Analysing EOG Signal Features for the Discrimination of Eye Movements with Wearable Devices

Mélodie Vidal
Lancaster University
m.vidal@lancaster.ac.uk

Andreas Bulling
University of Cambridge
Lancaster University
andreas.bulling@acm.org

Hans Gellersen
Lancaster University
hwg@comp.lancs.ac.uk

ABSTRACT
Eye tracking research in human-computer interaction and experimental psychology traditionally focuses on stationary devices and a small number of common eye movements. The advent of pervasive eye tracking promises new applications, such as eye-based mental health monitoring or eye-based activity and context recognition. These applications might require further research on additional eye movement types such as smooth pursuits and the vestibulo-ocular reflex as these movements have not been studied as extensively as saccades, fixations and blinks. In this paper we report our first step towards an effective discrimination of these movements. In a user study we collect naturalistic eye movements from 19 people using the two most common measurement techniques (EOG and IR-based). We develop a set of basic signal features that we extract from the collected eye movement data and show that a feature-based approach has the potential to discriminate between saccades, smooth pursuits, and vestibulo-ocular reflex movements.

Author Keywords
Eye Tracking, IR-Based Eye Tracking, Electrooculography (EOG), Smooth Pursuit, Vestibulo-Ocular Reflex (VOR)

ACM Classification Keywords
I.5.3 Pattern Recognition: Clustering; C.3 Special-Purpose and Application-Based Systems: Real-time and embedded systems

General Terms
Experimentation, Measurement

INTRODUCTION
Eye tracking has been an increasingly researched topic in the last decades, especially in human-computer interaction (HCI) and experimental psychology. In both fields, researchers traditionally use stationary eye tracking to investigate the main eye movement types: fixations (when the eyes are still), saccades (the fast movement of the eyes to a new location) and blinks. Typically, these movements are the basis for extracting a small number of specific characteristics such as the blink rate or the fixation duration.

The advent of pervasive eye tracking, that is tracking movement of the eyes for several hours or even days using mobile and wearable eye tracking systems, opens up new applications areas, particularly at the crossroads of ubiquitous computing and HCI. It is now conceivable to imagine eye trackers to become part of our everyday lives within the next 10 years. Such systems will be integrated into ordinary glasses, interactive displays or devices used for a specific task [3]. Potential applications of pervasive eye tracking include activity and context recognition [4], feedback systems for improving reading performance or eye-based mental health monitoring.

There has been extensive work on the effect of mental illnesses, such as Alzheimer’s, on eye movements, particularly on smooth pursuit movements [6, 11, 7, 10]. Smooth pursuit are the movement of the eyes when following a moving target, such as a bird or a car. Being able to monitor and automatically analyse smooth pursuit movements may allow us to detect an early onset of such mental diseases. Another type of movement that has not been studied much is the visual reflex to keep a precise point in sight while moving the head, known as Vestibulo-Ocular Reflex (VOR). We could imagine a TV set or a mobile phone that would answer to eye gestures [5] where the user moves their head (thus their eyes, staying focused on the screen) to command it without looking away.

With the growing potential of pervasive eye tracking comes the need to develop algorithms and methods specifically geared towards mobile settings. These settings require a more general framework for eye movement analysis including new algorithms to detect the eye movements and more immediate interaction. The requirements for such a framework are low computational costs, real-time analysis, robustness against artefacts and a good recognition performance. When considering eye movement detection, evidence is there is a lack of algorithm to detect all of these movements altogether in real time. There has been extensive research on saccades, fixations and blinks detection [13, 8, 9] and recent developments to detect them online [12, 2], yet no algorithm detects smooth pursuit or VOR in a generic fashion alongside with these well-studied movements.
mean aged 18 to 40 years ($= 27$

We collected data from 19 participants (13 male, 6 female),

Participants

sampling frequency than stationary ones.

get in the visual field and portable eye-trackers have a lower

fluences both the user and the data. Video based eye-trackers

everyday life environment. Although the study is stationary,

cessions of eye movements from the user, as expected in an

stimulus. The stimulus was designed to generate rapid suc-

was conducted in a stationary environment with a controlled

this task was to get ground truth data. For this reason, al-

use in mobile environments. With this in mind, the goal of

our aim was to analyse eye movement features to provide

DATA COLLECTION

Goal and Contributions

This work-in-progress paper represents the first step towards a complete online algorithm for the detection of saccades, fixations, blinks, smooth pursuit and VOR from mobile eye tracking data. In this paper we present the results of a user study gathering naturalistic ground truth data from all these movements, using wearable eye trackers. We analysed the movements collected by extracting basic signal features from them and show that feature analysis is promising to characterise the different types of eye movements. The contributions of this paper are: 1) the identification of potentially useful eye movements for pervasive eye tracking and 2) the characterisation of these eye movements using basic signal features extracted from the eye movement data.

