CHALMERS
UNIVERSITY OF TECHNOLOGY

Automatic detection of saccadic eye movements using EOG for analysing effects of cognitive distraction during driving

Master's thesis in Biomedical Engineering

ALESSANDRO CAFASSO

SEBASTIAN KARLSSON

Department of Applied Mechanics
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2017

Master's thesis 2017: Biomedical Engineering

Automatic detection of saccadic eye movements using EOG for analysing effects of cognitive distraction during driving

ALESSANDRO CAFASSO
SEBASTIAN KARLSSON

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Applied Mechanics
Division of Vehicle Safety
Chalmers University of Technology
Gothenburg, Sweden 2017

Automatic detection of saccadic eye movements using EOG for analysing effects of cognitive distraction during driving
ALESSANDRO CAFASSO
SEBASTIAN KARLSSON

Supervisor: Emma Nilsson, Volvo Car Corporation
Supervisor: Per Lindén, Volvo Car Corporation
Examiner: Robert Thomson, Department of Applied Mechanics

Master's Thesis 2017:62
Department of Applied Mechanics
Division of Vehicle Safety
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Gothenburg, Sweden 2017

Automatic detection of saccadic eye movements using EOG for analysing effects of cognitive distraction during driving.
Master's Thesis in Biomedical Engineering
ALESSANDRO CAFASSO and SEBASTIAN KARLSSON
Department of Applied Mechanics
Division of Vehicle Safety
Chalmers University of Technology

Abstract

Driver distraction is a relevant driving safety issue and an ongoing field of research. A particular distraction is cognitive distraction, which refers to when the driver is mentally engaged in a task unrelated to driving, e.g. talking to a passenger. Eye movements can be analyzed to study effects of cognitive distraction during driving, and are typically recorded using video-based eye tracker systems. An alternative technique that might be suitable for eye movements measurements during driving is the electro-oculography (EOG). EOG is a method for recording the electrical signal of the eyes as they move. One interesting eye movement in cognitive distraction studies is the saccade, the rapid movement of the eye from one point of interest to another.

The primary purpose of this thesis is to develop an algorithm for automatic detection of saccades using EOG. The resulting algorithm is a combination of two modified existing eye detection algorithms, namely Continuous Wavelet Transform Saccade Detection (CWT-SD) and Shape Features. It is found that the developed algorithm can be used in driving environments if good signal quality can be assured. The secondary purpose of this thesis is to investigate how cognitive distraction affects saccadic rate and amplitude during driving. The findings suggest a statistically significant decrease in saccadic rate during cognitive load but not in saccade amplitude. However, further research on bigger datasets and different driving scenarios is needed to verify the results.

Keywords: saccades, eye movements, EOG, saccade detection, classification, cognitive distraction, driver distraction

Contents

List of Figures

List of Tables

1
Introduction

Drivers' behavioral abilities and limitations need to be thoroughly studied in order to develop optimal driver support systems. The level of drowsiness, distraction and stress may significantly affect the driver's ability to drive safely and handle traffic situations. A particular driver state is cognitive distraction, which refers to when a driver is mentally engaged in something else than driving. When a driver is cognitively distracted, his/her eyes might be on the road and hands on the steering wheel, but driving abilities might be impaired since the mental attention is allocated to another task. An example of this is using a speech-to-text system or talking on the phone. The effects of cognitive distraction have been heavily debated, and in order to learn more about the subject and understand the effects of driver distraction, reliable measuring techniques are needed. Deeper understanding of cognitive distraction during driving could potentially increase traffic safety and consequently lead to societal benefits.

One way to study cognitive distraction is through measuring the driver's eye movements. They can both be used to detect cognitive distraction (e.g. through an increased blink frequency), and to understand effects on driver performance (e.g. through increased gaze concentration wherein drivers look more in the forward road direction). The most interesting eye movements in cognitive distraction studies are saccades, the rapid movement of the eye from one point of interest to another.

Traditionally, eye movements in car drivers are detected by camera-based eye tracking systems. However, eye trackers do not always provide sufficient precision in time resolution in studies where a precise time detection of the eye movements is necessary. Moreover, the light conditions, as well as large and sudden movements of the driver's head, can affect tracking performance.

An alternative measurement technique for recording eye movements is electro-oculography (EOG), which is used to measure the electrical signal of the eye. The signal varies depending on the eye's position. A typical measurement setup consists of four electrodes placed around the test subject's eyes, measuring the potential between the cornea and retina as the eyes move.

The main purpose of this thesis is to develop an algorithm in MATLAB for automatic detection and classification of saccades using EOG data. The goal is to detect as many and small saccades as possible. The main challenges are noise and artifacts, limiting the detection capability. Therefore, limitations related to the qual-

ity of the signal have to be taken into account and an amplitude threshold below which the detection algorithm cannot perform reliably, has to be found. Moreover, a consistent definition of saccade start- and end point is needed to determine the amplitude of saccades. The thesis aims to answer the question "Can EOG be used for saccade detection in driving environments in a research context?".

The intended application of the developed algorithm is primarily cognitive distraction studies. In this thesis, a limited analysis of how cognitive distraction affects car drivers' saccadic eye movements is performed. Saccadic rate and saccade amplitude are considered for the analysis. The difference between drivers that are cognitively distracted and not cognitively distracted is studied in order to answer the second question; "How are frequency and amplitude affected by cognitive distraction during driving?".

2
Background

In this chapter, background information about the eye, eye movements and the implemented algorithm is presented. First, the relevant anatomy of the eye and eye movements is provided. In connection to this, EOG and other eye measuring techniques are presented. Next, current eye movement detection algorithms are provided together with the core theory behind the implemented algorithms. Finally, background information regarding the cognitive distraction study is given.

2.1 General structure of eye

The eye is the organ of the body responsible for detection, localization, and analysis of the light, and represents the core of one of the sensory systems, called "vision" [1]. In a general broad way, the eye can be thought as a camera, which mainly consists of the lens, the shutter (eyelid), the film (retina) and the cable (optic nerve) connecting the camera (eye) to the computer (brain). A rough illustration of the anatomy of the eye is shown in Figure 2.1.

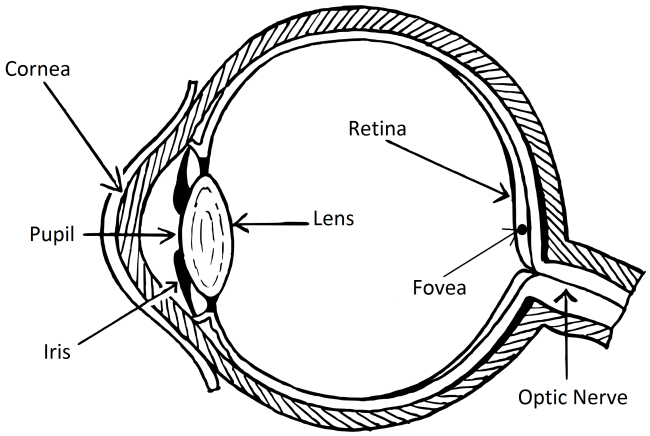


Figure 2.1: Gross anatomy of the eye

The retina is the light-sensitive innermost coat of the eye and it plays a fundamental role both for vision and eye movements recording. It consists of a neural layer containing photo-receptor cells specialized in converting light into images, and a pigmented layer, called "retinal pigment epithelium", that contributes to absorbing the stray light rays [2] and it also represents the origin of EOG [3]. The accumulation of electrical charges in the pigment layer of the retina creates an electrical potential in

the eye and it makes EOG recording possible. More details about EOG are further presented in Section 2.3.1.

The retina has also a paramount importance due to the presence of the fovea in its center. The fovea is a small depression of about 2 mm in diameter [1] and it is the area of highest visual acuity (sharpness of vision). Its presence is the reason why the eyes move, so that the light rays (the real world images) can be focused on the fovea.

2.2 Eye Movements

The eyes move in order to extend the visual range that can be covered without moving the head. The oculomotor system is responsible for the eyeballs to move and rotate inside their sockets in the skull with the help of the extraocular muscles. There exist different types of eye movements that can be distinguished based on their physiological properties and how they contribute to vision [4]. These basic types are saccade, fixation, smooth pursuit, vergence, vestibular and optokinetic nystagmus.

Saccades are quick movements of the eye performed in order to bring images of objects of interest onto the fovea. Fixations consist of miniature movements that stabilize the image of a stationary object on the fovea in order to acquire visual information, while smooth pursuits occur when the eyes are tracking a smooth moving target. Thus, they stabilize the image on the fovea as the eyes are rotating. Vergence movement refers to the situation when the eyes move in opposite directions in order to place the image of a object onto the fovea of each eye. Nystagmus ocular movements consist of smooth pursuits interspersed with saccades, appearing as a sawtooth-like time series pattern, and they can be distinguished in two types, vestibular and optokinetic nystagmus. Vestibular nystagmus occurs in order to hold the images steady on the retina during short head movements, whereas optokinetic nystagmus has the same function but during faster movements of the head [4, 5]. However, only three of these eye movements, namely saccade, fixation and smooth pursuit, represent the fundamental ones that need to be modelled because of their unique functionality [5]. Therefore, in the sections 2.2.1, 2.2.2, 2.2.3 more detailed information about these three eye movements is given.

The rotation of the eyes is due to the action of three pairs of extraocular muscles, which are displayed in Figure 2.2. These work together as agonist-antagonist muscles allowing rotations in the horizontal, vertical and torsional directions. Antagonist refers to that the activity of the antagonist muscle opposes the one of the agonist, and follow the principle of reciprocal innervation [6]. These three pairs of extraocular muscles are: the **lateral rectus** and **medial rectus**, the **superior rectus** and **inferior rectus**, the **superior oblique** and **inferior oblique** [6]. The first two aforementioned muscles (medial and lateral recti) are entirely responsible for moving the eye horizontally, while the other four are used to move the eye in a vertical and torsional motion [6]. Indeed, elevation (upwards eyeball movement) is

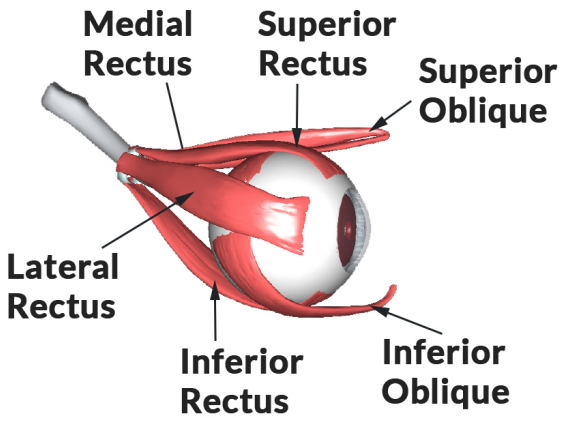


Figure 2.2: Oculomotor system showing the six extraocular muscles. To note that the medial rectus muscle is not visible in the figure. As its counterpart it looks like the lateral rectus but located on the opposite side of the eye.

performed by the superior rectus and the inferior oblique muscle, while depression (downwards eyeball movement) is actuated by the inferior rectus and the superior oblique [7]. Three cranial nerves innervate the six extraocular muscles, namely the abducens (cranial nerve VI), the trochlear (cranial nerve IV) and the oculomotor nerve (cranial nerve III), which are originated in two different regions of the brainstem. The abducens nerve supplies the lateral rectus and has its associated nuclei in the pons, while both the trochlear nerve and the oculomotor nerve exit from the midbrain. The trochlear nerve innervates the superior oblique, whereas the oculomotor nerve triggers the superior rectus, the inferior rectus, the medial rectus and the inferior oblique [7][2].

2.2.1 Saccades

A saccade is a rapid eye movement with the purpose of moving the fovea from one point of interest to another. Consequently, saccades carry out a vital function of the eye, since the fovea is where the visual acuity is sharpest. Saccades can be both reflexive or voluntary. Normal triggers of saccades are the appearance of an object seen or heard, but they can also be triggered by memory or be part of visual search. There exists a direct link between the presence of a fovea in primates and voluntary saccades, and primates without a fovea, i.e. rabbits, mainly performs saccades during head movements [4].

The saccade signal of the ocular motoneurons has the shape of a pulse-step, as shown in Figure 2.3 and it encodes the characteristics of the saccade in terms of their temporal discharge. The duration and the size of the saccade are proportional to the total number of the pulses (discharge spikes). This burst of activity of the ocular motoneurons is originating in the brainstem and it starts about 8 ms before the eye begins to move [4].

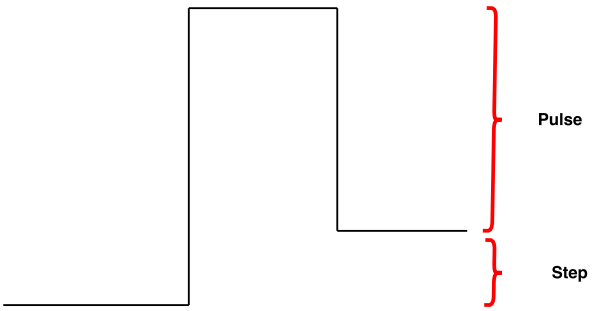


Figure 2.3: Pulse-step waveform. To generate a saccade the ocular motoneurons send a sequence of signal with a shape of a pulse-step.

Once the programming of a motor command is completed and the saccade is initialized, the target point cannot be altered. Saccades rarely last longer than 100 ms, which is about the same time needed for the visual system to process new information and turn it into a new motor command. This is why visual information cannot be used to alter the target point after a saccade is triggered. The latency of saccades (time between target presentation and the start of eye movement) depends on several factors. Examples of factors are properties of the stimulus (visual or auditory, target size/intensity etc.), the amount of available information and the urgency to make a decision. In studies where subjects were asked to track a target that moves from one point to another, a saccade is normally initiated with a latency of around 200 milliseconds [4]. It is believed that the latent period represents the time interval during which the decision whether to generate a saccade or not is taken by the central nervous system (CNS) [6].

Speed, amplitude, and duration of normal saccades are related to each other, where greater amplitude results in higher peak velocities and longer duration. The peak velocity V_{peak} of a saccade normally increases approximately linear with the amplitude up to 20° and starts to saturate around 500°/s for larger saccades [4, 8]. The relationship is commonly described by equation (2.1), where V_{max} is the maximum velocity of a saccade (around 500°/s), A is the amplitude and C is a constant normally set to the amplitude where the velocity reaches 63% of the saturation value [8].

$$V_{peak} = V_{max}(1 - e^{-A/C}) \tag{2.1}$$

The period before a saccade reaches peak velocity is the acceleration period, and the period after peak velocity until the end of the saccade is the deceleration period. The asymmetry between acceleration and deceleration period duration is called skewness (which is calculated as [duration of acceleration period]/[duration of acceleration period]+[duration of deceleration period]). The skewness ranges between 0.2 and 0.5, where 0.5 corresponds to the smallest saccades and 0.2 to the largest. Another measurement of saccades is the ratio between peak velocity and mean velocity. This ratio is normally around 1.6 [4].

Start- and end points of a saccade are commonly decided by a velocity criterion.

The criterion can be either fixed or dynamic. Fixed criterion is when the start point is defined as the point where the velocity exceeds a fixed threshold, i.e. 30 °/s, and the end point when the velocity falls below the same threshold [4]. Another similar approach is a dynamic threshold, where the thresholds are set to 10% of the peak velocity of the saccade [9].

The duration of saccades is approximately linear to the amplitude A. This relationship is often described by a linear function described in equation (2.2) or by a power function described in equation (2.3):

$$Duration = D_0 + dA \tag{2.2}$$

$$Duration = D_1 A^P \tag{2.3}$$

where D_0 and d are the constants of the linear function, while D_1 is the duration of a 1-degree saccade and P is a constant. The power function relationship more accurately describes the duration of smaller saccades.

After a primary saccade, a short, oppositely directed small saccade can occur. This is called "dynamic overshoot" and the phenomenon is illustrated in Figure 2.4. Dynamic overshoots normally have an amplitude in the range of 0.25° - 0.5°, but larger and smaller dynamic overshoots can occur [4]. The reasons for dynamic overshoots are not clear. One theory is that the reversal breaking pulses that bring the eye to an abrupt stop occasionally are too large, causing a dynamic overshoot to occur [10]. The reason can also be mechanical properties of the orbital tissues [4].

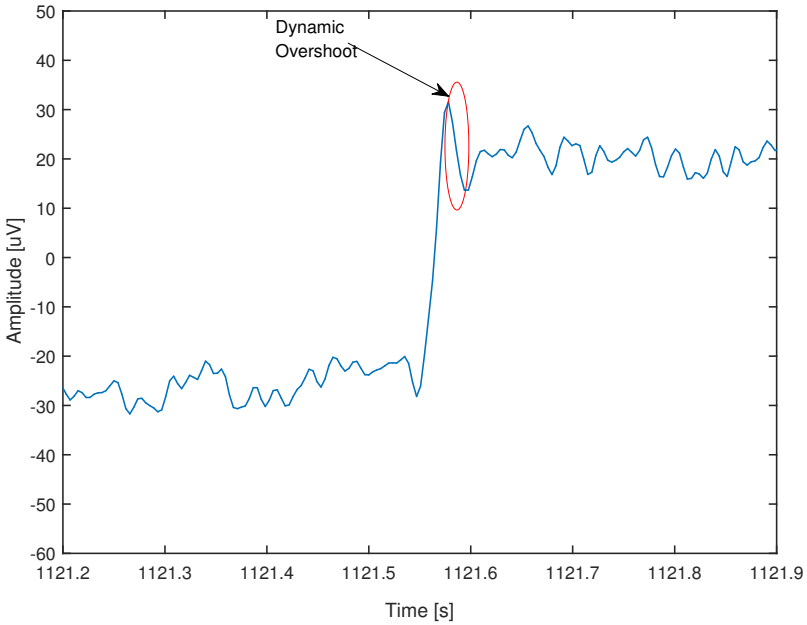


Figure 2.4: Example of a dynamic overshoot, a small saccade occurring after a primary saccade and it is oppositely directed

If the primary saccade fails to reach the desired target point, a secondary saccade can be triggered to reduce the error. These secondary saccades are commonly called "corrective saccades". This type of saccades can occur even if no visual feedback is available after the primary saccade is completed (target is no longer visible). However, when visual feedback is present, the corrective saccade are more accurate and have lower latency. Furthermore, the tolerance for error is also lower when visual information is available. The latency of corrective saccades is dependent on the size of the position error, where greater errors are corrected faster [11]. In studies where test subjects track a moving target point, the average latency of visually guided corrective saccades is in the interval of 100-130 ms [4]. The latency of saccades was reduced when the target point was available in the deceleration phase of the saccade but turned off before the saccade was completed. This is an indication that visual information available during the deceleration phase of the saccade is used to trigger corrective saccades [4, 11].

