0.8^\circ$
R	1.0$^\circ \pm 0.5^\circ$	1.1$^\circ \pm 0.4^\circ$	0.9$^\circ \pm 0.6^\circ$	**1.8**$^\circ \pm 1.0^\circ$	1.1$^\circ \pm 0.6^\circ$

For the **NDE data** and **CDE - O** data (both with OA) where calibration data fixation inaccuracies are present, the accuracies for the $\mathcal{A}_1$, $\mathcal{B}$, $\mathcal{G}$ and $\mathcal{P}$ mapping functions are approximately the same while the $\mathcal{A}_4$ yields a considerably worse accuracy. The fact that the accuracies are worse for the **NDE** database than for the **CDE - O** database indicates that the true Nystagmus calibration errors are more severe than the simulated ones. If good accuracies are defined as being smaller than or equal to 0.5°, it is difficult to achieve good accuracy with inaccuracies in the calibration data.

4.2 Calibration plane distortion

The calibration plane distortion CDFs are presented in Figure 7 and the A_{CDF} results are listed in Table 4. The differences between the results for the vertical and horizontal OKN data within each MF are small. The performance of the $\mathcal{A}_1$, $\mathcal{B}$ and $\mathcal{P}$ MFs are quite similar. The results for the other two MFs are worse. This is confirmed by Figure 7.



(b) Horizontal OKN

Figure 7: The calibration plane distortion plotted for the five mapping functions as cumulative distribution functions. The calibration plane distortion of the $\mathcal{A}_1$, $\mathcal{B}$ and the $\mathcal{P}$ are almost the same. The $\mathcal{G}$ and $\mathcal{A}_4$ polynomials performs worse compared to the three aforementioned MFs.

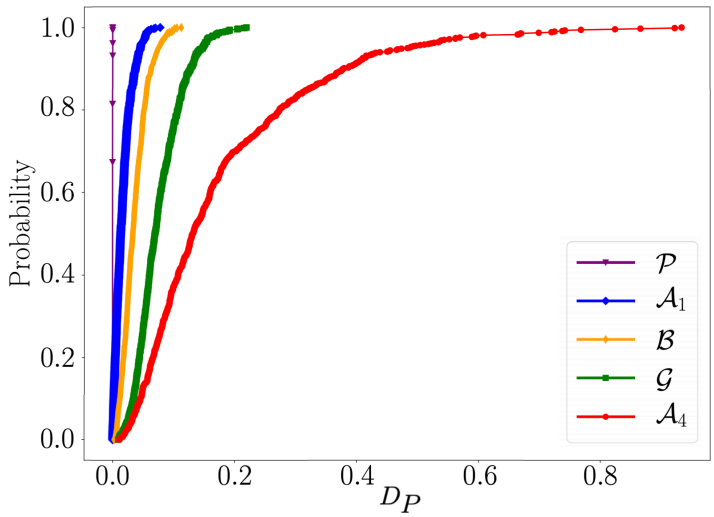
4.3 Waveform robustness

The waveform robustness CDFs are presented in Figure 8 and the corresponding A_{CDF} results are presented in Table 4.

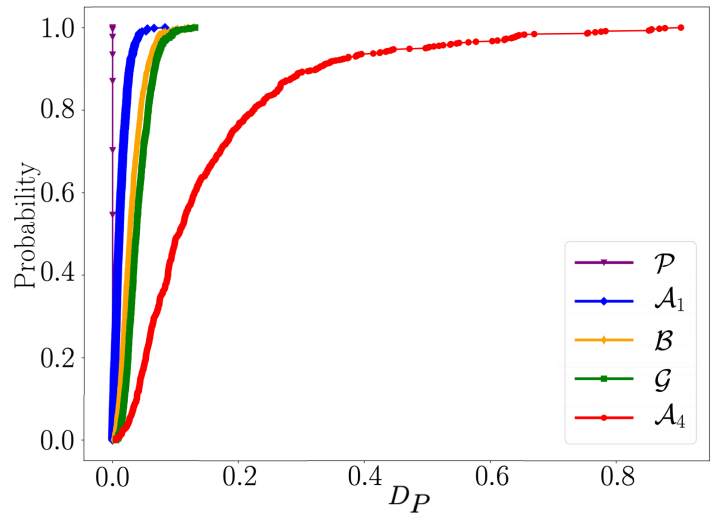
Table 4: A_{CDF} Results. The $\mathcal{A}_1$ MF generates the best calibration plane distortion scores whereas the $\mathcal{P}$ MF generates a perfect waveform robustness score (1.0). The difference in μ-value for the $\mathcal{A}_1$, $\mathcal{B}$ and $\mathcal{P}$ MF are quite small. The analysis was conducted on OKN data.

Dataset	$\mathcal{A}_1$	$\mathcal{B}$	$\mathcal{G}$	$\mathcal{A}_4$	$\mathcal{P}$
Calibration Plane Distortion (μ)					
Vertical	0.61	0.57	0.29	0.23	0.59
Horizontal	0.59	0.57	0.31	0.18	0.58
Waveform Robustness (ξ)					
Vertical	0.91	0.82	0.64	0.36	1.00
Horizontal	0.93	0.85	0.80	0.43	1.00

The results in Figure 8 show that the Procrustes calibration method performs the best and the $\mathcal{A}_4$ performs the worst for both the vertical and the horizontal OKN tasks. This is quantified in Table 4. The waveform robustness seems to be linked to the non-linearity of the MF; a higher degree of non-linearity causes worse waveform robustness performance and vice versa.



(a) Vertical OKN

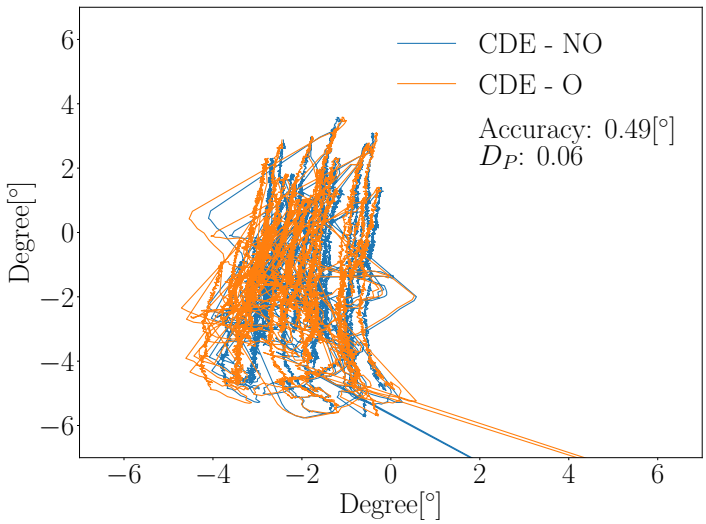


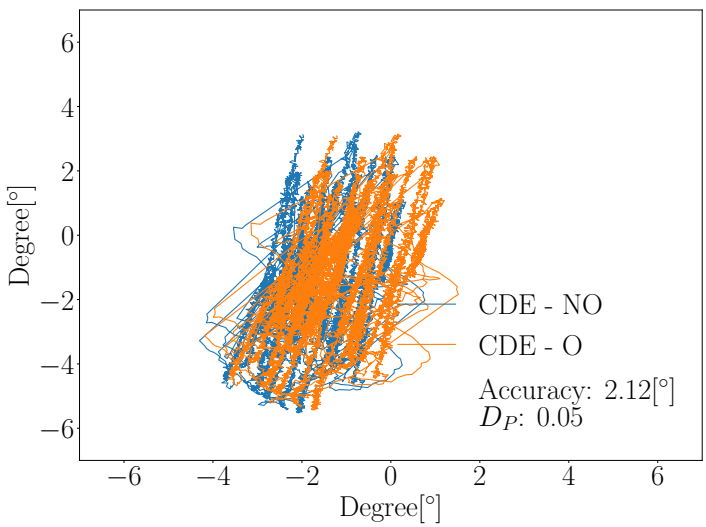
(b) Horizontal OKN

Figure 8: Waveform Robustness CDF. The vertical and horizontal OKN data CDF:s for waveform robustness. The $\mathcal{P}$ generates the best results and the $\mathcal{A}_4$ generates the worst results.

4.4 Waveform robustness and accuracy examples

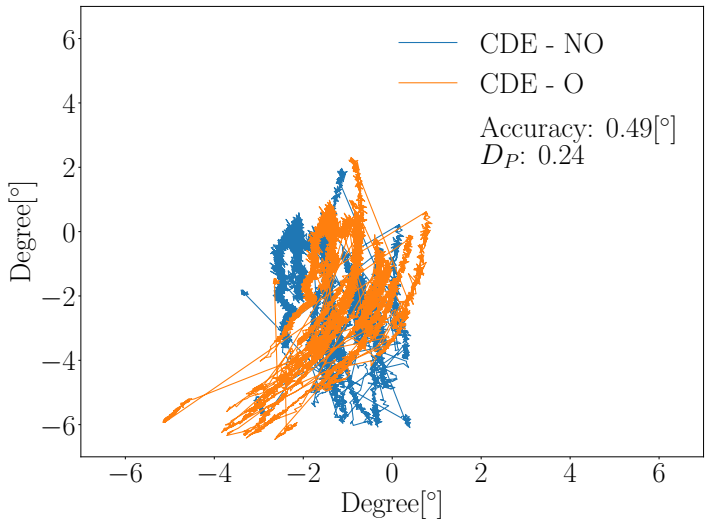
A few examples illustrating the relationship between accuracy and waveform robustness for the $\mathcal{G}$ MF are presented in Figures 9 and 10. As can be seen in Figure 9, it is possible for an MF to produce small waveform robustness values, $D_P = 0.05$, with a relatively large accuracy value, $2.12°$. On the other hand, Figure 10 illustrates that a good accuracy does not guarantee a small waveform robustness value. A D_P value larger than 0.2 is high, since only 1 % of the waveforms generates a higher value in the **CDE - O** dataset. All waveforms estimations were made using the $\mathcal{G}$ MF.

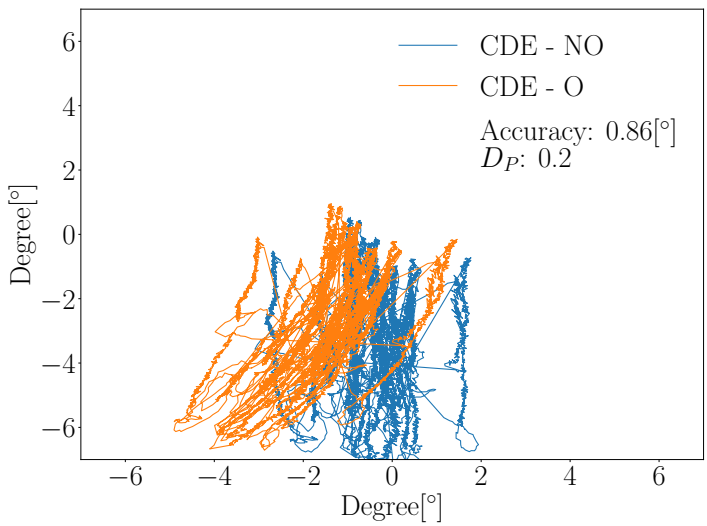




(b) Poor Accuracy, small D_P

Figure 9: Small D_P. Examples of waveforms where the **CDE - O** estimation is similar to the **CDE - NO** estimation. The probabilities of the D_P-values 0.06 and 0.05 are $Prob(D_P > 0.06) \approx 0.62$ and $Prob(D_P > 0.05) \approx 0.73$ for **CDE - O** dataset, respectively. As illustrated by the plots, it is possible to achieve good waveform robustness even if the accuracy is **poor**. The $\mathcal{G}$ MF was used for all estimations. The analysis was conducted on OKN data.





(b) Fair accuracy, large D_P

Figure 10: Large D_P. Examples of data generating good accuracy 10a and fair 10b accuracy, but with poor waveform robustness values. The probabilities of the D_P-values of 0.2 and 0.24 are $Prob(D_P > 0.2) = 0.01$ and $Prob(D_P > 0.24) < 0.01$ for **CDE - Offset** dataset, respectively. The $\mathcal{G}$ MF was used for all estimations. The analysis was conducted on OKN data.

5 Discussion

In this paper, we investigated the suitability of commonly used calibration mapping functions for data from people with nystagmus and proposed a new approach for calibration of these participants. The new method utilises an outlier correction algorithm based on the experiment geometry and calibrates the eye tracker using Procrustes analysis. Our method was compared to different calibration MFs previously used in nystagmus research. Accuracy and Procrustes distance were used to study the properties of the various MFs. Procrustes distance was used to study waveform robustness, i.e., how well waveform PoR data can be repeated within the same participants despite fixation inaccuracies during the calibration, and calibration plane distortion, i.e., how close, in absolute terms, data with simulated fixation inaccuracies were to data without simulated fixation inaccuracies. Data from people with nystagmus (**NDE**), visually healthy participants (**CDE - NO**) and participants with simulated fixation inaccuracies (CDE - Offset) were included in the study.

The accuracy data show that there is little difference between the $\mathcal{A}_1$, $\mathcal{B}$, $\mathcal{G}$ and $\mathcal{P}$ MFs for the **NDE** and **CDE - O** when using the outlier algorithm. However, when studying the calibration plane distortion presented in Figure 7 and Table 4 it becomes apparent that the $\mathcal{G}$ polynomial performs worse compared to the $\mathcal{A}_1$, $\mathcal{B}$ and $\mathcal{P}$ MFs. This observation is likely explained by poor performance on interpolated data (the OKN dataset) by the $\mathcal{G}$ polynomial. The calibration plane distortion thought as an accuracy measure for interpolated data, using the **CDE - NO** as reference. Finally, the results from the waveform robustness in Figure 8 show that the $\mathcal{P}$ MF has the best performance. Since the Procrustes calibration method is based on linear operations only, the waveform robustness is 1.0 by default. The performances of the other MFs are ordered by their non-linearity; the more non-linear, the worse performance. The overall results show that it is not beneficial to use non-linear mapping functions when working with difficult to calibrate participants. Therefore, Procrustes analysis is the best choice when repeatable calibrations are desirable.

The outlier correction algorithm improved the validation accuracies in all cases. This suggests that there is a potential value in modelling the experiment geometry. Even though our results show that the accuracy alone is not a reliable measure for evaluation of an MF it is still desirable to improve the accuracy as long as it does not affect other properties, such as the waveform. It should be noted that if the distribution of the calibration targets is different from the one presented in this paper, the algorithm needs to be adapted for the specific target constellation. One could try to find the geometric relationship between data and targets for calibration target distributions as well, but that would likely demand a more in-depth analysis of the geometry of the experimental setup. The threshold for detecting an outlier, described in Stage I of the outlier correction algorithm in Section 2.1, is an important parameter

for the correction performance. This parameter reflects the maximum deviation that is accepted from the theoretical horizontal distribution of the calibration data. As can be seen in Figure 4, the foveation position varies spontaneously for people with nystagmus. If the threshold value is set too low, there is a risk to affect the structure of the calibration data. On the other hand, if the threshold is set too high, there is a risk to not detect outliers in the data.

The reason why accuracy is not considered as a good indicator of calibration performance for people with nystagmus are the following: 1) It is difficult to know if the validation data were recorded when the participant looked at the corresponding validation target. The accuracy analysis does not make sense if the participant did not look at the presented target, since the entire point of the validation is to test how well the mapping function transforms PCRV data to some known position. Since gaze estimation is dependent on the calibration, it is not possible to know if poor validation results originate from the calibration or the validation. 2) Data distortion effects, as shown in Figure 10a, may occur even if the accuracy is considered to be *good*. This is a problem because one will think that the calibration went well, when in reality gaze data do not correspond to the actual eye movements generated by the participant. However, accuracy is a good measure in the sense that it is a unit (degree) that can be compared between recordings and systems.

The distance measure was included to complement the accuracy and it was used to study how the waveform is affected by the calibration. A problem with the calibration plane distortion and waveform robustness measures is that the value may be difficult to interpret. In this paper, we have computed them on the same PCRV dataset for each mapping function, which makes it possible to compare the distance values between the MFs. The results can only be used to find *that* there are differences in the waveform, not the nature of these differences. For the nystagmus case, more specific differences such as foveation duration, amplitude, frequency and the nystagmus waveform, are of interest but are not possible to find using D_P.

The **CDE - O** used in this work is likely not representative of fixation inaccuracies caused by nystagmus, which the results also indicate; the accuracy of **CDE - NO** is better compared to that of the **NDE**. The idea of making random errors of fixed magnitude does have its limitations and a continuous distribution may possibly be a more realistic representation of the fixation errors for some participants. Signal (1) in Figure 4 shows that it is possible for the position after the fast phase to vary as much as 4° between cycles. The fixation errors introduced in the CDE - O database are therefore considered reasonable.

The creation of the **CDE - NO** and **CDE - O** databases serves a useful purpose in the sense that we have created two identical PCRV datasets, but with different estimates of the mapping functions. This allowed us to study differences between the tested calibration mapping functions. It is not possible to turn off the nystagmus

oscillations for the affected patients, causing this analysis to be impossible to carry out for nystagmus data, since there is no reference waveform to compare the estimations with.

In this work, we tested the EyeLink 1000 Plus system, which is frequently cited in nystagmus research. The applicability of the proposed method for other eye trackers has not been studied.

Finally, the calibration data selection has not been central to the analysis in this paper. It is reasonable to assume that a poor calibration data selection method does have a negative impact on the PoR results, especially considering the results presented in this paper. The adjustments to Dunn's method [4] may have influenced the results in this paper. But based on the data we recorded, the method adjustments are considered reasonable. An updated version of the method has recently been developed [27]. This method may further improve the accuracy of the algorithm. It should also be noted that the calibration data selection implemented in this work is designed for nystagmus with foveation periods or at least waveforms with a distinct fast phase. As can be seen in Figure 4, there were no pendular waveforms present in this dataset. For pendular nystagmus waveforms, it is, however, still possible to use the method since the outlier correction algorithm estimates missing data. It is, however, necessary to have at least three recorded data points, one in each row and one in each column, in order for the algorithm to work.

6 Conclusion

The Procrustes analysis calibration method was shown to be the best when working with data from participants who have a decreased ability to fixate their gaze during the calibration. The principal difference between the Procrustes calibration method and the other investigated methods was the ability to generate repeatable waveform estimations regardless of the calibration recording condition. The choice of calibration mapping function may have a significant impact on the resulting eye movement estimations, which in turn may decrease the reliability of subsequent data analysis.

7 Acknowledgements

This work has been funded by the Swedish Research Council [grant number VR 2015-05442]. We gratefully acknowledge the Lund University Humanities Laboratory. We would also like to thank all the participants.

References

[1] R. W. Hertle, "Nystagmus in infancy and childhood: characteristics and evidence for treatment," *American Orthoptic Journal*, vol. 60, no. 1, pp. 48–58, 2010.

[2] N. Hussain, "Diagnosis, assessment and management of nystagmus in childhood," *Paediatrics and Child Health*, vol. 26, no. 1, pp. 31–36, 2016.

[3] R. McLean, F. Proudlock, S. Thomas, C. Degg, and I. Gottlob, "Congenital nystagmus: Randomized, controlled, double-masked trial of memantine/gabapentin," *Annals of Neurology*, vol. 61, no. 2, pp. 130–138, 2007.

[4] M. Dunn, *Quantifying perception and oculomotor instability in infantile nystagmus*. PhD thesis, Cardiff University, 2014.

[5] M. Theodorou and R. Clement, "Classification of infantile nystagmus waveforms," *Vision Research*, vol. 123, pp. 20–25, 2016.

[6] L. Dell'Osso and R. Daroff, "Congenital nystagmus waveforms and foveation strategy," *Documenta Ophthalmologica*, vol. 39, no. 1, pp. 155–182, 1975.

[7] A. Kumar, S. Shetty, P. Vijayalakshmi, and R. W. Hertle, "Improvement in visual acuity following surgery for correction of head posture in infantile nystagmus syndrome," *Journal of Pediatric Ophthalmology and Strabismus*, vol. 48, no. 6, pp. 341–346, 2011.

[8] J. R. Naegele and R. Held, "The postnatal development of monocular optokinetic nystagmus in infants," *Vision Research*, vol. 22, no. 3, pp. 341–346, 1982.

[9] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer, *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford, 2011.

[10] N. Barot, R. J. McLean, I. Gottlob, and F. A. Proudlock, "Reading performance in infantile nystagmus," *Ophthalmology*, vol. 120, no. 6, pp. 1232–1238, 2013.

[11] D. Sheena and B. Borah, "Compensation for second-order effects to improve eye position measurements, pp. 257-268, in," *Eye movements: Cognition and Visual Perception?, University of Illinois Press*, 1981.

[12] J. J. Cerrolaza, A. Villanueva, and R. Cabeza, "Taxonomic study of polynomial regressions applied to the calibration of video-oculographic systems," in *Proceedings of the 2008 symposium on Eye Tracking Research & Applications*, pp. 259–266, ACM, 2008.

[13] P. Blignaut and D. Wium, "The effect of mapping function on the accuracy of a video-based eye tracker," in *Proceedings of the 2013 Conference on Eye Tracking South Africa*, pp. 39–46, ACM, 2013.

[14] Z. Tai, R. W. Hertle, R. A. Bilonick, and D. Yang, "A new algorithm for automated nystagmus acuity function analysis," *British Journal of Ophthalmology*, pp. bjo–2010, 2010.

[15] L. A. Abel, Z. I. Wang, and L. Dell'Osso, "Wavelet analysis in infantile nystagmus syndrome: limitations and abilities," *Investigative Ophthalmology & Visual Science*, vol. 49, no. 8, pp. 3413–3423, 2008.

[16] L. Dell'Osso, R. W. Hertle, R. J. Leigh, J. B. Jacobs, S. King, and S. Yaniglos, "Effects of topical brinzolamide on infantile nystagmus syndrome waveforms: eyedrops for nystagmus," *Journal of Neuro-Ophthalmology*, vol. 31, no. 3, pp. 228–233, 2011.

[17] R. W. Hertle, D. Yang, K. Adams, and R. Caterino, "Surgery for the treatment of vertical head posturing associated with infantile nystagmus syndrome: results in 24 patients," *Clinical & Experimental Ophthalmology*, vol. 39, no. 1, pp. 37–46, 2011.

[18] G. Taibbi, Z. I. Wang, and L. F. Dell'Osso, "Infantile nystagmus syndrome: broadening the high-foveation-quality field with contact lenses," *Clin Ophthalmol*, vol. 2, no. 3, pp. 585–589, 2008.

[19] S. Thomas, F. A. Proudlock, N. Sarvananthan, E. O. Roberts, M. Awan, R. McLean, M. Surendran, A. A. Kumar, S. J. Farooq, C. Degg, *et al.*, "Phenotypical characteristics of idiopathic infantile nystagmus with and without mutations in frmd7," *Brain*, vol. 131, no. 5, pp. 1259–1267, 2008.

[20] F. Behrens, M. MacKeben, and W. Schröder-Preikschat, "An improved algorithm for automatic detection of saccades in eye movement data and for calculating saccade parameters," *Behavior Research Methods*, vol. 42, no. 3, pp. 701–708, 2010.

[21] L. Dell'Osso, "*Recording and Calibrating the Eye Movements of Nystagmus Subjects*. Omlab report 011105, 1-4, 2005."

[22] D. M. Stampe, "Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems," *Behavior Research Methods, Instruments, & Computers*, vol. 25, no. 2, pp. 137–142, 1993.

[23] J. C. Gower, "Generalized procrustes analysis," *Psychometrika*, vol. 40, no. 1, pp. 33–51, 1975.

[24] SR-Research, "*EyeLink 1000 User Manual*.," 2010.

[25] J. W. Peirce, "Psychopy psychophysics software in python," *Journal of Neuroscience Methods*, vol. 162, no. 1-2, pp. 8–13, 2007.

[26] D. W. Hansen and Q. Ji, "In the eye of the beholder: A survey of models for eyes and gaze," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 3, pp. 478–500, 2010.

[27] M. J. Dunn, C. M. Harris, F. A. Ennis, T. H. Margrain, J. M. Woodhouse, L. McIlreavy, and J. T. Erichsen, "An automated segmentation approach to calibrating infantile nystagmus waveforms," *Behavior research methods*, pp. 1–11, 2019.

Paper II

Modelling and quality assessment of nystagmus eye movements recorded using an eye-tracker

Abstract

Mathematical modelling of nystagmus oscillations is a technique with applications in diagnostics, treatment evaluation, and acuity testing. Modelling is a powerful tool for analysis of nystagmus oscillations, but, quality assessment of the input data is needed in order to avoid misinterpretation of the modelling results.

In this work, we propose a signal quality metric for nystagmus waveforms, the *normalised segment error* (NSE). The NSE is based on the energy in the error signal between the observed oscillations and a reconstruction from a harmonic sinusoidal model called the *normalised waveform model* (NWM). A threshold for discrimination between nystagmus oscillations and disturbances is estimated using simulated signals and receiver operator characteristics (ROC). The ROC is optimised to find noisy segments and abrupt waveform and frequency changes in the simulated data that disturb the modelling. The discrimination threshold, ϵ, obtained from the ROC analysis, is applied to real recordings of nystagmus data in order to determine whether a segment is of high quality or not. The NWM parameters from both the simulated dataset and the nystagmus recordings are analysed for the two classes suggested by the threshold.

The optimised ϵ yielded a true positive rate and a false positive rate of 0.97 and 0.07, respectively, for the simulated data. The results from the NWM parameter analysis show that they are consistent with the known values of the simulated signals, and that the method estimates similar model parameters when performing analysis of repeated recordings from one subject.

William Rosengren, Marcus Nyström, Björn Hammar, Markus Rahne, Linnea Sjödahl, and Martin Stridh,
"Modelling and Quality Assessment of Nystagmus in Eye Tracking Signals,"
in. *Behavior research methods*, pp. 1 - 15, 2020

1 Introduction

Nystagmus is a symptom expressed as involuntary oscillating eye movements with a reported prevalence of 24 per 10 000 people in the general population [1]. The nystagmus symptoms may lead to decreased visual acuity and in certain cases to *oscillopsia*, which is a sensation that the world is in motion [2]. For some of those who are affected, the symptoms are persistent. Traditionally, nystagmus is divided into acquired nystagmus, which develops later in life due to for example trauma, and early-onset nystagmus, which is developed before or a few months after birth [3, 4].

Diagnostics of nystagmus is difficult and often requires long clinical experience. In order to evaluate the condition of a person with nystagmus, various methods, both automatic and manual, for diagnostics and evaluation of the nystagmus signal patterns have been developed [2], [5], [6]. These methods may involve classification of the eye movement dynamics, often referred to as the *waveform*. The waveform may provide insights into the underlying cause of the nystagmus symptoms [2]. In this work, the waveform, as measured by an *eye tracking* system, is considered on a cycle–to–cycle basis.