DATA COLLECTION

Our aim was to analyse eye movement features to provide a base for a future eye movement detection algorithm for use in mobile environments. With this in mind, the goal of the user study was to collect ground truth annotated, naturalistic data using wearable eye-trackers; The difficulty of this task was to get ground truth data. For this reason, although we were using mobile eye-trackers, the user study was conducted in a stationary environment with a controlled stimulus. The stimulus was designed to generate rapid successions of eye movements from the user, as expected in an everyday life environment. Although the study is stationary, the use of wearable eye-trackers is significant because it influences both the user and the data. Video based eye-trackers get in the visual field and portable eye-trackers have a lower sampling frequency than stationary ones.

Participants

We collected data from 19 participants (13 male, 6 female), aged 18 to 40 years ($mean = 27$ years, $sd = 6.9$ years).

None of the participants wore glasses during the experiment and three of them wore contact lenses. 11 participants had brown or dark eye color and eight had green or blue eyes. Participants were given £10 as a reward.

Apparatus

The experimental system consisted of three devices (see Figure 1): 1) Electrodes connected to Mobi from Twente Medical Systems International (TMSI), which was recording Electrooculography (EOG) at 128Hz, 2) the Dikablis infrared eye tracker from Ergoneers GmbH, which records gaze with a sampling rate of 25Hz, and 3) a cap with an inertial measurement unit from Xsens connected to an XBUs master, recording head movements at 50Hz. The signals from these three devices were sent to a computer running the Context Recognition Network Toolbox (CRNT) [1] to handle synchronisation and data storage. In addition to these signals, the CRNT received the ground truth labels from the graphical user interface showing the stimulus to the participants. These labels correspond to the eye movements the users were supposed to perform in reaction to the stimulus, including their amplitude and duration.

Both signals were recorded using mobile devices. This implies they have a lower sampling rate than other devices designed for use in a fixed environment, especially considering the infrared-based video eye trackers. We chose to run the user study with mobile devices since we aim at designing an algorithm for use in an everyday environment. Although this means the lower sampling rates influence the accuracy of the movements recorded and the quality of the signal, it is still sufficient for context recognition or eye movement gestures.

Stimulus

The visual stimulus was designed to be fast, to collect out-of-the-lab-like movements (very short fixations, fast and brief smooth pursuits, nods and shakes) but still ensure ground truth. The total duration of the stimulus was 21 minutes, with seven identical experimental rounds of three minutes each. The seven rounds ensured that the participants learned the stimulus and were used to its speed after the second round. The first two rounds were thus left out during the data analysis.

Participants were instructed to follow a red point of 0.25° of visual angle presented in front of a light grey background. Each round consisted of a series of horizontal saccades of amplitudes 14, 11, 6, 3 and 1° of visual angle followed by a signal to which participants were instructed to nod during a two seconds interval after each saccade. The nodding records vertical VOR. This was followed by a similar series of vertical saccades that have the same amplitudes as the horizontal, with a signal to shake the head during two seconds after each saccade, to record horizontal VOR. Each saccade was preceded and followed by a 700ms fixation. The stimulus continued with horizontal smooth pursuits of amplitudes 28, 22, 12, 6 and 2° of visual angle at 30°/s, as this speed has been previously noted as the fastest people can follow without having to produce catch-up saccades. Vertical smooth pursuits of similar amplitude followed, with

Figure 1. Experimental setup including (1) the EOG electrodes (the green electrode is the reference), the Dikablis eye tracker (2), and the inertial measurement unit from Xsens (3).
a speed of $15^\circ$/s. Each smooth pursuit ended with a 700ms fixation point at the same location of the end of the smooth pursuit, to compensate for the gain of the movement. Eventually, participants were presented smooth pursuits in circles of diameter of 14, 11, 8, 5 and 3° at 3rad/s.

In addition to providing movements that are as quick as we would expect from an unconstrained and outdoor environment, the stimulus also provides us with data to verify the smallest saccade and smallest smooth pursuit detectable by the two measurement techniques. This is made in order to later assess the limits of the future algorithm to be designed on results from this set of data.

**Procedure**

Participants were placed 60 cm from a 23” (1680x1050 pixels) screen. An array of five EOG electrodes was placed on their face about 45 minutes before the start of the recording. The two horizontal EOG electrodes were placed on both temples, the two vertical ones were placed above and below the left eye (the same eye the Dikablis eye trackers records) and the reference electrode was placed on the top-center of the forehead. Each participant had the stimulus explained to them prior to the experiment. The lighting in the room was dimmed and participants were instructed to not talk or smile during the study to minimise artefacts in the recorded EOG signals. The Dikablis was attached to their head, followed by the cap with the inertial measurement unit attached on top. After calibrating the Dikablis, the study began with a custom 9-point calibration procedure to calibrate the EOG. This calibration was repeated at the end of the experiment. After the study, participants were asked to provide feedback on the comfort of the general setup, the electrodes and the eye tracker in a questionnaire.