2.2.2 Fixations

Fixations are considered as the time intervals of apparent ocular immobility between two saccades, during which visual and cognitive information is acquired and processed [12]. On average the fixations have a duration that lies between 100 ms and 200 ms [13] but they can also last longer [14]. An example of fixation occurring between two saccades can be seen in Figure 2.5, where the saccades are marked by violet ellipsoids and the fixation by a gray ellipsoid.

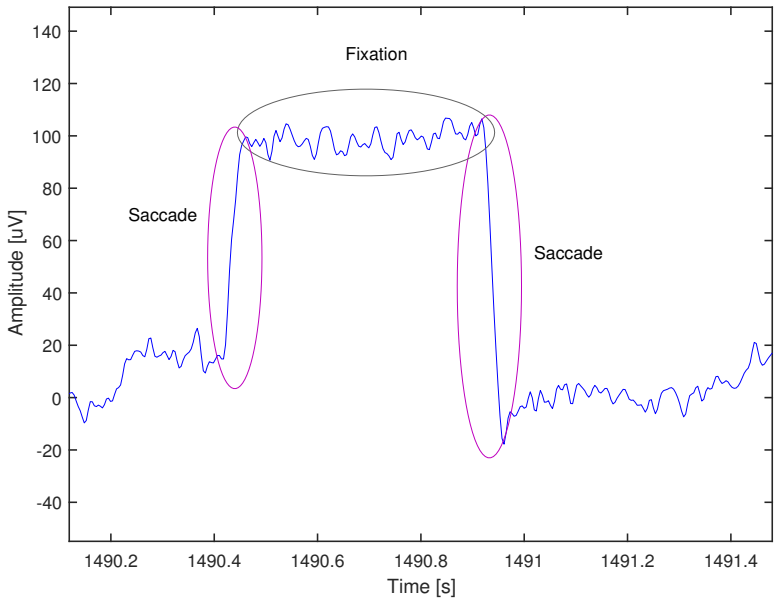


Figure 2.5: Saccades and fixation. Example of a sequence of saccade-fixation-saccade taken from a horizontal EOG signal.

The eye is not completely still during fixation since it performs microscopic gaze shifts, commonly known as "fixational eye movements", which help to hold the central foveal vision in place. These involuntary fixational eye movements prevent fading of images and consist mainly of three components: tremor, microsaccades and ocular drift. Ocular tremor has a frequency range up to 150 Hz and an amplitude less than 0.01°, which is smaller than the size of one photo-receptor in the fovea. Microsaccades are the small involuntary saccades that occur during fixation and can be considered the miniature replicas of the saccadic gaze shift presenting an amplitude range between 0.02° and 2° [4][12]. Ocular drift is the low-frequency component (<40 Hz) that resembles the random motion of a particle in a fluid (Brownian motion) [12].

2.2.3 Smooth Pursuits

Smooth pursuit movements allow clear vision of the object as it moves within the visual environment [4]. They are designed to maintain the image of the smooth moving object on the fovea [7]. Actually, saccades are able to catch the image of a moving target but, in absence of pursuit, the image would slide off after each saccade [4]. Moreover, it is a voluntary eye movement and the observer can choose if to track the target or not, but it is not possible to perform a smooth pursuit without tracking a moving target [7]. Smooth pursuit eye movement has an initiation latency of 100-150 ms, which is shorter than for a saccade. Although smooth pursuits are considered "slow" when compared to saccades, in some rare cases their velocities can reach more than 100°/s [15].

2.2.4 Eyelid Movements (Blinks) and their effects on eye movements

Eyelid closure, also referred as "blink", is the passive movement of the levator palpebrae muscle (LPM), as its innervation ceases. Eyelid and eye movements are coordinated in order not to interfere with each other. During upwards gaze (when the eye looks/moves upwards) the eyelid does not cover the pupil, while during downwards gaze (when the eye looks/moves downwards) the eyelid protects the upper part of the eyeball. Eyelid movements are commonly categorized as spontaneous (voluntary), passive (following saccades) or reflexive (triggered by sounds or visual stimuli) [16]. On average, blink duration lies between 100 and 400 ms [13].

During blinks, visual sensitivity decreases and this temporary blanking goes practically unnoticed even though it can be for a relative large amount of time. However, the eyeball does not stay put while blinks occur, but it moves inwards, downwards and torsionally, hence activating all oculomotor muscles. These are called "blink-associated eye movements" and are slower than saccades [16, 4].

Often blinks accompany horizontal and vertical saccades, and influence the saccadic velocity, acceleration, deceleration and they increase saccade duration without affecting saccade amplitude [16]. Typically, the occurrence of blinks is approximately

20 times per minute and the likelihood of a blink occurring is proportional to the size of the gaze shift. Moreover, saccades made with blinks tend to present an higher rate of dynamic overshoots while saccades occurring under closed lids are much slower compared to the visually-guided or memory-guided saccades [4].

When blinks occur during fixation of a stationary object, the eyes move down approaching the nose [4], while during smooth pursuits blinks tend to reduce the pursuit velocity [16].

2.3 Eye movements sensing systems

Eye movements have been a subject of study since the end of the 19th century and diverse techniques have been developed [17]. The three most popular methods for monitoring the eyes are magnetic search coil, electro-oculography and video-based eye tracking. These techniques are based on different working principles, which are presented in the following sections 2.3.1, 2.3.2, 2.3.3.

2.3.1 Electro-oculography (EOG)

The electro-oculography (EOG) is a technique used to measure the electrical signal of the eye, where the signal varies depending on the eye's position. The eye is the origin of a standing electrical potential field and can therefore be modeled as a dipole, where the negative pole represented by the retina (posterior part of the eye) is oriented towards the positive pole, represented by the cornea (anterior part of the eye) [18]. The potential between positive and negative charged parts of the eye is called corneo-retinal potential (CRP). The origin of the potential is ion permeability changes through the retinal pigment epithelium (RPE) membrane, situated behind the photo-receptor layer, which is sensitive to retinal illumination [19]. However, the CRP varies so slowly that it can be assumed steady and this assumption is the basis for the EOG. Such CRP lies in the range of 0.4 - 1.0 mV [3]. CRP produces an electrical field, which is the source of the EOG signal. By recording changes in this electrical field due to the moving (rotation) of the dipole, the eye movement can be tracked and its characteristics, such as amplitude, duration, and speed, can be obtained.

The EOG is measured by two pairs of skin electrodes placed at the periorbital positions of the eye, plus an additional reference electrode on the forehead. This set up allows identification of two different signal components corresponding to the vertical movement and the horizontal movement of the eye [18]. When the eye is at rest, the direction of the electrical field altered by the dipole is perpendicular to the electrodes and thus no voltage is recorded [3]. However, as the eye moves from the center to periphery, a difference of potential is detected because the retina approaches one electrode while the cornea gets closer to the opposite one. Hence, the polarity of the signal is positive at the electrode to which the eye is moving [3]. This change in orientation of the dipole, and therefore in electrical potential, corresponds to a change in the measured EOG signal. Consequently, eye movements

can be studied by analyzing these changes. Moreover, EOG is usually included in the electro-encephalography (EEG) setups in order to allow elimination of eye movement artifacts [20]. Typical amplitudes of the signal range from 5 μV/° to 20 μV/° and the essential frequency content spans from 0 Hz to 30 Hz [13, 5]. EOG is considered linear to the eye position in the range of ±25°[9].

2.3.1.1 Baseline drift and noise in the EOG signal

EOG signal tends to be strongly affected by different sources of noise, such as residential power-line, measurement circuitry, electrodes, muscle activity, etc [13]. For instance, one physiological source of noise derives from the fact that the CRP is not actually stable, especially in high illumination conditions [3]. Moreover, the activity of facial muscles introduces a considerable number of artifacts in the EOG trace due to, for example, eyelid closure, movement of the opposite eye, and the muscle action potential spike at the onset of a saccade [4]. In addition, thermal noise generated by the contact resistance of the skin electrodes, and unstable attachment of the electrodes to the skin represent a relevant cause of disturbances that impair the signal [9, 13].

Furthermore, a particular interference, which distorts the raw EOG signal is something called baseline drift. It may be caused by background signals, electrode polarization or varying contact pressure of the electrodes. Essentially, it appears as a slow signal change that is mostly uncorrelated to eye movements but it superimposes the signal. Baseline drift affects the signal of all eye movements, except for the saccades, since their duration is so short that the drift cannot be noticed [13].

Nevertheless, electro-oculography is a relatively inexpensive and noninvasive method that can be applied in conditions of total darkness and it is also able to measure the signal when the eyes are closed.

2.3.2 Scleral search coil

The most precise technique used to record and measure eye movements is the scleral search coil technique, which is, at the same time, the most invasive. This method consists of placing a silicon contact annulus/lens on the eye, precisely over the cornea and sclera. The contact lens contains copper wire coil that is connected to an external plug. By placing the subject in a magnetic field, the movements of the eyes induce a current in the coil, from which the position of the eye can be determined. Advantages of this technique are the high precision (up to 0.1°) of the measurements in the horizontal, vertical and torsional dimension, and the robustness against eyelid artifacts, which, affect the quality of the signal recorded with other methods (EOG, video-based techniques). Besides its intrusiveness and consequent discomfort, the scleral search coil method has the drawback to be unsuitable for recordings longer than 30 minutes since it can cause irritation due to rubbing of the device against the eyelid [21, 5].

2.3.3 Video-based eye tracking systems

Video technology used to track and record eye movements, also defined as video-oculography (VOG), has become increasingly in vogue as the electronic data processing has been progressing. Nowadays, video-based trackers consist of affordable cameras and image processing tools with more and more robust algorithms. Typically, the principle of VOG techniques is to track the position of eye-fixed markers in a 2-D image, which has to relate to the head-fixed coordinates. High precision of these positions is difficult to reach, therefore either the camera must be firmly fixed to the head, or multiple ocular features must be determined in order to distinguish head and eye movements [5, 9]. The most common ocular features used to compensate for head translation are the pupil center corneal reflections (PCCRs). An infrared light source illuminates the eye, and a camera captures the reflection patterns of the light bouncing back from cornea and pupil. These two are the points of reference on the eye needed to differentiate eye movements from head translations. Hence, a vector containing the angle between the cornea and pupil center reflections is computed, and it is then combined with other features to calculate the gaze direction [22].

The major advantage of VOG technique is its non-invasiveness because there is no contact with the eye. The only discomfort might be if the device is mounted on the head, or if it limits the field of view, which might happen when wearing mobile eye tracker glasses [9]. VOG offers a temporal resolution that normally ranges between 50 and 400 Hz [9], but a sampling rate of 1000 Hz can also be reached [23]. However, a series of drawbacks affects this technique. Firstly, using the reflections of the cornea adds more noise that has to be filtered out. Secondly, this method relies on the profile of the cornea, which differs between people. Furthermore, VOG is not able to track the eye during eyelid closure, such as while blinking, and neither when the eyes stay closed for a prolonged time, which prevents this technique from being used for sleep-related eye movements studies [9]. Finally, a disadvantage related to the usage of eye tracker during driving task is that it loses tracking, e.g. during large head movements.

2.4 Current methods for baseline drift removal and EOG denoising

As mentioned before, noise and baseline drift contamination represent a big problem that impairs the quality of the EOG signal. In the following two sections, some of the related works done on baseline drift and noise removal are presented.

2.4.1 Baseline Drift Removal

As common for many physiological signals, the raw EOG signals are distorted by the aforementioned baseline drift, which is a temporal slow signal change that superimposes the EOG trace. However, it may affect the representation quality of the

eye movements, such as fixations, smooth pursuits. Saccades are not strongly influenced by baseline drift because of their high velocity [13]. Baseline drift needs to be removed in order to ensure the accuracy of any extracted feature, such as duration and amplitude. A number of novel techniques, which outperform the traditional high-pass filtering solution, has been developed to remove baseline drift from electrocardiogram (ECG), such as in [24, 25, 26, 27]. These algorithms perform fairly well since the ECG signal has repetitive patterns [13]. Nonetheless, baseline drift removal is still a field of ongoing research for non-repetitive signals such as EOG. For example, in [20] Pettersson used an approach, which consisted of modeling the drift with a polynomial function of degree 20, and then subtracting it from the raw signal.

Another well-used approach to extract and then remove the drift from the EOG signal is based on Discrete Wavelet Transform (DWT). It was proved by Von Borries in [28] to be very powerful to remove the drift from ECG data and later on a very similar approach gave great results by [13] when applied to EOG data.

2.4.2 EOG Denoising

The noise sources mentioned in the Section 2.3.1.1 may produce distortions in the signal that compromise the quality and detection of certain eye movements. Therefore, techniques, such as low pass filtering, wavelet denoising, and median filtering have been used in several EOG related studies to reduce the noise in the signal. For instance, low-pass filtering on the acceleration signal derived from the EOG was used by Behrens in [29] to suppress frequencies higher than 25 Hz. However, it results in a distortion of acceleration signal of the saccades, but it does not shift them in time.

On the other hand, in [20] the noise was removed by applying wavelet denoising. After decomposing the EOG signal, the coefficients with frequencies above 100 Hz, which represent the fine structure like noise, were set to zero. The only drawback of this method is that the slope of the saccades may be distorted.

Nevertheless, another method for filtering out the noise from EOG data is by using a median filter, which by definition takes the median value of the points in a sliding window that runs through the signal. Median filtering was chosen by for example Bulling in [13] since it preserves some of the saccade parameters, such as slope steepness.

2.5 Current eye movements detection approaches

Several methods have been implemented for eye movements detection from both EOG and eye tracker data, especially focusing on detecting fixations and saccades [30]. Usually, saccades are detected by velocity-based threshold methods, in which a fixed threshold determines whether a saccade is present or not [31]. However, it requires knowledge about the velocity in the conventional format, i.e. in degrees per second (°/s). Therefore, a calibration to find the relation between recorded

voltage and eye movement is needed. However, other methods for detection have been developed, such as by Niemenlehto in [32] and by Behrens in [29], in which adaptive thresholds are used. In [32] a technique commonly used in radar systems is applied and computes an adaptive threshold to the velocity waveform of the EOG. In the work by Behrens a saccade detection algorithm uses adaptive thresholds, calculated from the Gaussian distribution of the acceleration profile of the EOG. This acceleration-based method is stated to be suitable on uncalibrated EOG data and also on video-based eye tracker data [29]. Also in the work by Pettersson an auto-calibrating algorithm can automatically estimate threshold values for saccades and blinks, based on their first derivative, i.e. their velocities [20].

In [33] Vidal developed an approach for specifically detecting smooth pursuit. It is based on shape features, representing the characteristic low variance of this eye movement, such as the integral measure of the signal, and a machine learning classification. Although this mentioned method originally was applied to eye tracker data, it is stated by the authors that it possibly could be implemented for EOG data as well. Worth to mention is a recent work by Hoppe, in which an approach based on convolutional neural networks (CNN) is able to detect eye movements directly from raw data, without any thresholding, pre-segmentation or manual intervention. However, this method was only tested on video-based eye tracker data [34].

A different method for detecting saccades and fixations from EOG data was developed by Bulling in [35, 13] and it is based on the Continuous Wavelet Transform (CWT). The signal is transformed into wavelet coefficients, describing the changes in the signal. Since fixations and saccades differ significantly, they can be detected by setting specific thresholds on the coefficients. Moreover, a similar approach was proposed for blink detection. Both algorithms developed by Bulling showed an accuracy of 94% [13], but saccades outside the duration range of 20 - 200 ms were discarded.

2.6 Basic theory behind the implemented algorithms

In this section, the basic theories used in the developed algorithm are presented.

2.6.1 Wavelet Transform

In the past several years wavelets have been studied and applied in the area of signal processing because of their powerful abilities in manipulating the signal for the main purposes of denoising, data compression, wave detection and feature extraction [36, 37]. Wavelets are mathematical transform methods that have the fundamental property to provide an adaptive time/space-frequency resolution of the signal. Essentially, it means that small scale features of the signal are scanned with fine resolution in time/space and coarse resolution in frequency, while big-scale features are analyzed with coarse time/space resolution but fine resolution in frequency [38].

There exist several families of wavelets, and some commonly used are Haar wavelets, Daubechies wavelet, the bi-orthogonal spline wavelets and the reverse bi-orthogonal spline wavelets [38].

The basic working principle of the wavelet transform, and more specifically of the Discrete Wavelet Transform (DWT), is a decomposition/reconstruction process of the data. First, the smallest scale features (the finest details of the signal) are separated from the others by decomposing the signal into two parts, namely approximation and details. The approximation is created by applying low pass filtering followed by a down-sampling operation, where every other point (sample) is rejected. Likewise, the detail is obtained by high pass filtering and down-sampling. In the second decomposition level, the features appearing on a scale twice as large as in the first level are separated by filtering and then down-sampling the first level approximation coefficients. This operation creates a coarser approximation and a sequence of finer details. The decomposition procedure continues until the level in which the largest-scale features are differentiated. In this way, the decomposed signal consists of a coarse approximation and a series of finer and finer details as shown in Figure 2.6.