Eye tracking is a technology used to estimate the gaze direction or to measure eye movements [7]. The technology has previously been used in nystagmus research, and various metrics for evaluation of nystagmus, based on eye tracking data, have been constructed. Many of these methods utilise automatic processing of the recorded eye movements signals. Automatic processing of nystagmus signals is an established and alternative analysis method compared to manual inspection of nystagmus oscillations. The analysis techniques originate from a wide range of different disciplines such as control theory [8], dynamic systems modelling [9], time series analysis [10] and time-frequency analysis [11]. One may consider two different approaches to model nystagmus: *system–based* and *signal–based* modelling. System–based modelling is aimed at the mechanisms of the nystagmus itself. Such methods are often concerned with modelling of, e.g., the saccadic system, and investigates possible mechanisms behind nystagmus [8], [9], [12], [13]. Signal–based modelling of nystagmus data is aimed at modelling and classification of measured nystagmus oscillations, i.e., waveforms. Such models may be used for classification of recorded signals into different established waveform morphologies [10, 14], in order to determine the visual function of a person with nystagmus [6, 15], or to evaluate the effect of a treatment [16]. Manual classification of cycle–to–cycle waveform morphologies have previously been investigated [5], where 12 different waveform morphologies observed in people with nystagmus are described.

When modelling nystagmus eye movements, it is desirable to only include segments in the data that contribute to the overall understanding of the underlying condition. Therefore, an *exclusion criterion* for the data may be introduced. In this context,

an exclusion criterion refers to the method for which data in a study is excluded from further analysis. In this work, exclusion of data is considered on a sub–signal level, meaning that some segments in a single recording may be included whereas other segments may be excluded from further analysis. In order to illustrate the importance of data exclusion, consider the eye movement recording in Figure 1. Here, a slow phase detector [17] based on velocity leads to detection of slow phases in the non-oscillatory part of this data segment, due to the low velocity of this part. Even though a low velocity is an indicator or a slow phase, it is undesirable to use this segment to describe the nystagmus oscillations of this specific person, or to use this segment for calibration. It is therefore desirable to have a method, which automatically rejects segments in the recorded signals that do not exhibit an oscillating pattern and thus should not be included in the analysis.

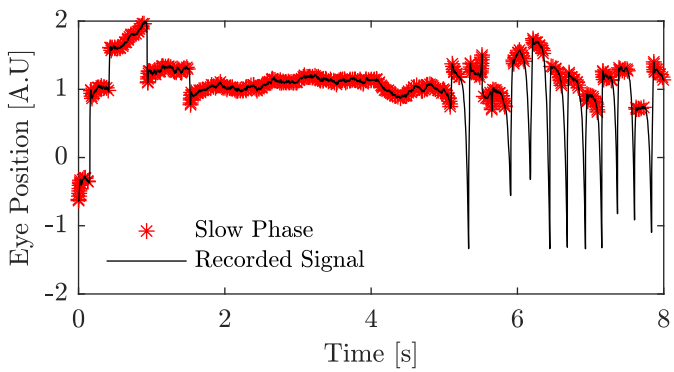


Figure 1: Example of slow phase detection. The slow phase detection method from [17] applied to a nystagmus signal. The red asterisks show the segments of detected slow phases. As can be seen, the segment without an oscillatory pattern in the recorded signal is erroneously labeled as a slow phase.

Removing segments of non-oscillating data from recordings is a process that has previously been used in nystagmus research. A method for removal of blinks using position, velocity and acceleration thresholds has been implemented for calibration purposes [18]. The threshold is based on the standard deviation of the recorded signals.

In order for this method to be operative, at least one blink needs to be present in the recording [18]. A threshold for the nystagmus cycle length has been used as an inclusion criterion on a signal segment level [10]. The cycle length was identified using a method using periodic orbits, where the main cycle length was determined as the peak in a histogram of the periodic orbits. All cycle lengths within ± 12.5 ms of

the main cycle length were included in the analysis.

Although some methods exist for detection of segments in nystagmus eye movement recordings that should be excluded from further analysis, no systematic investigation with the specific purpose to evaluate the performance of the data exclusion methods has to our knowledge been conducted. This means that there is very little information on how well these methods work for excluding undesirable data segments. Data exclusion is an important component when characterising nystagmus oscillations, since the reliability of automatic analysis methods may be significantly influenced by how the method handles disturbances and 'non-nystagmus' segments. If the exclusion criterion allows too many segments to enter further analysis, there is a risk of modelling segments that do not represent the nystagmus oscillation of the patient. On the other hand, if the exclusion criterion allows too few segments, there is a risk of not having enough data left for subsequent analysis, as well as of missing dynamic changes, such as waveform and frequency changes, in the recording.

In this work, we present a signal–based method for modelling of nystagmus signals with the purpose to select segments for further analysis. In order to do so, a metric called the *normalised segment error* (NSE) is introduced. The modelling method is based on a harmonic sinusoidal model, and the NSE is computed as the normalised error between the reconstructed signal and the recorded eye movement signal. An evaluation dataset consisting of simulated signals is created and used for performance evaluation. The method is also evaluated on a dataset with recordings from participants with nystagmus.

This paper is organised as follows: In Section 2, the model of the nystagmus oscillations is presented. The datasets used for evaluation of the proposed model are described in detail in Section 3. The evaluation strategy for the proposed method is presented in Section 4. Finally, the results are presented and the method is discussed in Sections 4 and 5, respectively.

2 Proposed model

The model considered in this work is based on a pseudo-stationary assumption of the nystagmus signal. Consider the harmonic sinusoidal model $s[n]$ with H harmonics [19],

$$s[n] = \sum_{h=1}^{H} s_h[n], \tag{1}$$

where

$$s_h[n] = a_h \sin[2\pi(f_1 h)n + \phi_h] \tag{2}$$

and, a_h is the amplitude of the h:th harmonic, f_1 is the first harmonic frequency and ϕ_h is the phase of harmonic h, and $n = 0, \ldots, N-1$.

The estimation of the model parameters consists of two steps. The first step describes the data preprocessing and the extraction of harmonic components, see Section 2.1. The second step, the parameter estimation, is described in Section 2.2. A description of the model features used for analysis of the nystagmus recordings is presented in Section 2.3. Finally, the NSE is defined in Section 2.4.

2.1 Preprocessing

The preprocessing stage consists of two parts. First, the signal is downsampled to 100 Hz (original sampling rate is 1000 Hz for the data presented in this work). The second step is to highpass filter the signal using a third order Butterworth filter with a cutoff frequency of 2 Hz, removing frequencies lower than the cutoff. The preprocessed signal is denoted $s_p[n]$.

In order to estimate the different harmonic components, $s_h[n]$, each component needs to be extracted from the preprocessed signal. After the preprocessing, the global first harmonic frequency of the signal, $\hat{F}_1$, is estimated using Welch spectrum estimation with an overlap of 50 % and a segment length of 512 samples (for a 100 Hz signal). The harmonic components, $s_h[n]$, are computed as $s_p[n]$ filtered through a Kaiser bandpass filter, $B_h(f, F)$ with the following design settings:

$$B_1(f, F) = \begin{cases} 1, & \text{if } \left|F - f\right| \leq f_{w1} \\ 0, & \text{if } \left|F - f\right| > f_{w2} \end{cases} \tag{3}$$

for the first harmonic, and

$$B_h(f, F) = \begin{cases} 1, & \text{if } Fh - (1 + \delta_h) \leq f \leq Fh + (1 + \delta_h) \\ 0, & \text{if } Fh - (2 + \delta_h) \leq f \leq Fh + (2 + \delta_h) \end{cases} \tag{4}$$

for $h > 1$, where $\delta_h = \frac{(h-1)}{2}$. In this work, $f_{w1} = 1.3$ and $f_{w2} = 2.3$ [19].

2.2 Block model parameter estimations

As stated above, the frequency of the nystagmus signal is in general not stationary. The signal is therefore divided into short segments of length N_b. The choice of the segment length is considered with the following tradeoff: if the segment length is too short, it may result in poor parameter estimates, and if it is too long, the stationarity assumption of each segment may not be valid. The latter problem is adressed by

using a method of overlapping segments, while reconstructing the signal for a shorter interval. The segment length was set to $N_b = 67$ (corresponding to 0.67 seconds).

Another issue is that the frequency estimate $\hat{F}_1$ is not necessarily representative for all intervals in the recorded signal. If the first harmonic frequency varies more than ± 1.3 Hz, the output energy of the affected segments may be severely reduced. In order to remedy this, two additional sets of harmonic components are computed for the frequencies $\hat{F}_0 = \hat{F}_1 - 2.6$ and $\hat{F}_2 = \hat{F}_1 + 2.6$. The frequency estimate for the time interval $\boldsymbol{n}_b = [n_0, \ldots, n_0 + N_b - 1]$ is determined by maximising the first harmonic energy,

$$\hat{F}[\boldsymbol{n}_b] = \underset{\hat{F}_i}{\operatorname{argmax}} \left[E(\hat{F}_0)[\boldsymbol{n}_b], E(\hat{F}_1)[\boldsymbol{n}_b], E(\hat{F}_2)[\boldsymbol{n}_b] \right], \tag{5}$$

where

$$E\big(\hat{F}_i[\boldsymbol{n}_b]\big) = \sum_{k=n_0}^{n_0+N_b-1} \left| s_1^{(i)}[k] \right|^2 \tag{6}$$

and $s_1^{(i)}[\boldsymbol{n}_b]$ is the resulting first harmonic signal after $s_p[\boldsymbol{n}_b]$ is filtered through $B_1(f, F)$, and where $F = \hat{F}_i$. The signal, however, is reconstructed for a time interval $n_c \in [l - c_0, l + c_0]$ where the overlap c_0 is computed as

$$c_0 = \frac{N_b}{2\hat{F}[\boldsymbol{n}_b]}, \tag{7}$$

n_0 is the start sample for the interval of length N_b and

$$l = n_0 + \frac{N_b}{2}. \tag{8}$$

This means that every approximate wave is reconstructed separately based on a window around it. The frequency, f_h, and phase, ϕ_h, of each harmonic are estimated according to [19]

$$\hat{f}_h = \underset{f}{\operatorname{argmax}} \left| \sum_{n_b=n_0}^{n_0+N} s_h[n_b] e^{-j2\pi f n_b} \right| \tag{9}$$

and

$$\hat{\phi}_h = \arctan\left(- \frac{\sum_{n_b=n_0}^{n_0+N} s_h[n_b] \sin(2\pi \hat{f}_h n_b)}{\sum_{n_b=n_0}^{n_0+N} s_h[n_b] \cos(2\pi \hat{f}_h n_b)} \right) \tag{10}$$

The amplitude, $\hat{a}_h$, for the h:th harmonic is estimated from the analytic signal transformation [20],

$$\hat{a}_h = \frac{1}{2c_0} \sum_{i=l-c_0}^{l+c_0-1} |\tilde{s}_h[i]| \tag{11}$$

where $\tilde{s}_h[i]$ is the analytical transformation of s_h.

The signal is not stationary unless $\hat{f}_h = h\hat{f}_1, \forall h$, which is generally not going to be the case. In order to create a stationary model, equation (3) is rewritten as

$$s[n_b] = \sum_{h=1}^{H} a_h \sin[2\pi(\hat{f}_1 h)n_b + 2\pi(\hat{f}_h - \hat{f}_1 h)n_b + \hat{\phi}_h] \tag{12}$$

where $\hat{f}_h$ is the frequency estimate of harmonic h. The second argument of the sinusoid, $2\pi(\hat{f}_h - \hat{f}_1 h)n_b$, may be viewed as a phase component. In order for this to be stationary, the index n_b is replaced by a fixed index value, for example the block center index l [19]. This results in the model

$$s'[n_b] = \sum_{h=1}^{H} \hat{a}_h \sin[2\pi(\hat{f}_1 h)n_b + \hat{\phi}'_h] \tag{13}$$

where

$$\phi'_h = 2\pi(\hat{f}_h - \hat{f}_1 h)l + \hat{\phi}_h. \tag{14}$$

2.3 Waveform features

When the model parameters have been estimated, it is possible to reconstruct and compare waveforms with different frequencies, amplitudes, or morphologies. The relationship between the amplitude coefficients in a harmonic model is a direct measure of the influence of each respective harmonic. For example, if the first harmonic amplitude is much greater compared to the amplitudes of the other harmonics, the resulting signal will be close to a pure sinusoid. Two waveforms may be similar in morphology, but where the absolute amplitudes are quite different. In order to compare waveform morphologies of a harmonic model, all amplitudes are normalised according to

$$\hat{R}_h = \frac{\hat{a}_h}{\hat{a}_1}. \tag{15}$$

The first harmonic phase is a measure of where the oscillation in a given segment begins. However, when comparing different waveform morphologies, it is not of interest to study the absolute phase, but rather to study the relative phase values. In

order to compare different waveform morphologies, the phases of all harmonics are adjusted to the first harmonic phase according to [19]:

$$\hat{\phi}''_h = \frac{\hat{\phi}'_h}{h}. \tag{16}$$

Once all phases have been rescaled to the first harmonic, the angular difference, $\delta_{\phi''_h}$, between the first harmonic phase and rescaled harmonic phases are computed according to [19]:

$$\hat{\delta}_{\phi''_h} = \hat{\phi}''_h - \hat{\phi}''_1 = \hat{\phi}''_h - \hat{\phi}'_1 \tag{17}$$

A new model, the *normalised waveform model* (NWM), may be written as

$$s_a[n_b] = \sum_{h=1}^{H} \hat{R}_h \sin(2\pi \hat{f}_1 h n_b + \hat{\delta}_{\phi''_h} h) \tag{18}$$

2.4 The normalised segment error (NSE)

As described in Section 1, the NSE is introduced in order to determine whether a signal segment should be considered for further analysis. The NSE of a segment with length N_s is computed as

$$NSE_s = \frac{\sum_{n_s=n_0}^{n_0+N_s-1} |s'_p[n_s] - s'[n_s]|^2}{\sum_{n_s=n_0}^{n_0+N_s-1} |s'_p[n_s]|^2} \tag{19}$$

where

$$s'_p[m] = s_p[m] - \frac{1}{N_s} \sum_{n_s=n_0}^{n_0+N_s-1} s_p[n_s] \tag{20}$$

and

$$s'[m] = s[m] - \frac{1}{N_s} \sum_{n_s=n_0}^{n_0+N_s-1} s[n_s]. \tag{21}$$

The signals $s_p[n]$ and $s[n]$ denote the preprocessed and reconstructed signals, respectively, and N_s is the segment length. If $NSE_s > \epsilon$ for some value ϵ, the segment should be excluded from further analysis. The choice of the ϵ value is further discussed in Section 4.1.

3 Datasets

In this section, the two datasets used for the evaluation of the proposed method are presented. In Section 3.1, the evaluation dataset (ED), which consists of simulated signals, is described in detail. The purpose of this dataset is to evaluate the proposed method on signals with known characteristics and known disturbances, which is important both for evaluation of the parameter estimation process and to set the threshold of the exclusion criterion. The content of the recorded participant dataset, PD, is presented in Section 3.2.

3.1 Evaluation dataset

In order to evaluate the performance of the proposed method, it is tested on a dataset with known reference signals. The analysis software for the evaluation dataset was written in Python (version 2.7) using the SciPy (version 1.0.0) signal processing library.

In order to use signals that resemble real nystagmus waveforms, the illustrations of nystagmus waveforms presented in [5] were digitised using image processing. The digitised signals were then parametrised using Fourier analysis in order to introduce signal modulations as well as to be able to track model parameter values. The different waveforms were captured as images using a print screen function on a 27 inch (5120 x 2880) resolution screen

and transformed into 1-dimensional signals. A detailed description of the parametrisation of the different waveforms are found in Appendix I together with the phase and amplitude values for each waveform type.

In order to study the performance of the modelling method, simulated template signals were created and corrupted by various signal modulations. Three different types of modulations were introduced: amplitude modulation, frequency shift and noise consisting of white Gaussian noise.

A total of 50 test signals, each consisting of the four template signals DJ-L, EF-R, PP_{FS} and T [5], with a total duration for each test signal of 24 seconds, were generated. The template waveforms were chosen as the signal parameters are reasonably different from each other, and it was desirable to test the method for different types of waveforms. Examples of these four waveforms are found in Figure 2. Each test signal comprised of four *waveform components*, where each waveform component consisted of one of the four waveforms DJ-L, EF-R, PP_{FS} or T. The duration of each waveform component was six seconds and the sampling frequency was set to 1000 Hz. The order of the different waveform component was randomised for each 24 second signal (4 waveform components $\times$ 6 seconds). The initial frequency of each of waveform component was set to 6 Hz. The model, $y[n]$, for the amplitude modulation and

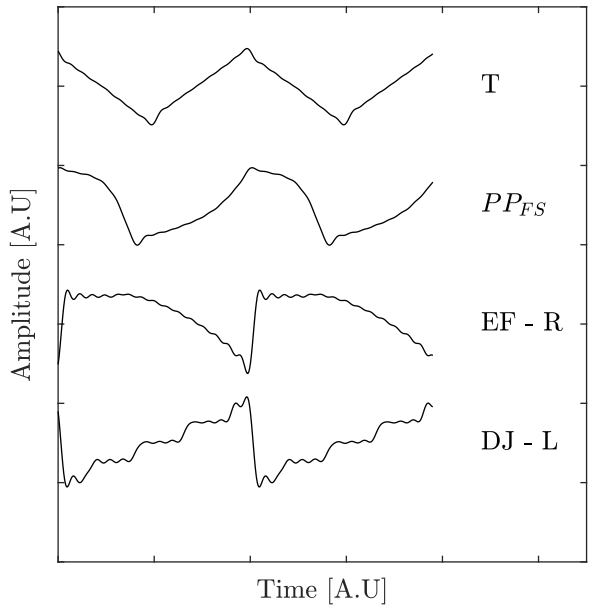


Figure 2: Evaluation Dataset Waveforms. The four different waveforms used to create the evaluation dataset were Dual Jerk - Left (DJ-L), Extended Foveation - Right (EF-R), Pseudo Pendular with Foveating Saccades (PP_{FS}) and Triangular (T). All waveforms have been reconstructed using Fourier analysis.

frequency shift is given by

$$y[n] = A[n] \sum_{h=1}^{H} a_h \sin[2\pi(f_1[n]h)n + \phi_h], \tag{22}$$

where $A[n]$ is the amplitude variation and $f_1[n]$ is the time–dependent frequency for each waveform component.

$$A[n] = \sin[2\pi f_A n + \cos(2\pi \frac{f_A}{5} n)] \tag{23}$$

$$f_1[n] = \begin{cases} 6, & \text{if } n \leq n_1 \\ f[u], & \text{if } n_1 < n \leq N, \end{cases} \tag{24}$$

For each waveform component there was one frequency shift, occurring between 1.5

seconds and 4.5 seconds after the onset of the waveform component. The new frequency, f[u], was randomly sampled from a uniform distribution, $\mathcal{U}\big(8, 10\big)$.

Table 1: The random variable values for the evaluation dataset.

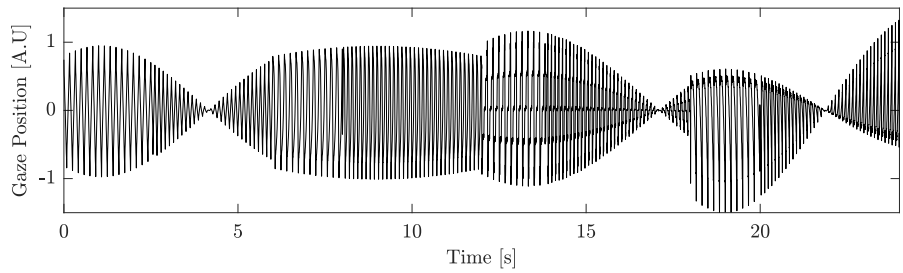
$f[u]$	n_1	f_A	w_a	d_w
$\mathcal{U}\Big(8, 10\Big)$	$\mathbb{Z} \in [1500, 4500]$	$\mathcal{U}\Big(10^{-2}, 10^{-1}\Big)$	$\mathcal{N}\Big(0, 0.05\Big)$	$\mathbb{Z} \in [500, 1000]$

The number of noise episodes in each simulated signal was randomised where the maximum number of occurrences was six times for every waveform component. The distribution of the noise was a zero mean white gaussian distribution, $\mathcal{N}\big(0, 0.05\big)$. The time of occurrence for each noise episode was also randomised, and it was possible for multiple episodes to occur simultaneously. The onset and offset times for the frequency shifts and noise episodes were stored for all waveform components.

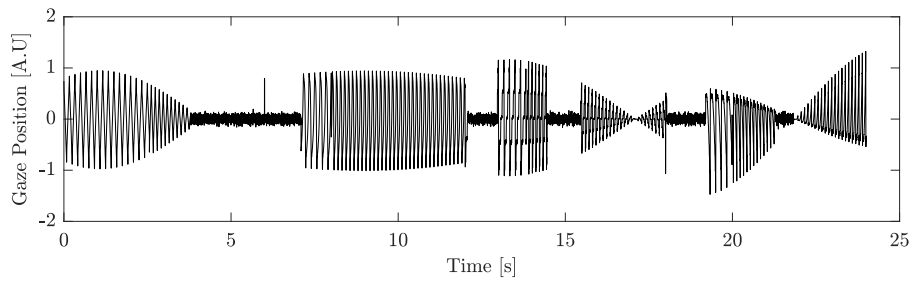
A *disturbance vector*, $D[n]$ is defined as

$$D[n] = \begin{cases} 1, & \text{if } t_{on} - \delta_t \leq n \leq t_{off} + \delta_t \\ 0, & \text{if } otherwise \end{cases} \tag{25}$$

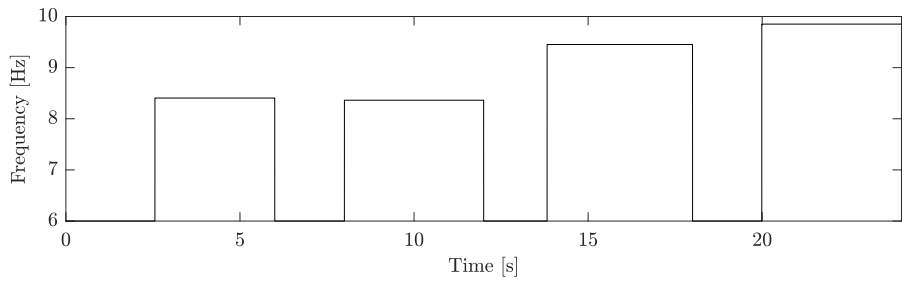
where t_{on} and t_{off} are the onset and offset times of either a frequency modulation or a noise segment. The term δ_t is an additional disturbance time added due to the moving window of the modelling method. For this work, $\delta_t = 100$ samples. An example of a simulated signal is presented in Figure 3. The original signal without any disturbances, $y_o[n]$, is plotted in Figure 3a. The change in frequency, waveform and the noise segments are illustrated in Figure 3b. The corresponding disturbance vector, $D[n]$, is plotted in Figure 3d.



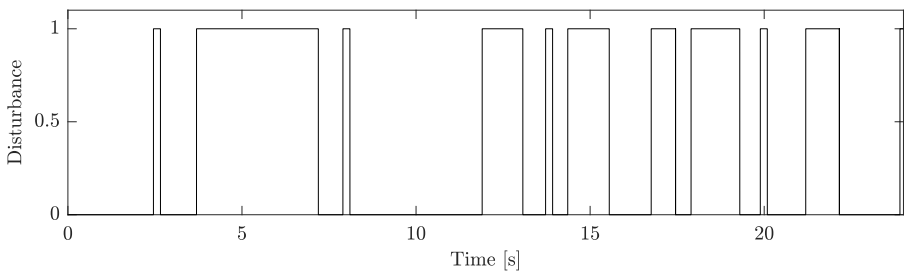
(a) $y_o(n)$



(b) $y(n)$



(c) $f_1(n)$



(d) $D(n)$

Figure 3: Illustration of the simulated signals. (3a) The original simulated signal $y_o[n]$, (3b) the 'observed' simulated signal $y[n]$, (3c) the frequency $f_1[n]$ and the (3d) disturbance vector $D[n]$ are illustrated.

3.2 Participant dataset

The dataset containing nystagmus recordings, PD, consists of two sub-datasets denoted PD_1 and PD_2. In short, these datasets were recorded with the following setup. Uncalibrated pupil and corneal reflection data were recorded binocularly at 1000 Hz using an EyeLink 1000 Plus eye tracker in desktop mode, with the host software v. 5.09 and the DevKit v. 1.11.571. The eyes were tracked using the center of mass mode. The geometry of the experiment setup was in accordance with the manufacturer recommendations [21]. The stimuli software was written in Python and PsychoPy (version 1.83) [22]. An ASUS VG248QE monitor with a resolution of 1920×1080 pixels, with dimensions 53×30 cm was used for stimuli presentation and all participants were seated 80 cm from the screen. The head was stabilised using a chin and forehead rest. The study is approved by the ethics board at Lund University and all experiments are in accordance with the Declaration of Helsinki.

All participants were subject to calibration and validation before the experiment started. During the calibration, the participants were instructed to focus on a set of nine calibration targets, appearing in a randomised order. The vertical target positions were $\pm 10^\circ$ and 0° and the horizontal target positions were $\pm 18^\circ$ and 0°. The same procedure was implemented during the validation using four different targets appearing in random order. Both the horizontal and vertical positions of the validation targets were $\pm 5^\circ$. During both the calibration and the validation, each eye was first recorded monocularly (the other eye was covered) followed by a binocular recording. A black circle with radius 0.6° with a red circle of radius 0.15° in the center were used as the calibration target. The color of the background was gray.

During the experiment the participants were presented with fixation targets in five different positions: $(-16^\circ, 0^\circ)$, $(16^\circ, 0^\circ)$, $(0^\circ, 0^\circ)$, $(0^\circ, -8^\circ)$ or $(0^\circ, 8^\circ)$, where the first coordinate is the horizontal position and the second is the vertical position and the coordinate $(0^\circ, 0^\circ)$ is at the center of the screen. The same target composition (black circles with red centers) as for the calibration was used and each target was shown for 15 seconds, except for the center target which was shown for 30 seconds. After each target had been shown, a target at the central position, $(0^\circ, 0^\circ)$, was shown for 5 seconds. The analysis of these segments has not been included in this work. In between each target, the participant was allowed to rest the eyes and blink for five seconds. This dataset is referred to as PD_1. In total, recordings from 5 male participants with diagnosed early onset nystagmus were included ($M = 34.9$ [years], $SD = 14.7$ [year]). The participants were diagnosed by a senior neuro–ophthalmologist [BH] at Skåne University Hospital, Lund, Sweden.

A second dataset, PD_2, was created by performing repeated experiments for one of the participants from PD_1 (8 times on different days). The data from one session were excluded due to equipment malfunction. The same calibration, validation and

fixation recording procedure as were used for PD_1 were also used for PD_2, where the horizontal calibration positions where $\pm 16^\circ$ and 0°. The age of this male participant was 25 years. PD_2 was created in order to investigate the repeatability of the estimated model parameters and thereby the measured nystagmus pattern. All participants were calibrated using the foveation detection and the Procrustes calibration method presented in [17].