**DATA ANALYSIS**

The user study provided us with three different sources of data: The infrared-based eye-tracker, the EOG signals and head movements. As a first step here we only analyse data from the EOG signals in order to get an idea of possible promising features. In the following analysis we also focus only on the main types of movements that are difficult to differentiate from one another: saccades, smooth pursuits and VORs. Blinks and fixations will thus be added to the feature analysis in future work.

**Feature Extraction**

Using the recorded labels as ground truth, we isolated each movement according to its type (saccade, smooth pursuit, VOR) and relevant characteristic (amplitude, duration, speed). The goal of this first analysis was to extract several basic features from these movements and determine which are key to discriminate between these movements. The work presented here is done using 14° horizontal saccades ($n = 358$), 28 and 22° horizontal smooth pursuits ($n = 752$) and all VORs ($n = 1880$). These are the most likely movements to show significant differences since they are the ones with the largest amplitude.

Figure 2 shows an example of the types of eye movements that we collected. As expected due to the lower sampling rate, temporal resolution of the video-based gaze data is lower compared to EOG. Although the onset and offsets of movements differ by several milliseconds, the overall signal characteristics are similar.

From the different types of movements we extracted a first set of 15 features (see Table 1).

We chose to extract features concerning range, velocity and acceleration both from raw EOG data and from EOG data filtered with a 5-point median filter. In addition to being easy and quick to extract from online data, these basic features also allow us to overview the results filtered data shows compared to raw data. Velocity and acceleration were estimated as:

$$v = \frac{d(x)}{dt} f_s \gamma$$  \hspace{1cm} (1)

$$\dot{v} = \frac{d^2(x)}{dt^2} f_s \gamma$$  \hspace{1cm} (2)

where $x$ is the EOG signal recorded, $f_s$ is the sampling frequency (here 128Hz) and $\gamma$ is the conversion factor between
The method was applied to extract the slope, by fitting a linear to the data using the least squares procedure. The same coefficients \( a_3 \) were used it for the VORs of this participant. Having estimated velocity and acceleration is sufficient for the purpose of the classification. We chose kNN because it is lightweight and fast, which are good characteristics for a future online use. First, we shuffled the data and standardized it. The data was then separated into five segments containing each 20% of the saccades, 20% of the smooth pursuits and 20% of the VORs, in order to achieve a 5-fold cross-validation.

The kNN classification is sensitive to redundant features, so we separated the datasets into two subsets that we tested independently, one being with velocity, acceleration and amplitude features from filtered data and the other from raw data.
data. We ran the classification with $k = 1, 3, 5$ and $7$. The precision and recall were calculated for each case, and the results are presented in Figure 4.

Figure 4 shows that filtered data performs generally better than raw data both in term of precision and recall. Filtered data’s results show quite good results, all of them being above 80%. As the number of neighbours $k$ increases, the only change is the classification of VORs gaining precision and losing recall.

**DISCUSSION**

The basic features selected show clusters of movements that overlap, which is an incentive to extract more features. A larger set of features should be able to separate the clusters and categorise movements accurately. With a larger amount of features, we can run a feature selection algorithm such as minimum redundancy-maximum relevance (mRMR) prior to running a classification in order to select the most relevant ones.

The kNN classification we used shows the features we extracted were enough to categorize saccades efficiently. It also indicates that more features are needed to classify VORs and smooth pursuits. Indeed, these will get harder to categorize when we consider a complete dataset, with saccades and pursuits with a smaller amplitude in addition to fixations and blinks. Even though the median filter we used showed good results, future work should include other types of filters as well as different classification methods.

**CONCLUSION**

The algorithm to be based on this work will help developing mobile and pervasive eye tracking applications that need interaction and immediate reaction. In this paper we presented our work in progress towards an online algorithm to detect saccades, smooth pursuits, blinks, fixations and VORs. We described the user study we conducted to collect naturalistic data, how we isolated each movement and demonstrated that basic features extraction shows promising results to discriminate three eye movements.

We showed that saccades appear easy to discriminate with basic features. Smooth pursuits and VORs seem to require further feature extraction and analysis to characterise them, although a simple classification already gave good results. Future work should include: a filtering method comparison, a larger set of features to analyse and select with a feature selection algorithm and results from several classification algorithms. It also implies to run the feature extraction on data collected from the IR-based eye-tracker and the head movement data and extending the analysis to blinks and fixations.

**REFERENCES**


8. D. H. Koh, S. Munikrishne Gowda, and O. Komogortsev. Real time eye movement...