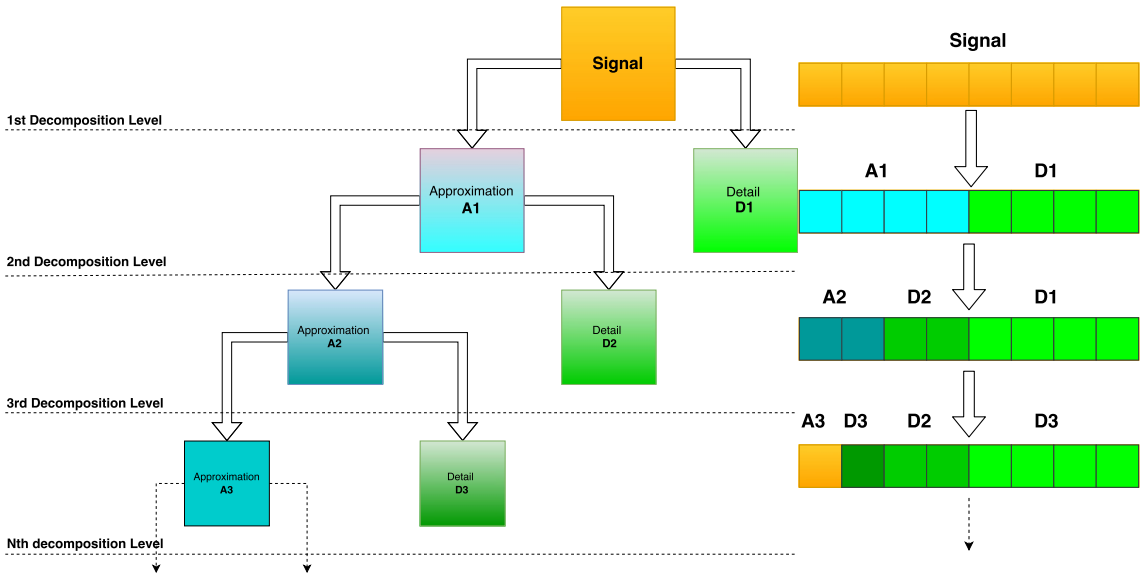


Figure 2.6: Working principle of Signal Decomposition with the Discrete Wavelet Transform.

The analyzing wavelet function can be mathematically represented by the Equation (2.4),

$$\psi_{ab}\big(t\big) = \frac{1}{\sqrt{a}}\psi\Big(\frac{t-b}{a}\Big) \tag{2.4}$$

where $\psi(t)$ is the mother wavelet, a is the scale parameter, which causes dilation (stretching), if $a > 1$, or contraction, if $a < 1$, and b is the position parameter responsible for translation [37]. The original signal can be retrieved by applying the inverse wavelet transform, which consists of up-sampling, i.e. inserting zeros at

even-indexed points, and then filtering the coarsest approximation and detail coefficients [38].

Another type of wavelet transform is the Continuous Wavelet Transform (CWT). The main difference between CWT and DWT is that the former discretizes scale more finely than the latter. In fact, in the DWT the scale parameter is always discretized to integer powers of 2, such as 2^j, j=1,2,3,..., while the scaling parameter for CWT varies as $2^{\frac{j}{v}}$, with j=1,2,3,... and v is an integer greater than 1 [39]. By continuously varying the scale parameter a (stretching or shrinking) and the position parameter b (shifting along the signal), the CWT returns the wavelet coefficients of the EOG signal $s(t)$, and these coefficients are defined as in equation (2.5).

$$C_{ab}\big(s\big) = \int_{-\infty}^{\infty} s\big(t\big)\frac{1}{\sqrt{a}}\psi\big(\frac{t-b}{a}\big)dt \tag{2.5}$$

By analyzing the coefficients at different scales, the characteristics of the signals can be detected. Generally, at low scales, i.e. at high frequencies, the analyzing wavelet is compressed, and sharp discontinuities can be more precisely tracked, while at high scale, thus at low frequencies, the dilated wavelet is better capable of detecting slow changing and coarse features.

2.6.2 Bag-Of-Patterns (BOP) and Symbolic Aggregate approXimation (SAX)

The "Bag-Of-Patterns" is a method for representation of time series data and it stems from the "Bag-Of-Words" representation used for text data. It was firstly proposed by [40] as a novel approach that outperforms leading current methods for time series data clustering, classification, etc. Unlike common shape-based methods, the concept is to find structural similarities in the signal based on the histogram representation of different patterns. The histogram contains the recurrence of certain Symbolic Aggregate approXimation (SAX) words in the time series that differs depending on the characteristics of the data. The creation of these SAX words is the first part of the algorithm, and it aims at representing the pattern as strings (i.e. words).

Originally proposed by Lin in [41], SAX is a symbolic approach that transforms time series T of arbitrary length j into a lower dimensional representation, such as words of arbitrary length w. The alphabet size α can be chosen by the user and it has to be an integer $\alpha > 2$. For example, if the alphabet is {**a**,**b**,**c**} the size is 3. First, the time series T is normalized using standard score as defined in equation 2.6:

$$T_{norm} = \frac{T-\mu}{\sigma} \tag{2.6}$$

producing a normalized time series T_{norm} with zero mean (μ) and standard deviation (σ) of one. The next step is to perform discretization by applying Piecewise Aggregate Approximation (PAA). This means that the time series is divided into w equal-sized segments, and the mean for each segment is calculated. Thus, the

dimensionality of the time series is reduced from j to w. The normalized time series is assumed to follow a Gaussian distribution. Thus discretization can produce symbols with equiprobability. The equiprobable regions are divided according to a set of breakpoints $\beta_{\alpha-1}$ depending on the specified alphabet size α. For example, with $\alpha = 3$ the distribution space is divided into three equiprobable regions by the breakpoints $\beta_1 = -0.43$ and $\beta_2 = 0.43$. Then, the PAA coefficients that are bigger than β_2 are mapped to the symbol "**c**", whereas the ones that are below β_1 correspond to the symbol "**a**" and the coefficients lying between β_1 and β_2 are mapped to "**b**". The sequence of these symbols represents a *word*, e.g. *aabcc*. Figure 2.7 illustrates how a time series, in this case containing a saccade, is mapped into a SAX string.

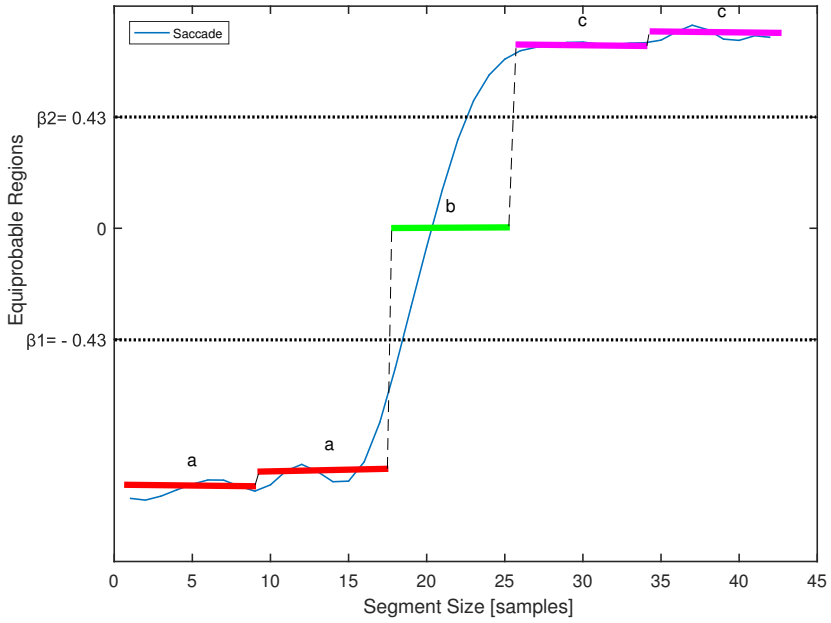


Figure 2.7: SAX Word: the time series, a saccade in this example, is transformed to the string *aabcc*. The parameters used here are $w = 5$, $\alpha = 3$ and segment size $j = 41$

The sequence of SAX strings is extracted using a sliding window of size n that runs over the segment. It results in a set of strings corresponding to a subsequence in the time series, i.e. the segment. Afterward, the word-sequence matrix can be constructed, so that the number of rows is the same as the size of the SAX dictionary and each column denotes a time series segment. The SAX dictionary is the collection of all SAX words (i.e. the patterns), and its size is expressed as α^w. Figure 2.8 shows an example of word-sequence matrix.

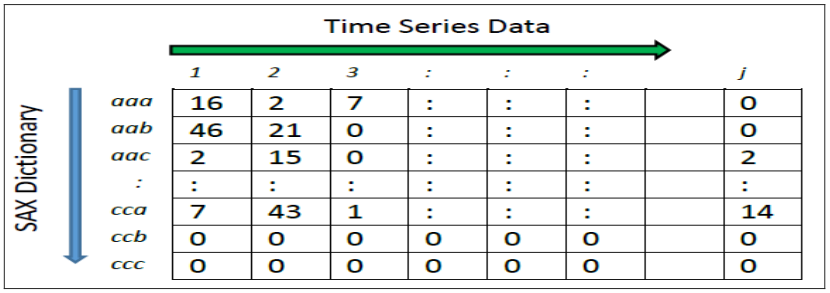


Figure 2.8: SAX word-sequence matrix

The word sequence matrix shows the occurrence of each word in each time series, hence each time series can be represented as a histogram. Once produced, the Bag-Of-Patterns, i.e. the histograms, can be sent to a classification or clustering machine learning algorithm. The authors in [40] tried this approach with many different datasets and they proved it to achieve superior clustering and classification results compared to other methods.

2.6.3 Shape Features

In [33], Vidal introduces an algorithm for detection of smooth pursuits. The algorithm utilizes a sliding window approach for detection, in which shape measures that capture aspects of different eye movements are calculated for each window. The shape measures used are:

- Slope
- Range
- Mean velocity
- Variance
- Integral
- Energy
- Waveform length

These measures are used as the basis for calculation of shape features, where the features are set to 1 or 0 depending on if the values are within boundaries associated with smooth pursuits. The shape features are then used as input to a k-nearest neighbor (kNN) classifier.

2.6.4 k-Nearest Neighbor (kNN) classifier

The k-nearest neighbor is a machine learning algorithm used for classification of objects based on the characteristics of the neighbor objects. An unknown object in the feature space is classified by a majority vote of its neighbors, whose number is defined by the parameter k. The kNN classifier commonly uses the Euclidean Distance as distance metric, to measure the position of the unknown object from the neighbors. Therefore, given a set of n objects (or points) and a distance function,

such as the Euclidean distance, the kNN classifier finds the k closest points to the unknown point (or query point).

Figure 2.9 shows a simple representation of the "majority voting" working principle done by the kNN classifier. The labeled training examples (data objects), illustrated by the red circles for class 1 and the blue squares for class 2 are represented in features space. An unlabelled data object, the violet triangle, is to be classified either to class 1 or class 2. By measuring the Euclidean distance between the query point and the neighbors, the k closest are found. In the example in Figure 2.9, when $k = 3$ the closest three labeled data objects consist of two blue squares and one red circle. Therefore, the unknown object will be classified as class 2, since the class 2 (blue squares) represents the majority. On the other hand, in the case of $k = 10$, the query point (the violet triangle) will be labeled as class 1 because the number of red circles is bigger than the number of blue squares.

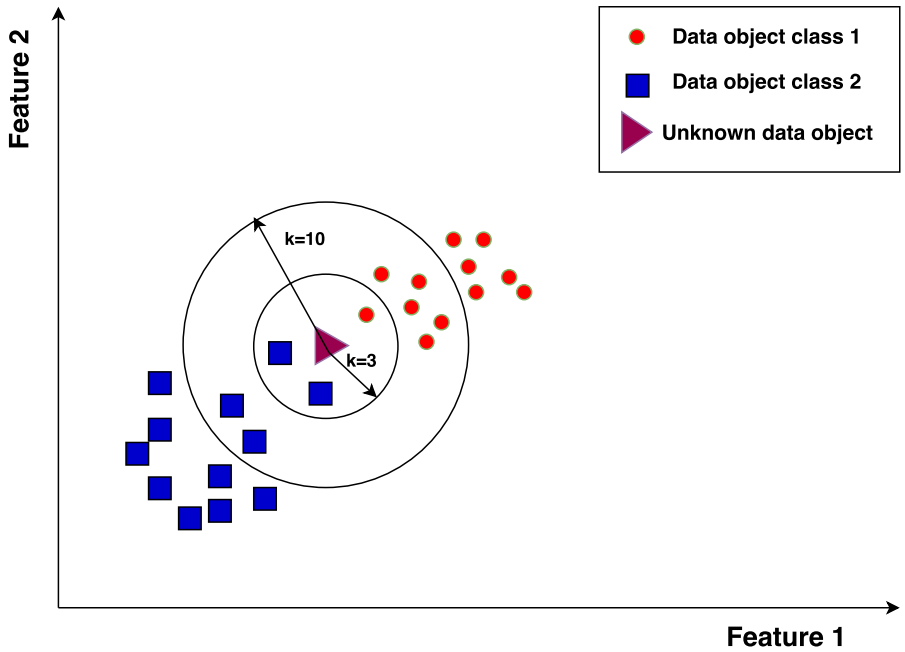


Figure 2.9: kNN classification: if $k = 3$ the unknown data object (violet triangle) is assigned to class 2 because the blue squares are the majority of the 3 neighbouring objects. However, if 10 neighbours are considered ($k = 10$), the number of red circles is bigger than the blue squares. Hence, the majority is represented by class 1 and thus the unknown data is classified to class 1.

2.7 Cognitive Distraction during driving

Driver distraction can be defined as *a diversion of attention away from activities critical for safe driving towards a competing activity* [42]. Typically, a differentiation is made between visual, manual and cognitive distraction [43]. Visual distraction is when the driver takes his/her *eyes off the road* to look either at a device or a person. Manual distraction is the action of the driver taking his/her *hand/hands off the steering wheel* to manipulate an object (e.g. a phone or fruit) [44]. Cognitive distraction refers to a more general diversion of attention, and it can be considered the most difficult type to assess because it relates to what the driver's brain is engaged in.Therefore, it can be defined as the driver's condition when he/she takes his/her *mind off the road* [43, 44], due to, for example, talking to a passenger, having a hands-free phone conversation [45] or simply thinking about something else than driving. However, while visual distraction is proved to increase the risk of accidents, the effect of cognitive distraction on traffic safety is less clear.

Eye movements and glance behavior analysis can be exploited to study cognitive distraction. As found in [46], under cognitive distraction the drivers make fewer saccades per unit time, with a consequent reduction in glance frequency and exploration of the driving environment. During increasing cognitive load the gaze concentration is higher on the road center, and thus more time is spent looking straight ahead [42, 46].

Moreover, in previous studies, it was shown that the amplitude of saccadic eye movements either increases [47, 48] or decreases [49] as cognitive workload increases. A similar discrepancy was also found regarding fixations, which have been reported to be either shorter [48] or longer [49] during cognitive load as compared to the baseline. This disagreement could be attributed to the fact that the study in [48] was carried out in a driving simulator, while the study in [49] was a field trial. In addition, this difference might be due to the visual complexity of the scene [42].

Furthermore, it is found that cognitive distraction during driving causes an increase in blink frequency, as a result of higher memory demand [42].

2.8 Driver Simulation Study on Cognitive Distraction

The collection of the electro-oculographic data used in this thesis was part of a project, called Vehicle Driver Monitoring (VDM), in which different physiological measurements, such as ECG, Electroencephalography (EEG), EOG, skin conductance and respiration, were taken during simulated driving. These were collected, before the beginning of our thesis work, from 36 volunteering test subjects while driving in a simulated rural area traffic environment using a moving base simulator.

In order to study the effects of cognitive distraction, the participants performed two kinds of cognitive task several times during driving. The following subsections 2.8.1 and 2.8.2 briefly explain how the cognitive workload was assessed and what were the relevant components for eye movements measurement.

2.8.1 Cognitive Tasks

Two cognitive tasks had to be performed by the test subjects during driving, one easy task called "1-Back" and the other more mentally demanding called "2-Back". During the 1-Back task, a number was orally read out every two seconds. The test subject had to respond by pressing a button mounted on his/her finger whenever the last presented number was the same as the previous one. The 2-Back task was essentially similar but with a higher cognitive load. Instead of pressing the button whenever the same number was presented twice in a row, the test subject had to respond whenever the currently presented number was the same as two numbers before. The duration of each task was one minute every time it was performed. For each scenario, one baseline (no cognitive load) and three cognitive tasks were performed (one 1-Back and two 2-Back).

2.8.2 Eye movements measurement setup

The EOG data, as well as the other physiological signals, was recorded using a plugin in DeweSoft software for data acquisition on main S-box. The sampling rate was 256 Hz. The sensors were four electrodes, which were connected to the EEG cap and shared the same reference and ground with the EEG. The setup also included an eye tracking equipment to find out where the test subject focused the eyes during driving. This was done using wired eye tracking glasses from SMI, "SMI Eye Tracking Glasses 2", whose sampling frequency for the tracking is 60 Hz.

3
Methods

The developed algorithm is presented in the following sections. The goal of the algorithm is to find and classify the saccades in the input data. This includes finding the start- and end points of the saccades, handling saccades during blinks and merging of oblique saccades. Firstly, we describe the signal processing techniques used to remove baseline drift and reduce noise from the EOG data (Section 3.1). Secondly, the algorithms for saccade detection and classification are presented and their working principles are in details explained in Section 3.2. This section also includes the calculation of start- and end points of saccades and how the segments that are to be classified are extracted. Since saccades during blinks needed to be detected, an algorithm for blink detection was implemented based on a previous work, and is described in Section 3.3. Next, the flowchart of the final algorithm is illustrated in Section 3.4 and the general description of the process is presented. Section 3.5 reports the procedure of creating training and validation datasets for evaluating the developed algorithm. Finally, in Section 3.6 we introduce the methodology used to perform a limited study on the effects of cognitive distraction, in which the developed saccade detection algorithm is applied.

3.1 EOG Signal Pre-Processing

3.1.1 Wavelet-based approach for baseline drift removal

As the baseline drift removal technique used in [13] revealed to be very effective, the decision was to replicate the same wavelet-based approach. A MATLAB function was created, which inputs the raw EOG signal, the number of decomposition levels and the chosen wavelet function. The algorithm performs a multilevel 1D wavelet decomposition at level 12 using a reverse bi-orthogonal wavelet "rbio6.8". Thus, the signal consists of approximation coefficients followed by a sequence of finer and finer detail coefficients. The low-frequency components of the baseline drift appear in the approximation coefficients (scaling coefficients) [28]. Therefore, the baseline drift estimation can be obtained, by first setting the detail coefficients to zero and then by reconstructing the remaining coefficients, i.e. the approximation coefficients, with the Inverse Discrete Wavelet Transform (IDWT). The obtained baseline drift estimation is then subtracted from the original EOG signal, resulting in a signal with removed or reduced drift offset [13].