4 Model performance evaluation

The performance evaluation is divided into two different parts. The first part, presented in Section 4.1, is the analysis of the ED. In Section 4.2 examples of the modelling performance for participant dataset PD_1 and PD_2 are presented. All signals are analysed using the NSE. The ϵ estimated from the ED, see Section 4.1, is used in the analysis of the PD_1 and PD_2. The NSE is computed in 200 ms segments, which corresponds to half of a cycle of an oscillation at 2.5 Hz. Segments in the data for which the EyeLink system cannot track the pupil are marked by the system. At each such occurrence, 200 ms before and after the occurrence were removed from the signal. This occurs for example during a blink. All segments for which the time between two episodes of missing data was less than 5 seconds were excluded from analysis.

4.1 Evaluation dataset analysis

ROC Analysis

As described in Section 3.1, three different signal modulations have been introduced in the ED. Two of these, the frequency shift and the noise leads to segments in the dataset that should be excluded from further analysis. Just before and after an abrupt change in frequency, there is no 'pure' waveform, which is why these segments in the signals should be excluded from further analysis.

The goal of the modelling is to find these episodes in the recorded signal by utilising the NSE of each reconstructed signal segment. In order to use the NSE as an inclusion criteria, an NSE threshold that maximises the number of segments that contains nystagmus waveforms (true positive rate) and at the same time minimises the number of segments with the unwanted waveform modulations (false positive rate) is constructed. For this purpose, the *receiver operating characteristics* (ROC) is used as follows:

$$\mathcal{C} = \begin{cases} 1, & \text{if } NSE_s > \epsilon \\ 2, & \text{if } NSE_s \leq \epsilon \end{cases} \tag{26}$$

A true positive (TP) is defined if $D[n_0] = \ldots = D[n_0 + N_d - 1] = 1$ and $\mathcal{C} = 1$ in equation (26). A false positive (FP) is defined if $\mathcal{C} = 1$, where $D[n_0] = \ldots = D[n_0 + N_d - 1] = 0$. A true negative (TN) is defined if $\mathcal{C} = 2$ and $D[n_0] = \ldots = D[n_0 + N_d - 1] = 0$ whereas a false negative (FN) is defined if $\mathcal{C} = 2$ and $D[n_0] = \ldots = D[n_0 + N_d - 1] = 1$. In this work, the length of each disturbance segment, N_d, equals 20 samples (200 ms).

The true positive rate (TPR) is defined as

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{27}$$

and the false positive rate (FPR) is defined as

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \tag{28}$$

All segments that contain both undesired episodes (noise, frequency or waveform change) and desired episodes (nystagmus oscillations) were excluded from the analysis. The optimal ϵ was determined by

$$\hat{\epsilon} = \underset{d_\epsilon}{\text{argmin}} \left(d_\epsilon \right) \tag{29}$$

where

$$d_\epsilon = \sqrt{(1 - TPR(\epsilon))^2 + FPR(\epsilon)^2}. \tag{30}$$

Frequency analysis

The frequency estimation is a crucial part of the model parameter estimation. If the frequency estimation is poor, there is a risk that the bandpass filters (see Section 2.1) use the wrong passband, which in turn may lead to poor estimation of the other model parameters. The bandpass filters have a theoretical binary gain of 1 if the frequency is within the passband and 0 if the frequency is outside. This means that if

$$|\hat{f}_1 - f_{w1}| \leq 1.3, \tag{31}$$

the spectral energy is large. The filtering process is evaluated by determining the percentage of segments for each class of the two classes $\mathcal{C} = 1$ and $\mathcal{C} = 2$, where the gain in equation (3) is equal to 1.

Amplitude analysis

The distributions of the amplitude ratios $\hat{R}_2$ and $\hat{R}_3$ are estimated using *kernel density estimation* (KDE) and are compared to the true values of R_1 and R_2. The KDEs are estimated using a Gaussian kernel with a bandwidth 0.05. If the modelling is accurate, it is expected that the energy of the distribution for the accepted values are centered around the true parameter values. At the same time, it is expected that there is a larger variance of the estimated values for the rejected segments.

4.2 Participant dataset analysis

There are three properties of the proposed model that are desired to study using the participant datasets. The first property is the ability to model different waveforms from different individuals. The second property is the ability to replicate results from one individual over multiple recordings. The last property is the ability to differentiate waveforms recorded from different spatial positions, i.e., fixation targets. The PD_1 is used for evaluation of the first property whereas the PD_2 is used for evaluation of all three properties.

In order to study these properties, the parameters of the normalised waveform model are analysed. Contrary to the analysis performed for the ED, there are no reference parameters to compare with for these two datasets. Instead of comparing the parameter values to known reference values, the distributions of the estimated model parameters are studied. The parameters $\hat{R}_2$ and $\hat{\delta}\phi_2''$ for each recording are used to compute a polar coordinate pair according to:

$$\begin{aligned} x &= \hat{R}_2 \cos(\hat{\delta}\phi_2''), \\ y &= \hat{R}_2 \sin(\hat{\delta}\phi_2''). \end{aligned} \tag{32}$$

Using the ϵ estimated in Section 4.1, each polar coordinate pair is classified as either accepted or rejected.

In order to study the spatial performance of the normalised waveform model, a *representative waveform* from each fixation target recording is reconstructed from the model parameters. An example of the variance in spatial waveforms from PD_2 is illustrated in Figure 4. The representative waveform is reconstructed using the densities of the $\hat{f}_1$, $\hat{R}_2$, $\hat{R}_3$, $\hat{\delta}\phi_2''$ and $\hat{\delta}\phi_3''$ parameters. For the first three parameters, KDEs are estimated from all estimated parameter values from the recording of one fixation target. The parameter values maximising each respective KDE are used for the representative waveform. The KDEs were estimated using a Gaussian kernel with a bandwidth of 0.5 for the estimation of $\hat{f}_1$ and a bandwidth 0.05 for the estimation of the ampli-

tudes ratios. For the angular parameters, $\hat{\delta}\phi_2''$ and $\hat{\delta}\phi_3''$, a circular histogram was used instead. The binwidth of the circular histogram was set to $5°$.

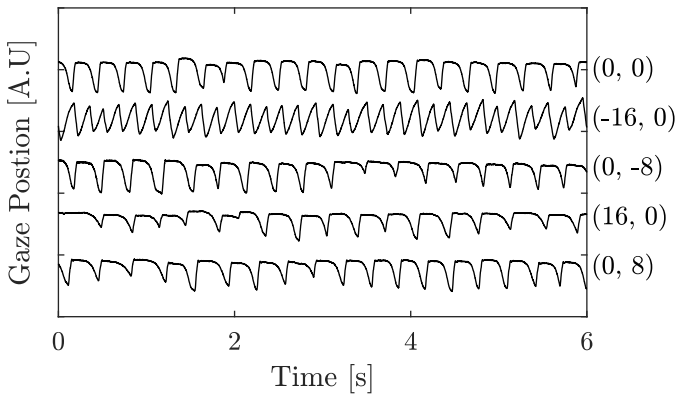


Figure 4: Example Signals from PD_2. The signals represents recordings from five different fixation positions (for one participant). The coordinate pair after each signal describes at which fixation target position the signals were recorded. The frequency and waveform of the signal recorded at position $(-16°, 0°)$ is different compared to the other four.

5 Results

The result section is divided into two parts where the ED results are presented in Section 5.1 and results for PD_1 and PD_2 are presented in Section 5.2.

5.1 Evaluation dataset

The results for the ED are organised in the following structure: First, the choice of ϵ, determined by equation (29), is presented. Based on this ϵ, the performance of the frequency and amplitude ratio estimations are analysed.

Choice of ϵ

The ROC analysis of ϵ is presented in Figure 5. Out of the blocks that were considered for the ROC analysis, the prevalence of corrupted episodes (excluding amplitude modulation) was 44 %. The total number of analysed segments was 4295.

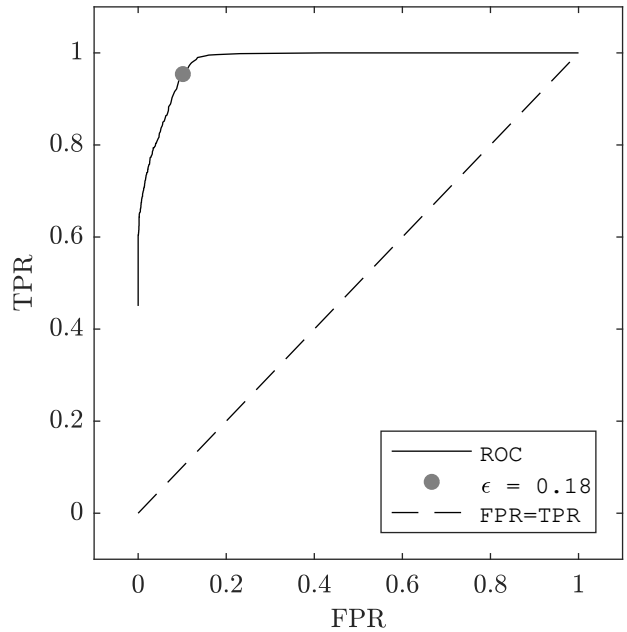


Figure 5: Receiver Operating Characteristics. The false positive rate (FPR) is plotted against the true positive rate (TRP) for various values of the error threshold. The position on the ROC minimising equation (30) is plotted as a gray circle.

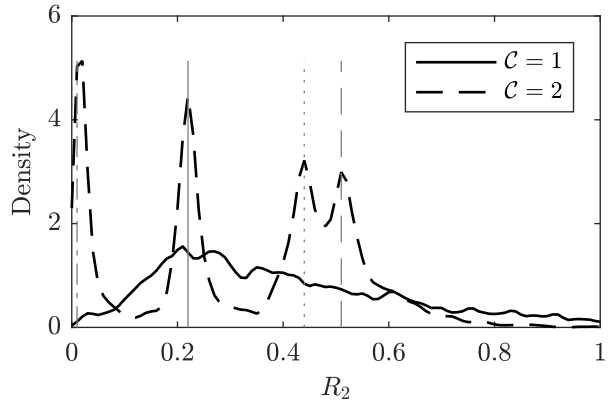
The ϵ that minimises the distance d in equation (30), is marked with a circle. The corresponding TPR and FPR values for $\epsilon = 0.18$ are TPR $= 0.97$ and FPR $= 0.07$. The interpretation of this is that 97 % of all segments of noise, frequency or waveform changes are detected and that 7 % of all segments where a non-oscillating signal was detected, are true nystagmus oscillations. This ϵ is used for all subsequent analysis, both for the ED and the PD datasets. All segments where $\epsilon > 0.18$ are labeled as rejected ($\mathcal{C} = 1$), and all other segments are label as accepted ($\mathcal{C} = 2$).

Frequency and amplitude ratio estimation performance

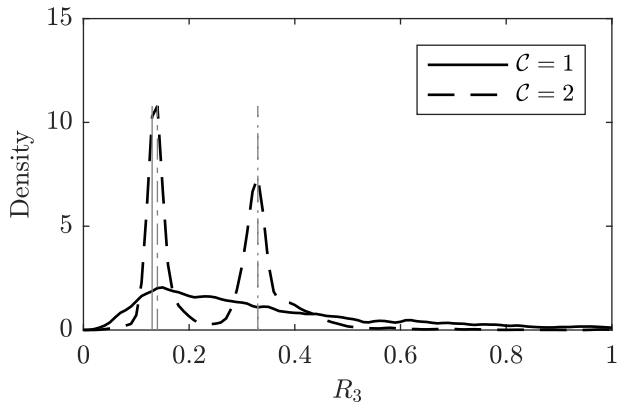
The results show that out of the accepted segments, 99.7 % of the blocks satisfy the criterion in equation (31). Only 65 % of the rejected segments satisfy this criterion. This means that 99.7 % of the accepted segments are within the allowed bandwidth of the bandpass filters.

The distributions of the estimated $\hat{R}_2$ and $\hat{R}_3$ are plotted in Figure 6. The distribution of $\hat{R}_2$ and $\hat{R}_3$ for accepted segments ($\mathcal{C} = 2$, dashed line) matches the

known references values well. The distribution of the estimated $\hat{R}_2$ and $\hat{R}_3$ for rejected segments ($\mathcal{C} = 1$, full line) has a larger variance compared to the accepted data. The peaks at the positions of the known values are not clearly distinguishable for the rejected segments.



(a) Distribution of $\hat{R}_2$



(b) Distribution of $\hat{R}_3$

Figure 6: KDE of $\hat{R}_2$ and $\hat{R}_3$ for the ED. The estimated distributions of the amplitude ratios $\hat{R}_2$ and $\hat{R}_3$ are presented in Figures 6a), 6b) for the two classes rejected segments ($\mathcal{C} = 1$) and accepted segments ($\mathcal{C} = 2$). The vertical gray lines represent the known reference values for each of the four waveforms. Note that the R_3 value for the DJ-L and EF-R are almost identical.

Table 2: The proportion of rejected segments for ED, PD_1 and PD_2. The lowest proportion of rejected segments is quite similar for the two datasets PD_1 and PD_2. The highest proportion of rejected segments, however, is quite different 0.9 and 0.41, respectively. The results are presented on a participant level, meaning that the total proportion of the rejected segments for one participant (five fixation target recordings for each participant) is presented for each of the datasets.

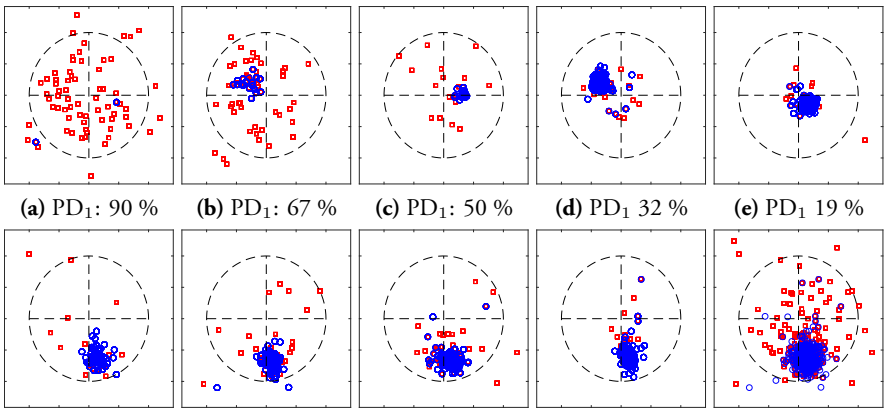
Dataset	Average	Lowest	Highest
ED	0.44	0.27	0.58
PD_1	0.47	0.19	0.90
PD_2	0.29	0.22	0.41

5.2 Participant data

The percentage of rejected segments for the participant datasets PD_1 and PD_2 are presented in Table 2. As can be seen, the average percentage of rejected segments are 47 % and 29 % for PD_1 and PD_2, respectively. For one participant in PD_1, 90 % of the total number of segments for the five fixation target recordings were rejected. The average rejection rate of PD_1 (47 %) data is similar to the average rejection rate of the ED data (44 %).

The polar representation of $\hat{R}_2$ and $\hat{\delta}\phi_2''$ are plotted in Figure 7, where the top and bottom rows show example recordings from PD_1 and PD_2, respectively. The rejected segment parameter estimations are plotted as red squares and the accepted segments parameters are plotted as blue circles. Two general trends can be observed in these figures. First, high rejection rates lead to a larger spread of values within the unit circle. For example, all estimated values in Figure 7e (19 % rejection rate) are more concentrated than the parameter values presented in Figure 7a (90 % rejection rate). This means that the waveform morphology for the data presented in Figure 7e varies to a lower degree compared to that in Figure 7a. Note that Figure 7j illustrates the aggregated parameter estimations for all recordings in PD_2. Second, the repeatability of the two parameter values $\hat{R}_2$ and $\hat{\delta}\phi_2''$ appears to be high for high quality signals. All of the included parameter values from the PD_2 dataset (the bottom row) are positioned at approximately the same location inside the unit circle. There is one important thing to note: the $\hat{R}_2$ values have been limited to $\hat{R}_2 \leq \sqrt{2}$ in these plots. In some cases, the $\hat{R}_2$ value is much larger than this restriction. If this is the case, the segment is most often classified as rejected ($\mathcal{C} = 1$). These results illustrates the individual variation in waveform morphology during one measurement and the reproducibility properties of the model between repeated measurements.

The polar coordinate representation of $\hat{R}_2$ and $\hat{\delta}\phi_2''$ of the signal shown in Figure



(a) PD_1: 90 % **(b)** PD_1: 67 % **(c)** PD_1: 50 % **(d)** PD_1 32 % **(e)** PD_1 19 %

(f) PD_2 41 % **(g)** PD_2: 34 % **(h)** PD_2 26 % **(i)** PD_2 22 % **(j)** Aggregated PD_2, 29 %

Figure 7: Polar coordinates from signals with varying rejection rates and waveform morphologies. The plots above show the estimated $\hat{R}_2$ and $\hat{\delta}''_{\phi_2}$ plotted as polar coordinates for five recordings from PD_1 (top row) and four recordings and the aggregated estimations (Figure (7j)) from PD_2 (bottom row). The percentage in each caption shows the overall exclusion rate for each participant recording. The parameters have been estimated from a recording of the primary position $(0°, 0°)$. The blue circles represent the accepted segments and the red squares represent the rejected segments.

1, is presented in Figure 8. The diamonds represent the oscillations after 5 seconds whereas the crosses represent the oscillations up until 5 seconds. As illustrated in Figure 8, there is a relatively large variance of the parameter values for the crosses compared to the diamonds. All diamonds have similar radius and angle, whereas the crosses are spread out over the entire unit circle. This indicates that when there is an oscillatory pattern with a repetitive waveform in the data, the model parameters are clustered together. However, when there is not such a pattern, the model parameters diverge. These results may be used to determine whether there is a stable oscillating pattern in the data or not.

The reconstruction of the waveforms for the PD_2 data, is illustrated in Figure 9. As is shown in Figure 4 (the original signals), the eye movement recorded at position $(-16°, 0°)$ is significantly different compared to the other four fixation recordings. This is true for all the repeated recordings in PD_2. The change in nystagmus characteristics on the left side of the screen for this patient has been captured in the signal reconstruction (Figure 9). This implies that the normalised waveform model is useful for analysis of and comparison between different nystagmus waveforms. These results

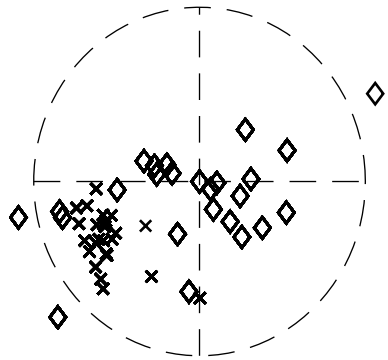


Figure 8: Illustration of polar coordinates. The $\hat{R}_2$ and $\hat{\delta}_{\phi''_h}$ values from the signal in Figure 1) are plotted as the radius and angle, respectively. The parameters estimated from the first five seconds of the signal are plotted as crosses, whereas the rest of the parameters are plotted as diamonds.

are an illustration of the spatial properties of the model.

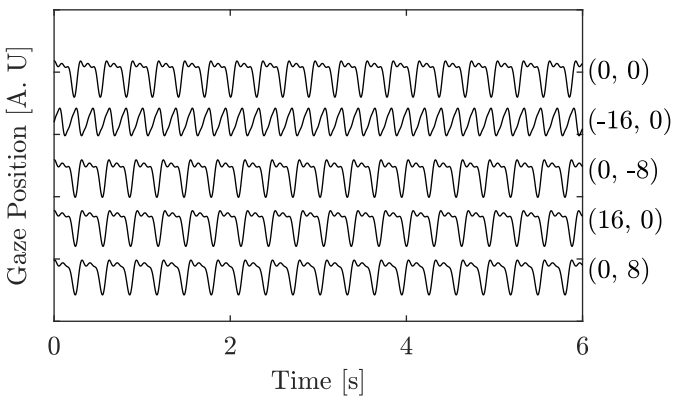
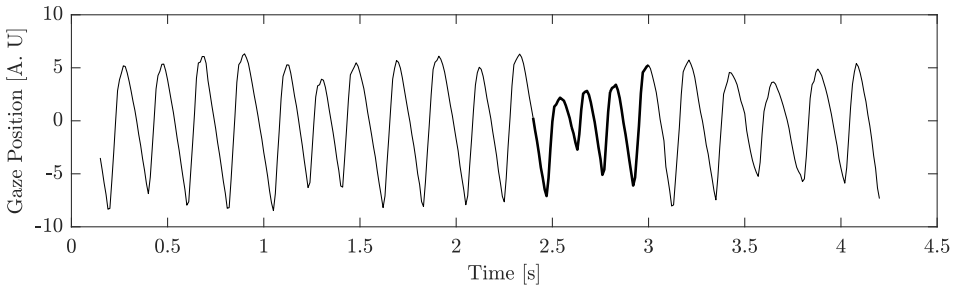


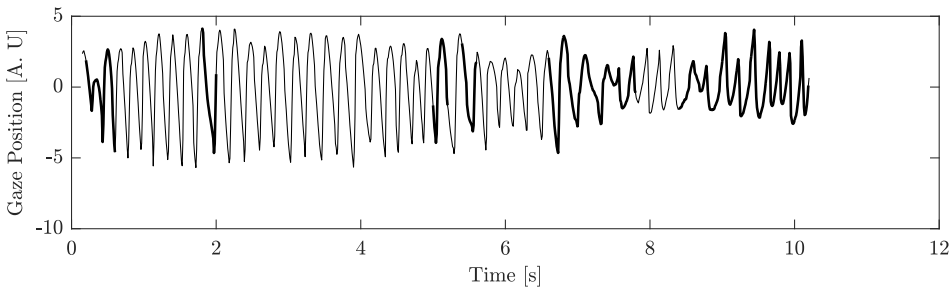
Figure 9: Example Reconstruction of the Participant dataset (PD_2) Waveforms. The reconstruction of the representative waveforms in Figure 4) for each of the five spatial positions. The x-axis represents time and the y-axis represents the position of the fixation targets where the data were recorded.

In Figure 10, examples of rejected cycles from the PD_1 and PD_2, are presented. The thick full line intervals in the signals represent segments that are rejected from further analysis. Figure 10a illustrates that segments that are too non-stationary are rejected from analysis. The same type of pattern is found in Figure 10b. In Figure

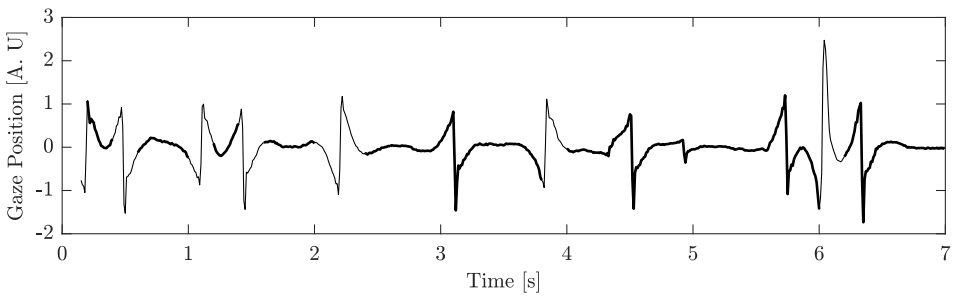
10c, most of the signal is rejected. This is likely due to a too low frequency of the oscillation that is highly affected by the highpass filter used in the preprocessing stage.



(a) Low Rejection Rate, 15 %



(b) Medium Rejection Rate, 37 %



(c) High Rejection Rate, 75 %

Figure 10: Examples of different rejection rates. Three different levels of rejection rates are presented in Figures 10a) - 10c). The rejected segments are plotted as thick lines. Note that the length of the different signals are varies. The title states the rate of rejection for each signal.

6 Discussion

In this work, we have presented a method for assessment of signal quality, referred to as the NSE, for use in eye tracking recordings from nystagmus patients. The NSE is used to exclude data segments that are undesirable to model due to for example recording artefacts. The nystagmus oscillations are modelled using a harmonic sinusoidal model. The signal is divided into segments and the NSE is assessed for each segment where a high NSE value suggests that the segment should be excluded from further analysis. The method is validated by analysis of simulated signals, which have been synthesised from previously reported waveform templates [5]. Various frequency and amplitude modulations as well as abrupt waveform changes and noise, have been introduced and the presented method is evaluated based on its ability to identify these segments. The accuracy of the frequency and amplitude estimations have been evaluated for the simulated signals. Finally, model parameter estimation has been evaluated on recordings from nystagmus patients.

There are four aspects of the method presented in this work that are of significance for the modelling of nystagmus oscillations. First, the NSE allows automatic detection of segments that should be rejected from further analysis. As illustrated in Figure 6 and by the ROC computations, the NSE is able to capture segments that represent the underlying nystagmus oscillation, while rejecting segments that do not contain an oscillatory signal. Note that the signal quality and NSE discussed in this work is different compared to the *waveform quality* presented in [6]. In that work, waveform quality is related to the results obtained from a visual acuity test for different nystagmus waveforms.

Second, the NSE allows the users of this method to compare different segments to each other, and use the segments that are of 'high quality', .i.e., that has a low NSE. This metric may be used to rate different segments in terms of their quality, and focus the analysis to the segments with the highest quality. Third, the harmonic model has proved itself to be useful when reconstructing waveforms. As can be seen in Figure 9, the presented model is able to capture differences in the waveforms of the signals observed in Figure 4. This suggests that the method can be used as a diagnostic tool and also suggests that the method can be used for evaluation of treatment effects.

Finally, the method does not need calibrated data in order to work. This is a great advantage, since calibration of nystagmus patients is often hard [17], and the results of calibration are often unreliable. The method may also be used as a preprocessing step in order to find segments with a high signal quality to be used for calibration.

The frequency estimation results that are presented in Section 5.1 suggest that inaccurate frequency estimation for segments with high NSE is 35 %, and at the same time, that almost all (99.7 %) of the segments with a low NSE provide accurate frequency estimates. This is important since the method relies on accurate frequency

estimations in order to work properly. The filtering method presented in equation (6) was developed specifically for nystagmus signals. In PD_1 and PD_2, the frequency range of most oscillations were between 3 and 6 Hz. It is, however, possible to find nystagmus oscillations with frequencies outside this range [2]. A possible alternative to the filter bank approach presented in this work, was presented in [23]. This method was developed for electrocardiogram (ECG) analysis and would likely need to be adjusted for nystagmus analysis.

The results of the ROC analysis suggest that the NSE assessment works well to separate oscillatory segments from corrupt segments. In this work, the optimum ROC value was estimated by computing the distance to the top left corner of the ROC coordinate system. The optimisation of ϵ may, however, be performed to maximise other desirable features. For example, in order reduce the number segments corrupted by noise considered for analysis, one may want to use a lower ϵ. This comes to the cost of fewer segments accepted for analysis and of more excluded segments with nystagmus oscillations.

The waveform reconstruction of the data presented in Figure 4, and illustrated in Figure 9, captures a representative waveform. By using only three harmonics (five features) it is possible to capture the main properties of the waveform. In this case, such properties include the presence of a foveation period, different frequency for the various waveforms and the direction of the fast phase. The reconstruction method used in this work is based on the assumption that there is only one type of waveform (according to the classification by [5]) recorded at each fixation target, which is not necessarily true. In order to use this method for reconstruction of recordings with multiple waveforms, an alternative waveform reconstruction approach would be needed.

