

BACHELOR THESIS  
ARTIFICIAL INTELLIGENCE

**Radboud University**



---

**Eye movement classification in aid  
of the cVEP speller.**

---

*Author:*  
Dimitar Dimitrov  
s1018291

*First supervisor:*  
dr. P.W.M. Desain  
Artificial Intelligence  
p.desain@donders.ru.nl

*Second supervisor:*  
dr. S. Ahmadi  
Artificial Intelligence  
sara.ahmadi@donders.ru.nl



June 25, 2021

## **Abstract**

All electroencephalography-based Brain-Computer Interfaces have to deal with signal artefacts produced by motor movements, such as jaw muscle clenching and eye movement. Instead of attempting to remove or ignoring these artefacts, recent studies have shown that meaningful information can be extracted from them. Ocular artefacts in particular, possess distinct features depending on the direction of the eye movement they originate from. These features have been used to extract the direction of eye movement using electrodes placed around the eyes, however recent studies have shown that the extraction is also possible with signal from the frontal and temporal electrodes of a standard 10-20 EEG cap. Information about the direction is useful to visual modality BCIs that require classification of simultaneously presented stimuli such as the code-modulated Visual Evoked Potentials speller. The current paper proposes a pipeline that uses ocular artefacts' direction and early stopping in order to increase the speed and subsequently the characters spelled per minute of the code-modulated Visual Evoked Potentials speller.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Research Questions . . . . .	4
<b>2</b>	<b>Preliminaries</b>	<b>5</b>
2.1	Eye Movement Artifacts: The Eye as a Dipole . . . . .	5
2.2	Code-modulated Visual Evoked Potentials Speller . . . . .	7
2.2.1	Dynamic stopping . . . . .	8
<b>3</b>	<b>Related Work</b>	<b>9</b>
<b>4</b>	<b>Methods</b>	<b>10</b>
4.1	Dataset . . . . .	10
4.2	Gaze position estimation . . . . .	10
4.2.1	Pre-processing . . . . .	10
4.2.2	Feature Extraction . . . . .	11
4.2.3	Feature selection . . . . .	13
4.2.4	Gaze estimation . . . . .	13
4.2.5	Validation strategy and exploration of the parameter space . . . . .	14
4.3	Modified cVEP Speller . . . . .	14
4.3.1	Evaluation . . . . .	14
4.4	Software . . . . .	15
<b>5</b>	<b>Results</b>	<b>16</b>
5.1	Gaze estimation . . . . .	16
5.1.1	Parameters . . . . .	16
5.1.2	Obtained accuracy . . . . .	16
5.2	cVEP speller . . . . .	17
5.2.1	Parameters . . . . .	17
5.2.2	cVEP Speller runs . . . . .	17
<b>6</b>	<b>Conclusions</b>	<b>20</b>
6.1	Research questions . . . . .	20

<b>7</b>	<b>Discussion and future work</b>	<b>21</b>
7.1	Dataset . . . . .	21
7.2	Additional directions . . . . .	21
7.3	Improved Saccade Classification . . . . .	21
7.4	Continuous Gaze Tracking . . . . .	22
7.5	EOG-based Gaze Tracking . . . . .	22
7.6	Language Model Prior . . . . .	22
<b>A</b>	<b>Appendix</b>	<b>26</b>

# Chapter 1

## Introduction

Electroencephalography (EEG) is a non-invasive method for recording the brain's electrical activity, using electrodes placed on the scalp. Every Brain-Computer Interface (BCI) that uses EEG to infer the user's intent has to deal with one common problem - signal artefacts. These artefacts stem from other potential-producing activities such as motor movements [11]. Typically BCIs would attempt to filter out those artefacts from the data during the pre-processing phase using methods such as Independent Component Analysis [10]. Another way to deal with eye movement-related artefacts is to use electrooculography (EOG) - a recording of ocular electrical activity using electrodes placed on the face, near the eyes. The recorded EOG signal is then linearly subtracted from the EEG signal to produce cleaner EEG recordings [17]. Studies in the past have shown that instead of simply removing the artefacts, meaningful information can be extracted from EOG, as the artefacts originating from a specific type of rapid eye movements, called saccades, exhibit specific shapes depending on the direction of the eye movement they stem from [1, 6, 13]. By using the directional information from EOG, gaze estimation becomes possible without the need for a camera-based eye-tracker, which has the requirement of fixed positioning relative to the head of the user, and as such introduce a level of discomfort. A BCI that depends on the visual modality can benefit from information about the user's gaze position. One such BCI is the code-modulated Visual Evoked Potentials (cVEP) Visual Speller BCI [21]. It presents 36 characters in a 6 by 6 grid, each cell in the grid starts flashing in a particular pattern, with as little as possible correlation to other patterns. As the user focuses on a particular character, the flashing pattern of that character causes a broadband visual evoked potential (BBVEP) which is picked up by EEG. As data starts coming in, the BCI calculates a correlation coefficient for each flashing pattern, representing how certain the BCI is that the user is focusing on that pattern. When one of the coefficients reach a certain threshold, a classification is made. If such BCIs can use the directional shift

to estimate a gaze position, the correlation coefficients of patterns within the user's estimated gaze can be boosted, so that they reach the classification threshold faster. The downside is that EOG electrodes also introduce a certain level of discomfort - they require time to be placed on the face, sit close to the eyes which obstructs a small part of the visual field and might leave temporary marks after removal. Recent studies have shown that the same eye movement direction information can be extracted from frontal and temporal electrodes of the 10-20 EEG standard [3–5].

## 1.1 Research Questions

1. Is it possible to use the spelled characters of the cVEP speller to optimize a gaze estimation algorithm based on saccade classification from EEG signal?
2. Is it beneficial to use said gaze estimation algorithm to pass gaze position as a prior information to the cVEP speller, in order to help it reach a dynamic-stopping threshold earlier for increased characters per minute spelled?