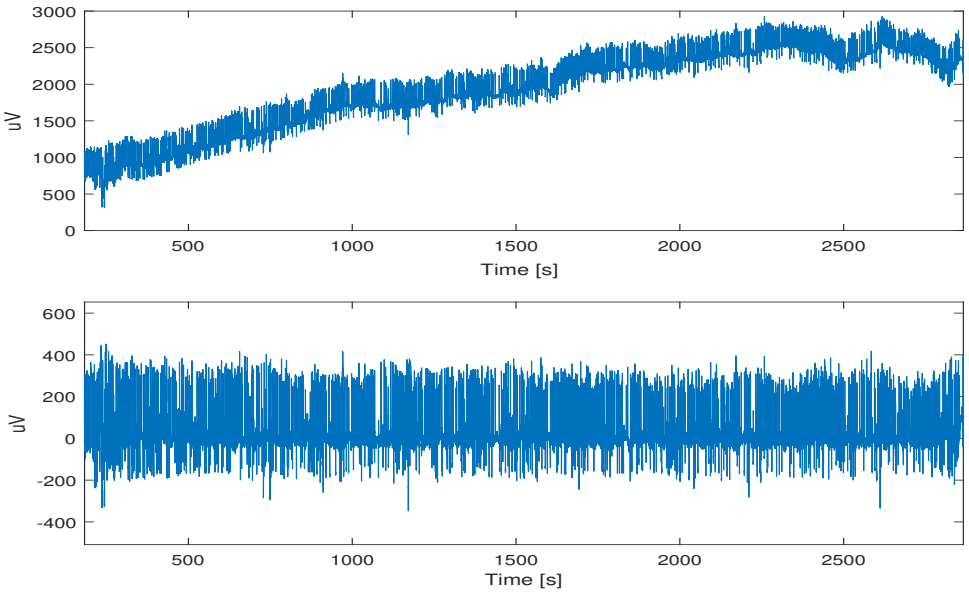


Figure 3.1: Vertical EOG with baseline drift (top) and without baseline drift (bottom)

Since baseline drift commonly is present in both channels, the algorithm is applied on both vertical and horizontal EOG signals. Figure 3.1 shows the difference between the EOG signal contaminated with baseline drift (top) and after baseline drift removal (bottom).

3.1.2 Noise Reduction

The noise corruption in EOG signal, due to, e.g. bad placement of the electrodes, power line interference, and muscle activity might significantly affect the quality of the signal, making the recognition of some eye movements, particularly saccades, very problematic. The provided EOG datasets for our work are especially noisy mainly because they were collected during driving and thus, head movements are common. Therefore, the choice of a filtering technique capable of preserving the signal characteristics of the saccades had to be made. As a result of the literature research, median filter and wavelet denoising appeared to be the most promising ones. Thus both methods were applied, but for different applications.

3.1.2.1 Median filtering

As also stated by the literature [13], median filtering is suitable for denoising EOG data, since it preserves the edge steepness of the saccades. However, if the chosen window size, or the filter order, is too large some saccade parameters, such as the start- and end points and amplitude, might be distorted. A large median filter can also remove small saccades. Consequently, a trade-off between noise reduction and detection precision exists.

In our work, median filtering is applied to the EOG input data to both wavelet-based algorithms for saccade detection and blink detection, which are described in Sections 3.2.1 and 3.3. Filter orders between 10 and 40 were considered, and the quality of the filtered signal was visual inspected and evaluated. For the saccade detection algorithm, the chosen filter order was 19, corresponding to a window size of about 74 ms. This window size reduces the noise without suppressing the small saccades of interest. The choice of the filter order is explained further in Section 3.2.1.

On the other hand, a median filter of order 39, i.e. window size of 150 ms, was used on the EOG data for the wavelet-based blink detection algorithm. This higher filter order suited the application because it did not impair the capability of detecting the blinks, which have a very high amplitude compared to the other eye movements. Moreover, a consequence of this sharp median filter is that the peaks of the blink signals (see the right plot in Figure 3.2) were cut off. Nonetheless, the amplitudes of the wavelet coefficients corresponding to the blinks did not differ from the amplitudes of coefficients produced with a median filter of order 19. In the left plot of Figure 3.2 it can be seen that the steepness of the saccades is preserved but the noisy spikes are chopped off.

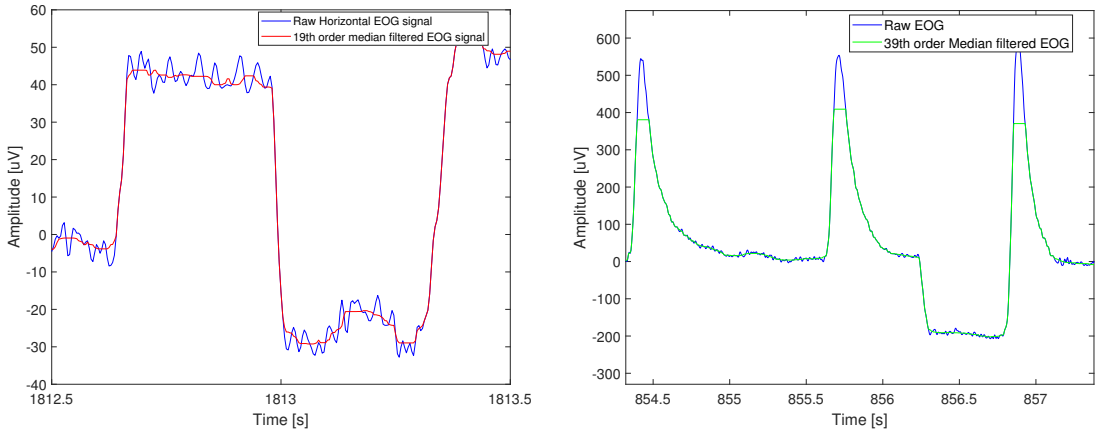


Figure 3.2: Median filtering of the horizontal EOG for the wavelet-based saccade detection algorithm (left) and of the vertical EOG for wavelet-based blink detection algorithm (right).

3.1.2.2 Combination of median filter and wavelet denoising

In the other parts of the algorithm, i.e. in the BOP and Shape Features classification (see Sections 3.2.4 and 3.2.5), the EOG signals were denoised by applying a combination of the median filtering and wavelet denoising methods.

The wavelet denoising function was performed using the MATLAB command "wden", which carries out an automatic denoising process [50]. The signal was denoised by

decomposing it at only level 1 with a "Symlet" wavelet. A soft thresholding on the wavelet coefficient was performed by setting the universal threshold $th = \sqrt{\ln(N)}$, which estimates the noise level [38] depending on the length of the time series (in this case number of samples in the EOG signal) N. This wavelet denoising method is very effective, but it might distort the steepness of saccades slopes. Therefore, decomposition level 1 is used to maintain the slope while suppressing high-frequency noise.

The filtered EOG output signal consists of the sum of the 19th order median filtered EOG and the wavelet denoised EOG signal, which is then divided by two, as shown in Equation 3.1.

$$FilteredEOG = \frac{medianfilteredEOG + waveletdenoisedEOG}{2} \tag{3.1}$$

This is done to further reduce the amplitude of the spikes of the high-frequency noise. In this way, the properties of the saccade, i.e. start- and end points, steepness and amplitude are maintained while the noise is reduced. This is illustrated in Figure 3.3.

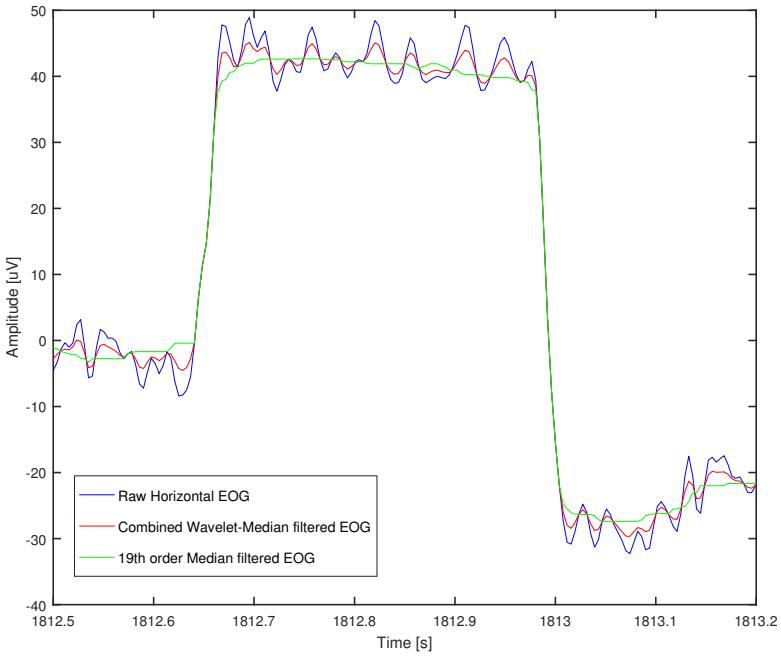


Figure 3.3: Filtering of raw EOG (blue line): filtered EOG using only median filter (green line), the combined median filter and wavelet denoising filtered EOG (red line)

3.2 Saccade detection and classification

3.2.1 Wavelet-Based Saccade Detection

We replicated the saccade detection method developed by [13], which is called *Continuous Wavelet Transform - Saccade Detection* (CWT-SD). One major reason why this wavelet-based approach was chosen is that Bulling used it in another study [51], in which EOG data were collected in a less controlled environment, namely while walking.

The input of the CWT-SD algorithm are the EOG signals (vertical and horizontal components) after baseline removal and denoising is performed. The CWT is computed using 'Haar' mother wavelet at scale $a = 20$. Among all the available wavelets, 'Haar' wavelet was selected due to its shape similarity with the saccade, which resembles a step function [13].

A scale parameter of 20 was chosen because, as can be seen in Figure 3.4, abrupt transitions (mainly corresponding to saccades) can be better detected at low scales with a scale parameter range of 10-22. Moreover, the choice of using $a = 20$ and mother wavelet 'Haar' was supported by the fact that it was used in [13].

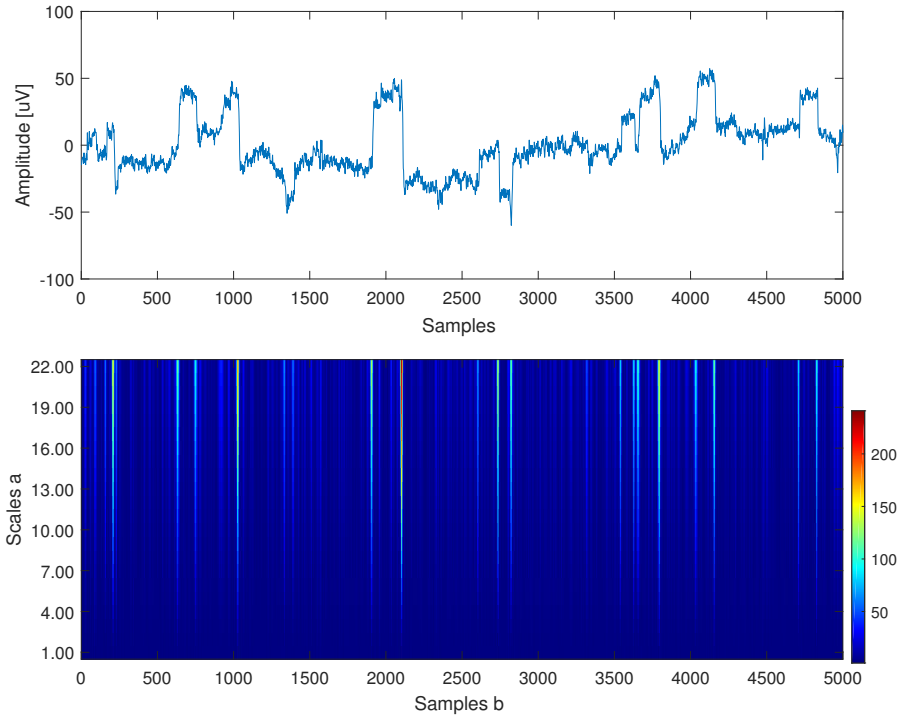


Figure 3.4: Continuous Wavelet coefficients (bottom) at different scales over a short part (5000 samples) of the horizontal EOG (top). The brighter vertical stripes, which go from color bright blue to red, corresponds to the sharpest and largest signal discontinuities, i.e. saccades. These abrupt transitions are localized at different scales.

Since the CWT is very sensitive in detecting abrupt changes in the signals, the coefficients amplitude will appear large in correspondence of saccades. Their sharpness and amplitude depend on the changes in the signal. Thus small saccades will have small, but rapidly varying coefficients, while big saccades will show big and steep coefficients peaks. Therefore, by applying two thresholds ($\pm th_{sd}$), one positive and one negative, on the coefficients, most saccades can be detected.

The choice of the threshold's value depends on the chosen scale since the amplitude of the coefficients varies along with the scale parameter. To find the optimal threshold value above which the highest number of saccades can be detected, several values of median filter orders and thresholds were tried, and the number of missed saccade candidates and false candidates were counted.

As one can see from Figure 3.5, a good coefficients threshold value lies in the range of 10 to 30 depending on whether precision or recall is prioritized. Precision and recall are performance measures that are further explained in Section 3.5. With the threshold set to 17, a median filter with order 19 was found to have the best trade-off regarding missed and false candidates.

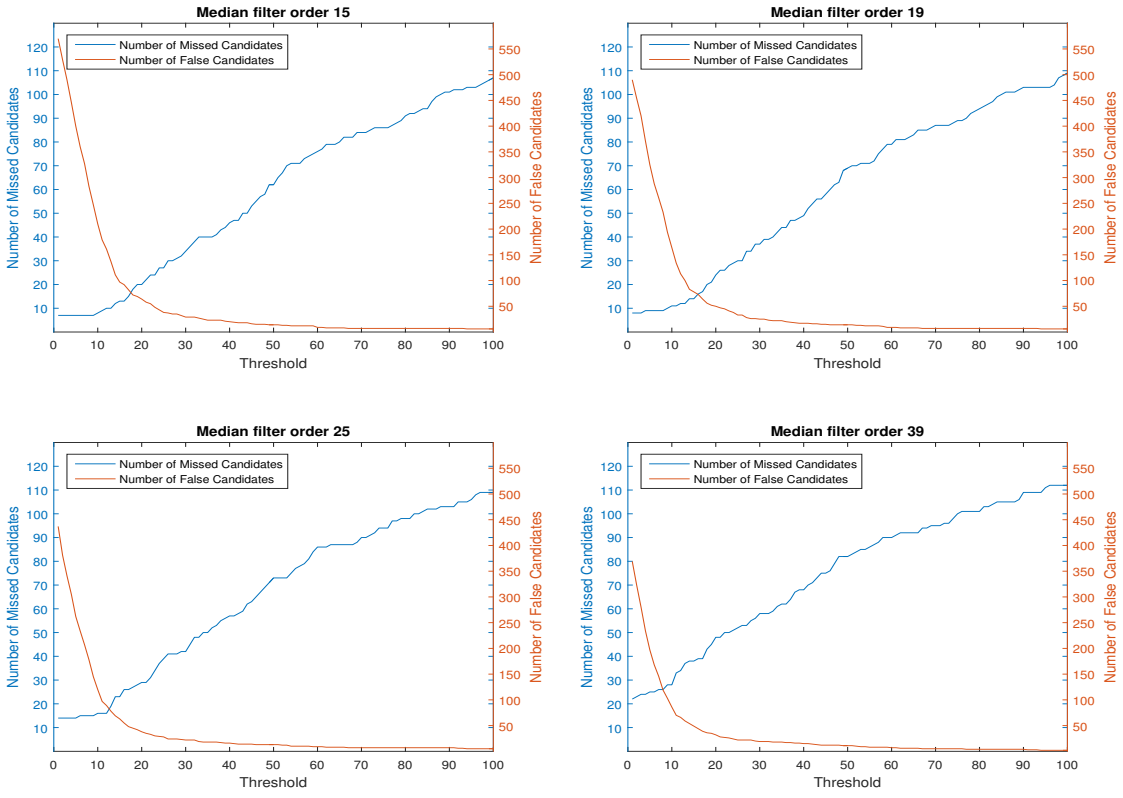


Figure 3.5: The plots shows how the number of missed candidates and of false candidates depending on the wavelet coefficients threshold and the order of the median filter.

The principle of the saccade detection algorithm is illustrated in Figure 3.6, in which a threshold $th_{sd} = \pm 17$ (one positive and one negative threshold) was applied. The coefficient peaks that cross the thresholds correspond to the saccades (red stars in the bottom plot).

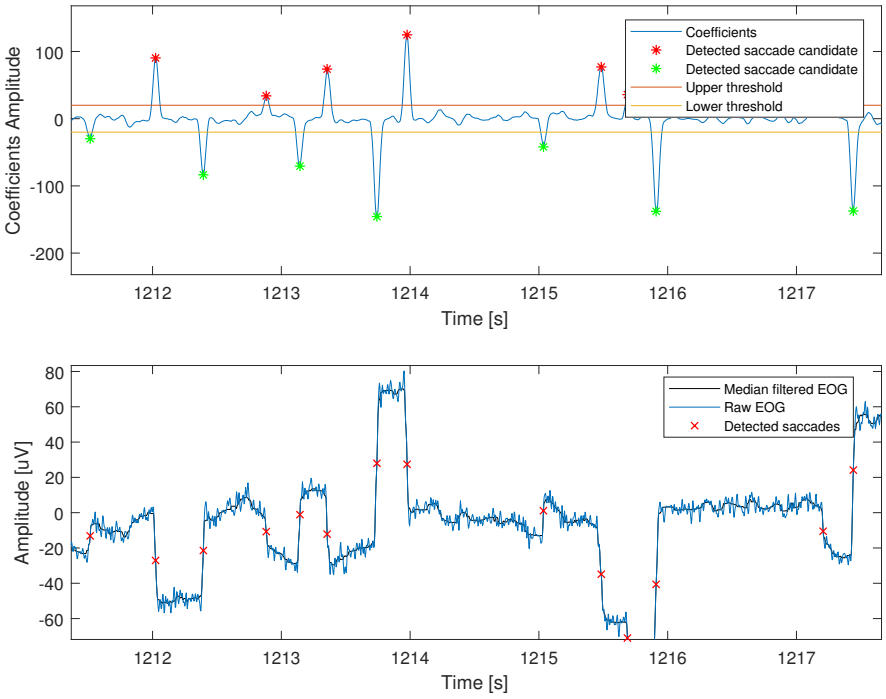


Figure 3.6: Wavelet coefficients (top) and horizontal EOG signal with the detected saccades (bottom)

3.2.2 Nominate Candidates

The CWT-SD is an effective algorithm for detecting saccades. However, as mentioned before, there exists a trade-off between precision and recall. A high threshold results in a high precision (detected saccades are saccades), but a low recall (many saccades are missed), while a low threshold has a high recall (almost all saccades detected) but a low precision (many false positives). From this relationship, the idea of using secondary machine learning algorithms to remove false saccades was conceived. In this way, recall could be increased without reducing precision.