The examples representing the estimated R_1 and $\delta_{\phi''2}$ parameters as polar coordinates, Figures 7 and 8, illustrate that ϵ can be used to separate oscillations from disturbances. There are a few things to take into consideration when evaluating the performance of the presented method on real data. First, it is not possible to know what the true parameter values should be. In order to perform an evaluation of the method on real signals, some assumptions are required. In this case, the assumption is that for a recording of a specific fixation target, the model parameters should be reasonably close to each other for the accepted segments, and they should be separated for the rejected segments. Second, the rejection of segments should be justified. For example, if only one segment is accepted, and all other segments are rejected, then the spread of the parameter values for the accepted segments would be zero. However, this would likely not be useful, since there would be no data left to analyse. As is illustrated in Figure 7, the parameter values of the accepted segments (blue circles) are more concentrated to a certain position inside the unit circle compared to the parameter values of the rejected segments (red squares). As illustrated in Figure

10, the rejection of segments is focused on finding segments in the signal where the waveform morphology is difficult to define, e.g., due to a change in waveform morphology. Although there are cases of nystagmus where the waveform changes, e.g., periodic alternating nystagmus (PAN), the time it takes for a waveform to change in PAN is usually several minutes [2]. The ability to detect waveform changes is determined by the segment length, N_b, which in this work equals 0.67 s, and the ϵ-value. As can be observed in Figure 10b, there are multiple waveforms embedded in the signal, and the model is able to capture some of the changing waveforms, although some are rejected. If a shorter segment length is chosen, a better waveform change resolution is obtained, however, this will lead to a decrease in frequency estimation performance. If the ϵ is increased, waveform changes will be easier to detect, but at the expense of a higher inclusion rate of non-nystagmus segments. Depending on the desired application of the method proposed in this work, the values of N_b and ϵ could vary from the suggestions made in this work.

There are some limitations in terms of the simulated signals generated for the ED. The reconstruction of these waveforms from the original work [5] introduced some artefacts, especially ringing. This is observed for the EF-R waveform in Figure 2. While this may look like a deviation from the original illustration, these waveforms still serve their purpose for modelling. One of the important characteristics of the nystagmus cycle are whether a foveation period exists and the proportion of this period relative to the cycle duration. These characteristics are well captured by the modell presented in this work.

Overall, this method can be used to model and detect changes in waveform morphology, which may be useful information when evaluating different treatment strategies for nystagmus patients, as well as for comparison of eye movement patterns between different patients.

7 Conclusion

The normalised segment error (NSE) and the normalised waveform model (NWM) have been shown to capture the most important features of the nystagmus oscillations. The NSE produced good results for the simulated signals, and the NWM is useful for describing real nystagmus recordings. The method presented in this work may be used to model entire nystagmus signals, to be used as a preprocessing step for other models or as a tool to classify different nystagmus waveforms.

Acknowledgements

This work has been funded by the Swedish Research Council [grant number VR 2015-05442]. We gratefully acknowledge the Lund University Humanities Lab. We would also like to thank all the participants. None of the data reported in this work are publicly available, and no experiments were preregistered.

References

[1] N. Sarvananthan, M. Surendran, E. O. Roberts, S. Jain, S. Thomas, N. Shah, F. A. Proudlock, J. R. Thompson, R. J. McLean, C. Degg, *et al.*, "The prevalence of nystagmus: the leicestershire nystagmus survey," *Investigative ophthalmology & visual science*, vol. 50, no. 11, pp. 5201–5206, 2009.

[2] R. J. Leigh and D. S. Zee, *The neurology of eye movements*. Oxford University Press, USA, 2015.

[3] N. Hussain, "Diagnosis, assessment and management of nystagmus in childhood," *Paediatrics and Child Health*, vol. 26, no. 1, pp. 31–36, 2016.

[4] R. McLean, F. Proudlock, S. Thomas, C. Degg, and I. Gottlob, "Congenital nystagmus: Randomized, controlled, double-masked trial of memantine/gabapentin," *Annals of Neurology*, vol. 61, no. 2, pp. 130–138, 2007.

[5] L. Dell'Osso and R. Daroff, "Congenital nystagmus waveforms and foveation strategy," *Documenta Ophthalmologica*, vol. 39, no. 1, pp. 155–182, 1975.

[6] L. Dell'Osso and J. Jacobs, "An expanded nystagmus acuity function: intra-and intersubject prediction of best-corrected visual acuity," *Documenta Ophthalmologica*, vol. 104, no. 3, pp. 249–276, 2002.

[7] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer, *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford, 2011.

[8] D. S. Broomhead, R. A. Clement, M. R. Muldoon, J. P. Whittle, C. Scallan, and R. V. Abadi, "Modelling of congenital nystagmus waveforms produced by saccadic system abnormalities," *Biological cybernetics*, vol. 82, no. 5, pp. 391–399, 2000.

[9] O. Akman, D. Broomhead, R. Abadi, and R. Clement, "Eye movement instabilities and nystagmus can be predicted by a nonlinear dynamics model of the saccadic system," *Journal of mathematical biology*, vol. 51, no. 6, pp. 661–694, 2005.

[10] M. Theodorou and R. Clement, "Classification of infantile nystagmus waveforms," *Vision Research*, vol. 123, pp. 20–25, 2016.

[11] M. Hosokawa, S. Hasebe, H. Ohtsuki, and Y. Tsuchida, "Time-frequency analysis of electronystagmogram signals in patients with congenital nystagmus," *Japanese journal of ophthalmology*, vol. 48, no. 3, pp. 262–267, 2004.

[12] R. Abadi, D. Broomhead, R. Clement, J. Whittle, and R. Worfolk, "Dynamical systems analysis: a new method of analysing congenital nystagmus waveforms," *Experimental brain research*, vol. 117, no. 3, pp. 355–361, 1997.

[13] C. M. Harris and D. L. Berry, "A distal model of congenital nystagmus as nonlinear adaptive oscillations," *Nonlinear Dynamics*, vol. 44, no. 1-4, pp. 367–380, 2006.

[14] R. Abadi and R. Worfolk, "Harmonic analysis of congenital nystagmus waveforms," *Clinical vision sciences*, vol. 6, no. 5, pp. 385–388, 1991.

[15] J. Felius, V. L. Fu, E. E. Birch, R. W. Hertle, R. M. Jost, and V. Subramanian, "Quantifying nystagmus in infants and young children: relation between foveation and visual acuity deficit," *Investigative ophthalmology & visual science*, vol. 52, no. 12, pp. 8724–8731, 2011.

[16] Y.-H. Young and T.-W. Huang, "Role of clonazepam in the treatment of idiopathic downbeat nystagmus," *The Laryngoscope*, vol. 111, no. 8, pp. 1490–1493, 2001.

[17] W. Rosengren, M. Nyström, B. Hammar, and M. Stridh, "A robust method for calibration of eye tracking data recorded during nystagmus," *Behavior research methods*, pp. 1–15, 2019.

[18] M. Dunn, *Quantifying perception and oculomotor instability in infantile nystagmus*. PhD thesis, Cardiff University, 2014.

[19] M. Stridh, D. Husser, A. Bollmann, and L. Sörnmo, "Waveform characterization of atrial fibrillation using phase information," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1081–1089, 2009.

[20] L. Sörnmo and P. Laguna, *Bioelectrical signal processing in cardiac and neurological applications*, vol. 8. Academic Press, 2005.

[21] SR-Research, "*EyeLink 1000 User Manual*.," 2010.

[22] J. W. Peirce, "Psychopy psychophysics software in python," *Journal of Neuroscience Methods*, vol. 162, no. 1-2, pp. 8–13, 2007.

[23] A. Buttu, E. Pruvot, J. Van Zaen, A. Viso, A. Forclaz, P. Pascale, S. M. Narayan, and J.-M. Vesin, "Adaptive frequency tracking of the baseline ecg identifies the site of atrial fibrillation termination by catheter ablation," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 969–980, 2013.

Paper III

Waveform characterisation and comparison of nystagmus eye-tracking signals

Abstract

Pathological nystagmus is a symptom of oculomotor disease where the eyes oscillate involuntarily. The underlying cause of the nystagmus and the characteristics of the oscillatory eye movements are patient specific. An important part of clinical assessment in nystagmus patients is therefore to characterize different recorded eye-tracking signals, i.e., waveforms.

A method for characterisation of the nystagmus waveform morphology is proposed. The method extracts local morphologic characteristics based on a sinusoidal model, and clusters these into a description of the complete signal. The clusters are used to characterise and compare recordings within and between patients and tasks. New metrics are proposed that can measure waveform similarity at different scales; from short signal segments up to entire signals, both within and between patients.

The results show that the proposed method robustly can find the most prominent nystagmus waveforms in a recording. The method accurately identifies different eye movement patterns within and between patients and across different tasks. In conclusion, by allowing characterisation and comparison of nystagmus waveform patterns, the proposed method opens up for investigation and identification of the underlying condition in the individual patient, and for quantifying eye movements during tasks.

Based on
William Rosengren, Marcus Nyström, Björn Hammar and Martin Stridh
"Waveform characterisation and comparison of nystagmus eye-tracking signals,"
Under review

1 Introduction

Nystagmus is a condition affecting the ability to control the eyes, and is manifested as uncontrolled oscillating eye movements [1]. Characteristics of nystagmus include an inability to fixate the gaze, which may lead to oscillopsia and decreased visual acuity [1]. Nystagmus may also lead to poor visual accommodation, which is a phenomena commonly seen in patients with ocular albinism [2]. Patients suffering from nystagmus may also show decreased reading speed [3]. Pathological nystagmus is usually divided into two categories: *early onset* and *acquired* nystagmus [4, 5]. The former is developed before or shortly after birth, and acquired nystagmus is developed due to external factors, such as substance abuse or neurological disease. There are a variety of pathological explanations for nystagmus, and by studying the eye movements of individual patients it is possible to improve diagnostics and more accurately determine the underlying condition [1, 6]. Nystagmus diagnostics, however, is by many clinicians considered difficult, and careful examination of the symptoms is required in order to correctly diagnose each individual patient [4, 7].

Analysis of nystagmus eye movements patterns is difficult to conduct due to various factors. Some nystagmus eye movement patterns are similar to each other, for example the asymmetric pendular and jerk - right waveforms are often mistaken for each other [8]. In addition, the oscillation patterns in some patients changes over time, for example periodic alternating nystagmus [1]. Because of these difficulties, strategies to aid clinicians working with nystagmus patients have been developed. *Eye tracking* has been used to record eye movements of nystagmus patients. An eye tracker is a device for capturing eye movements, and video–based eye trackers are the ones most commonly used [9]. The main advantages of using eye-tracking signals instead of manual observations of eye movements are that the eye tracker provides more objective measurements, that the signals contain details the human eye cannot perceive, and that the resulting data can be stored and analysed offline using various algorithms.

A common approach to analyse nystagmus oscillations has been to study the eye movement pattern, known as *waveform*, of the recorded eye movements signals [6, 10, 11]. This eye movement pattern can further be descibed by its *waveform morphology*, which is the detailed shape of the waveform within each *oscillation cycle*. The oscillation cycle is one period of a nystagmus oscillation in the waveform signal. Typically, the data in these studies have been recorded from participants with stabilised heads looking at points in a static world, i.e., performing a fixation task. An early method for manual characterisation of the nystagmus waveform morphology resulted in a description of eighteen different morphologies [8]. Mapping of signals onto a sine and a sawtooth waveform has been proposed for classification of nystagmus waveform morphologies [11]. A sum of harmonic sinusoids has also been used to model nystagmus waveforms [12].

Some studies investigate waveforms recorded during *smooth pursuit* [13, 14]. Analysis of smooth pursuit eye movements has been suggested to improve the understanding of the underlying cause of nystagmus [15]. Assessments of the pursuit gain, i.e., the ratio between the eye velocity and the target velocity, found that it was comparable to recordings from healthy controls [6, 13]. This analysis was based on the *superposition model*, which states that during smooth pursuit, the nystagmus is added to the underlying smooth pursuit eye movement [6].

When nystagmus develops, the eye movement pattern may change over time. For example, it has been reported that the development of nystagmus during infancy results in changes in the waveform over the course of weeks and months, and that this development may differ between patients [16]. Detecting differences in waveform morphologies for different gaze angles, or during different pursuit directions is an important step when identifying individual characteristics of nystagmus [1]. Comparison of eye movements before and after treatment can be used to evaluate the individual response to treatment.

Previous methods for analysing smooth pursuit signals have focused on differences in pursuit gain between nystagmus patients and healthy controls [6, 13]. However, little attention has been directed towards the analysis of nystagmus waveform morphologies during smooth pursuit. According to the superposition model, smooth pursuit in nystagmus consist of two components [6], which results in a more complex analysis of smooth pursuit compared to analysis of attempted fixations.

In this work, a sinusoidal model referred to as the normalised waveform model [17] is used to model the nystagmus waveforms. Each waveform is divided and modelled in short segments and the model parameters are then clustered together. The clustering is done in relation to predefined template waveform morphologies and requires the use of a waveform distance measure. For this purpose, a *waveform distance index* is designed based on the features of the normalised waveform model, and is used to measure the similarity between different waveform morphologies.

2 Methods

In this work, nystagmus eye movements from participants viewing both fixation stimuli and smooth pursuit stimuli are analysed. In order to analyse signals from the smooth pursuit tasks, the signals are subject to smooth pursuit stimuli subtraction, see Section 2.1. After the smooth pursuit stimuli subtraction, the eye movement signals are analysed using the normalised waveform model. Based on the model, the waveform morphology classification and the waveform distance index are derived and presented in Section 2.2. The observed waveform morphologies are then compared to the template morphologies described in [8], in order to create a *morphological trend*

and a *morphological profile* for each recorded nystagmus signal. The morphological trend and morphological profile are a time series and the distribution of the different waveform morphologies present in a single signal, respectively. The most frequent waveform morphology in each morphological profile is referred to as the typical waveform. The process to create the morphological trend and profile is referred to as *template clustering*, and is described in Section 2.3. The morphological trends from two different signals are compared using a method referred to as *signal segment matching*, presented in Section 2.4. A block diagram of the proposed analysis steps is shown in Figure 1.

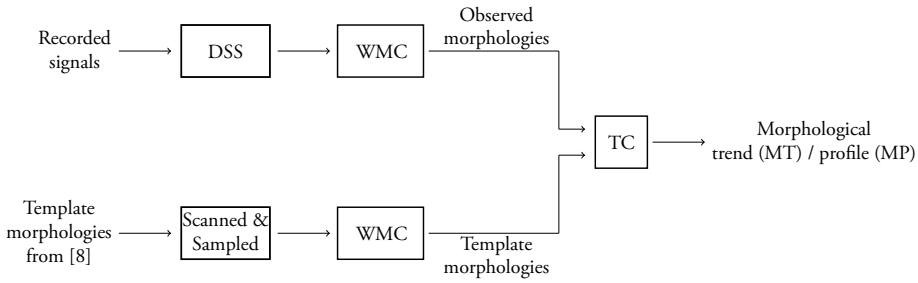


Figure 1: Scheme of the proposed method. The signals recorded from nystagmus patients are subject to the smooth pursuit stimuli subtraction (SmPSS). These signals are then fed into the waveform morphology characterisation (WMC) [17], resulting in the observed waveform morphologies. These morphologies are compared with the template morphologies using the proposed template clustering (TC) method in order to create a morphological trend (MT) and a morphological profile (MP). The most frequent waveform for each morphological profile is called the typical waveform (TW). The template waveforms have been created from the morphologies in [8]. These were scanned and sampled from the original publication, and characterised using Fourier analysis [17].

2.1 Smooth pursuit stimuli subtraction

The superposition model describes the relationship between eye movements and nystagmus waveform during smooth pursuit. In this work, the smooth pursuit component, $s_{smp}[n]$, is estimated and subtracted from the recorded signal, $s_{eye}[n]$ in order to estimate the nystagmus waveform, $s_{nys}[n]$. In order to only capture eye movements related to nystagmus, it is important to remove the smooth pursuit stimuli component from the eye movement signal. The superposition model is defined as

$$s_{eye}[n] = s_{smp}[n] + s_{nyst}[n] \tag{1}$$

where $s_{eye}[n]$ is the recorded eye movement, $s_{smp}[n]$ is the smooth pursuit component, $s_{nys}[n]$ is position component of the nystagmus eye movements and $[n]$ is the sample index. If the frequency of the nystagmus oscillations is higher than the smooth pursuit frequency, $s_{smp}[n]$ is estimated using a lowpass filter and the resulting nystagmus component is estimated as

$$\hat{s}_{nys}[n] = s_{eye}[n] - \hat{s}_{smp}[n]. \tag{2}$$

where $\hat{s}_{nys}[n]$ and $\hat{s}_{smp}[n]$ are the estimations of the nystagmus position and smooth pursuit position components, respectively.

2.2 Waveform morphology characterisation

The model used to describe the various nystagmus waveform morphologies is referred to as the *normalised waveform model* [17]. The normalised waveform model is used to parametrise the different observed nystagmus oscillations, as well as to assess the quality of the observed signal. In short, this model is designed to describe short segments of nystagmus waveforms as a sum of sinusoidal signals, i.e., a harmonic series. The model is stated as [17],

$$s_m[n] = \sum_{h=1}^{H} s_h[n], \tag{3}$$

where

$$s_h[n] = a_h \sin[2\pi(f_1 h)n + \phi_h], \tag{4}$$

and a_h, $f_1 h$ and ϕ_h are the amplitude, frequency and phase of each harmonic h, respectively. The advantage of this model is that it allows for a normalisation of the amplitude, frequency and phase, leading to a robust description of different signal segments [17]. A metric called the normalised segment error (NSE) has previously been introduced to evaluate the goodness of fit between the observed signal and the signal model. If the goodness of fit is poor, it implies that the data are not oscillatory, which means that it is unlikely that the modelled segment is a nystagmus waveform. The definition of the NSE is given by,

$$NSE_s = \frac{\sum_{n_s=n_0}^{n_0+N_s-1} |\bar{s}[n_s] - \bar{s}_m[n_s]|^2}{\sum_{n_s=n_0}^{n_0+N_s-1} |\bar{s}[n_s]|^2}, \tag{5}$$

where $s[n]$ and $s_m[n]$ are the recorded signal and the signal model, respectively. The mean subtracted signals $\bar{s}[n]$ and $\bar{s}_m[n]$ are computed as

$$\bar{s}[n] = s[n] - \frac{1}{N_s} \sum_{n_s=n_0}^{n_0+N_s-1} s[n_s] \tag{6}$$

and

$$\bar{s}_m[n] = s_m[n] - \frac{1}{N_s} \sum_{n_s=n_0}^{n_0+N_s-1} s_m[n_s], \tag{7}$$

respectively, and N_s is the length of the analysed segment [17].

Although the model in (3) is useful for reconstructing and normalising various waveforms, it is complicated to use its parameters to compute distances between two waveform morphologies. This is due to the structure of the harmonic series, where the resulting waveform morphology is dependent on the relative amplitude, frequency and phases of each harmonic. Since the phases are circular properties, and their influence on the resulting waveform morphology is linked to the corresponding amplitude, it is difficult to use the model features directly for waveform similarity measurements. In order to solve this problem, the nystagmus signal model is reconstructed in the following way: First the signal amplitude and frequency are normalised by multiplying the signal with a factor $\frac{1}{a_1}$, and by setting the frequency $f_1 = 1$ Hz. The relative phase between the first harmonic phase and all other harmonic phases, δ_h'', are computed as [18]

$$\hat{\delta}_{\phi_h''} = \hat{\phi}_h'' - \hat{\phi}_1', \tag{8}$$

where $\hat{\phi}_h''$ is computed as [18],

$$\hat{\phi}_h'' = \frac{2\pi(\hat{f}_h - \hat{f}_1 h)\frac{N_s}{2} + \hat{\phi}_h}{h}, \tag{9}$$

h is the harmonic index and $\hat{f}_h$ and $\hat{\phi}_h$ are the frequency and phase estimations for each harmonic. After these steps, all signals have a first harmonic phase of 0 and first harmonic amplitude of one. This means that the first harmonic of the signal, $\sin(2\pi f_1 n)$, is identical for all waveforms and is therefore omitted from the analysis. For the other harmonics, they may be rewritten as

$$\frac{a_h}{a_1} \cdot \sin(\omega_h n + \hat{\delta}_{\phi_h''} h) = R_s(h) \cdot \sin(\omega_h n) + R_c(h) \cdot \cos(\omega_h n) \tag{10}$$

where the solutions to $R_s(h)$ and $R_c(h)$ are obtained by:

$$R_s(h) = \frac{a_h}{a_1} \cdot \cos(\hat{\delta}_{\phi''_h} h) \tag{11}$$

$$R_c(h) = \frac{a_h}{a_1} \cdot \sin(\hat{\delta}_{\phi''_h} h), \tag{12}$$

and $\omega_h = 2\pi h$. The waveform distance index is based on the adjusted model features $R_s(h)$ and $R_c(h)$ of four harmonics, $h = [2, 3, 4, 5]$. Thus, using these eight values, a vector $\boldsymbol{\theta}^{1\times 8}$ is created, where the elements are the four pairs of coefficients:

$$\boldsymbol{\theta} = \begin{bmatrix} R_s(2) & R_c(2), & \ldots, R_s(5) & R_c(5) \end{bmatrix}^T, \tag{13}$$

The waveform distance index between two waveforms, with feature vectors $\boldsymbol{\theta}_m$ and $\boldsymbol{\theta}_n$, is then computed as

$$\mathrm{WDI}(\boldsymbol{\theta}_m, \boldsymbol{\theta}_n) = \sqrt{\sum_{i=1}^{8} \big(\boldsymbol{\theta}_m(i) - \boldsymbol{\theta}_n(i)\big)^2}. \tag{14}$$

If the two waveforms have a similar feature vector, the WDI between them will be close to zero. Since there is little energy above 50 Hz, all signals were downsampled to 100 Hz [17]. If the frequency of the nystagmus oscillation is too high, there may not be parameters values for higher order harmonics, in which case these parameter values are set to zero.

2.3 Template clustering

In order to categorise the observed nystagmus waveforms, a method called *template clustering* is introduced. This method uses the waveform distance index when comparing waveform morphologies and is based on a set of eighteen template waveform morphologies that previously have been used to classify waveform morphologies on a cycle–to–cycle basis [8]. These waveform morphologies have been digitised from the original publication and parametrised using Fourier analysis. In this work, the parametrised waveforms from [8] are used as non-orthogonal base functions, onto which each observed waveform is mapped using the waveform distance index.

Some of these waveforms are similar, which may cause problems when mapping the observed morphologies onto the template waveforms. In order to address this issue, an analysis of the correlation between the different template morphologies is performed. If the correlation is above 0.95, one of the waveforms is considered redundant and is removed from the set of base functions. In total, this leads to set of

elven morphologies considered in the template clustering stage: Asymmetric Pendular (AP), Bidirectional Jerk - Left (BDJ-L), Bidirectional Jerk - Right (BDJ-R), DJ-L, Dual Jerk - Right (DJ-R), Extended Foveation - Left (EF-L), EF-R, Pendular with Foveating Saccades (P_{FS}), Pseudo Cycloid - Left (PC-L) Pseudo Cycloid - Right (PC-R) and Pseudo Jerk - Left (PJ-L).

The template clustering is implemented in the following way: Each signal segment is mapped onto the bases of the ten template morphologies described above using the waveform distance index. In order to reflect the observed data, a typical waveform morphology, $\bar{\boldsymbol{\theta}}_i$, for each template cluster i, is computed. The typical waveform is used for outlier detection, where feature vectors that deviate too much from the average feature vector are considered as outliers. In order to determine the threshold for an acceptable deviation from the typical waveform, bootstrap statistics are utilised. The steps for creating the typical waveform and the outlier detection protocol is based on bootstrapping and is presented in Figure 2 and listed below:

1. First, the average of the feature vectors of the original observations, $\bar{\boldsymbol{\theta}}_o$, is computed for each template category i.

2. For each of these categories, the feature vectors are sampled with replacement from the original observations N times, where N is equal to the number of observations.

3. After each iteration, the typical waveform, $\bar{\boldsymbol{\theta}}_b$, is computed from the N feature vectors, and this process is iterated 1000 times.

4. The highest accepted waveform distance index is computed as the largest waveform distance index between the set of 1000 $\bar{\boldsymbol{\theta}}_b$ vectors and the $\bar{\boldsymbol{\theta}}_o$. If the waveform distance index between any observed waveform feature vector and the typical waveform $\bar{\boldsymbol{\theta}}_o$ is greater than this waveform distance index value, the observed waveform is considered an outlier.

5. The final typical waveform is computed as the mean feature vector using only the inliers.

For those template waveforms where no observed waveforms has been assigned, the original template waveform morphology is used instead. Once this process is finished, all signal segments are assigned to one of the ten clusters based on the waveform distance index. The morphological trend is the time series of morphology classes, where each class is an integer between 1 and 11, corresponding to the typical waveform. The morphological profile is the histogram of the morphological trend.

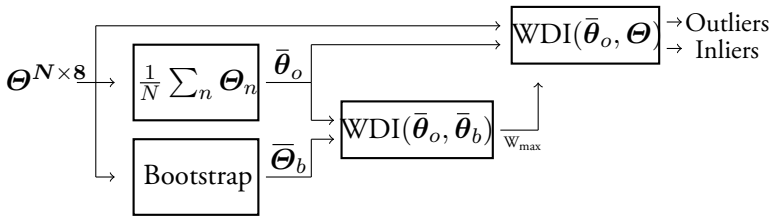


Figure 2: Flow chart of the outlier rejection process. The $\boldsymbol{\Theta}^{N\times 8}$ matrix contains feature vectors for each waveform class. The features in the matrix are used to compute the $\bar{\boldsymbol{\theta}}_o$, which is used to classify each individual feature vector as either an outlier or an inlier. The matrix $\bar{\boldsymbol{\Theta}}_b$ represents the mean feature vectors from the bootstrapping procedure.

2.4 Signal segment matching

For some signals, especially the smooth pursuit recordings, there are subsegments within entire signals that are similar to other subsegments, both within the same signal, as well as subsegments of other signals. Analysing these similarities may show patterns that are of interest when, for example, comparing smooth pursuit recordings to fixation recordings. In order to conduct this analysis, it is desirable to develop a method for comparison of subsegments from different recordings.

For this purpose, a method called signal segment matching is proposed. The method is based on sampling the morphological trend, where each sample is an integer between 1 and 10, representing each respective template class. The four sample subsegment is then compared to another four sample subsegment either from the same signal or from a different signal. Using the normalised waveform model, four samples from the morphological trend corresponds to 2 consecutive seconds from a signal [17]. If at least three (chosen empirically) out of the four consecutive samples of the two morphological trends have equal values, these two signal subsegments are considered to consists of the same waveform morphology pattern. For example, the two sequences [1 1 3 2] and and [1 1 2 4] would be considered as a match. For the analysis in this work, the samples are analysed using a 50 % overlap between the subsegments, i.e., the signal is shifted one second. Note that the order of the samples within each subsegment is not considered here.