In the following chapters I will first expand on the nature of eye movement artefacts and explain the early-stopping cVEP Speller. After that I will describe the methods used to build a pipeline that attempts to answer the research question. Finally I will discuss the obtained results, the limitations of this research and possible future research.

## Chapter 2

# Preliminaries

### 2.1 Eye Movement Artifacts: The Eye as a Dipole

Due to the concentration of neurons in the retina, the back of the eye exhibits a negative electric charge, while the cornea on the opposite end exhibits positive charge[14, 16]. As such the eye can be seen as a corneo-retinal dipole that produces offsets in signals recorded near it. For example in EOG each eye movement produces an offset the polarity of which depends on the direction of the eye movement as seen in Figure 2.1.

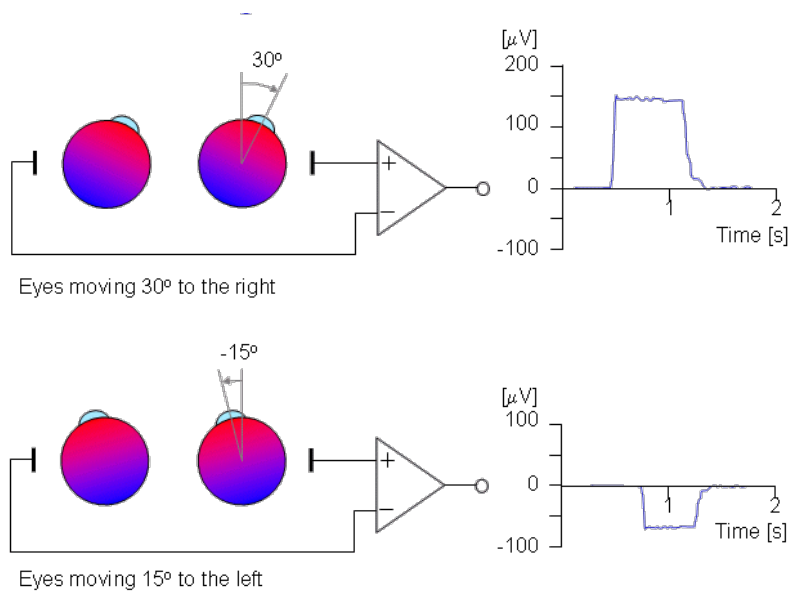
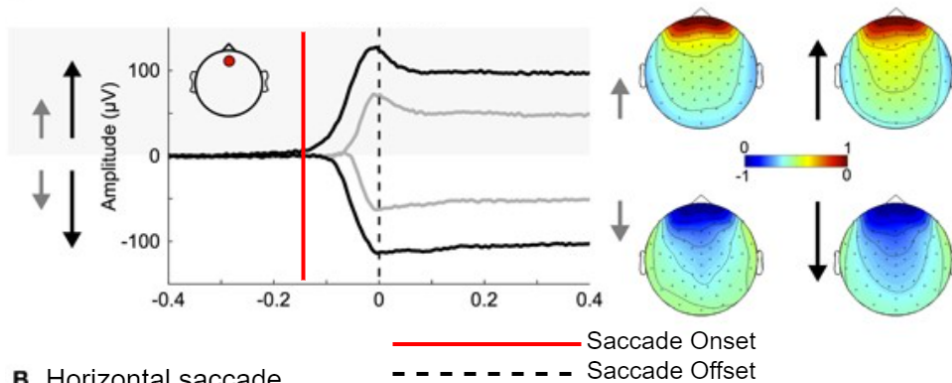


Figure 2.1: The eye as a dipole [14]

There are a few different types of eye movements, but in this paper we are considering only saccades - quick simultaneous movements of both eyes between 2 phases of fixation. Due to the speed and distance travelled

during saccadic movements, signals produced by them are much stronger than those emitted by cortical activity. As such they can be observed clearly from EEG [Figure 2.2]. Other types of eye movements are either too small to measure consistently from EEG (fixations/micro-saccades) or too slow to be detectable with the proposed pipeline (smooth pursuits). That is not a problem, as the task at hand requires the shifting of the gaze position within a grid. This is essentially a visual search task and during a visual search saccades are the type of ocular movement utilized to scan the visual field. While saccades are the thing we can detect, we are actually more interested

### A Vertical saccade



### B Horizontal saccade

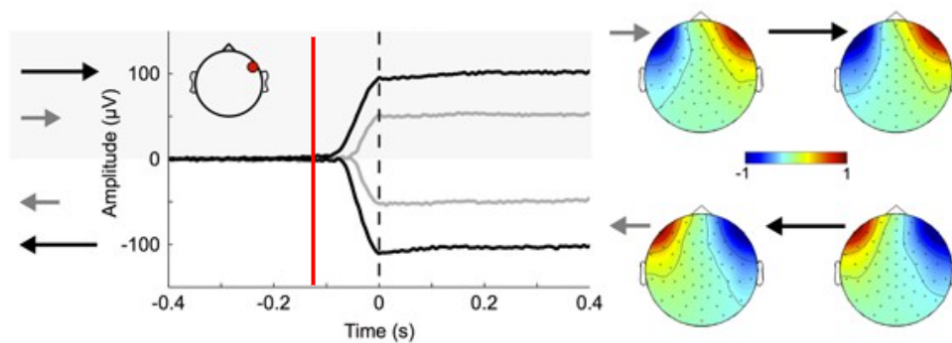


Figure 2.2: Saccadic potentials observed from EEG signal [16]

in the position of the gaze. As the electrical charge of the eye is constant, the offset produced while the eye is rotated away from the neutral position is constant, until the eye returns back to looking forward[20]. This can be observed by the illustrations on the left side of Figure 2.2 where after a saccade the amplitude stays at the offset position.