Instead of defining the peaks over a threshold as saccades, they are from now on defined as saccade candidates. These saccade candidates are then used to extract segments around the candidates. Bag-of-Patterns and Shape Features classification are applied to the segments. In Figure 3.5, the number of missed saccades and number of false candidates is illustrated for thresholds in the range of 1-100. From this figure, it is clear that the most interesting range to investigate if precision can be increased is in the threshold range of 10-30. This approach has several advantages over using the machine learning algorithms in e.g. a sliding window approach,

mainly:

- Since only parts of the signal with a high probability of containing saccades are to be classified, the precision is likely to be higher.
- Wavelet detection is less computationally heavy compared to Bag-Of-Patterns and Shape Features Classification. By only classifying parts of the signal with a high probability of containing saccades, computational time is significantly reduced.
- The peak of the wavelet coefficients is located in close proximity to the peak velocity of the saccade. By extracting the segment around the peak, the saccade movement will be in the center of the segment. This results in a higher similarity between segments containing saccades, potentially simplifying training and increasing the performance of the verification algorithms.

3.2.3 Extract Candidate Segments

Before the candidates are classified, the boundaries of the segments containing the candidate need to be identified. First, the start- and end points of the saccade is calculated. Secondly, the center of the saccade is used as the center point of the segment. The start- and end of a saccade are typically defined by either a static or dynamic velocity threshold. The variable amplitude and noise level limit the performance of a static threshold. Therefore, the dynamic velocity threshold definition is used, where the threshold is 10% of the peak velocity.

In order to calculate the saccade start- and end points, the peak velocity of the candidate is identified by locating the closest local extrema in the derivative of the signal associated with the saccade direction. Consequently, a saccade candidate with a positive slope direction is related to the closest local maxima and a saccade with a negative slope direction is related to the closest local minima. Secondly, the peak velocity is used for calculation of the velocity threshold. This is illustrated in Figure 3.7.

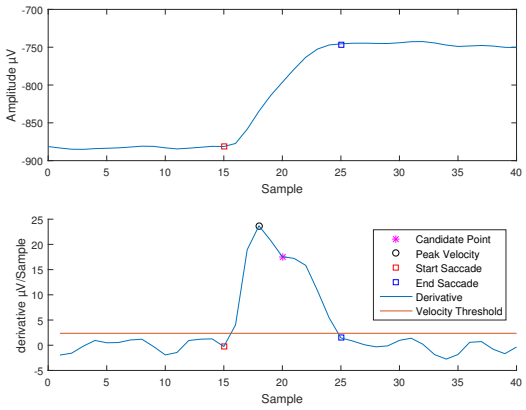


Figure 3.7: EOG signal of a saccade segment (top) and derivative of the same segment (bottom).

The first sample which crosses the threshold closest to the peak velocity before and after the peak velocity is regarded as the start- (closest crossing before the peak velocity point) and end point (closest crossing the after peak velocity point). If the duration of the crossing is limited to one sample, the crossing is ignored. This is done to reduce the number of incorrect start- and end points caused by interrupted saccades or noise.

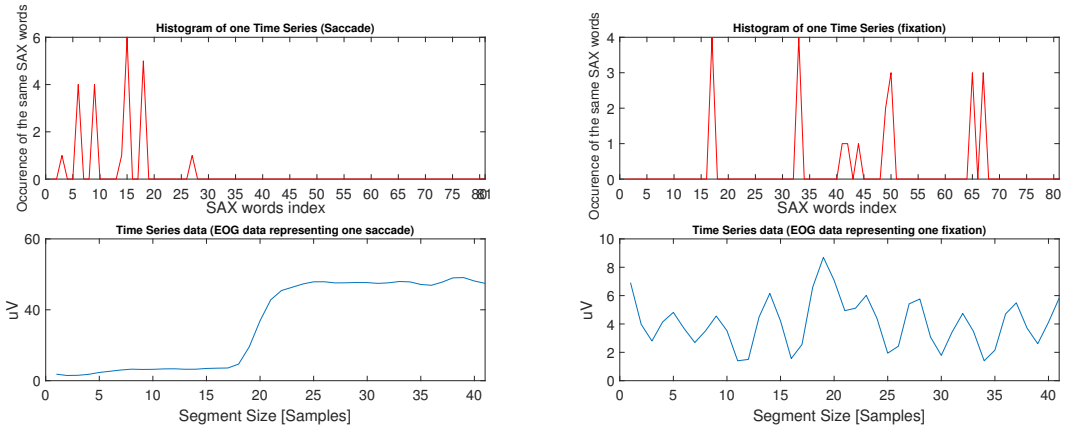
3.2.4 Bag-Of-Patterns (BOP) Classification of Saccades

Bag-Of-Patterns was selected as one of the classification algorithms since it has a good ability to compare structures in time series. The algorithm presented in [40] was therefore implemented. In our work, the algorithm is applied to the "saccade candidates", which were previously extracted using the CWT-SD algorithm.

The main parameters of the BOP algorithm are the segment size S_{size} (i.e. the time series length, containing the nominated saccade), the alphabet size α, which also is the number of equiprobable regions, the word length w and the length of the sliding window n. The algorithm is used to classify nominated candidates using different values of the parameters. Three segment sizes S_{size}, namely 21, 31 and 41 samples, were chosen to cover a reasonable range of saccade duration. In fact, these sizes correspond to a duration value of 82 ms, 121 ms and 160 ms, respectively. When transforming the time series into SAX strings, two alphabet sizes, recommended in [40] were tested, namely $\alpha = 3$ and $\alpha = 4$, with two different word lengths, $w = 4$ and $w = 5$. In this way, the possible SAX words collection varied between 81 ($\alpha = 3$ and $w = 4$) and 1024 words ($\alpha = 4$ and $w = 5$). Finally, the tested sliding window sizes n were 12, 15, 16, 20, 28 and 30 samples.

Figure 3.8 illustrates two typical examples of BOP representation. As one can see, the histograms differ significantly depending on the type of eye movement. The BOP of an upwards saccade (caption 3.8a) presents a localized concentration of SAX words only in the first half of the dictionary, while the BOP representing a fixation (3.8b) shows a spread distribution of SAX words along the whole dictionary. Generally, the majority of the histograms for saccades and fixations looks similar to those in Figure 3.8, but when the signal is strongly affected by noise the BOP representation might be quite different.

Finally, the word sequence matrix is input to a kNN classifier with k=1, as done by Lin in [40]. MATLAB's built-in kNN classifier is used and it was trained using the MATLAB Classification Learner application.



(a) BOP representing a saccade.

(b) BOP representing a fixation.

Figure 3.8: Bag-Of-Patterns representation of a saccade (a) and a fixation (b). The parameters used here are $\alpha = 3$, $w = 4$, thus the dictionary contains $\alpha^w = 81$ SAX words.

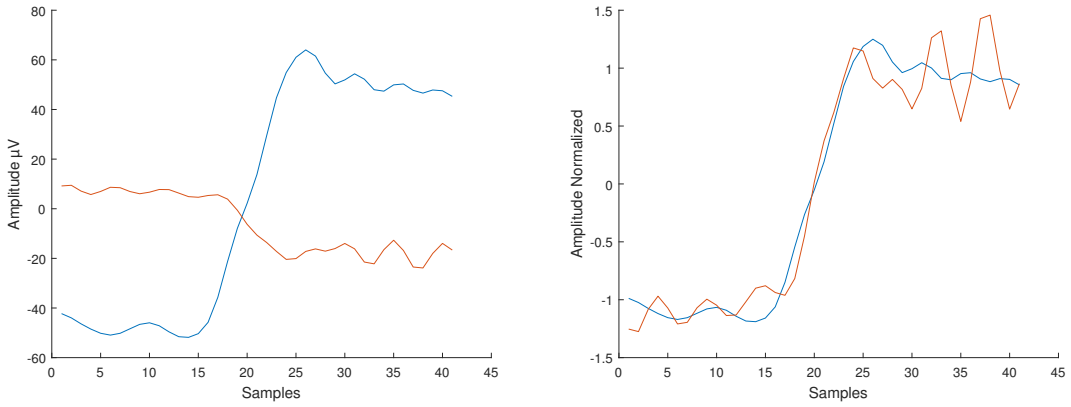
3.2.5 Shape Features Classification of Saccades

The developed Shape Features Classification algorithm is based on Vidal's approach in [33]. The main similarity between the methods is that both are based on the same shape measures mentioned in Section 2.6.3. An additional measure that was used in our algorithm is the zero crossing rate. This feature measures the number of times the signal crosses the x-axis (zero crossing) and is commonly used in e.g. speech recognition [52].

The change in amplitude properties from subject to subject in EOG signals (see 2.3.1) makes it hard to define strict boundaries for the features. Therefore, classification is performed directly on the shape measures instead of having features depending on boundaries. This is the main difference between the developed algorithm and Vidal's approach. Since the shape measures are used as features in the developed algorithm, these will be referred to as shape features from now on.

Each segment is normalized according to equation 2.6 before the shape features are calculated. This reduces the effects of change in amplitude properties. Since the saccades are re-scaled to similar amplitude levels, the change over time is given more significance, resulting in a greater similarity between saccades of different amplitudes. Segments containing an assumed saccade with a negative derivative are also inverted to further increase similarity in shape features. The inversion and normalization of two saccades of different amplitude and direction are shown in Figure 3.9.

In order to better capture differences between saccades, smooth pursuits, and fixations, a growing window size is applied. The window sizes are seven samples, 1/3 of segment length, 2/3 of segment length and the full segment length. Segment sizes



(a) Before normalization and inversion **(b)** After normalization and inversion

Figure 3.9: Illustration of normalization and inversion of saccades with negative derivative in order to make saccades more similar before feature calculation.

in similar range as for BOP was tested for the same reasons (see 3.2.4). For each window size, a set of shape features is calculated, resulting in a feature set of size $7a$, where $a = 4$ is the number of windows. Furthermore, the features are standardized before they are used as input to the classifier, in order to have all the features in a similar range.

Finally, a kNN classifier with Euclidean distance measure is utilized, such as in Vidal's approach. MATLAB's built-in kNN classifier is used for increased computing efficiency. Number of neighbors k was set to 10. The choice of k is explained further in Section 4.2. The classifier was trained using 25% holdout validation using the MATLAB Classification Learner application.

3.3 Wavelet-Based Blink Detection

The same principle as for the CWT-SD is employed for the blink detection. The method developed in [35, 13] was implemented, which also consists of thresholding the wavelet coefficients obtained by applying the CWT on the signal. Similar to the CWT-SD, the algorithm is named *Continuous Wavelet Transform - Blink Detection* (CWT-BD). The blinks can be distinguished from the saccades because they are characterized by two large consecutive peaks in the coefficient vector, one negative followed by one positive. However, the most important parameter separating saccades from blinks is the time interval between these two peaks, which is smaller than the minimum time between two consecutive saccades occurring in opposite direction [13]. Therefore, blink detection can be accomplished by applying a maximum time threshold on this time interval. As suggested by [35], the best detection performance is reached when the time difference threshold is 500 ms.

The choice of the coefficient thresholds $\pm th_{bd}$ is very arbitrary and large values can be chosen since blinks correspond to very big coefficient amplitudes. As shown in

Figure 3.10, the thresholds applied on the coefficients are -150 and 150, respectively for the negative and positive coefficients values. Moreover, since blink detection was a secondary task of our work, visual inspection of the detected blinks was considered a sufficient verification method.

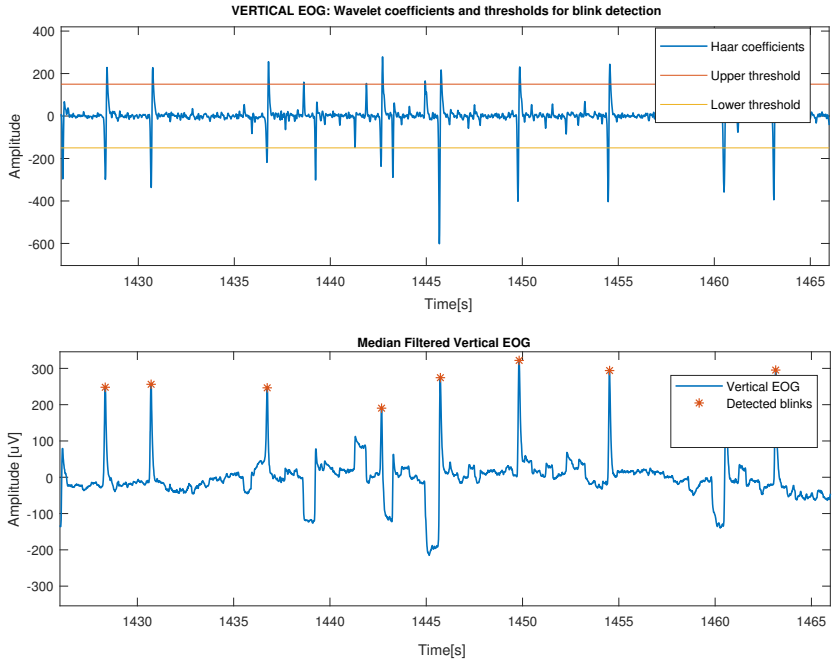


Figure 3.10: Detected blinks using the CWT-BD algorithm. By setting a negative and a positive threshold on the wavelet coefficients and a time difference limit on the interval between negative and positive peaks of the coefficients, the blinks can be detected and displayed (bottom figure).

3.4 Final Algorithm

The flowchart of the final algorithm is shown in Figure 3.11. First, the EOG signals are pre-processed, meaning that the baseline drift is removed, and the median and the combined wavelet-median filtered signals are generated. The median filtered signals are then used as input to CWT-SD to find and nominate the saccade candidates. The corresponding candidate segments are then extracted, followed by the calculation of start, end, and center of each candidate.

In order to reduce the risk of false positives, candidates not meeting certain criteria are removed. Duration of saccades is rarely shorter than 10 ms and commonly less than 100 ms. Therefore, saccades outside the range of 10-150 ms are removed. Furthermore, a minimum amplitude of 10 μV was set since the amplitude of the noise normally is equal to or greater than this. After evaluating the results of the tuning, minimum amplitude was later increased to 20 μV. This is discussed further in Section 4.3.

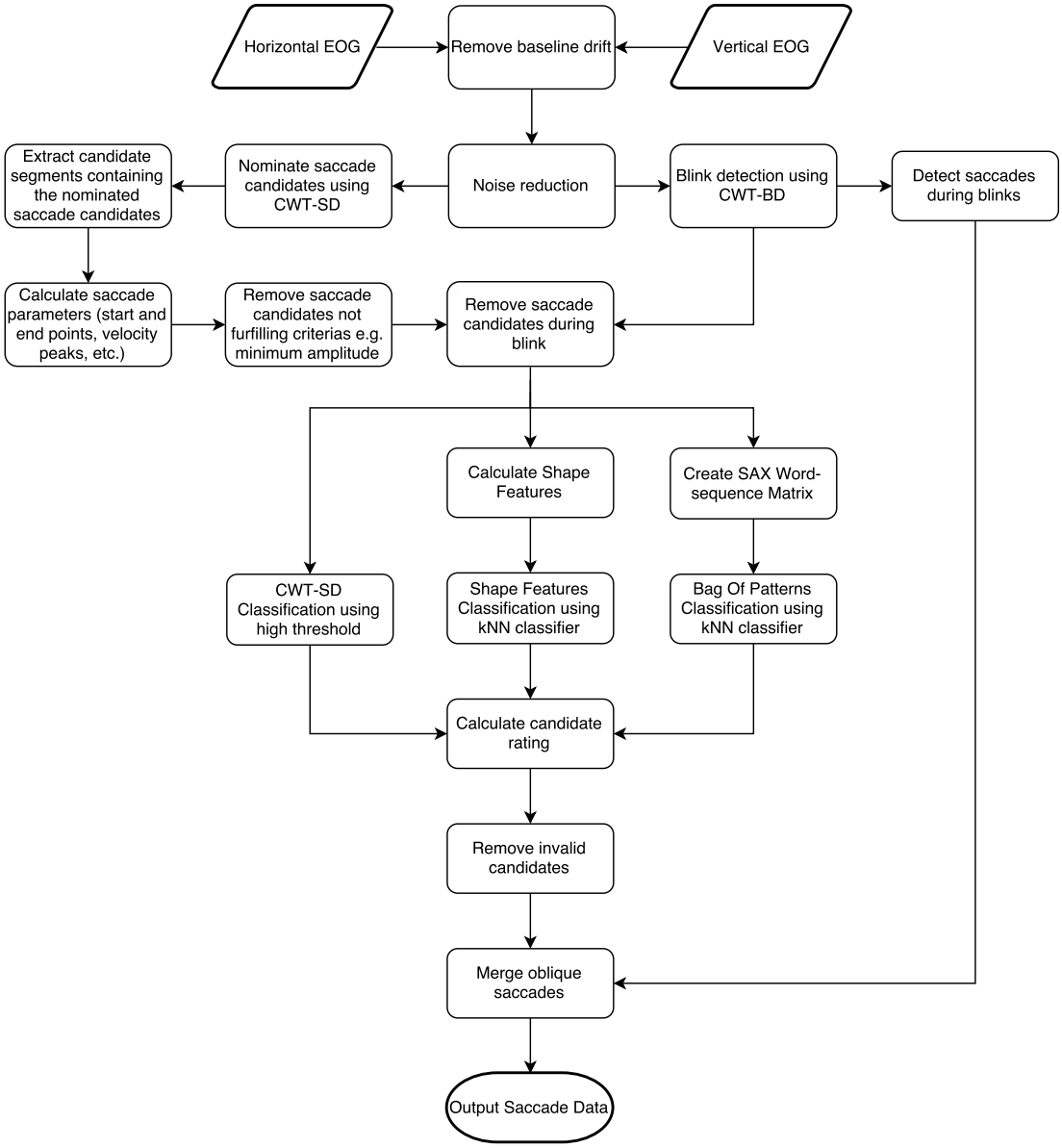


Figure 3.11: Flowchart of the implemented algorithm

Vertical saccades cannot be detected during blinks. Artifacts in the horizontal channel are also common during blinks. All saccades detected during blinks are therefore removed. Instead, saccades are detected by calculating the difference in amplitude before and after the blink. If a significant change is detected, a saccade during the blink is assumed. Amplitude differences of more than 30 μV in the horizontal channel and 70 μV in the vertical is detected as saccades. A larger difference is needed in the vertical channel since the "discharge period" of the blink often is longer than the duration of the blink (signal has a raised potential after the eyes are open, without any indication of eyelid or eye movement in the video).