3 Stimuli and data description

This section contains a description of the equipment used for data recording, see Sec. 3.1, and the various eye movement tasks performed during the experiments, see Sec. 3.2.

3.1 Recording equipment

The data for the experiment in this work were recorded using the EyeLink 1000 Plus eye tracker (SR Research, Ottawa, ON, Canada) with binocular uncalibrated pupil data and corneal reflection data sampled at 1 kHz. The eye tracker was used in desktop mode and the configuration of the eye tracker, stimuli computer screen and participant seating were in accordance with the manufacturer manual [19]. The eye tracker used the host software v. 5.09 and the DevKit v. 1.11.571 and the software for presenting the experiments was written in PsychoPy (version 1.83) [20]. All stimuli were presented on an ASUS VG248QE monitor with 1920×1080 pixels resolution and with dimensions 53×30 cm. A chin and forehead rest were used to stabilise the head for all participants. The study was performed in accordance with the Declaration of Helsinki and was approved by the ethics committee at Lund University. All participants signed an informed consent form before the experiment started.

3.2 Fixation and smooth pursuit data data

Fixation data and smooth pursuit data were recorded both for control participants and nystagmus participants. Fixation data from the control participants were not used in this work. The data were recorded at five different fixation positions: $(0°, 0°)$, $(0°, -8°)$, $(0°, 8°)$, $(16°, 0°)$ and $(-16°, 0°)$, where the center of the computer screen is defined as $(0°, 0°)$.

Four smooth pursuit tasks were presented to the participants in this study. An illustration of the stimuli patterns are shown in Figure 3, and the settings are presented in Table 1. These stimulus patterns were: purely horizontal (SMP - H) with an increase in frequency or purely vertical (SMP - V) with an increase in frequency across the range $\in [0.4, 1]$ [Hz], a horizontally propagating sinusoid (sinus) and a rectangular moving stimulus (box) with constant propagation speed.

In order to evaluate the method presented in Section 2, two different datasets have been used:

- **NY-Fix**: Fixation data from nystagmus participants are added to the smooth pursuit data recorded from healthy control participants. This is a simulation of the superposition model described in equation (1), and serves as a control group when evaluating the proposed method.
- **NY-SMP**: In this dataset, repeated recordings $(n = 8)$ of a single participant are also included. The box and sinus tasks were recorded only for this participant. The repeated recordings are used to study the repeatability of the method. Both smooth pursuit and fixation data from nystagmus participants were included in this dataset.

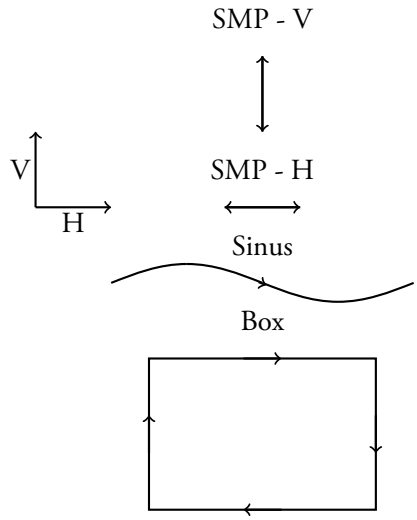


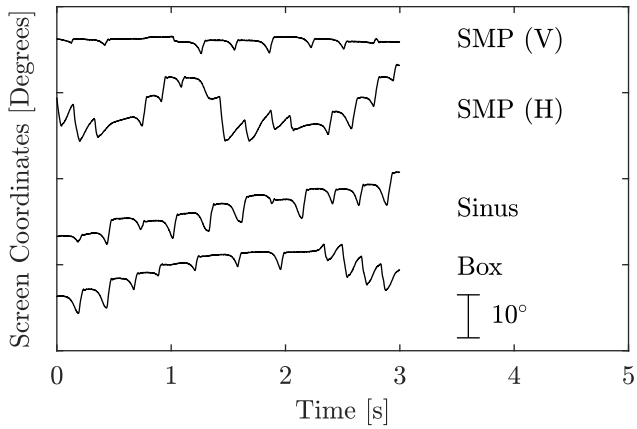
Figure 3: Examples of the four smooth pursuit stimuli movements presented to the participants. Each stimulus was generated by moving a red dot across the screen. The screen coordinate system to the left shows the horizontal direction (H) and the vertical direction (V).

A total of fourteen recordings were made for these six participants and the average age of the six nystagmus participants was 39.6 ($\pm$15.5) years. For the healthy controls, 9 participants were recorded a total of 17 times and the average age was 38.9 ($\pm$12.4) years. All nystagmus participants were male, four were diagnosed with infantile nystagmus, one with acquired nystagmus and one was unidentified. All participants, both controls and nystagmus participants, were calibrated using the calibration protocol proposed in [21]. Nine calibration targets and four validation targets were presented to each participant. The calibration mapping function was the linear Procrustes mapping function.

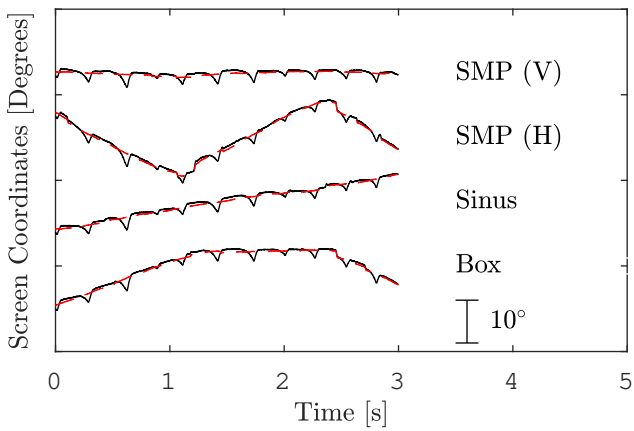
Examples of the recorded waveforms from the two datasets are presented in Figure 4. In Figure 4a, horizontal smooth pursuit data from one nystagmus participant for the four different tasks are presented. An example of simulated nystagmus smooth pursuit data that constitutes the **NY-Fix** dataset is presented in Figure 4b. The same nystagmus fixation data (from the left eye) have been added to four different smooth pursuit tasks from the same control participant.

Table 1: Description of the four smooth pursuit tasks presented to the participants. The variables x_{min}, x_{max}, y_{min} and y_{max} describe the horizontal and vertical minimum and maximum screen coordinates for each smooth pursuit task in degrees. The frequencies f_{min} and f_{max} states the minimum and maximum frequencies of the smooth pursuit stimuli for each task.

	SMP - V Box	SMP - H	Sinus	Box
x_{min}	$0°$	$-8°$	$-20°$	$-10°$
x_{max}	$0°$	$8°$	$20°$	$10°$
y_{min}	$-8°$	$0°$	$-7°$	$-5°$
y_{max}	$8°$	$0°$	$7°$	$5°$
f_{min} [Hz]	0.4	0.4	0.12	-
f_{max} [Hz]	1	1	0.12	-
Duration [s]	23.5	23.5	8.33	12.5



(a) Nystagmus pursuit waveforms



(b) Simulated nystagmus smooth pursuit

Figure 4: In 4a), an illustration of different eye movement patterns observed when nystagmus participants pursue moving targets during recording of nystagmus smooth pursuit data. In 4b), illustration of simulated nystagmus smooth pursuit data. The simulated signals were created by combining the nystagmus fixation data and control smooth pursuit data according to the superposition model (1). The red dashed curve in 4b) represents the original smooth pursuit data recorded from a control participant. The bar in each figure represents an amplitude of 10°.

4 Evaluation of the proposed method

In this section, the evaluation strategies for the different parts of the proposed method are presented. First, in Section 4.1, the evaluation of the smooth pursuit stimuli subtraction method is described. In Section 4.2, the evaluation of the template clustering is described. Finally, the evaluation of the segment matching for smooth pursuit and fixation data is provided in Section 4.3.

4.1 Smooth pursuit stimuli subtraction

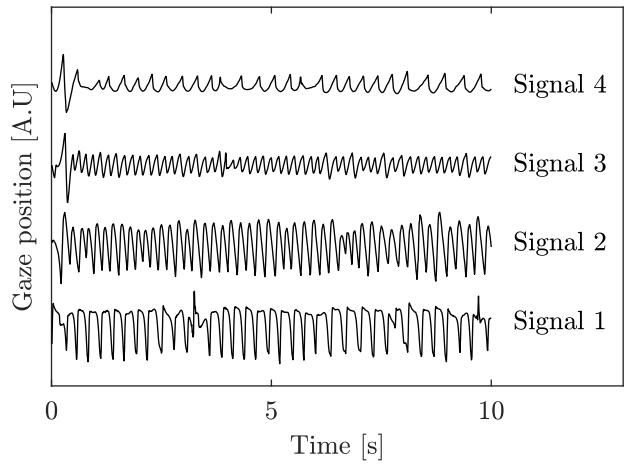
The purpose of the smooth pursuit stimuli subtraction is to reduce the influence of the smooth pursuit eye movement on the nystagmus waveform modelling. In order to evaluate the smooth pursuit stimuli subtraction as a tool for analysis of nystagmus eye movements recorded during smooth pursuit, the residuals from the stimuli subtraction method need to have a low impact on the overall waveform morphology. In order to determine the impact of the stimuli subtraction method, the waveform distance index between the estimated nystagmus fixation component, $\hat{x}_{nys}$, using the superposition model, and the known reference fixation component, x_{nys} from the **NY-Fix** datasets are computed for each of the four smooth pursuit tasks analysed in this study.

The evaluation consists of the following steps: Both the estimated nystagmus waveforms, $\hat{x}_{nys}$, and the original nystagmus waveforms, x_{nys}, are modelled using the normalised waveform model [17]. Since the nystagmus signal component from both these signals are identical, a high performance of the stimuli subtraction corresponds to a small waveform distance index between corresponding blocks in the two signals $\hat{x}_{nys}$ and x_{nys}. In order to determine the performance of the stimuli subtraction, the average waveform distance index is computed between the blocks of the x_{nys} signal and the corresponding blocks of the $\hat{x}_{nys}$ signal.

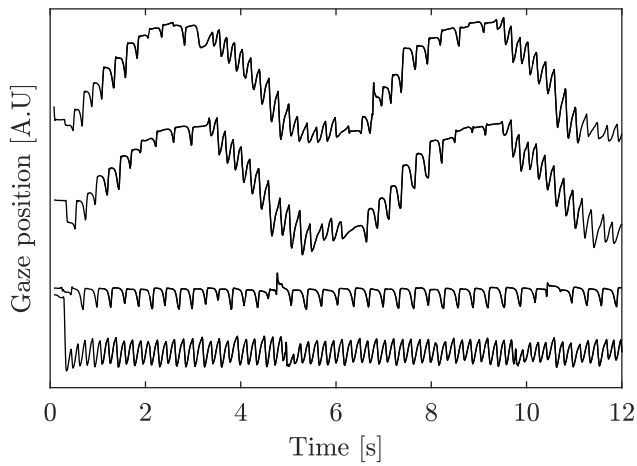
4.2 Template clustering

The template clustering method was developed with two requirements: First, different waveforms should be assigned to different template waveforms, i.e., different types of morphologies should be distinguishable. Second, the template clustering method should repeat classification of waveforms in signals where the waveforms do not change.

The evaluation of how the method distinguishes between waveform morphologies is based on four example signals, which are presented in Figure 5a. The presented signals are 10 second subsegments of longer recordings, and the waveform morphologies in each individual signal are representative for each the longer recording. In this



(a) Multiple participant signals



(b) Fixation and Box signals

Figure 5: In 5a), an illustration of four fixation signals from multiple participants, and in 5b), two repeated box signals and two fixation signals recorded at different gaze position from the same participant.

evaluation, the signals S1 - S4 are considered to consist of one waveform mixed with some noisy segments, and each of these four waveforms are considered to be different from each other. As may be observed from the data, there are some segments in each signal where there are no oscillations or no clear waveform morphology.

In order to evaluate the waveform distance index for these different signals, the

average waveform distance indices within and between all signals are computed. For each signal, the waveform distance index is computed between each segment and all other segments in the signal as well as all segments in the three other signals. The average is then used to represent the distance between the different signals.

In order to evaluate the consistency of the template clustering method, six repeated recordings from two fixation targets from the same participant are analysed. The signals were for fixation positions $(0°, 0°)$ and $(-16°, 0°)$, respectively. The same analysis of the waveform distance index as is described above is performed for the all the twelve signals recorded from positions $(0°, 0°)$ and $(-16°, 0°)$. The distributions of the average waveform index are compared for three different cases: The first is based on all the recordings from the first gaze angle, the second is based on all the recordings from the second gaze angle and the last is the distribution when comparing signals from the first gaze angle to the second gaze angle.

4.3 Signal segment matching

The signal segment matching is aimed at detecting subsegments in recorded signals that resembles subsegments in other signals. In Figure 5b two box signals and two fixation signals from the same participant are presented. In order to test the signal segment matching, two different cases are considered. First, the box stimuli data for two signals recorded at different times from the same participant are analysed. The second analysis compares the box data with fixation data recorded from two different positions: $(0°, 0°)$ and $(-16°, 0°)$, for the same participant.

5 Results

The results section is divided into four sections. First, the smooth pursuit stimuli subtraction performance is reported in Section 5.1. This is followed by an analysis of whether different waveform morphologies are distinguishable and whether repeated recordings are consistently similar, reported in Section 5.2. An example of the outlier rejection model is presented in Section 5.3. Finally, in Section 5.4, two examples of the signal segment matching method are presented.

5.1 Dynamic stimuli subtraction

In Figure 6, the results for the evaluation of the dynamic stimuli subtraction are presented. Here, corresponding to segments without an oscillatory pattern, data from segments with an NSE of more than 0.18 were omitted from the analysis. For all

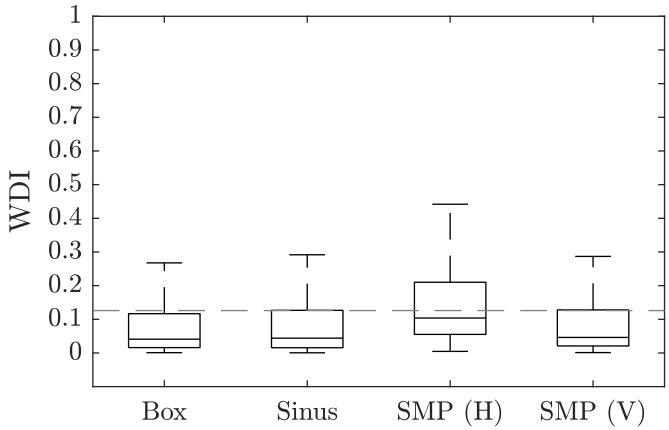
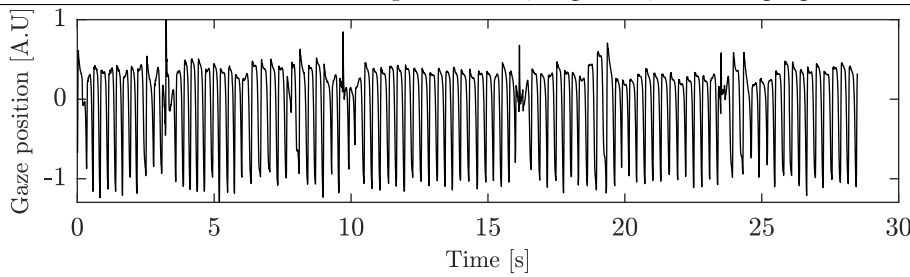


Figure 6: Distribution of waveform distance index between the simulated smooth pursuit signals and the corresponding known fixation signals. Each signal segment has been analysed using the normalised segment error, and segments that were above the threshold 0.18 were omitted from this analysis. The dashed gray line represents the threshold for when two waveform morphologies are considered to be the same (waveform distance index $= 0.13$).

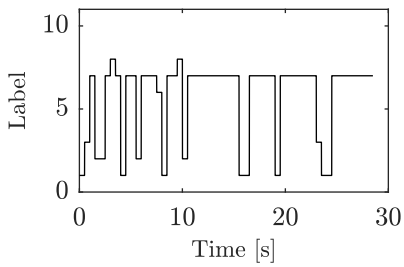
four smooth pursuit tasks, the median waveform distance index between the original and the estimated nystagmus signals are approximately 0.1 or below, which is lower than the waveform distance index between the two most similar template waveforms (P_{FS} and PP_{FS}), which is approximately 0.13. This threshold is plotted as a gray dashed line in the figure.

5.2 Template clustering

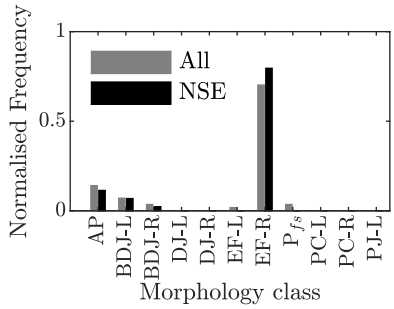
In Figure 7, the classification results for four different fixation signals are presented. In the leftmost column, the original signals are plotted. The middle column shows the morphological trends of the signal in the left column. In the rightmost column, the morphological profiles for each respective signal are presented. The gray bars represent all segments, whereas the black bars represent the segments with an NSE of 0.18 or below. As is illustrated in Figure 7, the four different signals in the left most column result in different typical waveforms.



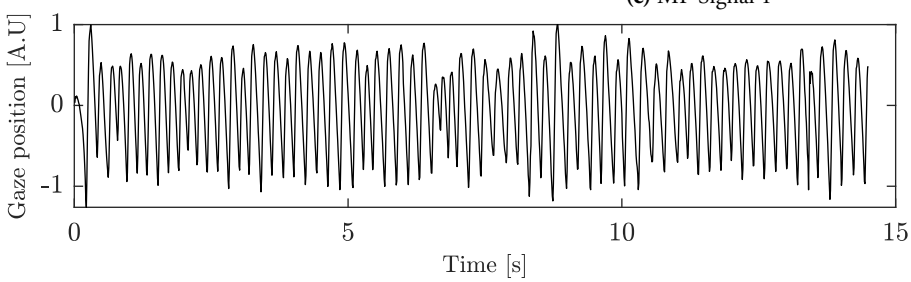
(a) Signal 1



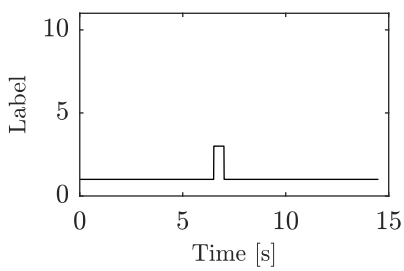
(b) MT Signal 1



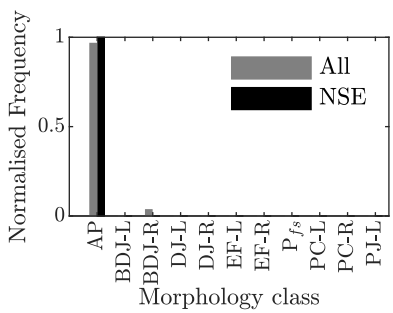
(c) MP Signal 1



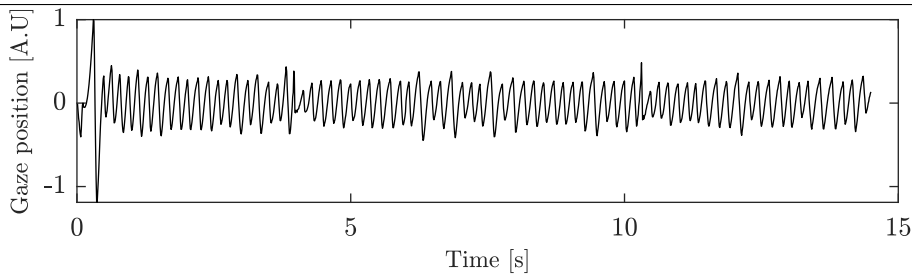
(d) Signal 2



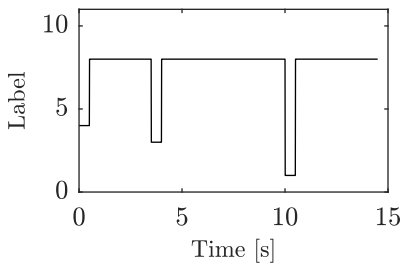
(e) MT Signal 2



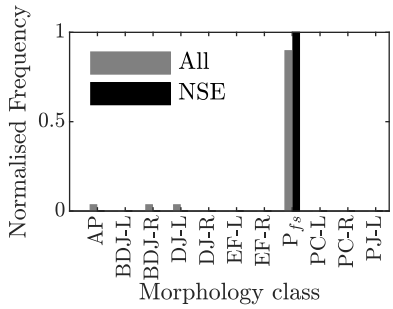
(f) MP Signal 2



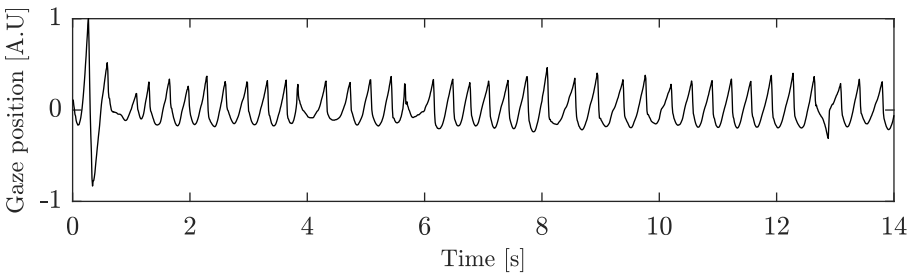
(g) Signal 3



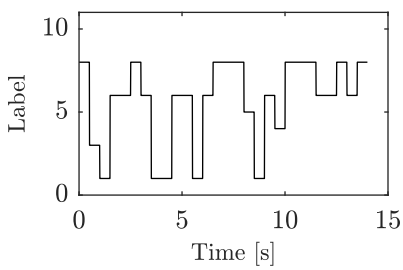
(h) MT Signal 3



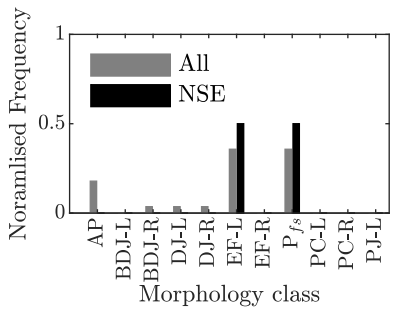
(i) MP Signal 3



(j) Signal 4



(k) MT Signal 4



(l) MP Signal 4

Figure 7: Illustration of the analysis of four different signals. The morphological trend (MT) is represented as a time series and the morphological profile (MP) is represented by a histogram, for each respective signal. The four signals S1 - S4 all have different morphological profiles.

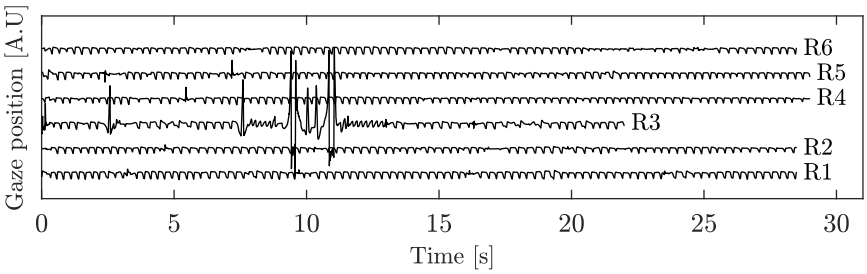
Table 2: Average waveform distance index within and between the four signals presented in Figure 5a. The normalised segment error has been used to omit poor quality signal segments. The bold numbers show the average internal waveform distance index for each respective signal.

Signal	1	2	3	4
1	**0.31**	0.54	0.75	0.88
2	0.54	**0.10**	0.51	0.58
3	0.75	0.51	**0.10**	0.36
4	0.88	0.58	0.36	**0.19**

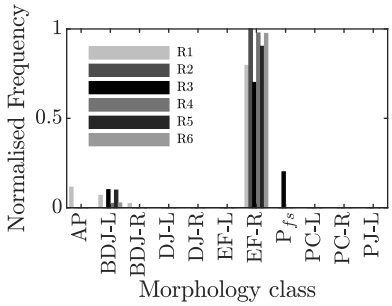
For example, the most common waveform class for the signal presented in Figure 7a is EF-R, which can seen in Figure 7c. Similarly, the most common waveform classes for the two signals in Figures 7d and 7g are AP and P_{fs}. For the signal presented in Figure 7j the waveforms EF-L and P_{fs} are equally occurring. In Table 2, the average waveform distance index within and between the four signals in Figure 5a are presented. The bold numbers on the diagonal represent the waveform distance index between all segment combinations within one signal. For all four signals, this is the lowest average waveform distance index when compared to all the other signals. This suggests that each respective signal is more similar to itself than to any of the other signals.

As described in Section 4.2, this result is expected and indicates that the the proposed method can used to identify different waveform patterns observed in nystagmus signals. Note that the normalised segment error is applied in order to remove poor quality signal segments.

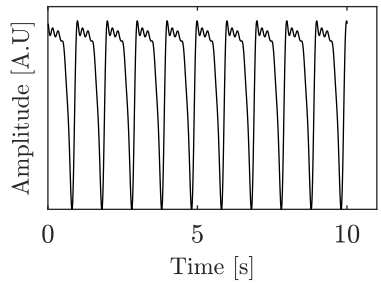
The results for the repeated recordings for one participant are presented in Figure 8.



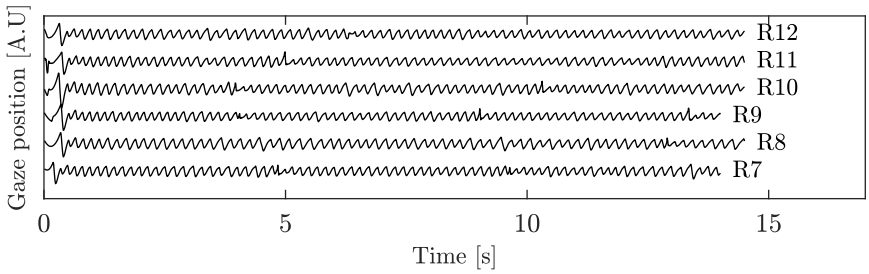
(a) Repeated $(0°, 0°)$



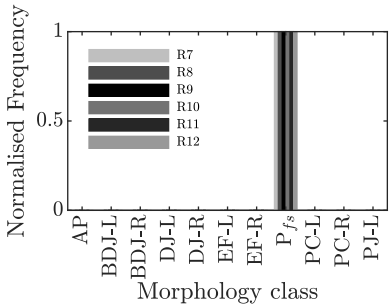
(b) MP $(0°, 0°)$



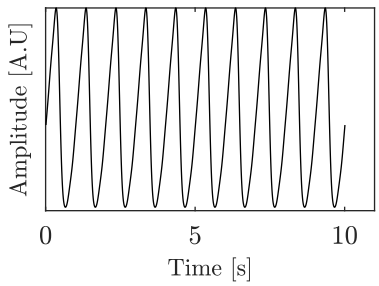
(c) TW $(0°, 0°)$



(d) Repeated $(-16°, 0°)$



(e) MP $(-16°, 0°)$



(f) TW $(-16°, 0°)$

Figure 8: Illustration of the six repeated recorded signals from two gaze angles $(0°, 0°)$ 8a; $(-16°, 0°)$ 8d. The histogram of the corresponding classes: 8b and 8e. The typical waveform for each of the six recordings: 8c and 8f . The recordings R1-R6 were recorded as position $(0°, 0°)$ and the recordings R7-R12 were recorded at position $(-16°, 0°)$. Notice that for each respective gaze angle, the most frequent morphology class is the same for all recordings.