## 2.2 Code-modulated Visual Evoked Potentials Speller

The cVEP Visual Speller [21] is a BCI that allows a user to input letters using only their gaze. The cVEP speller presents a 6 by 6 table with a single character in each cell. During a trial each cell flashes with a particular pattern, called a code. These codes are optimized so that they are orthogonal to each other, meaning their correlation is minimized. Upon focusing their gaze on a character, the user's visual cortex starts exhibiting broadband visual evoked potentials (BBVEPs) corresponding to the code behind that character. Reconvolution is used to create templates that predict the EEG response to each code, based on design matrices describing long and short flashes and their corresponding signal responses [Figure 2.3]. A correlation coefficient is calculated for each template based on how well it matches the observed signal. When a criterion, calculated from these correlation coefficients, reaches a predefined confidence threshold, the classification is made.

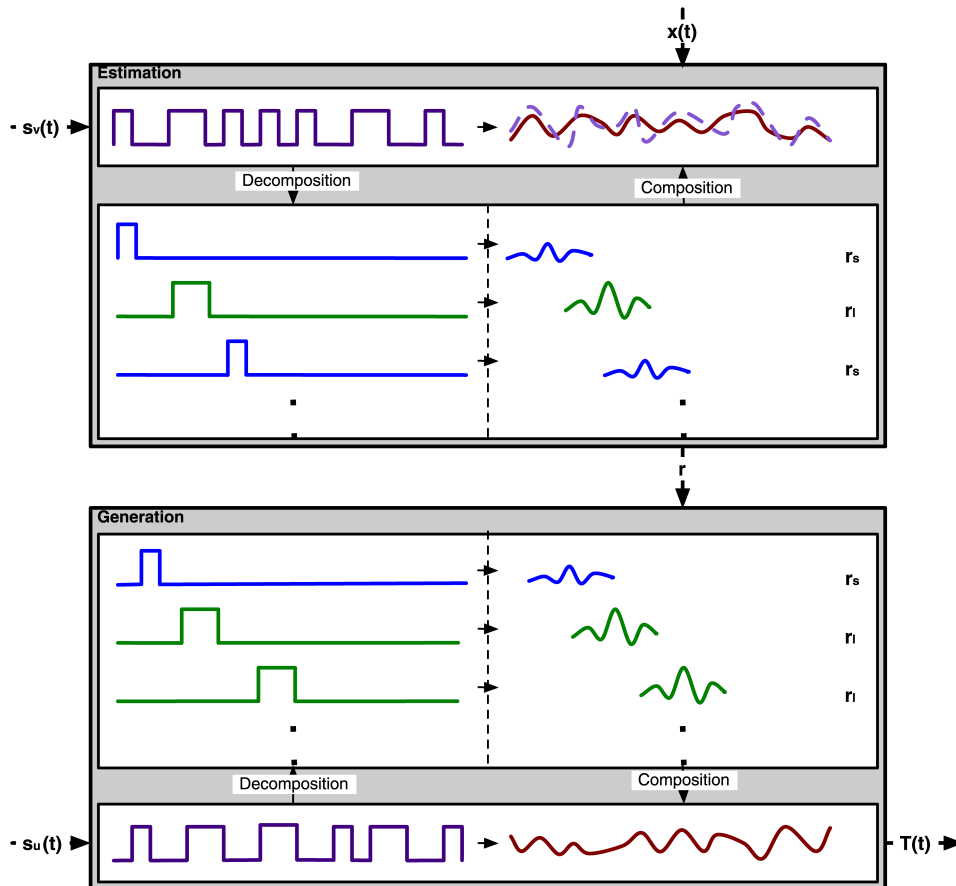


Figure 2.3: Reconvolution, Thielen et al [21]

### 2.2.1 Dynamic stopping

BCIs can choose between one of 2 paradigms - static and dynamic stopping. Static stopping entails the collection of a set amount of data, i.e. fixed trial duration, and making a classification based on all the data. Dynamic stopping in contrast is the practice of re-computing a criterion, based on which a decision is made, as new data comes in. This allows for the stopping of a trial earlier than the preset trial length, if the classifier becomes certain enough about the predicted label. Additionally if the end of the preset trial time is reached, a classification can be forced based on the label with highest confidence, even if it doesn't reach a preset confidence threshold. The early stopping of a trial results in higher Information Transfer Rate. There are a few ways to apply dynamic stopping[7, 18] in BCIs, but the implementation in Thielen et al [22] in particular is used as baseline in the current paper. After calculating the correlation coefficients of each flashing code, it takes all coefficients, except the largest one, and fits a Beta distribution with them. Then the probability that the largest coefficient is larger than the expected maximum of the Beta distribution is calculated and if it exceeds a set confidence threshold (e.g. 99%) a classification is made.

## Chapter 3

# Related Work

Multiple classification pipelines have been proposed for saccade direction classification from EOG signal, using various classification approaches, such as Neural Networks [2] and Hierarchical Models [15]. In contrast with those models, previously developed for EOG, Belkacem et al claims that significant patterns are more important than using a strong machine-learning based classifier. The classifier proposed in their papers[3–5] is relatively simple hierarchical classifier, depending on thresholds that can be tuned by hand.

Recently a paper was published, describing a deep neural network approach for classifying saccades to the left and to the right, using EEG data[12]. It describes aspirations for creating a fully EEG-based eye tracker, which, as that author suggests, will be useful for BCIs depending on the visual modality. The Deep Neural Network approach, unlike Belkacem et al’s work, does not require any specific feature extraction and feature selection.

Machine learning approaches such as deep neural networks require an extensive dataset of properly labelled saccade data to be trained on in order to perform well. As such dataset was not available, Belkacem et al’s approaches were chosen for saccade extraction part of the gaze estimation algorithm.

# Chapter 4

## Methods

### 4.1 Dataset

The dataset provided by Thielen et al [21] was used both for validation of the eye movement classification, as well as the overall proposed pipeline. It contains data collected from 12 subjects. 3 blocks of 36 trials were considered for each subject. The blocks contain a copy-spelling task, which means a symbol was highlighted, after which the user had to focus their gaze on it, in order for the cVEP speller to classify it as input. There is an inter-trial time of 2 seconds and each trial lasts 4.2 seconds.