Next, Shape Features and Bag-Of-Patterns are used to classify all remaining candidate segments. Each candidate is evaluated using the results of the classification.

If an algorithm classifies the candidate as a saccade, the rating is increased by one. Furthermore, if the peak value of the wavelet coefficients for the candidate is greater than 40, the candidate rating is also increased by one. A candidate is valid if the candidate rating is equal to or greater than two. After evaluating the results of the tuning, minimum accepted candidate rating was changed to three and rating given for Shape Features was increased to three. This is discussed further in Section 4.3.

Finally, all saccades occurring simultaneously in both EOG channels are merged into one oblique saccade. All saccades are separated into six different classes. These are horizontal saccades, vertical saccades, oblique saccades, horizontal saccades during blink, vertical saccades during blink and oblique saccades during blinks. This is used as output, together with the saccade duration, and amplitude in each channel.

3.5 Creation of Training and Validation Datasets

To be able to evaluate the performance of the algorithms, the results needed to be compared with manually classified (true) data. This data is classified using visual inspection of video obtained with the eye tracker and eye tracker data. In order to simplify this process, a Graphical User Interface (GUI) was developed using MATLAB. As shown in Figure 3.12, the GUI displays the video together with the horizontal and vertical eye tracker and EOG signals. Using the keyboard or the navigation buttons in the GUI, the user can step between the frames in the video. At the same time, a red and yellow line shows eye tracker and EOG signal between the previous and current frame. Using this information, the user can identify the eye movements. Using the vertical cursors, the user then marks the saccades and the blinks. The start- and end point of the saccades can be automatically calculated using the method in Section 3.2.3, but can also be marked manually if the automatic algorithm fails. Both the horizontal and vertical EOG channels were classified this way.

This verification approach is a potential source of error since small eye movements occasionally are hard to distinguish in the video. The tracking glasses can also move, creating an artifact that resembles an eye movement. The limited frame rate (10 Hz) in the provided video is also a potential source of error, particularly when the EOG signal contains a significant amount of noise and artifacts. Moreover, the eye tracker signal also contains artifacts and signal drops, further limiting the quality of the validation dataset. A sampling frequency of 60 Hz for the eye tracker is also a limitation when investigating small saccades since these could contain few (one or two) data points.

In controlled environments, where the head position is kept firm, the resolution of EOG is normally considered to be in the range of 0.5° to 1°. Since dynamic overshoots generally have an amplitude less than this, detection of dynamic overshoots is unreliable. Furthermore, in a driver simulation study, the head is not kept stationary, likely introducing more noise and artifacts. The separation between dynamic overshoots and eyelid artifacts are also unreliable in EOG data, especially in the

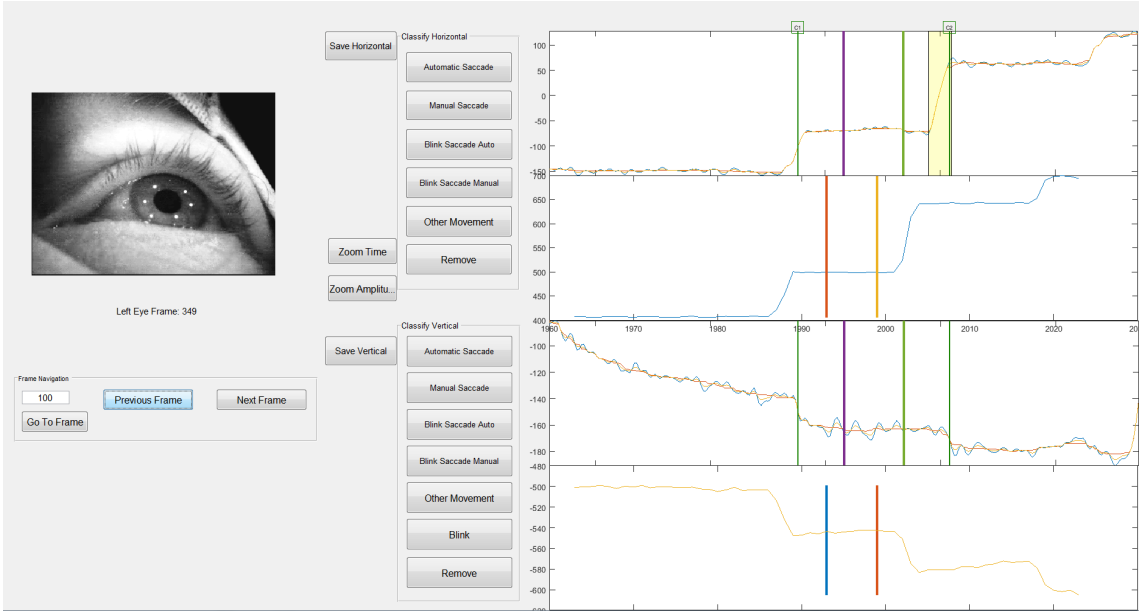


Figure 3.12: Screenshot of the developed GUI for creating the training and validation datasets. On the right side, the GUI displays (from the top to the bottom) the horizontal EOG, the horizontal eye tracker position signal, the vertical EOG and the vertical eye tracker position signal. On the left side, the GUI displays the video frame corresponding to the time segment delimited by the innermost vertical lines in the EOG and eye tracker signals.

vertical channel. The validation data is also not able to reliably validate a dynamic overshoot since the sampling rates of the video and eye tracker are too low. For the reasons described above, only the primary saccade movement is considered when determining start and end points of a saccade. Dynamic overshoots are therefore ignored to be able to get a consistent start and end position of saccades.

Another limitation related to EOG is the difference in signal properties between subjects. Since the sensitivity $\mu V/^\circ$ varies up to a factor 4 between subjects, the same movements can have up to a factor 4 of amplitude difference in the EOG signal. Without calibrating the signal, it is hard to define a strict limit between smooth pursuits and fixations. Since the provided dataset does not have calibration data, the decision was made only to classify saccades, consequently creating a two class problem, saccade, and other eye movements. Blinks and horizontal saccades during blinks were also classified as separate classes, but for secondary use.

Three datasets were created to train and test the algorithm. One is a rough classification of 4 subjects, intended to give many examples of saccades and other eye movements of different amplitude, duration and in different parts of the signal. The purpose of it was to create a large dataset used to train the algorithms and differentiate between saccades and other eye movements (examples of signal segments that should not be classified as saccades, such as smooth pursuits and fixations). The resulting dataset contains 1017 saccades and 1064 other movements.

The second dataset was created to tune the precision of the detection algorithm and to test the algorithm on untrained data. The dataset consists of precisely classified EOG data from 5 different subjects during a 10-second interval. Two of the subjects were the same as in the first dataset but gathered at another time interval. The resulting dataset contains 147 horizontal saccades, 66 vertical saccades, and 16 blinks. This dataset was created to tune the detection and classification performance of the CWT-SD, BOP and Shape Features classification. Horizontal saccades during blink was therefore only classified if the shape of the saccade was of "saccade shape" and vertical saccades during blinks was not annotated.

Finally, a validation dataset was created. This was done to test the final performance of the complete algorithm. In this dataset, movements were classified to match the output of the algorithm. In other words, saccades during blinks were classified as saccades if the level in the channels changed before and after the blink. Simultaneous saccades in both channels were also merged into one saccade (oblique). Classification of saccades is occasionally hard and subjective. Consequently, there is a significant risk for bias in this dataset. To reduce this risk, the following measures were taken:

- The saccades that are hardest to classify manually are in the range of 0-30 μV. Since all the detected saccades below 20 μV due to performance issues are discarded in the final algorithm (see Section 4.3), only clear cases (clear indication of saccade in EOG, video and eye tracker) were classified. Since performance in this range will always be zero, saccades under 20 μV were primarily annotated to indicate that saccades exist in this interval.
- Oblique saccades often have a large component in one direction and a smaller one in the other. This makes it difficult to distinguish if a saccade is oblique or not, particularly in the video. Because of this, all oblique saccades are classified as saccades (no separate oblique saccade class exists in the validation dataset, only the class saccade). The result of this is that performance of the separation of oblique and non-oblique saccades cannot be evaluated, but detection of saccades is evaluated more accurately. This was considered more important since number of oblique saccades is not used in the cognitive distraction analysis, but detection (number of saccades of any type) is.

The data was collected in a 20-second interval for three previously untested subjects. The resulting dataset contains 159 saccades, with 144 saccades with EOG amplitudes greater or equal to 20 μV.

The final algorithm is evaluated in six different EOG saccade amplitude ranges. These are 0-20, 20-30, 30-50, >50, >0 and >20 μV. For each range, F1 score, precision and recall is calculated. Precision is a measure of the fraction of detected instances that are true instances, in other words how many of the detected saccades are saccades. Precision is calculated according to

$$Precision = \frac{T_P}{T_P + F_P} \tag{3.2}$$

where T_P is the number of true positives (detected saccades are saccades) and F_P is the number of false positives (detected saccade is not a saccade). Recall measures the fraction of instances that are detected, or in other words, how many of the saccades are detected and classified as saccades. Recall is calculated as

$$Recall = \frac{T_P}{T_P + F_N} \tag{3.3}$$

where F_N is the number of false negatives (number of missed saccades). F1 score is the harmonic mean of precision and recall and is often used as an accuracy measure of binary classification. It is calculated according to Equation (3.4):

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \tag{3.4}$$

3.6 Limited analysis of saccades evaluated with and without cognitive load

The final part of this thesis work was to evaluate how the saccadic eye movements change during periods of driving with and without a cognitively distracting task. It was decided to only consider the scenario during which the driver approaches and then passes a four-way intersection. During the scenario, a second car also approaches the intersection but it stops once has reached it. The developed algorithm was applied to the provided 29 subjects' EOG signals. For this study, two saccade parameters were considered, namely the saccade rate and the saccade amplitude. The information about these two parameters during the time intervals corresponding to the cognitive tasks and the baseline was then extracted from the data output of the algorithm. For each time interval only the last 50 seconds out of 60 were considered in order to neglect the "surprise effect" in the beginning of the tasks (the first 10 seconds). In these time segments, the number of detected saccades for the baseline and for each task was obtained, as well as the saccade amplitudes. However, the amplitude of the oblique saccades needed to be calculated as the resultant amplitude of the horizontal and vertical component, such as in Equation 3.5

$$S_{oblique} = \sqrt{S^2_{horizontal} + S^2_{vertical}} \tag{3.5}$$

where $S_{horizontal}$ and $S_{vertical}$ are the amplitude values of horizontal and vertical saccades, while $S_{oblique}$ corresponds to the amplitude of the oblique saccade.

Furthermore, it was decided to carry out the comparison between cognitive load and baseline (no load) "subject-wise", in the sense that the saccadic rate and saccade amplitude in each task were compared within each individual subject, and not between all subjects. The reason lies in the fact that the correlation between the saccade parameters for different subjects is not reliable due to the lack of EOG calibration.

4
Results

In this chapter, the performance of the algorithms and the results of the cognitive distraction study are presented. First, parameter selection and performance of BOP and Shape Features are evaluated separately. Changes to the algorithm are then motivated together with the final results of the complete algorithm. Finally, the results of the cognitive distraction study are presented.

4.1 Bag-Of-Patterns

The parameters for the Bag-Of-Patterns algorithm were found by comparing the results when the algorithm was trained on the training set with 25 % holdout validation. An alphabet size (α) in the range of 3-4, word size (w) in the range of 4-5, and a window size about 1/2 - 3/4 of the segment size resulted in performance higher than other parameter settings. Table 4.1 shows the performance of the algorithm using parameters in the specified range for three different segment sizes. Four parameter settings were selected to be evaluated further on the second dataset.

Segment Size	α	w	n	Accuracy Test
21	3	4	12	0.89033
21	3	4	16	0.9052
21	3	5	15	0.9368
21	4	5	15	0.92193
31	3	4	16	0.88476
31	3	4	20	0.94052
31	3	5	20	0.92565
31	4	5	20	0.9052
41	3	4	20	0.92885
41	3	4	28	0.96154
41	3	5	30	0.94231
41	4	5	30	0.96923

Table 4.1: Results of Bag Of Patterns with kNN classifier for different parameters using 25 % holdout validation.

Table 4.2 shows the F1 score for the different parameter settings. The F1 score is presented for five different EOG saccade amplitude ranges, together with the threshold used in CWT-SD that generated the best results. When the best threshold is defined as 10-30, the threshold in that range did not have an impact on the results.

All algorithms perform poor in the range 0-20 μ V. The two parameter settings that perform the best for saccade amplitudes larger than 20 μV are parameter setting 1 and 4. The smaller segment size S_{size} and SAX dictionary for parameter setting 1 together with a reduced segment size make parameter setting 1 faster than parameter setting 4 in a computational point of view. The significantly greater dictionary length, 1024 against 81, also increases the risk that the classifier is overfitted. Based on this, parameter setting 1 was selected as the most appropriate.

	Parameter setting 1		**Parameter setting 2**		**Parameter setting 3**		**Parameter setting 4**	
	$S_{size} = 31$ n=20 w=4 $\alpha = 3$		$S_{size} = 31$ n=20 w=5 $\alpha = 3$		$S_{size} = 41$ n=28 w=4 $\alpha = 3$		$S_{size} = 41$ n=30 w=5 $\alpha = 4$	
Horizontal Saccade Amplitude in μV	**F1**	**BestTh**	**F1**	**BestTh**	**F1**	**BestTh**	**F1**	**BestTh**
0-20	0.49	17	0.54	17	0.51	17	0.48	17
20-30	0.95	18	0.83	18	0.90	10-30	0.88	18
30-50	0.98	10-30	0.95	10-30	0.89	10-30	0.98	10-30
>50	0.92	10-30	0.87	10-30	0.92	10-30	0.94	10-30
>0	0.84	18	0.78	17	0.81	19	0.81	26
Vertical Saccade Amplitude in μV	**F1**	**BestTh**	**F1**	**BestTh**	**F1**	**BestTh**	**F1**	**BestTh**
0-20	0.51	17	0.51	17	0.59	16	0.52	16
20-30	0.9	10-30	0.84	28	0.80	28	0.80	28
30-50	0.82	10-30	0.88	28	0.80	28	0.88	10-30
>50	0.92	10-30	0.93	10-30	0.93	10-30	0.93	10-30
>0	0.73	17	0.73	17	0.74	18	0.72	19

Table 4.2: Performance of the BOP algorithm for four different parameter sets. The table shows the F1 Score and the value of the best wavelet coefficients threshold in five different saccade amplitude ranges.

4.2 Shape Features

Examples of features plotted against each other are illustrated in Figure 4.1. These plots raised the question if all features should be included in the classifier. As seen in the figure, some features are fairly separable, while others are not, indicating that some possibly are redundant or irrelevant, thereby increasing computing complexity while not increasing or reducing performance. The results of the classifier were therefore also evaluated with a number of excluded features. Visual inspection

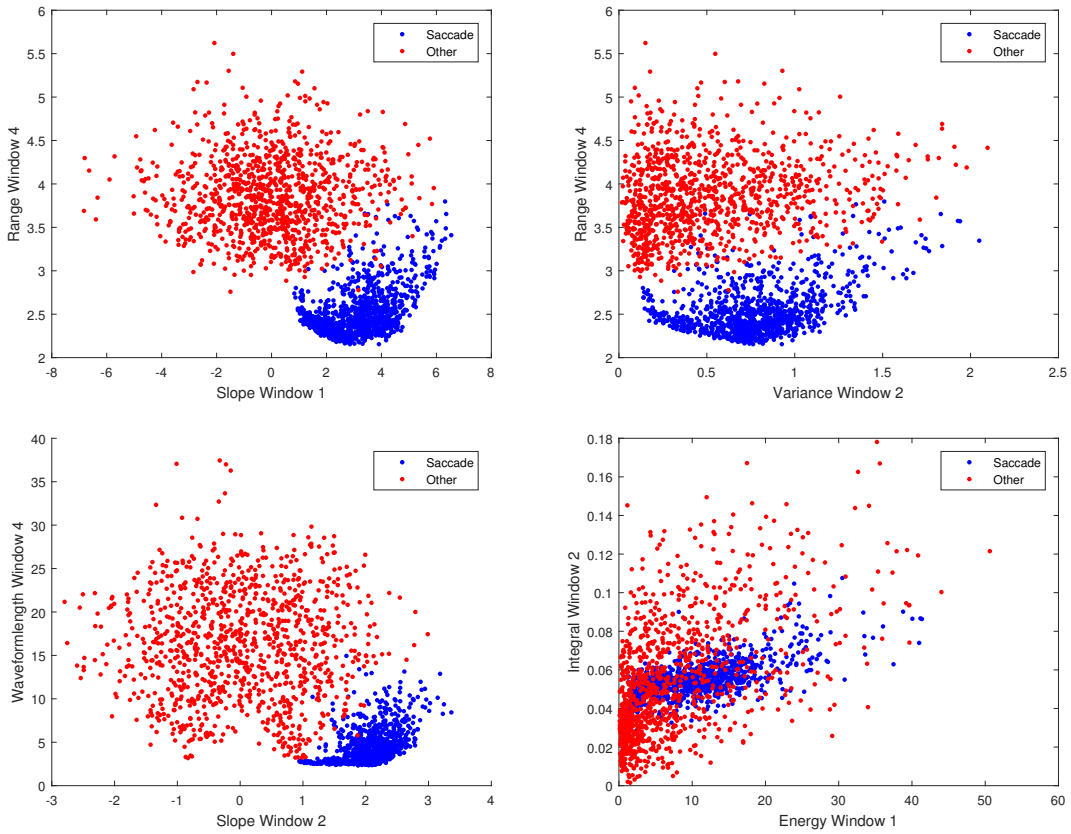


Figure 4.1: Examples of features plotted against each other.

of feature scatter plots and testing revealed that 20 of the features did not provide any information increasing performance. By removing the irrelevant features, the accuracy of the kNN classifier increased by approximately 0.05. The remaining features are slope$_{w1}$, waveform Length$_{w1}$, number Of Zero Crossings$_{w1}$, slope$_{w2}$, slope$_{w3}$, waveform Length$_{w3}$, range$_{w4}$ and waveformLength$_{w4}$.