Here, the leftmost column shows six signals recorded for the same position, the middle column shows the morphological profiles and the rightmost column shows the typical waveform. Figures 8a-8c show results for viewing angle $(0°, 0°)$, and Figures 8d-8f show the results from viewing angle $(0°, -16°)$. All signal segments have been classified using the normalised segment error, where segments with an NSE above 0.18 were omitted from analysis. For viewing angle $(0°, 0°)$ the most common waveform class is EF-R and for viewing angle $(0°, -16°)$ the most common waveform class is P_{FS}. For some signals, such as R3 in Figure 8a, the frequency of class EF-R is much lower compared to the other signals. This is due to a change in waveform morphology occurring at around 10 seconds. The results in these two figures show that the proposed method is able to repeat classification for signals that are morphologically similar, as well as to detect signals which deviates from the others, for example due to noise. The distribution of the average waveform distance index for the signals recorded in positions $(0°, 0°)$ and $(-16°, 0°)$ and the average waveform distance index between these two signal groups are presented in Figure 9. These results show that the average waveform distance index for each respective gaze angle is lower compared to the waveform distance index between the signals. This suggests that repeated measurements for each respective gaze angle are more similar to each other compared to recordings at these two gaze angles.

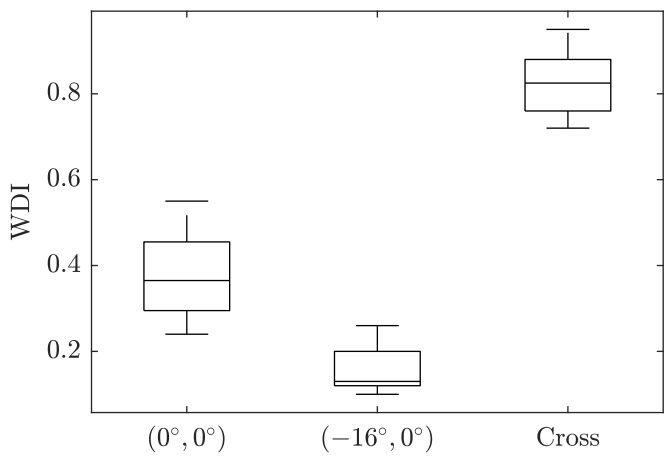
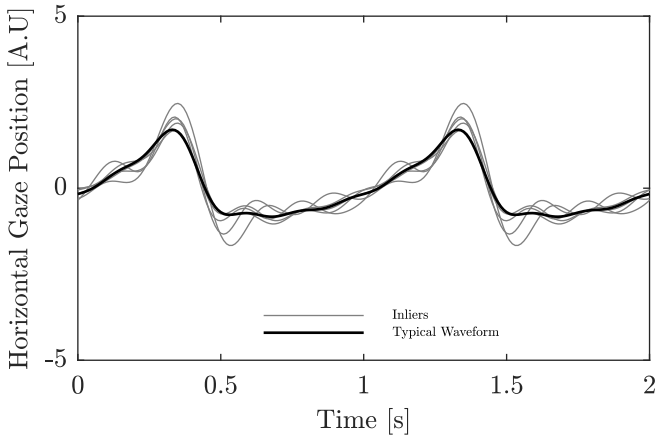


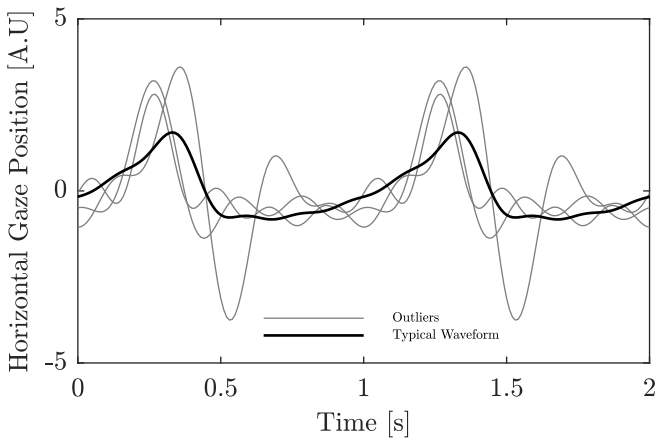
Figure 9: The distribution of the mean waveform distance for the signals in Figures 8a) $(0°, 0°)$ and 8d), as well as the waveform distance index between the two set of signals (Cross). The signals have been subject to normalised segment error analysis.

5.3 Outlier rejection model

The outlier rejection model is aimed at removing segments that do not fit to the class they were assigned to by the template clustering method. An example of the results of the outlier rejection method are presented in Figure 10. In Figure 10a, six inliers are



(a) Inliers



(b) Outliers

Figure 10: Illustration of inlier and outlier classification based on the bootstrap method. In 10a) six inlier waveforms are plotted and in 10b) two outlier waveforms are plotted. The black signal represents the typical waveform in both cases, and the gray waveform morphologies illustrates the observations for each category.

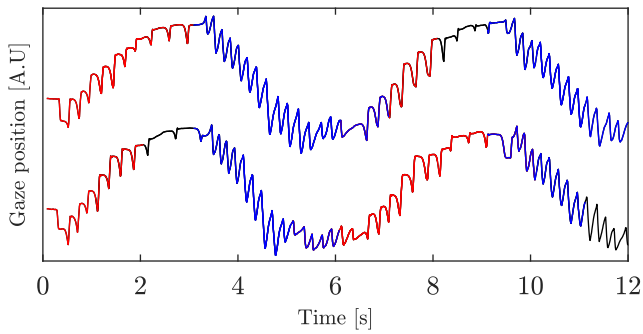
plotted in gray together with the typical waveform, plotted in black. The same typi-

cal waveform is plotted together with two outliers in Figure 10b. The reconstructed waveforms in Figure 10a are in good agreement with the typical waveform, whereas for the reconstructed waveforms in Figure 10b, this appear not to be the case. Note that the waveforms in these figures have not been subject to normalised segment error analysis. Even though the best fit for the gray signals in Figure 10b is the typical waveform, this is clearly not a good fit, and these waveforms are identified as outliers.

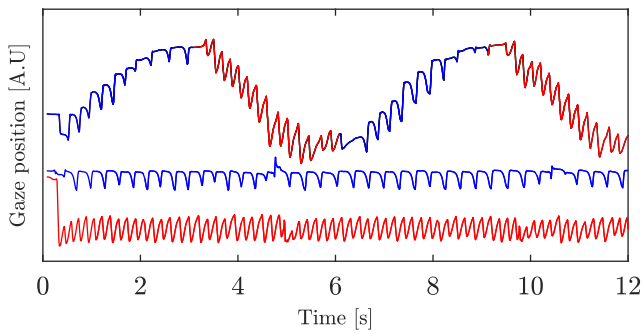
5.4 Signal segment matching

The signal segment matching method compares subsegments from two morphological trends with each other. In Figure 11a, the morphological trends of two smooth pursuit box signals have been compared. The figure shows the presence of two different frequently occurring morphologic classes (blue and red), corresponding to two different waveform morphologies, that occurs at exactly the same times in the two repeated signals. The black segment was not matched between the two signals. During the left to right eye movement (positive y-direction), the waveform morphologies are classified as the same, and the same holds for movements from right to left (negative y-direction).

In Figure 11b, the morphologic trends for one smooth pursuit box signal is compared to the morphologic trends of fixations at two different gaze angles from the same patients. As is illustrated, the red segments are matched to fixation data recorded at position $(0^\circ, 0^\circ)$, whereas the blue segments are matched to fixation data recorded at $(-16^\circ, 0^\circ)$. These two figures dings using the same stimuli are characterised similarly, and that segments of the smooth pursuit waveform morphology resembles fixations waveform morphologies recorded from the same person, depending on pursuit direction and fixation angle.



(a) Box to Box



(b) Box to fixations

Figure 11: Illustration of the similarity between two box recordings from the same participant at different times for horizontal eye movement data 11a) and similarity between box and two fixation recordings 11b). In both figures, two classes of waveform morphologies have been detected using the signal segment matching, and they are coloured red and blue, respectively. The black segment is not similar to any of the two morphology classes. In Figure 11b), two examples of fixation data matched with the box data. The blue (middle) and the red (bottom) signals have been recorded at positions $(0°, 0°)$ and $(-16°, 0°)$, respectively. A positive change in the y-direction corresponds to a rightward eye movement.

6 Discussion

In this work, a method for categorising and comparing nystagmus waveform morphologies between short signal segments and between entire recordings and tasks, has been proposed. A waveform distance index has been proposed as a metric to compare and find the similarities between nystagmus signals, and is in this work used to cluster observed waveforms. The waveform distance index is based on features of the nor-

malised waveform model [17], and it is evaluated on both fixation and smooth pursuit data recorded from nystagmus patients. A method referred to as template clustering has been proposed to group observed waveforms into ten different categories, which are derived from the template waveforms presented in [8]. From the template clustering, a typical waveform is derived, which is used for outlier detection. The waveform distance index is used in the template clustering to find the closest match between an observed waveform morphology and one of the ten template waveforms.

The main advantage of the proposed method compared to previous methods used to analyse nystagmus signals, is that it allows a comparison of nystagmus waveforms on different time scales and between different tasks. Previous methods have mainly been concerned with the modelling of nystagmus waveforms [11, 12], but not the quantification of differences between waveforms . By using a method that automatically analyses the shape of the oscillations, it is possible to improve the accuracy of nystagmus analysis and thereby nystagmus diagnostics. By applying this method to eye movement tasks other than fixations, the diagnostics of nystagmus eye movements can be improved [15].

The results in this work show that the proposed method is able to accurately repeat the waveform morphology classification for eye movement signals, both within and between recordings, as well as distinguish between different waveform morphologies. The proposed method for outlier rejection shows that signals that are morphologically different compared to the typical waveform, are identified as outliers. Note that the detection of outliers proposed in this work does not serve the same purpose as the classification of the signal quality, which is determined using the normalised segment error [17]. The results in Figure 11 show that the method is capable of comparing fixation and smooth pursuit recordings and identify similar waveforms between viewing angles and pursuit directions. Note that the detection of matching segments is more consistent when comparing the smooth pursuit data to fixation data (Figure 11b) than when comparing smooth pursuit data to other smooth pursuit data (Figure 11a).

The superposition model has previously been used to study smooth pursuit gain for nystagmus patients [6], and the method may be reliably used for early-onset nystagmus [6]. However, the model is not representative of smooth pursuit eye movements for all nystagmus patients [22, 14]. The application of the superposition model on the horizontal smooth pursuit stimuli presented to participants in this study, seems to perform worse in terms of the highest median and the largest variance of the waveform distance index. All analysed participants had horizontal nystagmus, and a possible explanation for worse performance of the smooth pursuit stimuli subtraction for the horizontal smooth pursuit is the increase in frequency of the horizontal smooth pursuit stimuli, which makes the target harder in the later stages of the task. The method is suitable for analysis when the frequency range of the nystagmus does not overlap with the frequency of the smooth pursuit stimuli. Since the highest frequency

of the smooth pursuit stimuli in this experiment was 1 Hz, and all participant exhibited nystagmus frequencies higher than 1 Hz, the nystagmus component of the eye movement pattern was not affected by the stimuli subtraction.

The results of the smooth pursuit stimuli subtraction, Figure 6, suggest that it is a good method for separating the smooth pursuit component from nystagmus oscillations, given that the superposition model is valid. In order to use the superposition model, it is required to assess the validity of the model for each individual patient.

It should be noted that the reliability of the waveform modelling of the smooth pursuit data is reduced compared to the fixation data, due the the uncertainty of the interaction between the nystagmus and smooth pursuit. During smooth pursuit there are more factors, e.g. catch-up saccades, that may affect the waveform modelling. If it is of interest to find representative waveforms in the smooth pursuit data, multiple recordings (or longer) seem to be an appropriate approach.

7 Conclusion

In this work, a novel method for analysis of nystagmus waveform morphologies is presented, together with a method for analysing nystagmus waveform morphologies during smooth pursuit. The results show that the method is suitable for grouping nystagmus waveform morphologies that are similar, that it is robust and consistent when analysing different and repeated morphologies. The analysis of the smooth pursuit data shows that the method is able to detect similarities between fixation and smooth pursuit data. The proposed method opens up for clinical studies investigating the relationship between eye movements during fixation at different viewing angles and smooth pursuit movements in different directions.

8 Acknowledgements

This work has been funded by the Swedish Research Council [grant number VR 2015-05442]. We gratefully acknowledge the Lund University Humanities Lab. We are thankful to Markus Rahne and Linnea Sjödahl, who have contributed with data collection. We would also like to thank all the participants. Dr Kristina Holm has been kind to refer patients to our study, which we are grateful for. None of the data reported in this work are publicly available, and no experiments were preregistered.

References

[1] R. J. Leigh and D. S. Zee, *The neurology of eye movements*. Oxford University Press, USA, 2015.

[2] E. Karlén, L. Milestad, and T. Pansell, "Accommodation and near visual function in children with albinism," *Acta ophthalmologica*, vol. 97, no. 6, pp. 608–615, 2019.

[3] L. Gantz, M. Sousou, V. Gavrilov, and H. E. Bedell, "Reading speed of patients with infantile nystagmus for text in different orientations," *Vision research*, vol. 155, pp. 17–23, 2019.

[4] N. Hussain, "Diagnosis, assessment and management of nystagmus in childhood," *Paediatrics and Child Health*, vol. 26, no. 1, pp. 31–36, 2016.

[5] R. McLean, F. Proudlock, S. Thomas, C. Degg, and I. Gottlob, "Congenital nystagmus: Randomized, controlled, double-masked trial of memantine/gabapentin," *Annals of Neurology*, vol. 61, no. 2, pp. 130–138, 2007.

[6] L. Dell'Osso, "Evaluation of smooth pursuit in the presence of congenital nystagmus," *Neuro-ophthalmology*, vol. 6, no. 6, pp. 383–406, 1986.

[7] M. Dunn, "Clinical assessment of nystagmus," *Optometry Today*, vol. 56, no. 7, pp. 80–85, 2016.

[8] L. Dell'Osso and R. Daroff, "Congenital nystagmus waveforms and foveation strategy," *Documenta Ophthalmologica*, vol. 39, no. 1, pp. 155–182, 1975.

[9] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer, *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford, 2011.

[10] R. Abadi, D. Broomhead, R. Clement, J. Whittle, and R. Worfolk, "Dynamical systems analysis: a new method of analysing congenital nystagmus waveforms," *Experimental brain research*, vol. 117, no. 3, pp. 355–361, 1997.

[11] M. Theodorou and R. Clement, "Classification of infantile nystagmus waveforms," *Vision Research*, vol. 123, pp. 20–25, 2016.

[12] R. Abadi and R. Worfolk, "Harmonic analysis of congenital nystagmus waveforms," *Clinical vision sciences*, vol. 6, no. 5, pp. 385–388, 1991.

[13] L. Dell'Osso, J. Van der Steen, R. Steinman, and H. Coleewijn, "Foveation dynamics in congenital nystagmus i: Fixation," *Documenta Ophthalmologica*, vol. 79, no. 1, pp. 1–23, 1992.

[14] L. McIlreavy, *Fixational and pursuit eye movements in infantile nystagmus: oculomotor control and perception*. PhD thesis, Cardiff University, 2016.

[15] T. Bakaeva, N. Desai, W. Dai, J.-R. Rizzo, and J. C. Rucker, "Increasing velocity slow phases in acquired nystagmus," *Journal of Neuro-ophthalmology*, vol. 38, no. 4, pp. 479–482, 2018.

[16] I. Gottlob, "Infantile nystagmus. development documented by eye movement recordings.," *Investigative ophthalmology & visual science*, vol. 38, no. 3, pp. 767–773, 1997.

[17] W. Rosengren, M. Nyström, B. Hammar, M. Rahne, L. Sjödahl, and M. Stridh, "Modeling and quality assessment of nystagmus eye movements recorded using an eye-tracker," *Behavior Research Methods*, pp. 1–15, 2020.

[18] M. Stridh, D. Husser, A. Bollmann, and L. Sörnmo, "Waveform characterization of atrial fibrillation using phase information," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1081–1089, 2009.

[19] SR-Research, "*EyeLink 1000 User Manual.*," 2010.

[20] J. W. Peirce, "Psychopy psychophysics software in python," *Journal of Neuroscience Methods*, vol. 162, no. 1-2, pp. 8–13, 2007.

[21] W. Rosengren, M. Nyström, B. Hammar, and M. Stridh, "A robust method for calibration of eye tracking data recorded during nystagmus," *Behavior research methods*, pp. 1–15, 2019.

[22] A. H. Weiss, J. P. Kelly, and J. O. Phillips, "Infantile nystagmus and abnormalities of conjugate eye movements in down syndrome," *Investigative ophthalmology & visual science*, vol. 57, no. 3, pp. 1301–1309, 2016.

Appendix A

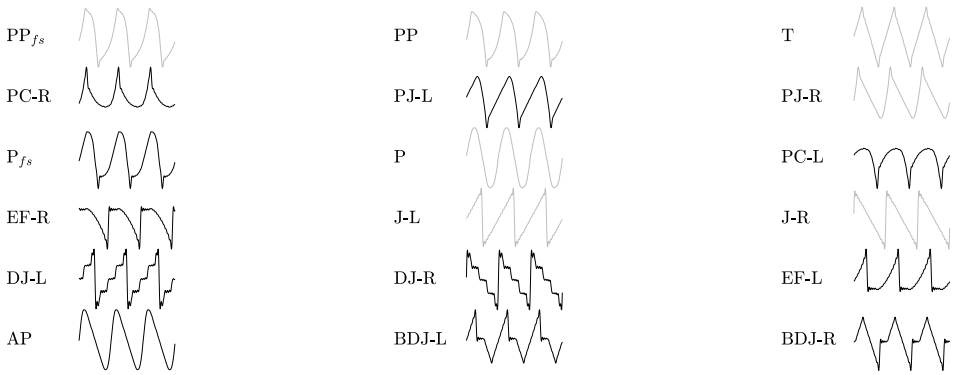


Figure 12: Illustration of the 18 template waveforms. The black waveforms are used in the template clustering and the gray waveforms are not used in the template clustering.

Paper IV

Cycle–to–cycle nystagmus waveform characterisation

Abstract

Nystagmus is a symptom of oculomotor disease, which causes the eyes to move involuntarily. The eye movement pattern expressed by nystagmus patients may vary greatly between individuals, and it is of clinical interest to analyse the details of these patterns. The goal of this work is to develop a method that classifies waveforms in eye-tracker signals recorded from nystagmus patients on a cycle–to–cycle level.

The proposed method divides each signal into cycles and performs cycle–to–cycle analysis in two steps. First, *morphological classification* based on principle component analysis is performed, and second, *cycle features* such as breaking saccades and foveation periods are identified. The outputs of these two methods are combined into a waveform class, which is used to describe the eye movement pattern in each cycle. The set of possible waveform classes is based on 18 template waveforms previously described in the literature. The method was evaluated by comparing detected waveform classes to data manually annotated by experts. The performance was determined both by comparison to annotations within recordings and by quantifying the similarity between waveforms across recordings.

The results show that about 70 % of the recorded cycles are classified into one of the 18 template waveforms. Analysis of waveforms without a matching annotation shows that 76 % of the detected waveforms are matched in terms of similarity to corresponding annotations in other recordings. Overall, the results show that the proposed method reliably performs nystagmus waveform classification.

Based on
William Rosengren, Marcus Nyström, Björn Hammar, and Martin Stridh,
"Cycle–to–cycle nystagmus waveform characterisation,"
Submitted for publication

1 Introduction

Pathological nystagmus is a condition where the ability to control eye movements is reduced, and the eyes move involuntarily in an oscillating pattern [1]. These symptoms may cause an inability to fixate the gaze, a decrease in visual acuity and oscillopsia, which is a sensation that the world is in constant motion [1].

There are different disorders associated with different nystagmus eye movement patterns, and it is of interest to study the eye movement pattern of each individual in order to facilitate diagnostics of the cause of nystagmus [1, 2, 3, 4]. In clinical practice, manual inspection of eye movements from nystagmus patients is typically performed. However, as *eye tracking* technology has become more accessible in recent years, recordings of nystagmus eye movements have become more common, both in clinical research and for diagnostic purposes. Eye tracking data allow a detailed analysis of the eye movements observed in nystagmus patients, and are therefore clinically useful. However, in order to interpret the eye–tracker signal in a clinical context, the main signal characteristics first need to be extracted.

When studying nystagmus eye movements recorded with an eye tracker, two different approaches may be used to extract signal characteristics: *cycle feature analysis* or *morphological classification*. Both approaches are based on identification of each *oscillation cycle* in the recorded eye–tracker signal. The nystagmus oscillation cycle is typically initiated with the *slow phase*, where the eye drifts away from the indented fixation target [1]. This drift is corrected through the *fast phase*, which is a saccade-like eye movement in the opposite direction of the slow phase.

The first approach, cycle feature analysis, is aimed at detecting fast phases, foveation periods and slow phases in each oscillation cycle. Cycle feature analysis has successfully been used for calibration purposes [5] and for detection of specific nystagmus waveform morphologies [6]. Cycle feature analysis is typically based on velocity and acceleration thresholds, and these thresholds have also been used to estimate the visual acuity in nystagmus patients [7].

The second approach, morphological classification, is aimed at analysing the overall shape of the oscillation cycle. Here, the objective is to classify each cycle into one of a finite set of possible morphological classes, corresponding to the waveforms presented in [3]. Typical for these modelling methods is to use a set of base functions, and then parametrise the various waveforms using these base functions [8, 9, 10].

For the cycle feature analysis, a novel approach to analyse each oscillation cycle is proposed in this work. Instead of only dividing the signal into a slow phase and a fast phase, each cycle is divided into four different components: breaking saccades, fast phases, foveation periods, and slow phases. Breaking saccades are short saccade–like eye movements that are found in some nystagmus waveforms, and they typically occur in the opposite direction of the fast phase [3]. For the morphological classi-

fication, principle component analysis is used to model the variation of oscillation patterns observed in nystagmus signals. A similar approach has previously been used for nystagmus waveform modelling [9].

The proposed method is evaluated using two datasets, where the first dataset is public and has been recorded using the magnetic scleral search coil eye–tracking technique [11]. Each recording in this dataset has been assigned with one or more waveform labels through expert annotation. The second dataset was recorded using a video–based eye–tracking technique, where no expert annotations are available. The proposed method was evaluated both by comparison to annotations within recordings and by quantifying the similarity between waveforms across recordings. This paper is organised as follows: In section 2, the proposed method is presented. The datasets and methods used for evaluation are described in Section 3. The results are presented in Section 4 and further discussed in Section 5. Finally, Section 6 contains the main conclusions of this work.

2 Methods

In this section, the proposed method for waveform classification is presented. The preprocessing and the cycle detection are described in Section 2.1. In Section 2.2, the method for the *morphological classification* is presented. Finally, the *cycle feature analysis* used to determine the presence of breaking saccades, fast phases, foveation periods and slow phases, is described in Section 2.3. The combined result of these two methods is the waveform class, which is expressed on a *waveform classification grid*. A block diagram of the proposed method is shown in Figure 1.

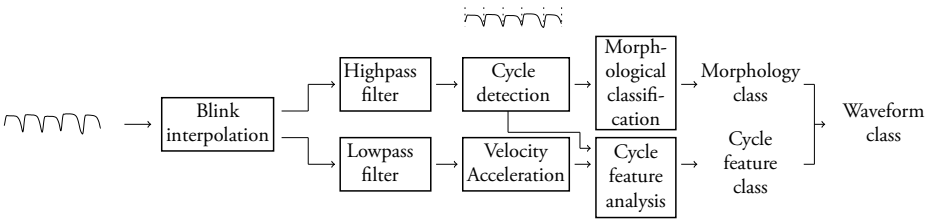


Figure 1: Block diagram of the proposed method. The samples that are missing due to blinks are estimated by linear interpolation. The resulting signal is analysed using two methods: First, the signal is high pass filtered in order to serve as input to the cycle detection algorithm. Once the cycles (plotted at the top of the figure) have been detected, each cycle is subject to morphological classification using the principle components obtained from template waveforms. A lowpass filtered signal is used to compute velocity and acceleration profiles of the oscillation. For each detected cycle, the cycle features are determined using velocity and acceleration thresholds. The results of the morphological and cycle feature classifications are combined to a waveform class.

2.1 Preprocessing

The first step of the oscillation cycle detection is to handle blinks and other missing data. For some recording techniques, such as video–based eye trackers, blinks may introduce data loss, or other artefacts that are undesirable. In this work, the missing samples are interpolated using linear interpolation, which ensures that both the left and right eye signals are of equal length, without missing values. The positions of the blinks are marked by the eye tracking system. If the difference in time between the offset and onset of two subsequent episodes is less then 50 ms, these episodes are considered as one blink. Before and right after a blink, the eye position signal changes drastically due to the occlusion of the pupil. Therefore, all data 500 ms before and after a detected blink are also included in the blink episode. Once the onsets and offsets of all blinks have been determined, the last sample before each blink is linearly interpolated to the first sample after the blink.

After the blink interpolation, the signal is separately filtered through two different filters. For the detection of local extreme points used in the cycle detection, the signal is filtered through a high–pass Butterworth filter of order three with a cutoff frequency of 2 Hz. The second filter is used in order to generate velocity and acceleration signals used for the cycle feature analysis. For this purpose a low–pass Butterworth filter of order three with a cutoff of 30 Hz is applied to the signal.

In the cycle detection, the signals are analysed in blocks corresponding to 3 seconds. This accounts for variations in the eye movement pattern, which commonly occur throughout a recording. The blocks are analysed with an overlap of 50 %, and

for each 3 second segment, the local extreme points are detected using a peak detection algorithm [12]. The cycles are divided using the local extreme points obtained by the peak detection.

2.2 Morphological classification

The morphological classification is based on principal components obtained from 18 template waveforms [3], that have been reconstructed from original drawings using Fourier analysis [10]. Two adjustments of the extracted cycles are required in order to use principle component analysis for classification. First, the cycles need to be resampled to the same length, e.g., a period of 100 samples, and this is performed using linear interpolation (SciPy v 1.2.1). Second, the template waveforms that are used to create the bases were all reconstructed with a sinusoidal model where the first harmonic phase was 0° [10]. In order to apply these bases to recorded data, the analysed cycles also need to be adjusted to a first harmonic sine phase of 0°. The phase adjustment is performed using the following method: The first harmonic phase of each cycle, α, is computed as:

$$\alpha = \arctan\left(\frac{s_{sin}}{s_{cos}}\right) \tag{1}$$

where

$$s_{sin} = \sum_{n=0}^{N-1} \sin(2\pi\frac{n}{100})s[n], \tag{2}$$

and

$$s_{cos} = \sum_{n=0}^{N-1} \cos(2\pi\frac{n}{100})s[n], \tag{3}$$

and where $N = 100$. The corresponding shift in samples is computed, and applied to each cycle. An example of this shift is illustrated in Figure 2. The solid black trace illustrates the original oscillation cycle and the dashed gray trace represents the shifted oscillation cycle with a first harmonic sine phase of 0°.