### 4.2 Gaze position estimation

As this paper was inspired by the work of Belkacem et al, we attempted to follow closely the preprocessing, feature extraction and classification steps described in the papers [3–5]. Some modifications were made to adjust the pipeline according to the datasets available at the time of conducting the research, as well as the goal - predicting which half of the screen the user is focusing on. The data considered was the inter-trial time before a trial as well as the first half a second after stimulation starts, i.e. from  $t-2$  until  $t+0.5$ .

#### 4.2.1 Pre-processing

Signals from channels F7 and F8 from the standard 10-20 EEG arrangement were considered, due to their physical proximity to the eyes and strong prevalence of ocular artefacts [3]. The signals were bandpass filtered in the range [0.5Hz - 100Hz] using a 4th order Butterworth filter to remove irrelevant high frequency components and bandstop (notch) filtered in the range [48Hz - 52Hz] to remove DC power line noise. In order to replicate Belkacem et al as closely as possible, data was resampled at 256Hz.

## 4.2.2 Feature Extraction

### Enhancing artifacts from horizontal eye movements

During an eye movement on the horizontal axis, the left and right side of the scalp experience opposite polarities [Figure 2.2], while a movement on the vertical axis, including blinks, produce a positive or negative spike across both sides of the scalp. This is exploited by subtracting the signal from the right side of the head from the signal from the left side of the head. By subtracting the right side from the left side, horizontal movement artefacts are enhanced while vertical movement artefacts are removed [Figure 4.1]

$$Y = F7 - F8 = X_{left} - X_{right} \quad (4.1)$$

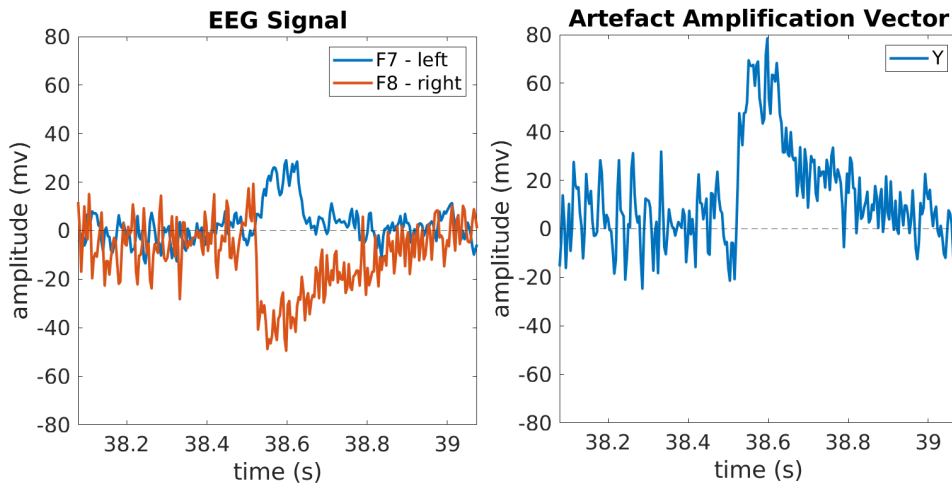


Figure 4.1: 1 second sample of EEG data (left) and the amplified artefact vector (right). The artefact is a saccade to the left.

The amplified horizontal artefact vector is then smoothed using a low-pass Butterworth filter with cutoff frequency of 10Hz, in order to remove spontaneous blink artefacts [4, 5].

### Sliding window

Once the enhanced horizontal artefact vector was calculated a 1 second sliding window was utilized to scan over it and extract relevant features, while simulating an online setup. The sliding window had temporal difference of 100ms, which coincided with the temporal amount of data delivered to the cVEP speller. This translated in each window being 256 samples long and each following window getting 26 new samples.

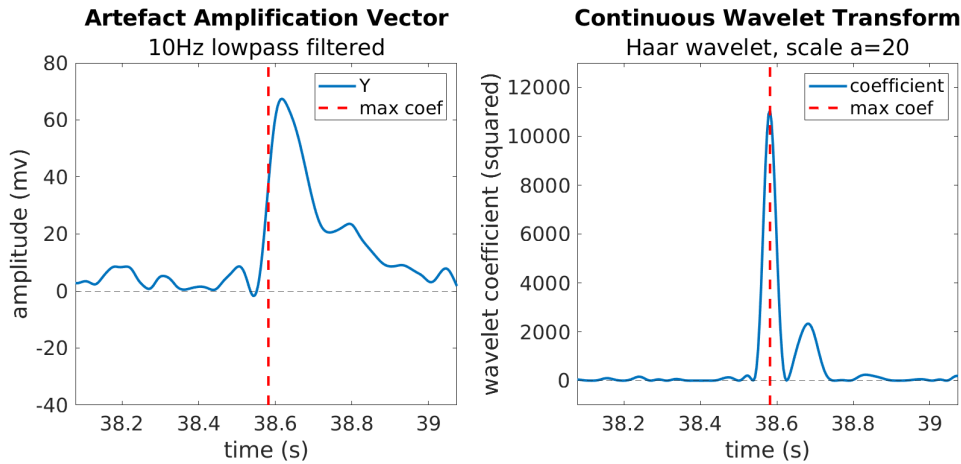


Figure 4.2: Left: 1 second window of the lowpass filtered amplified artefact vector containing a saccade to the left. Right: the squared coefficients obtained from Continuous Wavelet Transform applied on the same window, using the Haar Wavelet and scale parameter  $a = 20$ .

### Continuous Wavelet Transform

For each window a Continuous Wavelet Transform (CWT) using the Haar wavelet as mother wavelet was applied with a relatively large parameter  $a=20$ . The Haar wavelet was chosen due to saccades being dramatic offsets akin to a step function, just like the Haar wavelet [8]. The relatively large scale parameter was chosen in order to detect only larger offsets in the data [9]. The output of the continuous wavelet transform is a coefficients vector that shows how much the mother wavelet contributes to the signal, given the scaling factor. As such higher coefficients value corresponds a larger offset in the signal, and therefore a saccade occurring. The coefficient values were squared to ignore the sign and enhance the value. For each window the largest squared wavelet coefficient's value and temporal location is recorded.