Next, the number of neighbors (k) was investigated. A small k gives sharp classification boundaries, a property well fitted for completely separated data. However, as in this case when no exact boundary between the features exists, a larger k giving a softer boundary normally yields better performance. Using k = 10 resulted in good accuracy in this case.

The last parameter to be investigated was the segment size S_{size}. Table 4.3 shows the results of the classifier for segments sizes between 21-43 samples. As seen in the Table 4.3, the segment size does not significantly affect the results. Nevertheless, the validation accuracy appears to increase slightly with the segment size. Segment sizes 31 and 41 were selected to be evaluated further using dataset 2. The results from dataset 2 are shown in Table 4.4. Again, no major performance difference between the different segment sizes is present, but segment size 41 performs slightly better, especially in the horizontal channel for amplitudes 20-30 μV. Similar to BOP, performance is poor for saccades in the amplitude range 0-20. However, for the other

intervals, Shape Features have similar or better performance than BOP. F1 scores exceed 0.9 for all intervals except 20-30 in the vertical channel.

Segment Size (S_{size})	Validation Accuracy
21	0.9790
23	0.9792
25	0.9798
27	0.9769
29	0.9792
31	0.9864
33	0.9884
35	0.9898
37	0.9939
39	0.9941
41	0.9954
43	0.9915

Table 4.3: Shape Features algorithm: Dataset 1 validation accuracy of the kNN (k=10) classifier for segment sizes ranging from 21 to 43 samples.

	$S_{size} = 31$		$S_{size} = 41$	
Horizontal Saccade EOG Amplitude [μV]	**F1**	**Best Threshold**	**F1**	**Best Threshold**
0-20	0.5	16	0.54	17
20-30	0.91	10-30	0.97	10-30
30-50	0.96	10-30	0.98	10-30
>50	0.97	10-30	0.97	10-30
>0	0.87	17	0.89	17
Vertical Saccade EOG Amplitude [μV]	**F1**	**Best Threshold**	**F1**	**Best Threshold**
0-20	0.46	16	0.51	16
20-30	0.84	28	0.83	28
30-50	0.92	10-30	0.92	10-30
>50	0.92	10-30	0.92	10-30
>0	0.74	18	0.74	18

Table 4.4: Shape Feature performance on the second dataset

4.3 Limitations of the Algorithm

After evaluating the results from the first two datasets, two changes were made to the algorithm. Firstly, the minimum acceptance amplitude of predicted saccades was increased to 20 μV. This action was taken as a result of the poor performance in the range 0-20 μV and since precision, in this case, is prioritized over recall. Secondly, the candidate rating system was changed. The rating points given for Shape Features classification and the accepted candidate rating was increased to 3. Consequently, a saccade candidate needs to be accepted by Shape Features since the maximum points rewarded by BOP and wavelet classification is limited to 2. That is to say, the role of BOP and wavelet classification is reduced to providing a confidence rating and could thus be removed from the algorithm if computational efficiency is favored.

As mentioned further in section 4.6, about half of the subjects needed to be removed from the study due to inaccurate results from the algorithm. Some of these were removed due to measurement error (inaccurate test setup, signal loss, unreasonable high noise level), but most were removed due to large artifacts in the EOG or algorithm failure. This is a clear indication that both the measurement method and algorithm is unstable in driving environments.

Noise and artifacts are the largest limitations to the stability of the algorithm. The noise level and presence of artifacts often vary within the subjects. Part of the signal could be of good quality while other parts are heavily affected by noise and artifacts. One type of noise that the algorithm fails to handle is a type of (relatively) low frequency, high amplitude noise common in some subjects. It often appears in bursts of 1 - 10 seconds and causes complete breakdown of the algorithm. This is illustrated in Figure 4.2. The noise in this segment is high with peak to peak amplitudes approaching 100 μV, but the algorithm is very sensitive to this type of noise and amplitudes down to 20-30 μV can cause problems. High-frequency noise is generally handled better, but noise levels higher than 20 μV increases the risk of false triggers.

One of the biggest limitation of the algorithm is that it does not discover the blinks of subjects where the peak amplitude of the blinks are low (amplitudes around 100-150 μV). This was the case for about 10 % of the tested subjects. Blinks are instead often classified as two saccades in these situations.

As mentioned previously in 3.5, another big limitation is the difference in amplitude in μV/° between subjects. This is particularly clear for some subjects. A 20-second section was manually classified for one subject and over 50% of the saccades had amplitudes below 20 μV. Some of these saccades were small, but most would exceed 20 μV by far in other subjects (based on movements visible in the video). The algorithm performed well for saccades greater than 20 μV, but 50% of the saccades was missed. For the algorithm to function properly the following criteria must be met

- No big artifacts can be present.
- The noise level should preferably have an average peak to peak noise level lower than 20 μV. If low-frequency noise of the type described above is present, the amplitude must be lower than 20 μV.
- The blinks must have a peak amplitude greater than 200 μV, preferably larger.
- The amplitude μV/degree can not be to low. Movements must show a clear change in signal level.

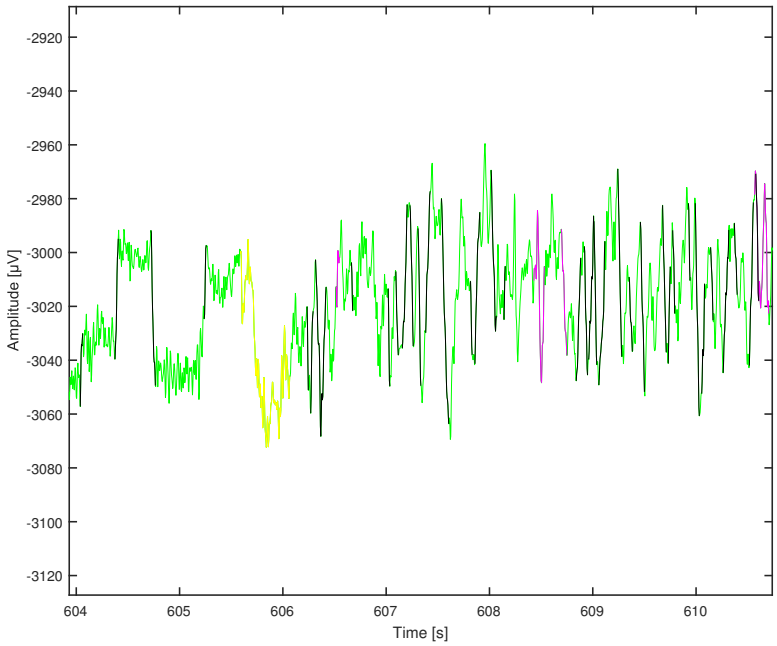


Figure 4.2: Example of the type of low-frequency noise normally causing algorithm failure. Colors other than green represent segments classified as saccades.

4.4 Performance of the algorithm

The created validation dataset follows the criteria described above. Results from the algorithm applied to the validation set are displayed in Table 4.5. Naturally, the scores for saccades with an amplitude less than 20 μV is 0, since these automatically are disregarded by the algorithm. For saccades in the range of 20-30 μV, the F1 score is 0.90, slightly less compared to the training set (horizontal channel). However, for saccades with an amplitude greater than 30 μV, the algorithm performs as well or better than on the training set with F1 scores of 1 for saccades in the range of 30-50 μV and 0.99 for saccades greater than 50 μV. This increased performance can potentially be explained by the merging of oblique saccades, where saccades that were missed in one channel are "saved" by detection in the other channel.

The overall F1 score for saccades greater than 20 μV is 0.98, with a precision of 0.97 and a recall of 0.99. However, the dataset is small and the analysis is performed on good signals. Thus, the results cannot be said to reflect realistic cases. The F1 score when all saccades (amplitudes less than 20 μV) are considered is 0.93 with a precision of 0.97 and recall of 0.81. This score should, however, be lower since not all saccades below this threshold were classified.

EOG Amplitude [μV]	F1 Score	Precision	Recall	Number of Saccades in the range
0-20	0	0	0	15
20-30	0.91	0.89	0.92	26
30-50	1	1	1	27
>50	0.99	0.99	1	91
>0	0.93	0.97	0.89	159
>20	0.98	0.97	0.99	144

Table 4.5: Algorithm performance: for each amplitude range F1 score, precision, recall and number of detected saccades is computed.

4.5 Saccades start- and end point performance

The start- and end points of the primary saccadic movement of the correctly classified saccades from the validation dataset was visually inspected. Only saccades not occurring during blinks were investigated since the saccades during blink have incorrect start- and end points (the start and end points are the start and end of the blink not the saccades). The results showed that 91% (out of 132 saccades) of the saccades had correct and consistent start- and end points of the primary saccade. This is, however, the assumed start- and end points in the EOG signal, and does not necessarily agree with the real start- and end point of the saccade. Furthermore, some start- and end points are hard to distinguish, and since no ground truth is available, it is hard to specify how well this function works. The results are however consistent in 91% of the validation set and are believed to be close to the truth.

4.6 Analysis of the effects of cognitive distraction during driving

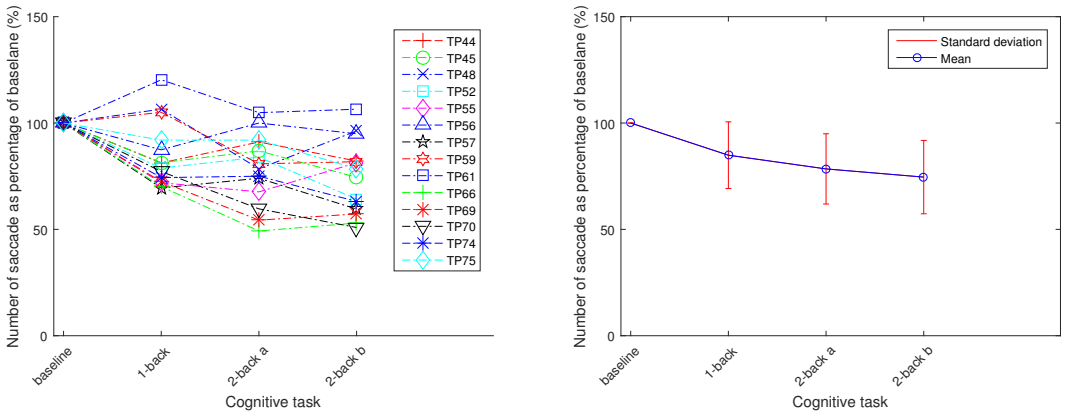
For this study, the EOG data of 14 out of 29 subjects were selected. The remaining 15 subjects were excluded due to the following reasons:

- Incomplete recording due to signal loss [1 subject]
- Presence of big artifacts [2 subjects]
- Small blink amplitudes compared to the other subjects, resulting in inaccurate blinks and thereby inaccurate saccade during blink detection. [3 subjects]

- High noise level (up to 100μV) resulting in algorithm failure [5 subjects]
- High blink rate (>1 per second) [1 subject]
- The horizontal and the vertical EOG signal were identical (assumed measurement setup error) [1 subject]
- Too many false triggered saccades (algorithm failure) [2 subjects].

Most of the accepted subjects meet the signal criteria described in Section 4.3. Some of them do contain big artifacts in the vertical channel, but visual inspection deemed these not to affect the results. It is worth noting that both false positives and false negatives are present in many of the accepted subjects but at an acceptable rate.

The parameters considered for the study is the saccadic rate, defined as the number of saccades per 50 seconds, and the average saccade amplitudes in each task. Figure 4.3 shows two plots, representing the saccadic rate for each subject during each task (left plot) and the mean saccadic rate of all subjects with the corresponding standard deviations (right plot). As mentioned before, each subject performed one 1-Back task (easy task) and two 2-Back tasks (difficult task) during the test. The two 2-Back tasks have been called "2-Back-a" and "2-Back-b", in order to distinguish them. The difference between these two is that 2-Back-a comes earlier in the drive than 2-Back-b. Moreover, the number of saccades is evaluated as the percentage of the saccadic rate of the baseline, defined as 100%.



(a) Mean number of saccades of each subject in each task.

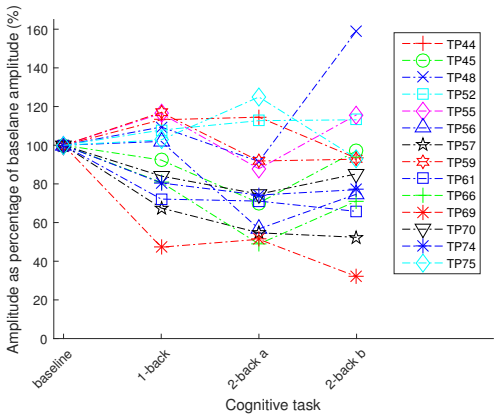
(b) Global means of number of saccades of all subjects in each task.

Figure 4.3: Saccadic rate for each cognitive task (1-back, 2-back-a, 2-back-b) and for the baseline (no cognitive task). After having defined the saccadic rate during baseline to be 100%, the number of saccades in the other tasks are evaluated as the percentage of the baseline saccadic rate. Saccadic rate of all subjects (on the left), mean values and standard deviations of the saccade rate for all subjects (on the right) are plotted.

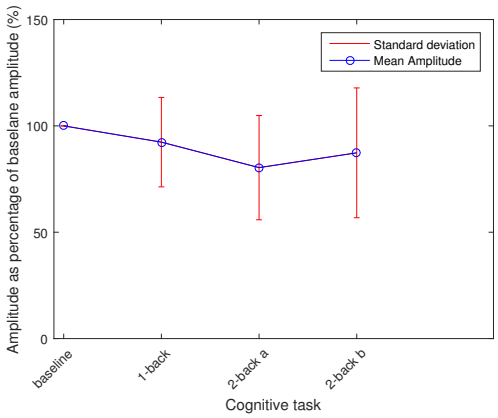
As seen in 4.3b, the mean saccade rate for saccades with amplitudes greater than 20 μV decreased as the cognitive load increased. Although Figure 4.3a shows that one subject (TP61-blue square) has a tendency to perform approximately the same number of saccades as the baseline during cognitive distraction, the overall trend

of the saccadic rate among most subjects decreases as the cognitive load increases. As illustrated in Figure 4.3b the mean of the number of saccades, expressed as the percentage of the baseline value, goes from 100% of baseline to 84.9% for 1-Back, and further to 78.4% and 74.6%, corresponding to 2-Back-a and 2-Back-b. A mixed-effects model ANOVA, using test subject as random factor was recommended to investigate if a statistically significant difference exists between the tasks. It revealed a significant difference of the number of saccades between the tasks ($F(2, 40) = 24.93$, $p < 0.0001$). Post hoc tests using Bonferroni correction showed significant differences in saccadic rate between baseline and 1-Back task ($p = 0.0006$), baseline and 2-Back task ($p < 0.0001$), but not between 1-Back and 2-Back tasks ($p = 0.0755$).

Figure 4.4 shows the average amplitude of the saccades performed by each subject during each task (Figure 4.4a) and the mean of the average amplitude of all subjects (Figure 4.4b). Having the baseline as 100%, the amplitude value of the tasks is expressed as the percentage of the baseline.



(a) Mean saccade amplitude of each subject in each task.



(b) Mean saccade amplitude values over all subjects in each task.

Figure 4.4: Average saccade amplitude for each cognitive task (1-back, 2-back-a, 2-back-b) and the baseline (no cognitive task), calculated as the percentage of the baseline amplitude. Average amplitudes of all subjects (on the left), mean values and standard deviation of the saccade rate for all subjects (on the right) are plotted.

There is a small indication that the amplitude decreases during cognitive load compared to baseline driving. However, some subjects make bigger saccades when cognitive load is present. Figure 4.4b shows that the general trend of the mean saccade amplitudes of all subjects is decreasing as the cognitive load increases. A mixed-effects model ANOVA, using test subject as random factor revealed a significant difference of the amplitude of saccades between the tasks ($F(2, 40) = 6.5$, $p = 0.0037$). Post hoc tests using Bonferroni correction showed significant differences in saccade amplitude between baseline and 2-Back task ($p = 0.003$) but not between 1-Back and 2-Back tasks ($p = 0.2692$), and between baseline and 1-Back ($p = 0.3744$).

In summary, the results from this limited study suggest that cognitive distraction causes the drivers to make fewer saccades as the cognitive load increases. These results are in agreement with the previously reported by [46]. A statistically significant reduction of saccade amplitude was revealed only between the difficult cognitive task and the baseline. Furthermore, the spread of the amplitudes are big and the reduction in amplitude is small. Therefore, no clear indication of how this saccade parameter is affected by cognitive load can be determined, which is in line with contradictory findings in [47], [48] and [49].