When using principle component analysis, the data are represented by the linear model:

$$\boldsymbol{c}_i = \boldsymbol{B}\boldsymbol{p}_i. \tag{4}$$

where $\boldsymbol{c}_i$ is the cycle data represented as a vector

$$\boldsymbol{c}_i = \begin{bmatrix} c_i[0], & \ldots, & c_i[N-1] \end{bmatrix}. \tag{5}$$

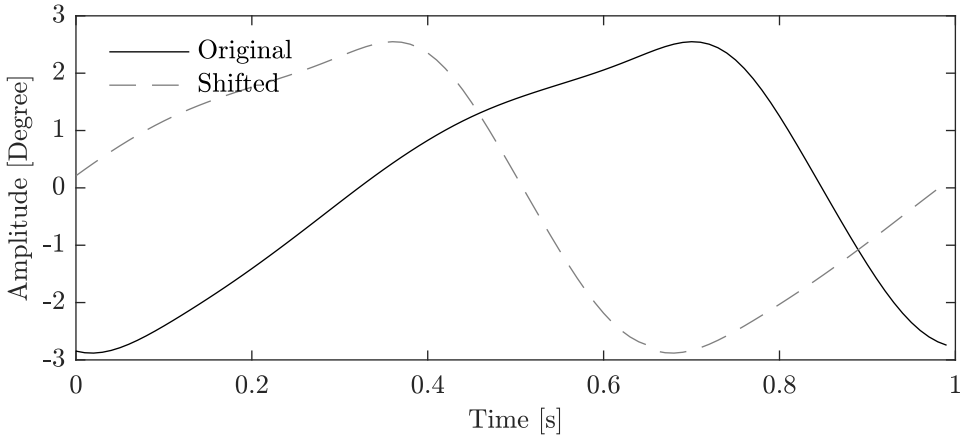


Figure 2: Illustration of segment alignment for one oscillation cycle. The solid black trace represents the original signal, and the dashed gray trace represents the signal after the phase shift.

The vector $\boldsymbol{p}_i$ contains the weight parameters for cycle i, and $\boldsymbol{B}$ is the matrix containing the bases obtained from the principal component analysis of the template waveforms. In this work, the components corresponding to 99 % of total the signal energy are used, and the weight parameters are estimated according to the least squares solution:

$$\hat{\boldsymbol{p}}_i = (\boldsymbol{B}^T\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{c}_i. \tag{6}$$

The assigned class for each cycle corresponds to the base for which the coefficient $p_{i,l}$ in $\boldsymbol{p}_i$ has the largest absolute value when reconstructing the waveforms using (4). If the coefficient is negative, it suggests that the waveform has the opposite beat direction compared to a signal with a positive coefficient. This means that there are two waveforms associated with each principle component, the original and the same waveform inverted (mirrored around the x-axis). The morphological class X_i for the i:th cycle is determined as

$$X_i = \begin{cases} \text{argmax}_l\, |\boldsymbol{p}_i|, \text{ if } p_{i,l} \geq 0 \\ \text{argmax}_l\, |\boldsymbol{p}_i| + N_b, \text{ if } p_{i,l} < 0, \end{cases} \tag{7}$$

where $p_{i,l}$ is the l:th element in $\boldsymbol{p}_i$ and N_b is the number of bases in $\boldsymbol{B}$. Since the negative bases are used as well, there are in total $2N_b$ possible morphological classes.

2.3 Cycle feature analysis

The next step is to detect the specific cycle features described in Section 1. In order to find the features in the signals, the velocity and acceleration signals are calculated. The velocity, $v(n)$, and acceleration, $a(n)$, signals are computed from the low–pass filtered signal described in Section 2.1, using the filter

$$h(n) = \frac{1}{8}\begin{bmatrix} -1 & -1 & -1 & -1 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}, \tag{8}$$

where the velocity is computed as

$$v(n) = z(n) * h(n), \tag{9}$$

and the acceleration as

$$a(n) = v(n) * h(n), \tag{10}$$

and where $z(n)$ is either the horizontal or the vertical low–pass filtered signal. An example of the cycle feature analysis based on these three signals is presented in Figure 3. The top panel shows the acceleration signal, the middle panel shows the velocity

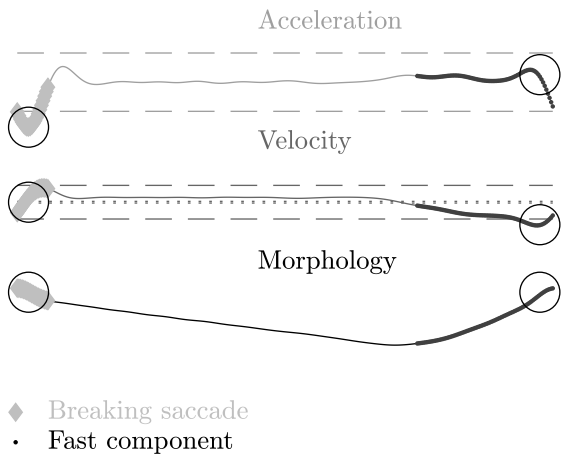


Figure 3: The top trace shows the acceleration trace, the middle trace shows the velocity trace and the bottom trace shows the morphology, ie., the position signal. The circle to the left shows a detection of a breaking saccade, and the circles to the right shows a detection of a fast phase.

signal and the bottom panel shows the position signal. The dashed lines show the $2\sigma_v$ and $3\sigma_a$ thresholds for each respective trace, and the dotted line shows the $0.1\sigma_v$ threshold for the velocity trace. For each respective curve, the number of local maxima / minima outside the respective thresholds are used as follows:

1. If there is a local extreme point outside the $2\sigma_v$ threshold for the velocity trace, a fast phase is detected.

2. If there is a local extreme point outside the $3\sigma_a$ threshold for the acceleration trace, a breaking saccade is detected, unless it coincides with a fast phase. In such a case, the segment is classified as a fast phase.

3. The foveations are defined as a segment where at least 20 ms of data are within $0.1\sigma_v$ in the velocity trace.

The duration of both the fast phase and the breaking saccade are defined as all samples that exceed 20 % of the amplitude of the extreme point for both their respective traces. For the example in Figure 3, a breaking saccade is detected in the beginning of the signal, and a fast phase is detected in the end of the signal. These detections are marked by black circles. No foveations were detected in this case. A summary of the definitions of the three cycle features breaking saccades, fast phases and foveation periods, are found in Table 1. All other parts of the cycle are defined

Table 1: Definition of the three cycle features: breaking saccade, fast phase and foveation periods. The slow phase is defined as all other parts of the cycle. The standard deviation of the velocity and acceleration, σ_v and σ_a respectively, are computed for each cycle.

Eye Movement	Velocity	Acceleration	Duration
Breaking saccade	$< 2\sigma_v$	$> 3\sigma_a$	-
Fast phase	$> 2\sigma_v$	-	-
Foveation Period	$< 0.1\sigma_v$	-	> 20 ms

as slow phases. The thresholds for the definitions of the cycle features were chosen empirically by analysing the template waveforms [3].

Each analysed cycle is categorised based on if there is a foveation period, a fast phase or a breaking saccade, and each cycle can have several of the three cycle features. The classification of each cycle with respect to these three cycle features is summarised in Table 2. For each of the classes A-H, a zero indicates that the component is not present in the cycle, and a one indicates that it is present. For example, if neither foveations, fast phases nor breaking saccades have been detected in a cycle, the corresponding class is A.

Table 2: Definitions of the eight cycle feature classes. A one represents that the cycle feature is present, and a zero that it is not present in the analysed cycle. Each of the eight possible combinations has been assigned a letter been A and H.

Class name	Foveation	Fast phase	Breaking Saccades
A	0	0	0
B	0	0	1
C	0	1	1
D	1	0	1
E	0	1	0
F	1	1	0
G	1	0	0
H	1	1	1

3 Method evaluation

In this section, the process to evaluate the proposed method is described. First, the patient recordings are described in Section 3.1. This is followed by an overview of the evaluation process and the assumptions made for the evaluation process, see Section 3.2. Next, the description of the template waveform classification is found in Section 3.3. Finally, a method to compare waveform classes between different recordings refered to as *average waveform matching*, is presented in Section 3.4.

3.1 Data

Two different datasets with patient recordings were used in the evaluation of the proposed method. The annotated magnetic scleral search coil and the unannotated video–based eye–tracking datasets are described below. All of the recorded data come from attempted fixations at the primary position $(0°, 0°)$, and both horizontal and vertical eye movement recordings have been used in the analysis.

Magnetic scleral search coil data

The magnetic scleral search coil data included in this study have previously been published in [11], and the original data files are publicly available at `http://www.omlab.org/OMLAB.html` [1]. In the original publication [11], expert annotations of observed nystagmus waveforms were included. As described in the original work, the

[1] Data downloaded: June 10 2020

oculomotor characteristics of each participant were described for the primary position. A total of 15 patients were recorded between one and five times over the course of several years. The recordings are available in the following formats: .DAT, .LAB, .bin or .txt. Since the other formats were difficult to read, only the .txt files were used in this work. The recordings of subject S5 were excluded as it was not possible to interpret which of the recordings that corresponded to the primary position data. For participant S10, recording 'a', the primary position data appeared to be corrupt and the recording was therefore excluded from the analysis. After the exclusion, 16 recordings from 8 participants remained. In order to match the sampling rate of the video–based eye tracker, the signals were resampled to 1000 Hz using linear interpolation before the analysis. An example of a signal segment from participant S7 is shown in Figure 4. The figure shows horizontal data from a patient recorded in year

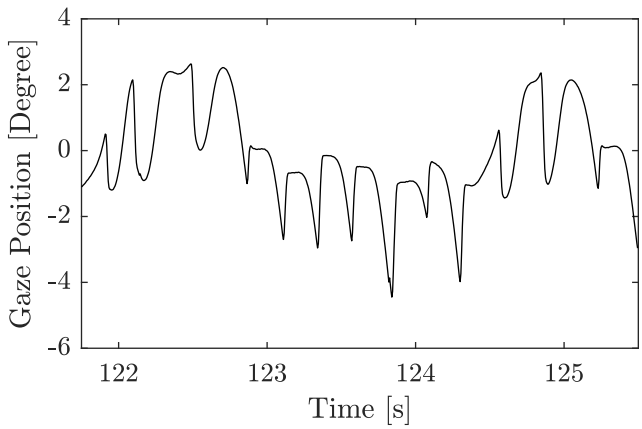


Figure 4: Illustration of a segment from a horizontal signal for the participant S7 recorded in 1999. This recordings was annotated with the single waveform EF-L. Note that a movement to the right corresponds to a movement in positive y-direction [11].

1999 which was annotated with the single label EF-L.

Video-based eye tracking data

The video–based eye tracking data were recorded according to the following protocol [10]: The eye movements were represented by pupil and corneal reflection data, which were collected binocularly with a sampling rate of 1 kHz. The recording system was an EyeLink 1000 Plus eye tracker, used in desktop mode. The eyes were tracked using the center–of–mass mode using the host software v.5.09 and the DevKit v.1.11.571. Patients were seated according to the manufacturer specifications [13], at 80 cm from

the screen and their heads were kept in position using a chin and forehead rest. The software for stimuli presentation was written in PsychoPy (version 1.83) [14]. The screen presenting the stimuli was an ASUS VG248QE monitor with a resolution of 1920×1080 pixels, with dimensions 53×30 cm and a refresh rate of 144Hz. The study was approved by the ethics board at Lund University and all experiments were in accordance with the Declaration of Helsinki. All patients signed a consent form before the experiment started. In this work, only attempted fixations at the primary position were included in the analysis. A total of 16 recordings from 8 participants were included and only data from the left eye were used.

3.2 Overview of method evaluation

The comparison between the expert annotations and the waveform classification of the proposed method is a classification problem, which typically is evaluated using detection performance metrics such as sensitivity and specificity. These metrics are, however, less applicable here, since only a general description of the most prominent waveforms is given by the experts; neither the prevalence nor exactly where these waveforms reside in the recordings are provided in the annotations. This means that waveforms other than those reported in the original work may be present in the signal. Consequently, it is not possible to assert the sensitivity and specificity for the waveform classification.

Instead of using detection metrics, the evaluation of the proposed method focuses on the similarity between the true positive waveforms and the false positive as well as unannotated waveforms. In this work, a true positive is defined as a waveform class that has been detected by the proposed method, and is also annotated in [11]. Correspondingly, a false positive is a detected waveform class that has not been annotated in [11]. True and false positive cycles refer to cycles belonging to true and false positive waveform classes, respectively. The following evaluation is performed:

1. For the true positive cycles in each signal, an average waveform is computed for each waveform class.

2. The average waveforms are computed also for the unannotated and false positive cycles, in the same way as for the true positive cycles, for each waveform class.

3. The average waveforms for the unannotated or false positive cycles are then compared to the set of true positive average waveforms, using the *waveform distance index* [15].

4. If the average waveform for an unannotated or a false positive waveform is sufficiently similar to the true positive waveforms in terms of waveform distance

index, the waveform classification of the analysed signal is considered to be correct, i.e., the deteced waveform is matched to the annotated waveform class.

All waveform classes with a prevalence of at least 5 % within each individual signal are included in the analysis. The above steps are illustrated in Figure 5.

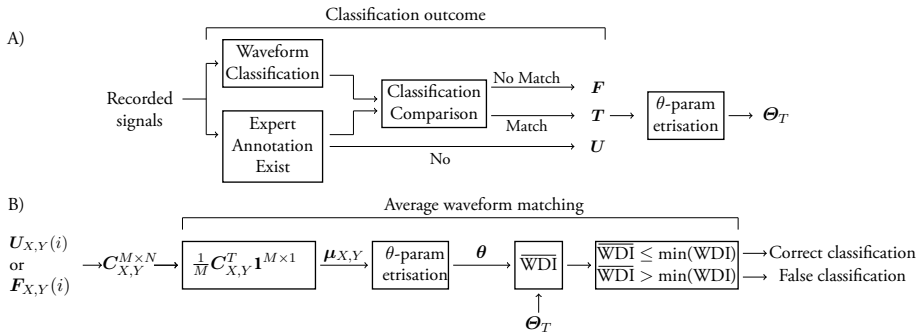


Figure 5: Flow chart of the evaluation process for the recorded signals. A) The cycles in each recorded signal are classified using the proposed method, and the results from this classification are compared to the expert annotation, in case such annotations exist. Three possible outcomes are possible for each classification: unannotated (U), false positive (F) or true positive (T). True positives are the waveforms where the waveform classification coincide with the expert annotation. For false positives, the detected waveform does not coincide with the expert annotation, and for the unannotated waveforms no expert annotation exist. The unannotated and false positive cycles are compared to the true positive cycles through average waveform matching, explained in panel B). The unannotated and false positive cycles for a signal (with index i) are denoted $\boldsymbol{U}_{X,Y}(i)$ and $\boldsymbol{F}_{X,Y}(i)$, respectively. All cycles that belong to the same waveform class are stacked as rows in the matrix $\boldsymbol{C}^T_{X,Y}$, where X and Y denote the morphological class and the cycle feature class, respectively. An average waveform, $\boldsymbol{\mu}_{X,Y}$, is created from the sample mean, and this is parametrised into the vector $\boldsymbol{\theta}$. The average waveform distance index, $\overline{\text{WDI}}$, between $\boldsymbol{\theta}$ and $\boldsymbol{\Theta}_T$ is computed, where $\boldsymbol{\Theta}_T$ contains the parametrisation of the average true positive cycles. If $\overline{\text{WDI}}$ is smaller than the smallest waveform distance index for the analysed waveform class, the waveforms in $\boldsymbol{U}_{X,Y}(i)$ or $\boldsymbol{F}_{X,Y}(i)$ are considered to be correct detections. The process is detailed in Section 3.4.

3.3 Feature classification of template waveforms

The cycle feature analysis is evaluated using the template waveforms presented in [3]. For each template waveform, the segments of the cycles corresponding to breaking saccades, fast phases and foveation periods are marked by the proposed algorithm.

The corresponding classes (A-H), which were introduced in Section 2.3, are used to describe the cycle features for each template waveform. The morphological classification is presented such that each template waveform is assigned to one of the potential $2N_b$ classes. This information is mapped onto a waveform classification grid, where the x-axis and y-axis represent the morphological and the cycle feature classes, respectively. The waveform classification grid is used to analyse the distribution of the waveform classes in the recorded signals.

3.4 Average waveform matching

The average waveform matching is designed to analyse the similarity between average waveforms from different signals for a specific waveform class, where $\boldsymbol{\mu}_{X,Y}$ is the average waveform, and X and Y represent the morphological and cycle feature classes, respectively. The average waveform is used to represent the waveforms of all cycles of the same waveform class in a single recording. All cycles from one recorded signal that are assigned to the same waveform class are stacked as rows in the $\boldsymbol{C}_{X,Y}$ matrix. The average waveform, $\boldsymbol{\mu}_{X,Y}$, for each waveform class in each signal is defined as:

$$\boldsymbol{\mu}_{X,Y} = \frac{1}{M}\boldsymbol{C}_{X,Y}^{T}\mathbf{1}^{M\times 1}. \tag{11}$$

Here, M is the number of cycles classified into the same waveform class, (X, Y), in a signal, and $\mathbf{1}$ is a column vector of all ones. The waveform matching is performed using the waveform distance index [15]. In order to use the waveform distance index, each average waveform $\boldsymbol{\mu}_{X,Y}[n]$ is parametrised into a $\boldsymbol{\theta}$-vector, where the waveform distance index between $\boldsymbol{\theta}_a$ and $\boldsymbol{\theta}_b$ is defined by [15]

$$\mathrm{WDI}(\boldsymbol{\theta}_a, \boldsymbol{\theta}_b) = \sqrt{\sum_{i=1}^{8}\big(\boldsymbol{\theta}_a(i) - \boldsymbol{\theta}_b(i)\big)^2}. \tag{12}$$

In order to remove outliers from each waveform class, an outlier rejection method is implemented as follows: The distribution of all the waveform distance indices between each cycle, and the average waveform $\boldsymbol{\mu}_{X,Y}$ is calculated. All cycles for which the waveform distance index to the average waveform is larger than the average waveform distance index plus two standard deviations, are considered as outliers. This threshold was chosen empirically. Based on the remaining inlier cycles, a new average waveform is constructed by calculating the sample–by–sample average described in (11), where the new matrix $\boldsymbol{C}'_{X,Y}$ contains only the inlier cycles.

In order to investigate whether an average unannotated or false positive waveform is sufficiently similar to the corresponding true positives in annotated recordings, the

waveform distance index is again utilised. For each average waveform that has been computed from either unannotated or false positive cycles, the average waveform distance index, $\overline{\text{WDI}}$, with respect to all average true positive waveforms from the same waveform class in other recordings is computed as

$$\overline{\text{WDI}} = \frac{1}{D}\sum_{d=1}^{D} \text{WDI}\Big(\boldsymbol{\theta}(i), \boldsymbol{\theta_T}(d)\Big), \tag{13}$$

where $\boldsymbol{\theta}(i)$ is the parametrisation of a single unannotated or false positive $\boldsymbol{\mu}_{X,Y}$ vector, and $\boldsymbol{\theta_T}(d)$ is the parametrisation of one of the D true positive $\boldsymbol{\mu}_{X,Y}$ vectors for a certain waveform class. The $\overline{\text{WDI}}$ is compared to the smallest waveform distance index $\min(\text{WDI}_{X,Y})$ for each waveform class. The waveform distances between different template waveforms are described in Table 5. The $\boldsymbol{\theta}(i)$ vector is considered to be matched to the the analysed waveform class if

$$\overline{\text{WDI}} < \min(\text{WDI}_{X,Y}). \tag{14}$$

Some horizontal and vertical waveforms are classified into the same waveform class, i.e., EF-R and EF-U are classified into F13. All horizontal and vertical waveforms that are classified into the same waveform class are analysed together. Only waveform classes for which there exist at least one true positive waveform are considered in this analysis.

4 Results

The results section is divided into two parts. First, the results of the cycle feature analysis and the morphological classification for the template waveforms are presented in Section 4.1. A summary of the analysis of the recorded signals is presented in Section 4.2.

4.1 Template waveforms

Cycle feature analysis

The cycle feature analyses for 12 out of the 18 original template waveforms are presented in Figure 6. The rightward beating versions of PC, PJ, EF, J, BDJ and PJ are omitted, since they produce the same results as the leftward beating waveforms in terms of the cycle feature class. The breaking saccades are represented by light gray diamonds, the fast phases by red circles, and the foveation periods by blue asterisks. In total, there are eight combinations of cycle features that are theoretically possible.

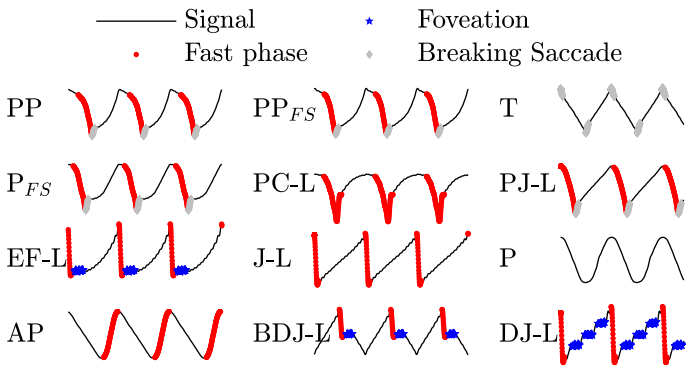


Figure 6: Illustration of the cycle features for the twelve waveforms described in [3] (right–beating versions are omitted in the figure). The black line represents the waveform, the red dots the fast phases, the blue asterisks the foveation periods and the gray diamonds the breaking saccades.

Five out of these eight are found when analysing the template waveforms: A, B, C, E, F. A summary of which template waveforms that belong to each respective cycle feature class is presented in Table 3. For some classes, there is only one matching tem-

Table 3: Classification of the template waveforms into one of the eight cycle feature classes. Five different cycle feature classes are used to describe the different waveforms.

Class	Waveforms
A	P
B	T
C	PJ, P_{FS}, PP; PP_{FS}
E	AP, J, PC
F	BDJ, DJ, EF

plate waveform, e.g., in classes A and B, while for others there are several matching waveforms, e.g., in class C, where both a fast phase and breaking saccade are present.

Morphological Classification

A summary of the morphological classification for the template waveforms is presented in Table 4. A total of 20 (10 positive and the corresponding 10 negative) principle components represent 99 % of the variance of the different template waveforms, after removing the first harmonic. The most important principle component

was selected to represent each template waveform. Only six out of the 20 principle components were selected as the most significant component for at least one template waveform, although other principle components also contributed to each waveform.

Table 4: Classification of the template waveforms into one of the twenty morphology classes. Six different morphology classes are used to represent all the different waveforms.

Class	Waveforms
1	DJ-L/D; J-L/D; P_{FS}, PJ-L/D; PP_{FS}
3	EF-L/D; PC-R/U
6	PP
11	AP; BDJ-L/D; DJ-R/U; J-R/U; P; PJ-R/U
12	BDJ-R/U; T
13	EF-R/U; PC-L/D

Waveform classification grid

The combined classifications of the template waveforms are shown in Figure 7. As can be seen, for three classes (C1, E11 and F11) there are more than one template waveform, but the majority of classes do not contain any template waveform. In total, 14 out of 160 positions (8.75 %) on the waveform classification grid are occupied by template waveforms. The waveform distance index between the template

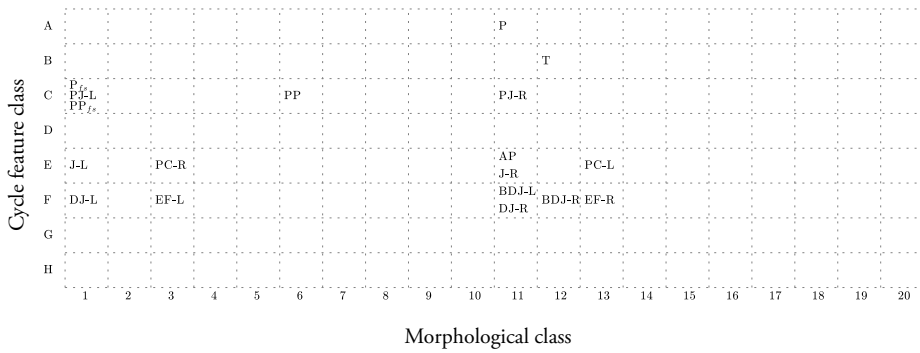


Figure 7: Mapping of all the 18 template waveforms onto the waveform classification grid. The x-axis represents the morphological class, and the y-axis represents the cycle feature class.

waveforms is presented in Table 5. The minimum waveform distance index for each

Table 5: Waveform distance index between waveform classes. The smallest waveform distance indices, representing the most similar waveforms, are marked in bold for each waveform.