### Event window

The largest wavelet coefficient's position is used to center a smaller event window around it [Figure 4.3 Left], with length of 200ms. The length of the event window was chosen based on the possible length of a saccade - spanning from 20 to 200ms. The surface under the curve described by the signal in the event window is calculated using the trapezoidal method [Figure 4.3 Right] and recorded at the same temporal location as the highest wavelet coefficient. The assumption is made that a negative trapezoidal area score is associated with a saccade that keeps the eye on the left side of the neutral

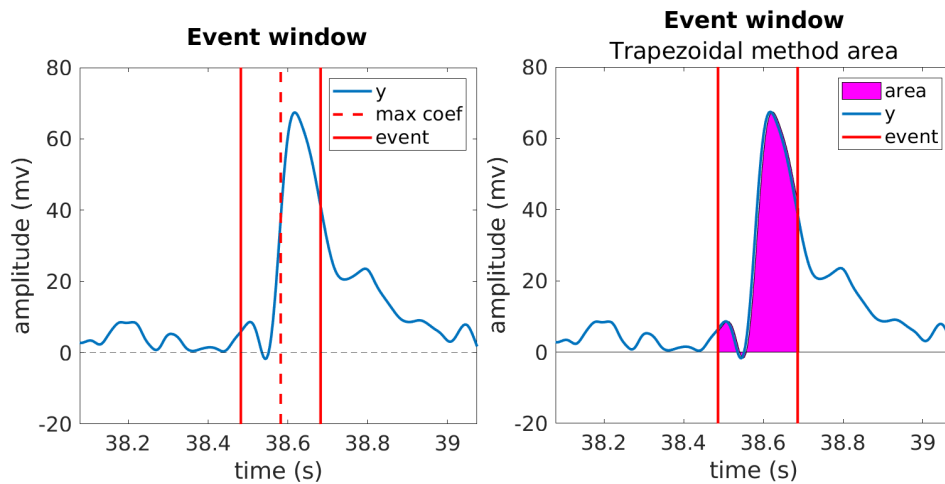


Figure 4.3: Left: 200ms event window, centered around the highest wavelet coefficient found.

Right: the area under the curve, calculated using the trapezoidal method.

position, and a negative trapezoidal area score is associated with a saccade that keeps the eye on the right side of the neutral position.

### 4.2.3 Feature selection

As described in the previous section, the two features collected were the temporal location of each highest wavelet coefficient and the trapezoidal area under the signal curve around it. Since one highest wavelet coefficient is extracted per time window, this doesn't guarantee that every recorded feature is an actual event. That's why a threshold is established, above which an ocular event occurs. Additionally another, higher threshold was found necessary, above which the ocular event is considered an outlier and thus discarded. Trapezoidal areas associated with wavelet coefficients that pass the 2 threshold criteria are collected and summed for each trial.

### 4.2.4 Gaze estimation

The estimation of the gaze position depends on the sum of the trapezoidal scores at the beginning of a trial ( $t-2$  until  $t+0.5$ ). It is assumed that larger saccades in one direction mean a shift of gaze towards that side of the screen. As such the linear sum of the trapezoidal scores would represent the final gaze position. If the sum is positive, the gaze is predicted to be on the left side of the screen. If the sum is negative, the gaze is predicted to be on the right side of the screen.

### 4.2.5 Validation strategy and exploration of the parameter space

The lack of properly labelled saccade data means that there wasn't a way to validate the saccade classification by other means except manual observation. The goal was a more general gaze position estimation pushed us to validate the gaze estimator according to the true label of a trial, instead of seeking validation of each saccade. E.g. if that label resided within the left half of the speller's grid, then the prediction we're looking for is an overall shift to the left, represented by a positive trapezoidal score sum. The optimal thresholds required to identify an ocular event were found by exploring the parameter space and calculating the mean accuracy of the saccade classifier across all blocks from all subjects.

## 4.3 Modified cVEP Speller

An offline simulation of dynamic stopping online experiment was ran using the Dataset and Toolbox provided in Thielen et al [21, 22] to obtain a baseline. Afterwards the function calculating the confidence from the correlation coefficients was modified to accept a prior probability mass array, multiply the prior with the outcome of the confidence coefficients. The prior was generated based on the outcome of the gaze estimator, by either being uniform if no saccades were detected, or biasing towards codes presented in either the left or right half of the speller's interface. The bias size is determined by a preset boost factor parameter. The cVEP speller only considered trial data upon start of stimulus presentation, i.e. inter-trial data was discarded. A minimum amount of 0.5 seconds of data was required before a classification can be made and also a classification was forced at 4.2 seconds (end of the trial), regardless of whether the stopping criterion had been reached.

### 4.3.1 Evaluation

Information Transfer Rate[23] (ITR) is a popular metric for evaluating BCIs. It is appealing due to its derivation from Information Theory and the fact it combines speed and accuracy in a single measure[19]. ITR makes the assumption that each class is equally likely, i.e. the probability is uniformly distributed across all possible classes. As our pipeline proposes a method of prior information extraction use during classification, this assumption is no longer valid. During classification half of the classes have an increase in probability compared to the other half. This is why we settled on a different metric - characters spelled per minute (CPM), which still takes into account accuracy and time but no longer uses the number of classes and therefore makes no assumptions about their uniformity. Mathematically CPM can be

expressed as:

$$CPM = \frac{60}{T} \cdot \max(0, (P - (1 - P))) \quad (4.2)$$

Where T is the average trial time and P is the achieved accuracy. Note that when the accuracy is below 50%, the CPM is equal to 0, as this accuracy implies the user will have to constantly correct miss-classifications.