5

Discussion

We have developed an algorithm that proved to work well in detecting saccades bigger than 20 μV in EOG signals that satisfy certain quality criteria. To the best of our knowledge, this is the first work that combines wavelet-based saccade detection with shape-based saccade detection. Even though no quantitative comparison has been made, we believe that our algorithm performs better than the existing *Continuous Wavelet Transform - Saccade Detection*. However, there are limitations that to some extent hindered the performance of the algorithm, and they will be discussed in the following paragraphs.

The lack of calibration data and an accurate ground truth dataset have been a significant limitation regarding the quality of the training and validation data sets. If calibration data was available, stricter limits connected to the velocity and amplitude definitions of what should be regarded as a saccade could have been derived. These limits would have been connected with the velocity definitions of saccades, giving more credibility to the created datasets, hence making the ground truth more accurate and improving the performance of the algorithm. Unfortunately, uncalibrated EOG data combined with unclear ground truth data made the classification hard, arbitrary and time-consuming. The biggest limitation with the ground truth data was the sampling rate of the video. One frame every 100 ms is not enough to determine if a movement was made rapidly enough to be a saccade. The motivation for why one movement is classified as a saccade while another similar movement is not is very arbitrary. A higher sampling rate could help with this, but since the differences between the frames would be very small, it might be difficult for a human to interpret these as well. An attempt of having an independent expert group creating a dataset was made, but since they had similar problems as us, this approach was abandoned. Consequently, the resulting dataset represents what we interpret as saccades, and not what necessarily are saccades.

The final dataset was made on subjects with good signal quality. Attempts to classify subjects with worse signals were done, but this turned out to be very hard. Because of this, no performance measure for subjects not fulfilling the criteria in 4.3 was obtained.

Currently, different saccade sizes will be detected or disregarded for different subjects. If calibration was available, it would be possible to define a limit for the smallest eye movement size that could be detected, significantly improving the usability of the algorithm, since data could be more reliably compared between sub-

jects. Furthermore, a strict limit and calibrated data would restrict the different signal properties of a saccade. This could create more clearly separable features, likely leading to better classification performance. Moreover, the current algorithm was optimized to be able to detect small saccades (10 μV). By finding the smallest reasonable saccade amplitude, the algorithm could be optimized for that amplitude (mainly in terms of filtering, thresholds and training data), further improving the results. Nonetheless, comparisons can still be made within subjects, even though differences exist between subjects.

The data used for BOP and Shape Features training was created by marking different sections containing saccades, smooth pursuits, and fixations. This was a mistake since the resulting dataset is "too nice". The created dataset has relatively separable features resulting in very high performance using 50% holdout validation for shape features. However, the saccade candidates nominated by CWT-SD are generally much more similar to saccades in their shape compared to the negative examples in the dataset. A better way of training the algorithm would have been to train it on candidates nominated by CWT-SD, and then manually classified. This would not only create a dataset that better represents the final performance of the algorithm but would likely better capture situations where the algorithm currently fails.

Robustness of the algorithm would likely be improved by calibration and improved training methods, but the main problem with noise, artifacts and reliability of EOG remains. EOG is traditionally used in controlled environments where the head is kept still to reduce artifacts. Since the head frequently moves in driving environments, we believe this is the main source of the big artifacts. Some subjects are also contaminated with significant noise with peak to peak amplitudes up to 100 μV. High amplitude noise could be present in the entire signal or come in bursts. This was made clear when inspecting the segments in the cognitive study. For some subjects, the performance of the algorithm was good in 2-3 out of 4 analysis segments for the investigated scenario but contained high amplitude noise or artifacts in the other 1-2 analysis segments.

Using an algorithm that adapts to the noise level either by adjusting filtering, nominating threshold or minimum accepted saccade amplitude was considered. This idea was however rejected since the saccade frequency was to be investigated. If an adaptable algorithm was used, more saccades would be detected in parts of the signal with low noise level compared to parts containing a high noise level. Consequently, comparing saccade frequency would not be appropriate.

As long as the signal is of good quality the performance is fairly high and the algorithm works well. However, it is highly sensitive to certain types of noise, and overall performance of the algorithm, when applied to all subjects, is low. The different noise properties of the signal also affect the stability within a subject, and the signal needs to be visually inspected to ensure that the signal does not contain significant noise or artifacts. In some cases, the algorithm simply fails since it is possible to see the saccades in the signal (but it is not classified as a saccade by the

algorithm), but often the signal cannot be used to reliably detect saccades even with manual classification. Unless the measurement setup can be significantly improved to a point where the signals are more stable, EOG is not a reliable signal for saccade detection in driving environments.

In our opinion, the eye tracker data from the SMI eye tracking glasses seems to be more stable and reliable. Tracking is occasionally lost, but only for short periods of time and mostly during blinks. The resolution is also higher than EOG and smaller movements are easier and more reliably distinguished in the eye tracker data. The algorithm should be able to be modified to work on eye tracking data, likely with improved and more stable results.

Our opinion is that EOG recordings have reached a technical limit while eye tracking has not. The currently used tracker has a limit in temporal resolution, but a 120 Hz version is already available. While EOG barely has developed since it was invented, eye tracking is a constantly developing and competitive field. Furthermore, eye tracking is becoming more common in consumer products and is believed to be a vital part of Virtual Reality (VR). Movement into the consumer space often accelerates development of electronic products leading to cheaper and better products. We believe this will happen to eye tracking technology as well. The result of this is likely that EOG will not be able to compete with eye tracking, except in studies where the eyes are kept closed.

5.1 Cognitive distraction analysis on saccades

The study of the effects of cognitive distraction on saccadic eye movements was planned to be a final and secondary aim of our thesis work. For this reason and due to a limited availability of time, the study was decided to be focused on the analysis of saccadic rate and saccade amplitude during only one specific scenario. For the considered "four-way intersection" scenario, a trend was a general decrease both in the number of saccades per unit time and saccade amplitude. Therefore, a tendency towards "tunnel vision", or "attentional narrowing", is indicated, which seems to be consistent with the literature. However, the results could possibly be improved regarding higher statistical significance if more test subjects and thus more data was available. Naturally, the quality of the signal should preferably be as good as possible in order for the algorithm not to trigger false saccades.

A clearer reduction in saccade amplitude was expected to be observed, since "attentional narrowing" implies that the gaze is concentrated straight ahead. Therefore, it is reasonable to assume that the driver during cognitive distraction should perform smaller saccades. Furthermore, the results are possibly affected by the inability to detect small saccades (<20 μV) that would have reduced the average amplitude.

5.2 Future Works

The recommendations for future works mainly aim at improvements in data collection. First of all, the calibration procedure of EOG needs to be done before the start of the test. It is necessary to be able to convert μV to degrees (°) and vice-versa, if a reliable saccade amplitude is desired.

Secondly, a higher sampling rate could be beneficial in gaining clearer information about small saccades. We had the opportunity to examine the EOG dataset used in [20]. In this study, the EOG was sampled at 500 Hz and the small saccades seemed to be more easily identified. Furthermore, a higher sampling rate means that the saccades are described by more data points. This could possibly improve the performance of the Bag-Of-Patters algorithm, since, according to [40], it works better on more data.

Finally, a minor recommendation is related to the test subject's preparation in terms of the measurement setup. Since the EOG signals of some test subjects were strongly affected by noise or artifacts, a good and stable placement of the electrodes must be ensured.

6
Conclusion

The thesis set out to develop an automatic method for detection of saccades using EOG in driving studies. The resulting algorithm is a combination of two modified existing eye detection algorithms, namely Continuous Wavelet Transform Saccade Detection (CWT-SD) and Shape Features. It is found that the developed algorithm can be used in driving environments if good signal quality can be assured. Results from the validation set reported an F1 score of 0.98 for saccades with amplitudes greater than 20 μV and an overall F1 score of 0.93. However, the limited dataset and lack of reliable ground truth make these results unreliable. Moreover, the instabilities regarding noise and artifacts discovered in the investigated EOG signals do limit the reliability and hence the usability of the algorithm. In general, therefore, it seems that EOG can be utilized for saccade detection during driving, but the experienced higher stability and resolution of the eye tracker likely make it a more suitable method.

The second aim of this thesis was to investigate the effects of cognitive distraction on saccades in driving environments. The findings of the study suggest a decrease in saccadic rate during cognitive distraction. The results were found to be statistically significant using a mixed-effects model ANOVA with subject as random variable and the results are in agreement with [46]. Average saccade amplitude was also found to decrease during cognitive distraction but a statistical significance was only observed between high cognitive load (difficult task) and no cognitive load (baseline). Since the study was limited to 14 subjects and one driving scenario, further research is needed to verify the results.

Bibliography

[1] M. F. Bear, B. W. Connors, and M. A. Paradiso, *Neuroscience : exploring the brain*. Baltimore : Lippincott Williams Wilkins, 2nd ed ed., 2001.

[2] G. J. Tortora and B. Derrickson, *Essentials of anatomy and physiology*. Chichester;Hoboken, N.J;: Wiley, 8. ed., 2010.

[3] J. Malmivuo and R. Plonsey, *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. Oxford University Press, USA, 1995.

[4] R. J. Leigh and D. S. Zee, *The neurology of eye movements*. Oxford University Press, 5 ed., 2015.

[5] A. Duchowski, *Taxonomy and Models of Eye Movements*, pp. 41–48. London: Springer London, 2007.

[6] J. Enderle, *Models of Horizontal Eye Movements: Early models of saccades and smooth pursuit*. No. pt. 1 in Synthesis lectures on biomedical engineering, Morgan & Claypool, 2010.

[7] F. D. e. a. Purves D, Augustine GJ, *The Actions and Innervation of Extraocular Muscles*. Sunderland (MA): Sinauer Associates, 2001.

[8] S. Lebedev, P. Van Gelder, and W. Tsui, "Square-root relations between main saccadic parameters," *Investigative Ophthalmology Visual Science*, vol. 37, no. 13, p. 2750, 1996.

[9] T. Eggert, "Eye movement recordings: methods," in *Neuro-Ophthalmology*, vol. 40, pp. 15–34, Karger Publishers, 2007.

[10] K. Bötzel, K. Rottach, and U. Büttner, "Saccadic dynamic overshoot in normals and patients," *Neuro-Ophthalmology*, vol. 13, no. 3, pp. 125–133, 1993.

[11] J. Tian, H. S. Ying, and D. S. Zee, "Revisiting corrective saccades: role of visual feedback," *Vision research*, vol. 89, pp. 54–64, 2013.

[12] M. Rucci, P. V. McGraw, and R. J. Krauzlis, "Fixational eye movements and perception," *Vision Research*, vol. 118, pp. 1 – 4, 2016. Fixational eye movements and perception.

[13] A. Bulling, J. A. Ward, H. Gellersen, and G. Troster, "Eye movement analysis for activity recognition using electrooculography," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 4, pp. 741–753, 2011.

[14] Tobii AB, "Types of eye movements." Available at `http://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/types-of-eye-movements/`. Accessed on 2017/03/02.

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

[16] C. Helmchen and H. Rambold, "The eyelid and its contribution to eye movements," in *Neuro-Ophthalmology*, vol. 40, pp. 110–131, Karger Publishers, 2007.

[17] R. I. Hammoud, *Passive eye monitoring: Algorithms, applications and experiments*. Springer Science & Business Media, 2008.

[18] M. Vidal, J. Turner, A. Bulling, and H. Gellersen, "Wearable eye tracking for mental health monitoring," *Computer Communications*, vol. 35, no. 11, pp. 1306–1311, 2012.

[19] P. A. Constable, M. Bach, L. J. Frishman, B. G. Jeffrey, A. G. Robson, I. S. for Clinical Electrophysiology of Vision, *et al.*, "Iscev standard for clinical electro-oculography (2017 update)," *Documenta Ophthalmologica*, vol. 134, no. 1, pp. 1–9, 2017.

[20] K. Pettersson, S. Jagadeesan, K. Lukander, A. Henelius, E. Hæggström, and K. Müller, "Algorithm for automatic analysis of electro-oculographic data," *Biomedical engineering online*, vol. 12, no. 1, p. 110, 2013.

[21] A. Sprenger, B. Neppert, S. Köster, S. Gais, D. Kömpf, C. Helmchen, and H. Kimmig, "Long-term eye movement recordings with a scleral search coil-eyelid protection device allows new applications," *Journal of Neuroscience Methods*, vol. 170, no. 2, pp. 305 – 309, 2008.

[22] Tobii AB, "How do tobii eye trackers work?." Available at `http://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/how-do-tobii-eye-trackers-work/`. Accessed on 2017/03/21.

[23] D. J. Mack, P. Schönle, S. Fateh, T. Burger, Q. Huang, and U. Schwarz, "An eog-based, head-mounted eye tracker with 1 khz sampling rate," in *Biomedical Circuits and Systems Conference (BioCAS), 2015 IEEE*, pp. 1–4, IEEE, 2015.

[24] V. S. Chouhan and S. S. Mehta, "Total removal of baseline drift from ecg signal," pp. 512–515, IEEE, 2007.

[25] N. Pan, V. Mang, M. P. Un, and P. S. Hang, "Accurate removal of baseline wander in ecg using empirical mode decomposition," pp. 177–180, 2007.

[26] H. Sharma and K. K. Sharma, "Baseline wander removal of ecg signals using hilbert vibration decomposition," *Electronics Letters*, vol. 51, no. 6, pp. 447–449, 2015.

[27] Y. Luo, R. H. Hargraves, A. Belle, O. Bai, X. Qi, K. R. Ward, M. P. Pfaffenberger, and K. Najarian, "A hierarchical method for removal of baseline drift from biomedical signals: Application in ecg analysis," *The Scientific World Journal*, vol. 2013, pp. 1–10, 2013.

[28] R. Von Borries, J. Pierluissi, and H. Nazeran, "Wavelet transform-based ecg baseline drift removal for body surface potential mapping," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pp. 3891–3894, IEEE, 2006.

[29] 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.

[30] D. D. Salvucci and J. H. Goldberg, "Identifying fixations and saccades in eye-tracking protocols," in *Proceedings of the 2000 symposium on Eye tracking research & applications*, pp. 71–78, ACM, 2000.

[31] A. Smith, P. Bittencourt, D. Lloyd, and A. Richens, "An efficient technique for determining characteristics of saccadic eye movements using a mini computer," *Journal of biomedical engineering*, vol. 3, no. 1, pp. 39–43, 1981.

[32] P.-H. Niemenlehto, "Constant false alarm rate detection of saccadic eye movements in electro-oculography," *Computer methods and programs in biomedicine*, vol. 96, no. 2, pp. 158–171, 2009.

[33] M. Vidal, A. Bulling, and H. Gellersen, "Detection of smooth pursuits using eye movement shape features," in *Proceedings of the symposium on eye tracking research and applications*, pp. 177–180, ACM, 2012.

[34] S. Hoppe and A. Bulling, "End-to-end eye movement detection using convolutional neural networks," *arXiv preprint arXiv:1609.02452*, 2016.

[35] A. Bulling, *Eye movement analysis for context inference and cognitive-awareness: Wearable sensing and activity recognition using electrooculography*. Lulu. com, 2010.

[36] M. A. Tinati and B. Mozaffary, "A wavelet packets approach to electrocardiograph baseline drift cancellation," *International journal of biomedical imaging*, vol. 2006, p. 97157, 2006.

[37] A. Romeo, C. Horellou, and J. Bergh, "N-body simulations with two-orders-of-magnitude higher performance using wavelets," 2003.

[38] A. B. Romeo, C. Horellou, and J. Bergh, "A wavelet addon code for new generation nbody simulations and data denoising," *Monthly Notices of the Royal Astronomical Society*, vol. 354, no. 4, pp. 1208–1222, 2004.

[39] Mathwork Nordic.

[40] J. Lin, R. Khade, and Y. Li, "Rotation-invariant similarity in time series using bag-of-patterns representation," *Journal of Intelligent Information Systems*, vol. 39, no. 2, pp. 287–315, 2012.

[41] J. Lin, E. Keogh, W. Li, and S. Lonardi, "Experiencing sax: a novel symbolic representation of time series," *Data Mining and knowledge discovery*, vol. 15, no. 2, p. 107, 2007.

[42] L. Dorn, G. Matthews, M. T. W. Victor, J. D. Lee, M. A. Regan, *et al.*, *Driver Distraction and Inattention: Advances in Research and Countermeasures*, vol. 1. Ashgate Publishing, Ltd., 2013.

[43] J. Engström, G. Markkula, T. Victor, and N. Merat, "Effects of cognitive load on driving performance: The cognitive control hypothesis," *Human factors*, p. 0018720817690639, 2017.

[44] D. L. Strayer, J. M. Cooper, J. Turrill, J. Coleman, N. Medeiros-Ward, and F. Biondi, "Measuring cognitive distraction in the automobile," 2013.

[45] N. Council, "Understanding the distracted brain: Why driving while using hands-free cell phones is risky behavior," 2010.

[46] J. L. Harbluk, Y. I. Noy, and M. Eizenman, "The impact of cognitive distraction on driver visual behaviour and vehicle control," tech. rep., 2002.

[47] J. G. May, R. S. Kennedy, M. C. Williams, W. P. Dunlap, and J. R. Brannan, "Eye movement indices of mental workload," *Acta psychologica*, vol. 75, no. 1, pp. 75–89, 1990.

[48] M. L. Reyes and J. D. Lee, "Effects of cognitive load presence and duration on driver eye movements and event detection performance," *Transportation*

research part F: traffic psychology and behaviour, vol. 11, no. 6, pp. 391–402, 2008.

[49] M. A. Recarte and L. M. Nunes, "Effects of verbal and spatial-imagery tasks on eye fixations while driving.," *Journal of experimental psychology: Applied*, vol. 6, no. 1, p. 31, 2000.

[50] Mathwork Nordic, "Automatic 1-d denoising." Available at `https://se.mathworks.com/help/wavelet/ref/wden.html`. Accessed on 2017/05/06.

[51] A. Bulling, J. A. Ward, and H. Gellersen, "Multimodal recognition of reading activity in transit using body-worn sensors," *ACM Transactions on Applied Perception (TAP)*, vol. 9, no. 1, p. 2, 2012.

[52] M. Wahde, *Interactive Partner Agents A Practical Introduction*. Chalmers University of Technology, 2017.