	AP	BDJ-L	BDJ-R	DJ-L	DJ-R	EF-L	EF-R	J-L	J-R	P_{fs}	P	PC-L	PC-R	PJ-L	PJ-R	PP_{fs}	PP	T
AP	0.0	0.38	0.66	0.97	0.58	0.8	0.58	0.87	0.43	0.52	0.27	0.72	0.55	0.66	**0.16**	0.51	0.4	0.36
BDJ-L	0.38	0.0	0.58	1.08	0.83	0.92	0.68	1.0	0.7	0.64	0.41	0.72	0.47	0.75	0.46	0.61	0.52	**0.33**
BDJ-R	0.66	0.58	0.0	0.83	1.08	0.68	0.92	0.7	1.0	0.48	0.44	0.47	0.72	0.46	0.75	0.46	0.47	**0.32**
DJ-L	0.97	1.08	0.83	0.0	1.36	0.93	1.11	**0.22**	1.21	0.57	0.8	0.98	1.02	0.6	1.04	0.57	0.7	0.84
DJ-R	0.58	0.83	1.08	1.36	0.0	1.11	0.93	1.21	**0.22**	0.94	0.76	1.02	0.98	1.04	0.6	0.96	0.86	0.84
EF-L	0.8	0.92	0.68	0.93	1.11	0.0	1.01	0.78	1.05	**0.55**	0.64	0.96	0.86	0.67	0.86	0.63	0.7	0.67
EF-R	**0.58**	0.68	0.92	1.11	0.93	1.01	0.0	1.05	0.78	0.83	0.61	0.86	0.96	0.86	0.67	0.82	0.74	0.67
J-L	0.87	1.0	0.7	**0.22**	1.21	0.78	1.05	0.0	1.09	0.44	0.68	0.85	0.97	0.46	0.94	0.46	0.59	0.72
J-R	0.43	0.7	1.0	1.21	**0.22**	1.05	0.78	1.09	0.0	0.82	0.64	0.97	0.85	0.94	0.46	0.83	0.73	0.73
P_{fs}	0.52	0.64	0.48	0.57	0.94	0.55	0.83	0.44	0.82	0.0	0.32	0.72	0.6	0.31	0.6	**0.13**	0.26	0.38
P	0.27	0.41	0.44	0.8	0.76	0.64	0.61	0.68	0.64	0.32	0.0	0.56	0.55	0.41	0.39	0.3	0.21	**0.17**
PC-L	0.72	0.72	**0.47**	0.98	1.02	0.96	0.86	0.85	0.97	0.72	0.56	0.0	0.95	0.56	0.81	0.66	0.58	0.51
PC-R	0.55	**0.47**	0.72	1.02	0.98	0.86	0.96	0.97	0.85	0.6	0.55	0.95	0.0	0.81	0.56	0.57	0.57	0.53
PJ-L	0.66	0.75	0.46	0.6	1.04	0.67	0.86	0.46	0.94	0.31	0.41	0.56	0.81	0.0	0.74	**0.27**	0.3	0.46
PJ-R	**0.16**	0.46	0.75	1.04	0.6	0.86	0.67	0.94	0.46	0.6	0.39	0.81	0.56	0.74	0.0	0.58	0.46	0.47
PP_{fs}	0.51	0.61	0.46	0.57	0.96	0.63	0.82	0.46	0.83	**0.13**	0.3	0.66	0.57	0.27	0.58	0.0	0.17	0.36
PP	0.4	0.52	0.47	0.7	0.86	0.7	0.74	0.59	0.73	0.26	0.21	0.58	0.57	0.3	0.46	**0.17**	0.0	0.3
T	0.36	0.33	0.32	0.84	0.84	0.67	0.67	0.72	0.73	0.38	**0.17**	0.51	0.53	0.46	0.47	0.36	0.3	0.0

template waveform is highlighted in bold. For example, the EF-R is most similar to the AP waveform, with a waveform distance index of 0.58. The smallest waveform distance index between two waveforms is observed for the pair P_{FS} and PP_{FS}, where the distance is 0.13.

4.2 Recorded data

General results

In this section, the results for the recorded signals are presented. i.e., the data recorded using the video–based or magnetic scleral search coil techniques. For the horizontal and vertical data, 71 % and 66 %, respectively, of all waveforms are assigned to one of the 14 positions on the waveform classification grid described in Section 4.1. These values change to 73 % and 69 % for the horizontal and vertical waveforms, respectively, when applying the outlier removal method described in Section 3.4. The distribution of the waveform classes for the combined horizontal and vertical data after outlier removal is shown in Figure 8. The numbers represent the relative prevalence of each waveform class rounded to the nearest percent. A gray square indicates that there exist at least one template waveform in the waveform class. As can be observed in the figure, the three most common waveforms in the analysed data are AP / Jerk - Right (J-R), PC-L and Jerk - Left (J-L).

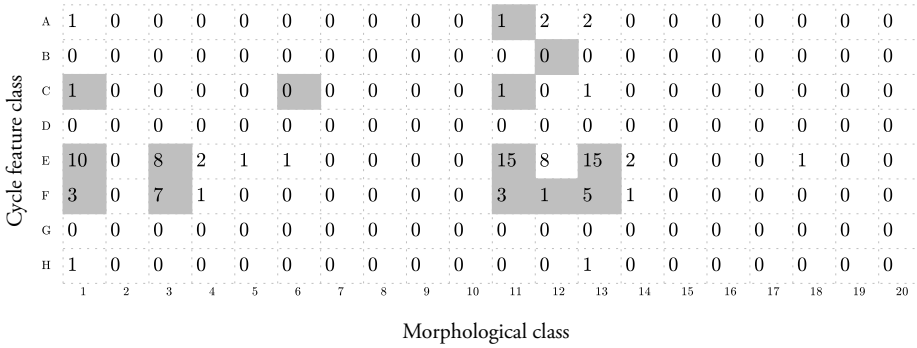


Figure 8: Distribution of the relative prevalence of waveform classification (after outlier removal) for horizontal and vertical eye–movement data, rounded to the nearest percent. The gray squares represent positions where a template waveform is present. The x-axis represents the morphological class, and the y-axis represents the cycle feature class.

Classification example

An example of the waveform classification for a segment of a recorded signal is shown in Figure 9. In this example, there are no breaking saccades, but there are both

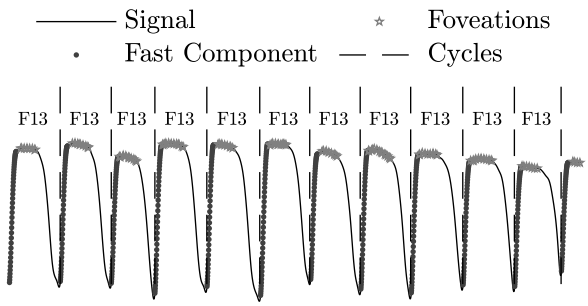


Figure 9: Illustration of the cycle feature detection and morphological classification for a recorded signal segment. The black line represents the waveform, the gray dots the fast phases, the gray asterisks the foveation period, and the dashed vertical lines the cycle boundaries. All cycles in this sequence are classified as F13 (EF-R).

foveation periods and fast phases. The resulting cycle feature class is F and the morphological class for these waveforms is 13, resulting in a class of F13. Based on the waveform classification grid, the resulting interpretation is EF-R.

Performance evaluation

The waveform classification results for each of the different participants are together with the expert annotations [11] summarised in Tables 8 – 11, Appendix A . The *participant* column shows the unique participant ID and the detected waveforms for the participant, the *percentage* column shows the relative prevalence of each waveform (after outlier rejection), the *CWL* column contains the *corresponding waveform label*, and finally, the *annotation* contains the expert annotation [11] for the participant. The corresponding waveform label is the name of the template waveform for a given position on the waveform classification grid, for example EF-R is represented by F13 on the grid. If there is no corresponding template waveform for that position in the waveform classification grid, it is marked by '-'. The comparison to the expert annotations resulted in six different waveforms for which there exists at least one true positive detection, i.e., a match between a detected waveform and an annotation. These are: EF-L, EF-R, J-L, J-R, PC-L and PC-R. A comparison between the detected waveforms and the available annotations can be found in Table 6. In total, for both the magnetic scleral search coil and video–based datasets, 32 signals were analysed (16 in each dataset). In the magnetic scleral search coil datasets, 26 false positive and 17 true positive average waveforms were found, and in the video–based dataset, 70 unannotated average waveforms were included in the analysis.

Table 6: Summary of the analysed average waveforms. A total of 32 signals were analysed, where 70 unannotated (UN), 26 false positive (FP), and 17 true positive (TP) average waveforms were detected.

Database	Number of Signals	Number of UN	Number of FP	Number of TP
Search coil	16	-	26	17
Video	16	70	-	-
Total	32	70	26	17

In Table 7, the results of the comparison between false positive/unannotated waveforms and the six true positive waveforms are shown. A total of 70 unannotated and 26 false positive average waveforms were analysed. The ratio denotes the proportion of average waveforms that are considered to match the true positive waveform for a specific waveform class in terms of waveform similarity. For example, all (2/2) unannotated average waveform for the EF-L/D class were matched to the corresponding true positive average waveforms.

For two of the waveform classes, EF-R/U and PC-R/U, all unannotated and false positive average waveforms are matched to the corresponding average true positive waveforms. The worst performance is seen for the J-R/U waveform, where 50 %

Table 7: Average waveform matching results. For each waveform class, the number of unannotated and false positive average waveforms that are matched to their corresponding true positive average waveforms are listed. This number is compared to the total number of average waveforms assigned to that waveform class. A total of 70 unannotated and 26 false positive average waveforms were analysed, respectively. The left column shows the waveform class, the middle column shows the performance for the unannotated data and the right column shows the performance of the false positive data for each specific waveform class, for each waveform class. In the EF-R/U waveform class for example, 16 / 16 unannotated and 3 / 3 false positives average waveforms are matched to the corresponding true positive average waveforms.

Waveform class	Unannotated	False positives
EF-L/D	2/2	0 / 0
EF-R/U	16/16	3/3
J-L/D	3/10	6/7
J-R/U	9/16	2/6
PC-L/D	19/23	7/7
PC-R/U	3/3	3/3
Total	52/70	21/26

(11/22) of the average waveforms are matched to the corresponding true positive average waveforms. In total, 21 out of 26 false positive waveforms were matched in terms of similarity to their corresponding true positive average waveforms in another signal.

In Figure 10, examples of the six analysed waveform classes presented in Table 7 are illustrated. The black signals represent the true positive waveforms, the gray signals the false positive waveforms and the light gray signals represents the unannotated signals. As indicated in the figure, the signals for the different waveform classes are morphologically similar to the true positives.

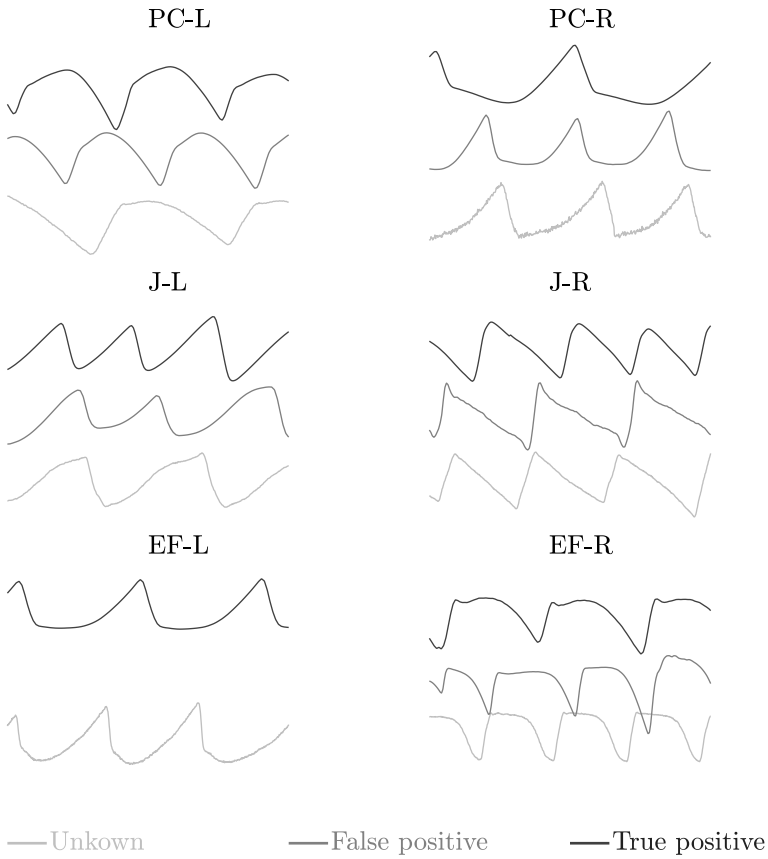


Figure 10: Illustration of representative waveforms for the six waveform classes presented in Table 7. For each waveform class, the true positive example is plotted in dark gray (upper), a false positive example is plotted in gray (middle), and an unannotated example is plotted in light gray (lower).

5 Discussion

In this work, a novel method for cycle–to–cycle classification of nystagmus oscillations is proposed. The method combines two approaches previously used to study nystagmus oscillations; cycle feature analysis and morphological classification. The cycle feature analysis identifies whether known features are present in a cycle, such as slow phases, fast phases, and foveation periods. A novelty compared to previous work is the inclusion of breaking saccades. For the morphological classification, principle component analysis is used to model the oscillatory patterns. The proposed method

was evaluated for idealised template signals as well as signals recorded with both video–based and magnetic scleral search coil eye tracking techniques. The magnetic scleral search coil recordings were accompanied by expert annotations. The main novelty of this paper is to classify nystagmus cycles on a waveform classification grid, which combines cycle feature analysis and morphological classification.

To analyse cycle features and waveform morphologies may be viewed as studying the waveforms on a micro and a macro scale, respectively. The analysis of cycle features is difficult to perform using only methods based on base functions. For example, the signal energy of a breaking saccade is small, which means that it is likely not captured by principle components or sinusoidal models. On the other hand, by adding the morphological classification, it is possible to identify the direction of the beats, and the overall shape of the oscillations, which may be more difficult to achieve using only cycle feature analysis, as it would require many features.

The usefulness of the proposed method becomes evident when considering the signal (S7) shown in Figure 4. In the figure, it appears that EF-R waveforms are present between seconds 123 and 124, even though these waveforms were not included in the expert annotation [11]. In general, annotation of nystagmus signals is difficult due to both subtle differences between different nystagmus types, and to temporal variations in the nystagmus waveforms. Since the magnetic scleral search coil signals are several minutes long, and were manually annotated, it was not feasible to annotate every single cycle. By using an algorithm to perform this analysis, signals can be analysed much faster and in more detail, even down to each oscillation cycle, as is done in this work.

As the results in Table 7 show, there are large similarities between the true positive, false positive and unannotated average waveforms. This shows that the method is capable of grouping together similar waveforms in nystagmus signals recorded with different eye tracking systems. Even though the performance vary between waveform classes, the results suggest that the proposed method is reliable for automatic waveform detection. For the false positive waveforms, i.e., those that did not match any expert annotations, 21 out of 26 waveform classes are still, based on similarity measures, considered to be correctly classified by the proposed method. The worst results were obtained for the J-R waveform, where only 50 % of the analysed averaged waveforms are matched to the corresponding true positive data. This means that categorising a cycle as J-R has the highest uncertainty of the analysed waveform classes in this work. For each of the three waveform classes EF-L/D, EF-R/ U and PC-R/U, there was a 100 % match between the true positives, false positives, and as well as unannotated data.

The results of a classification algorithm are usually presented in terms of sensitivity and specificity. These metrics are difficult to use in this work, since the waveform annotations are assumed to contain only the most prominent waveforms, and

no cycle–to–cycle annotations exist. In the original work, the annotations were used to describe the oculomotor characteristics for each participant [11], and not cycle–to–cycle oscillations. The results presented in Table 7 and Figure 10 show that the proposed waveform classification leads to high similarity within each waveform class. This is an important feature when performing repeated experiments, or when evaluating treatment effects.

An important consideration for this method is that the morphological classification is based on reconstruction of the template waveforms presented in [3]. These template waveforms are schematic drawings of different types of oscillatory patterns during nystagmus. It is possible that cycle features and morphological characteristics do not perfectly correspond to these templates when studying real data. For example, it is not clear whether the dual jerk (DJ) waveform actually contains any foveations (Figure 6). It is also likely that the signal–to–noise ratio in recorded data is an important factor in certain recordings. If the signal–to–noise ratio is too low, misclassifications may occur. The proposed method was designed to find patterns in the different signals that resemble these template waveforms, but it would be preferable to have a larger dataset of annotated waveforms to base the classification on. In some cases, multiple template waveforms may have the same classification, for example AP and J-R. Clinically, they are often mistaken for one another [3]. In cases where multiple waveform morphologies are mapped to the same waveform classification grid coordinate, a closer inspection of the signal is required.

There are some uncertainties related to the magnetic scleral search coil dataset that is used in this work: `http://www.omlab.org/OMLAB.html`. The annotations described in [11] lists a set of waveform morphologies for each participant. The available dataset, however, contains multiple recordings for most participants, and it is not clear which waveform morphologies that are associated with which recordings.

In the present study, the focus was mainly to divide the different oscillation patterns into separate waveform classes. The proposed method may, however, be used to study other important characteristics such as foveation duration, frequency and amplitude of the oscillations, etc. In addition, by using adaptive thresholds instead of absolute thresholds, it is possible to use the proposed method on uncalibrated eye-tracking data [5]. This is beneficial when working with nystagmus patients, who in general are difficult to calibrate [16].

6 Conclusion

The waveform classification method proposed in this work has demonstrated high performance for the classification of nystagmus eye movement recordings. The majority of the classified waveforms are in good agreement with the simulated template wave-

forms on the waveform classification grid. The results of the average waveform matching method shows that a majority of the non–true positive waveforms are matched in terms of waveform similarity to corresponding true positive waveforms in the analysed datasets.

7 Acknowledgements

This work has been funded by the Swedish Research Council [grant number VR 2015-05442]. We gratefully acknowledge the Lund University Humanities Lab. We are thankful to Markus Rahne and Linnea Sjödahl, who have contributed with data collection. We would also like to thank all the participants. Dr Kristina Holm is acknowledged for kindly referring patients to our study, which we are grateful for. None of the data reported in this work are publicly available, and no experiments were preregistered.

References

[1] R. J. Leigh and D. S. Zee, *The neurology of eye movements*. Oxford University Press, USA, 2015.

[2] L. Dell'Osso, "Evaluation of smooth pursuit in the presence of congenital nystagmus," *Neuro-ophthalmology*, vol. 6, no. 6, pp. 383–406, 1986.

[3] L. Dell'Osso and R. Daroff, "Congenital nystagmus waveforms and foveation strategy," *Documenta Ophthalmologica*, vol. 39, no. 1, pp. 155–182, 1975.

[4] L. Frisén, "Nystagmus lathund," *Läkartidningen*, vol. 78, no. 22, pp. 2229–2232, 1981.

[5] M. Dunn, *Quantifying perception and oculomotor instability in infantile nystagmus*. PhD thesis, Cardiff University, 2014.

[6] G. Pasquariello, M. Cesarelli, M. Romano, A. La Gatta, P. Bifulco, and A. Fratini, "Waveform type evaluation in congenital nystagmus," *Computer Methods and Programs in Biomedicine*, vol. 100, no. 1, pp. 49–58, 2010.

[7] J. Jacobs and L. Dell'Osso, "Extending the expanded nystagmus acuity function for vertical and multiplanar data," *Vision research*, vol. 50, no. 3, pp. 271–278, 2010.

[8] R. Abadi and R. Worfolk, "Harmonic analysis of congenital nystagmus waveforms," *Clinical vision sciences*, vol. 6, no. 5, pp. 385–388, 1991.

[9] M. Theodorou and R. Clement, "Classification of infantile nystagmus waveforms," *Vision Research*, vol. 123, pp. 20–25, 2016.

[10] W. Rosengren, M. Nyström, B. Hammar, M. Rahne, L. Sjödahl, and M. Stridh, "Modeling and quality assessment of nystagmus eye movements recorded using an eye-tracker," *Behavior Research Methods*, vol 52, pp. 1–15, 2020.

[11] L. Dell'Osso, J. Jacobs, and A. Serra, "The sub-clinical see-saw nystagmus embedded in infantile nystagmus," *Vision research*, vol. 47, no. 3, pp. 393–401, 2007.

[12] E. Billauer. "peakdet: Peak detection in MATLAB," url: *http://www.billauer.co.il/peakdet.html*, visited: 12/07/2017.

[13] SR-Research, "*EyeLink 1000 User Manual.*," 2010.

[14] J. W. Peirce, "Psychopy psychophysics software in python," *Journal of Neuroscience Methods*, vol. 162, no. 1-2, pp. 8–13, 2007.

[15] W. Rosengren, M. Nyström, B. Hammar, and M. Stridh, "Waveform characterisation and comparison of nystagmus eye-tracking signals," *in review*, 2020.

[16] W. Rosengren, M. Nyström, B. Hammar, and M. Stridh, "A robust method for calibration of eye tracking data recorded during nystagmus," *Behavior research methods*, vol 52, pp. 1–15, 2019.

Appendix A

In this appendix, detailed information about the waveform classification for all recorded signals are presented. For each table the following applies. The columns show the **participant** ID, the detected prevalence of the waveform class, the corresponding waveform label for the specified class, and the **annotation** for the recording. If there are no corresponding template waveform or no annotations for a specific signal, it is marked by '-'. The participants listed in Tables 8 and 10 have been recorded using the EyeLink system, and the participants listed in tables 9 and 11 have been recorded using the magnetic scleral search coil system, and this data is publicly available at `http://www.omlab.org/OMLAB.html`.

Table 8: Horizontal waveforms - unannotated. Eyelink system

Participant	Percentage	CWL	Annotation
p1			
E1	20.0	J-L	
F13	13.0	EF-R	
F1	11.0	DJ-L	-
H1	10.0	-	
C1	8.0	PFs, PJ-L, PPFs	
p2			
E11	67.0	AP, J-R	
C11	18.0	PJ-R	-
B11	5.0	-	
p3			
F3	57.0	EF-L	
F1	18.0	DJ-L	-
E3	9.0	PC-R	
p4-1			
E1	87.0	J-L	
E3	8.0	PC-R	-
p4-2			
E11	51.0	AP, J-R	
E13	33.0	PC-L	-
E12	7.0	-	
p5			
E13	19.0	PC-L	
E1	13.0	J-L	
E11	11.0	AP, J-R	-
C13	7.0	-	
C1	6.0	PFs, PJ-L, PPFs	
p6			
E11	24.0	AP, J-R	
E12	21.0	-	
E3	8.0	PC-R	-
E1	5.0	J-L	
E4	5.0	-	
p7			
F1	21.0	DJ-L	
E1	12.0	J-L	
E12	8.0	-	
E13	7.0	PC-L	-
E14	7.0	-	
E11	6.0	AP, J-R	
p8-1			
F13	73.0	EF-R	
E13	15.0	PC-L	-
p8-2			
F13	90.0	EF-R	
E13	6.0	PC-L	-
p8-3			
F13	73.0	EF-R	
E13	13.0	PC-L	-
p8-4			
F13	39.0	EF-R	
E1	18.0	J-L	-
p8-5			
F13	82.0	EF-R	
E13	11.0	PC-L	-
p8-6			
F13	83.0	EF-R	
E13	10.0	PC-L	-
p8-7			
F13	57.0	EF-R	
E13	23.0	PC-L	-
F12	5.0	BDJ-R	
p8-8			
F13	50.0	EF-R	
E13	31.0	PC-L	-

Table 9: Horizontal waveforms - Annotation. Magnetic scleral search coil system

Participant	Percentage	CWL	Annotation
s7a			
F1	18.0	DJ-L	
F13	17.0	EF-R	
E1	15.0	J-L	EF-L
E13	6.0	PC-L	
E3	5.0	PC-R	
s8a			
E13	27.0	PC-L	
E12	22.0	-	J-L, P, PC-R
E1	19.0	J-L	
E11	10.0	AP, J-R	
s9a			
E13	22.0	PC-L	
F3	19.0	EF-L	
E3	13.0	PC-R	
F13	10.0	EF-R	J-R, PC-R, EF-L, PC-L
E12	8.0	-	
E11	6.0	AP, J-R	
E1	6.0	J-L	
s11b			
E1	14.0	J-L	
E13	9.0	PC-L	PPFs
E11	9.0	AP, J-R	
E18	6.0	-	
s12a			
F3	44.0	EF-L	
E4	14.0	-	PC-L, EF-L, PJ-R
E3	12.0	PC-R	
F4	7.0	-	
s13a			
E11	48.0	AP, J-R	
E1	17.0	J-L	J-R, J-L, EF-L
E13	6.0	PC-L	
s14a			
E3	26.0	PC-R	
E12	10.0	-	PPFs, P, PFs, PC-L, EF-L
E4	9.0	-	
E11	9.0	AP, J-R	
s15a			
A12	59.0	-	PC-L, J-L, DJ-L PC-R, J-R, DJ-R, P
E3	8.0	PC-R	
E1	8.0	J-L	

Table 10: Vertical waveforms - unannotated. EyeLink system.

Participant	Percentage	CWL	Annotation
p1			
E13	24.0	PC-D	
A1	17.0	-	-
E1	7.0	J-D	
F11	7.0	BDJ-D, DJ-U	
p2			
E11	44.0	AP, J-U	
E13	16.0	PC-D	-
F11	7.0	BDJ-D, DJ-U	
p3			
A1	30.0	-	
E3	11.0	PC-U	
E1	10.0	J-D	
E13	10.0	PC-D	-
A3	6.0	-	
E6	6.0	-	
E11	6.0	AP, J-U	
A18	6.0	-	
p4-1			
E11	50.0	AP, J-U	-
F11	10.0	BDJ-D, DJ-U	
p4-2			
E11	69.0	AP, J-U	-
E13	15.0	PC-D	
p5			
p6			
E1	21.0	J-D	
E12	21.0	-	
E13	17.0	PC-D	-
E11	10.0	AP, J-U	
E14	7.0	-	
p7			
A3	25.0	-	
E11	19.0	AP, J-U	
A4	9.0	-	-
E3	8.0	PC-U	
A11	5.0	P	
p8-1			
E13	39.0	PC-D	
F13	17.0	EF-U	-
E11	8.0	AP, J-U	
p8-2			
E13	37.0	PC-D	-
F13	26.0	EF-U	
p8-3			
F13	47.0	EF-U	-
E13	32.0	PC-D	
p8-4			
F3	17.0	EF-D	
F12	14.0	BDJ-U	
F11	11.0	BDJ-D, DJ-U	
E11	8.0	AP, J-U	-
E13	8.0	PC-D	
E1	6.0	J-D	
F13	6.0	EF-U	
p8-5			
E13	40.0	PC-D	
F13	24.0	EF-U	-
E11	6.0	AP, J-U	
F12	5.0	BDJ-U	
p8-6			
F13	66.0	EF-U	
E13	17.0	PC-D	-
F12	6.0	BDJ-U	
p8-7			
E13	16.0	PC-D	
A1	11.0	-	-
B11	8.0	-	
E11	8.0	AP, J-U	
p8-8			
E13	27.0	PC-D	
F13	21.0	EF-U	-
E11	16.0	AP, J-U	
F11	5.0	BDJ-D, DJ-U	

Table 11: Vertical waveforms - Annotation. Magnetic scleral search coil system

Participant	Percentage	CWL	Annotation
s7a			
E1	13.0	J-D	
E13	12.0	PC-D	
F1	10.0	DJ-D	
F13	8.0	EF-U	-
E14	6.0	-	
C1	5.0	PFs, PJ-D, PPFs	
F14	5.0	-	
s8a			
E13	41.0	PC-D	
E12	11.0	-	P, EF-U
F13	8.0	EF-U	
E11	6.0	AP, J-U	
s9a			
E11	40.0	AP, J-U	
E3	10.0	PC-U	J-U, PC-D
E1	10.0	J-D	
s11b			
E11	36.0	AP, J-U	J-U
F11	10.0	BDJ-D, DJ-U	
s12a	-		
E11	25.0	AP, J-U	
F11	11.0	BDJ-D, DJ-U	
E1	9.0	J-D	EF-U, J-U
E13	7.0	PC-D	
F1	7.0	DJ-D	
s13a			
E1	17.0	J-D	
E13	11.0	PC-D	
F1	8.0	DJ-D	
E11	8.0	AP, J-U	J-D
C1	7.0	PFs, PJ-D, PPFs	
F13	6.0	EF-U	
F11	5.0	BDJ-D, DJ-U	
s14a			
E3	22.0	PC-U	
F3	14.0	EF-D	
E1	6.0	J-D	PPFs, P, PFs, EF-D
E12	6.0	-	
E11	5.0	AP, J-U	
F11	5.0	BDJ-D, DJ-U	
s15a			
A13	38.0	-	
A12	17.0	-	J-D, DJ-D, J-U DJ-U, P
E1	9.0	J-D	