## 4.4 Software

The analysis was done using MATLAB R2020b, using the Signal Processing toolbox and Wavelet toolbox, and FieldTrip daily release ca. April 2021. The cVEP speller's simulation was done using the MATLAB routines provided in Thielen et al [21]. The MATLAB routines produced by the current author according to the methods section can be found on GitHub ([https://github.com/mechachki/oa\\_cvep](https://github.com/mechachki/oa_cvep)).

# Chapter 5

## Results

### 5.1 Gaze estimation

#### 5.1.1 Parameters

The parameters used by the gaze estimation algorithm were set to the following values:

- sliding window size of 1 second (= 256 samples at sampling rate of 256Hz)
- data delivered in 0.1 second chunks ( $\approx$  26 new samples at sampling rate of 256Hz)
- Continuous Wavelet Transform:
  - Haar mother wavelet
  - scale parameter  $a = 20$
- minimal squared wavelet coefficient = 1300
- maximal squared wavelet coefficient = 3200

#### 5.1.2 Obtained accuracy

By optimized the values for the squared wavelet coefficient thresholds, using the desired true labels, a mean accuracy of 52.97% was obtained across all subjects. As seen in Figure 5.1, most of the time the gaze estimation is unfortunately chance-level with this pipeline. After experimentation with different threshold values, it was established that the non-stationary nature of EEG signal across users makes the determination of universal thresholds hard, as it just ends up trading accuracy between subjects. A few more factors influencing the signal will be discussed later as possible reasons for this poor performance, the most important being lack of ground truth.

subject	Block 1	Block 2	Block 3	mean
1	50%	66.67%	41.67%	52.78%
2	44.83%	53.85%	44%	47.58%
3	72.22%	46.15%	55.56%	57.98%
4	42.86%	41.67%	44.44%	42.99%
5	62.5%	58.33%	34.62%	51.82%
6	33.33%	48.39%	65.52%	49.08%
7	47.22%	50%	61.11%	52.78%
8	46.67%	70.37%	44.44%	53.83%
9	47.83%	64.52%	58.06%	56.80%
10	82.35%	60%	71.43%	71.26%
11	55.17%	61.54%	39.29%	52%
12	35.71%	42.11%	62.5%	46.77%
mean	51.72%	55.3%	51.89%	52.97%

Figure 5.1: Gaze estimation accuracy per subject per block, validated with true label position

## 5.2 cVEP speller

### 5.2.1 Parameters

The parameters used by the cVEP speller were set to the following values:

- target confidence = 0.99
- minimum trial time = 0.5 seconds
- data delivered in 0.1 second chunks (= 36 new samples at sampling rate of 360Hz)
- gaze prior boost factor = 0.01 (only used by modified speller)

### 5.2.2 cVEP Speller runs

In order to test the performance of the pipeline the cVEP speller experiment was simulated 3 times:

1. Original setup from Thielen et al - cVEP speller with Beta early stopping. Dubbed OG.
2. Modified setup with estimated gaze prior - cVEP speller with Beta early stopping confidence coefficients boosted according to the predicted gaze position. Dubbed MP.

3. Modified setup with idealized gaze prior - cVEP speller with Beta early stopping confidence coefficients boosted according to an idealized gaze position prediction, meaning that instead of using our gaze estimation pipeline, the true labels were used to always have a correct prediction of the screen half that the user is looking at. Dubbed MI.

### Accuracy

Unfortunately even with idealized labels, the modified speller did not perform as well as the original speller [Figure 5.2] in terms of accuracy. The reason for this is speculated to be the fact that there is still a certain level of correlation between different codes. While the layout was optimized in Thielen et al, such that nearby codes have minimized correlation, it is still possible that codes further away but on the same half of the screen could trigger false positives when signal-to-noise ratio is poor and their confidence is boosted by the gaze estimator.

### Speed

The trial time was improved significantly for both the gaze estimator pipeline ( $p < 0.001$ ) and the idealized gaze estimation pipeline ( $p < 0.001$ ) compared to the original Beta early-stopping cVEP speller.

### Characters per minute

The CPM showed a significant improvement, when comparing the original Beta speller with the idealized gaze speller ( $p < 0.001$ ). This wasn't the case when comparing it to the gaze estimator speller ( $p = 0.8337$ ). As such the current pipeline cannot be seen as a necessary improvement over the original cVEP speller, unless a better performing gaze estimation algorithm can be curated.

	OG	MI	MP
mean accuracy	84.3364%	81.4815%	81.0185%
mean trial time	3.30424s	2.37299s	2.86968s
mean CPM	13.9573 ch/min	20.0041 ch/min	14.9912 ch/min

Figure 5.2: Statistics collected from the three runs of the cVEP speller.

A more detailed look at per subject statistics can be found in the appendix, accuracy in Figure A.1, average trial time in Figure A.2 and characters per minute in Figure A.3.

### **Additional Insights**

It is interesting to observe what are the label differences between the original Beta stopping speller and the modified Beta with Gaze prior speller. In all 1296 trials across all subjects and all blocks:

- 7 trials were corrected by the idealized gaze speller and 4 by the predicted gaze speller, meaning that they were miss-classifications in the original Beta speller but correct classifications in the modified spellers.
- 44 trials were corrupted by the idealized gaze speller and 47 by the predicted gaze speller, meaning that they were correct classifications in the original Beta speller but miss-classifications in the modified spellers.

This could be interpreted as the pipeline requiring a more specific gaze estimation (e.g. a quadrant of a quarter of the speller grid), as the current one allows for more errors to happen than corrections.

## Chapter 6

# Conclusions

This project shows that there can be a benefit to the characters spelled per minute when additional information can be used to speed up classification. The approach to use the existing EEG measurement channels proved possible for detecting ocular movements, that can be used as a prior for the cVEP Speller. The proposed classification pipeline however is not robust enough to guarantee consistent performance and the lack of consideration for head movements in the experimental design of the dataset at hand makes it impossible to obtain a ground truth for the saccade classification's and gaze estimator's performance.

### 6.1 Research questions

1. With the present dataset it is not possible to sufficiently train the proposed gaze estimation algorithm.
2. It is beneficial to pass gaze position information to the cVEP speller, in order to improve the characters spelled per minute.

## Chapter 7

# Discussion and future work

### 7.1 Dataset

The main problem with the current work is the used dataset. It was collected with intentions of testing the cVEP speller alone, and as such does not contain ground truth about the phenomena we seek to find and classify. This means we don't know the true labels of any given saccade we find and any validation had to be done by hand, according to our knowledge of saccades and intuition. This makes it hard to train and validate a detected saccade's direction and the gaze's overall estimation. As such future work should attempt to create a new dataset that includes VOG eye-tracker recordings for eye movement ground truth and accelerometer data for head movement ground truth, alongside EEG. As most modern EEG headsets incorporate an accelerometer already, it is sensible to include that information in the pipeline.

### 7.2 Additional directions

The limitations of the current dataset pushed us to only work with the horizontal axis, however this should not limit research when a better dataset becomes available. Vertical axis movements have been shown to also be detectable from EOG and EEG ocular artefacts. By combining the direction and size of a saccade relative to both axes, more complex movements such as diagonal saccades could be classified as well.

### 7.3 Improved Saccade Classification

With a large enough dataset that includes properly labelled ocular movements the training and deployment of a deep neural network for robust saccade classification becomes possible.

## 7.4 Continuous Gaze Tracking

With a dataset better suited for the task at hand a continuous EEG-based gaze tracker could be attempted. By using EEG to detect saccades and accelerometer data to detect head movements, it should be possible to estimate the exact position of the user's gaze at any moment. In such a pipeline the last spelled letter can be incorporated as an intermediate ground truth.

## 7.5 EOG-based Gaze Tracking

As mentioned before, EOG is still better than VOG in terms of user friendliness, due to the lack of need for a camera in front of the user, which can be affected by shift of the relative position of the head and various environment conditions. As such the same pipeline idea could be attempted with EOG-based ocular movement classification. The advantage of such an approach would be that EOG has stronger representation of ocular movements and has been well studied with various pipelines available that can detect more complex movements such as clockwise and counterclockwise saccades.

## 7.6 Language Model Prior

Since the speller is inherently tied to the production of language, a language model could be added as an additional prior in the classification pipeline. By observing previously spelled letters a probability distribution can be constructed of the next letter, used to make a decision earlier.

# Bibliography

- [1] Siriwadee Aungsakun et al. “Robust Eye Movement Recognition Using EOG Signal for Human-Computer Interface”. In: *Software Engineering and Computer Systems*. Ed. by Jasni Mohamad Zain, Wan Maseri bt Wan Mohd, and Eyas El-Qawasmeh. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 714–723. ISBN: 978-3-642-22191-0. DOI: 10.1007/978-3-642-22191-0\_63.
- [2] R. Barea et al. “EOG-based eye movements codification for human computer interaction”. In: *Expert Systems with Applications* 39.3 (2012), pp. 2677–2683. ISSN: 0957-4174. DOI: <https://doi.org/10.1016/j.eswa.2011.08.123>. URL: <https://www.sciencedirect.com/science/article/pii/S0957417411012541>.
- [3] Abdelkader Nasreddine Belkacem et al. “Classification of Four Eye Directions from EEG Signals for Eye-Movement-Based Communication Systems”. In: *Journal of Medical and Biological Engineering* 34 (2014), pp. 581–588.
- [4] Abdelkader Nasreddine Belkacem et al. “Online classification algorithm for eye-movement-based communication systems using two temporal EEG sensors”. In: *Biomedical Signal Processing and Control* 16 (2015). 40, pp. 40–47. ISSN: 1746-8094. DOI: 10.1016/j.bspc.2014.10.005. URL: <https://doi.org/10.1016/j.bspc.2014.10.005>.
- [5] Abdelkader Nasreddine Belkacem et al. “Real-Time Control of a Video Game Using Eye Movements and Two Temporal EEG Sensors.” In: *Computational intelligence and neuroscience 2015* (2015), p. 653639. ISSN: 1687-5265. DOI: 10.1155/2015/653639. URL: <https://doi.org/10.1155/2015/653639>.
- [6] DP Belov et al. “Electrooculogram detection of eye movements on gaze displacement”. In: *Neuroscience and behavioral physiology* 40.5 (2010), pp. 583–591. DOI: 10.1007/s11055-010-9299-z.
- [7] Luigi Bianchi, Chiara Liti, and Veronica Piccialli. “A new early stopping method for p300 spellers”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27.8 (2019), pp. 1635–1643.

- [8] Andreas Bulling et al. “Eye Movement Analysis for Activity Recognition Using Electrooculography”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33.4 (2011), pp. 741–753. DOI: 10.1109/TPAMI.2010.86.
- [9] Alessandro Cafasso and Sebastian Karlsson. “Automatic detection of saccadic eye movements using EOG for analysing effects of cognitive distraction during driving”. MA thesis. 2017.
- [10] Rodney J Croft and Robert J Barry. “Removal of ocular artifact from the EEG: a review”. In: *Neurophysiologie Clinique/Clinical Neurophysiology* 30.1 (2000), pp. 5–19. DOI: 10.1007/978-3-540-89208-3\_300.
- [11] Suguru Kanoga and Yasue Mitsukura. “Review of artifact rejection methods for electroencephalographic systems”. In: *Electroencephalography* 69 (2017), pp. 69–89.
- [12] Ard Kastrati, Martyna Płomecka, and Nicolas Langer. “Using Deep Learning to Classify Saccade Direction from Brain Activity”. In: (May 2021). DOI: 10.1145/3448018.3458014.
- [13] D. Kumar and E. Poole. “Classification of EOG for human computer interface”. In: *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society [Engineering in Medicine and Biology]*. Vol. 1. 2002, 64–67 vol.1. DOI: 10.1109/IEMBS.2002.1134390.
- [14] Jaakko Malmivuo and Robert Plonsey. “Bioelectromagnetism”. In: Jan. 1995. Chap. 28. ISBN: 978-0195058239.
- [15] Pornchai Phukpattaranont et al. “Efficient feature for classification of eye movements using electrooculography signals”. In: *Thermal Science* 20.suppl. 2 (2016), pp. 563–572.
- [16] Michael Plöchl, José Ossandón, and Peter König. “Combining EEG and eye tracking: identification, characterization, and correction of eye movement artifacts in electroencephalographic data”. In: *Frontiers in Human Neuroscience* 6 (2012), p. 278. ISSN: 1662-5161. DOI: 10.3389/fnhum.2012.00278. URL: <https://www.frontiersin.org/article/10.3389/fnhum.2012.00278>.
- [17] Alois Schlögl et al. “A fully automated correction method of EOG artifacts in EEG recordings”. In: *Clinical neurophysiology* 118.1 (2007), pp. 98–104.
- [18] Martijn Schreuder et al. “Optimizing event-related potential based brain–computer interfaces: a systematic evaluation of dynamic stopping methods”. In: *Journal of Neural Engineering* 10.3 (May 2013), p. 036025. DOI: 10.1088/1741-2560/10/3/036025. URL: <https://doi.org/10.1088/1741-2560/10/3/036025>.

- [19] William Speier, Corey Arnold, and Nader Pouratian. “Evaluating True BCI Communication Rate through Mutual Information and Language Models”. In: *PloS one* 8 (2013). DOI: 10 . 1371 / journal . pone . 0078432.
- [20] Yasuo Terao et al. “Recording horizontal saccade performances accurately in neurological patients using electro-oculogram”. In: *Journal of visualized experiments: JoVE* 133 (2018).
- [21] Jordy Thielen et al. “Broad-Band visually evoked potentials: re (con) volution in brain-computer interfacing”. In: *PloS one* 10.7 (2015), e0133797.
- [22] Jordy Thielen et al. “Re (con) volution: accurate response prediction for broad-band evoked potentials-based brain computer interfaces”. In: *Brain-Computer Interface Research*. Springer, 2017, pp. 35–42.
- [23] Jonathan R Wolpaw et al. “Brain–computer interfaces for communication and control”. In: *Clinical neurophysiology* 113.6 (2002), pp. 767–791.

## Appendix A

# Appendix

subject	OG acc	MI acc	OG-MI ttest	MP acc	OG-MP ttest
1	97.2222%	93.5185%	0.184	95.3704%	0.423
2	92.5926%	91.6667%	0.423	89.8148%	0.225
3	77.7778%	74.0741%	0.270	75.0000%	0.225
4	94.4444%	89.8148%	0.199	91.6667%	0.225
5	97.2222%	91.6667%	0.184	91.6667%	0.074
6	65.7407%	64.8148%	0.423	62.9630%	0.423
7	49.0741%	48.1481%	0.808	47.2222%	0.529
8	92.5926%	92.5926%	NaN*	87.9630%	0.130
9	91.6667%	87.9630%	0.270	87.9630%	0.184
10	87.0370%	80.5556%	0.073	82.4074%	0.038
11	86.1111%	84.2593%	0.184	83.3333%	0.000
12	80.5556%	78.7037%	0.423	76.8519%	0.270

Figure A.1: Accuracy per cVEP run, average of 3 blocks per subject. acc means accuracy. ttest columns show the p-value produced by a paired two-tailed t-test. \*NaN means there was no difference between the two populations.

subject	OG time	MI time	OG-MI ttest	MP time	OG-MP ttest
1	3.0287s	1.9056s	0.0000	2.6537s	0.0000
2	1.7472s	1.0037s	0.0000	1.4500s	0.0000
3	3.8491s	3.0093s	0.0000	3.3778s	0.0000
4	2.7556s	1.5500s	0.0000	2.4000s	0.0000
5	3.1954s	2.0000s	0.0000	2.6667s	0.0000
6	3.8991s	3.3991s	0.0000	3.6037s	0.0003
7	4.1528s	3.4815s	0.0000	3.8074s	0.0003
8	3.4694s	2.3972s	0.0000	2.9204s	0.0000
9	3.0185s	2.0204s	0.0000	2.5167s	0.0000
10	3.7741s	2.7806s	0.0000	3.0500s	0.0000
11	2.9963s	1.8620s	0.0000	2.4685s	0.0000
12	3.7648s	3.0667s	0.0000	3.5213s	0.0022

Figure A.2: Trial time per cVEP run, average of 3 blocks per subject, excluding the 2 seconds of inter-trial time. ttest columns show the p-value produced by a paired two-tailed t-test. 0 means  $p < 0.0001$

subject	OG cpm	MI cpm	OG-MI ttest	MP cpm	OG-MP ttest
1	18.7746	27.6677	0.0131	20.5280	0.1718
2	29.2466	49.8358	0.0076	33.0213	0.0571
3	8.7300	9.7227	0.1034	8.9492	0.6923
4	19.3574	30.8152	0.0274	20.8622	0.1624
5	17.7348	25.0473	0.0592	18.7463	0.1312
6	4.8638	5.2682	0.4605	4.2849	0.5937
7	1.0582	0.6661	0.4226	0.8811	0.4226
8	14.8068	21.3318	0.0082	15.7193	0.3962
9	16.8683	22.9405	0.0571	18.7072	0.4070
10	11.8102	13.3280	0.3545	12.8433	0.3060
11	14.4801	22.1527	0.0086	16.2317	0.0933
12	9.7574	11.2732	0.0464	9.1204	0.4191

Figure A.3: Characters per minute per cVEP run, average of 3 blocks per subject, excluding the 2 seconds of inter-trial time. CPM columns contain measurements in characters per minute. ttest columns show the p-value produced by a paired two-tailed t-test .