



A supervised machine-learning method for detecting steady-state visually evoked potentials for use in brain computer interfaces: A comparative assessment

By

Kieran Eamon Duggan
DGGKIE001

SUBMITTED TO THE UNIVERSITY OF CAPE TOWN
In fulfilment of the requirements for the degree

MSc (Med) Biomedical Engineering

Faculty of Health Sciences

UNIVERSITY OF CAPE TOWN

Date of Submission: **July 2017**

Supervisors: **Prof Ernesta Meintjes**

Co-supervisor: **Dr Kylie de Jager and Dr Lester John**

Department of Human Biology

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

DECLARATION

I, Kieran Duggan, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signed by candidate

Signature:
.....

Date:2017/04/19.....

ACKNOWLEDGEMENTS

To all those who took up the quest for knowledge before me, thank you for striving to understand and better the world we live in and for inspiring me to take up a research question of my own.

Alan Turing for authoring 'Computing machinery and intelligence', you ignited my mind and changed the direction I chose to take through the universe.

Thank you to my supervisors, who worked tirelessly to help shape my work into a more polished and focused effort. This write up would not have been completed without the assistance which Prof Ernesta Meintjes and Dr Kylie de Jager provided.

My appreciation goes to the University of Cape Town for allowing this research to take place, and the National Research Foundation for funding a year of it.

Finally to my family; my partner Emma for supporting me while working full time and writing up. My parents for their never boundless love and encouragement. To my daughter Eliza who arrived into this world in April this year as I concluded my research – serving as a beautiful motivation to complete this portion of my academic life.

“May not machines carry out something which ought to be described as thinking
but which is very different from what a man does?” - Alan Turing¹

¹ Turing, Alan M. "Computing machinery and intelligence." *Mind* 59.236 (1950): 435.

Abstract

It is hypothesised that supervised machine learning on the estimated parameters output by a model for visually evoked potentials (VEPs), created by Kremláček et al. (2002), could be used to classify steady-state visually evoked potentials (SSVEP) by frequency of stimulation. Classification of SSVEPs by stimulus frequency has application in SSVEP-based brain computer interfaces (BCI), where users are presented with flashing stimuli and user intent is decoded by identifying which stimulus the subject is attending to. We investigate the ability of the model of VEPs to fit the initial portions of SSVEPs, which are not yet in a steady state and contain characteristic features of VEPs superimposed with those of a steady-state response. In this process the estimated parameters, as a function of the model for a given SSVEP response, were found. These estimated parameters were used to train several support vector machines (SVM) to classify the SSVEPs. Three initialisation conditions for the model are examined for their contribution to the goodness of fit and the subsequent classification accuracy, of the SVMs. It was found that the model was able to fit SSVEPs with a normalised root mean square error (NRMSE) of 27%, this performance did not match the expected NRMSE values of 13% reported by Kremláček et al. (2002) for fits on VEPs. The fit data was assessed by the machine learning scheme and generated parameters which were classifiable by SVM above a random chance of 14% (Range 9% to 28%). It was also shown that the selection of initial parameters had no distinct effect on the classification accuracy. Traditional classification approaches using spectral techniques such as Power Spectral Density Analysis (PSDA) and canonical correlation analysis (CCA) require a window period of data above 1 s to perform accurately enough for use in BCIs. The larger the window period of SSVEP data used the more the Information transfer rate (ITR) decreases. Undertaking a successful classification on only the initial 250 ms portions of SSVEP data would lead to an improved ITR and a BCI which is faster to use. Classification of each method was assessed at three SSVEP window periods (0.25, 0.5 and 1 s). Comparison of the three methods revealed that, on a whole CCA outperformed both the PSDA and SVM methods. While PSDA performance was in-line with that of the SVM method. All methods performed poorly at the window period of 0.25 s with an average accuracy converging on random chance - 14%. At the window period of 0.5 s the CCA only marginally outperformed the SVM method and at a time of 1 s the CCA method significantly ($p < 0.05$) outperformed the SVM method. While the SVMs tended to improve with window period the results were not generally significant. It was found that certain SVMs (Representing a unique combination of subject, initial conditions and window period) achieved an accuracy as high as 30%. For a few instances the accuracy was comparable to the CCA method with a significance of 5%. While we were unable to predict which SVM would perform well for a given subject, it was demonstrated that with further refinement this novel method may produce results similar to or better than that of CCA.

Contents

DECLARATION	iii
ACKNOWLEDGEMENTS	iv
Abstract	v
Contents	vi
List of figures	x
List of Equations	xv
List of Tables	xv
GLOSSARY	xvi
Abbreviations.....	xvii
1 Introduction	1
1.1 Research Rationale and Hypotheses.....	8
1.2 Objectives.....	9
1.3 Thesis Outline.....	10
2 Background.....	11
2.1 Visually evoked potentials.....	12
2.1.1 VEP and SSVEP responses	12
2.1.2 SSVEPs as repeated VEPs	14
2.1.3 SSVEP-based BCIs.....	15
2.1.4 10–20 system of electrode placement.....	16
2.1.5 Information transfer rate (ITR)	16
2.2 SSVEP Classification Methods.....	17
2.2.1 Pre-processing	17
2.2.2 Feature classification of power spectral density analysis (PSDA)	17
2.2.3 Feature classification of canonical correlation analysis (CCA)	18
2.2.4 Extraction and classification of time-domain features.....	20
2.3 Model-based Feature Extraction	22
2.4 Feature Classification of Support Vector Machine (SVM)	23
2.5 Assessing Multiclass Classifier Performance.....	24
2.5.1 Accuracy of classification	25
2.5.2 Precision of classification.....	25
2.5.3 Recall of classification	25

2.6	A Model of Visually Evoked Potentials (VEPs) as a Feature Extractor	26
2.6.1	Underlying principles of the model	26
2.6.2	Model implementation	26
2.6.3	Model parameters (feature descriptors).....	28
2.6.4	Model output.....	28
2.6.5	Model validation	28
3	Experimental methodology.....	30
3.1	EEG Data Acquisition	30
3.1.1	Recording equipment and electrode placement	30
3.1.2	Stimulus design and presentation	30
3.1.3	Synchronisation of stimulus display and EEG recording	32
3.1.4	Exploratory data set.....	32
3.1.5	Main data set used for hypothesis testing	33
3.2	Data Pre-processing.....	34
3.2.1	SSVEP window period selection	34
3.2.2	Removal of DC offset	35
3.2.3	Noise reduction	35
3.2.4	Eyeblink removal	35
3.3	Model Construction.....	36
3.4	Model Validation	37
3.5	Feature Extraction by Model Fitting	38
3.5.1	Initial parameter selection.....	38
3.5.2	Parameter constraints	40
3.6	Model Fitting to the Initial SSVEP Responses	41
3.7	Feature Classification	43
3.7.1	Multiclass SVMs	43
3.7.2	SVM construction	45
3.7.3	CCA feature extraction and classification.....	45
3.7.4	PSDA feature extraction and classification	46
3.8	Assessing Classifier Performance.....	46
3.8.1	Cross-validation	46
3.8.2	Statistical comparison of classifiers	47
3.9	Classifier Performance in the Context of BCIs	47

3.9.1	Probability of a true positive decoding	47
3.9.2	McNemar testing	47
3.9.3	Hypothetical ITRs	47
4	Results	48
4.1	Model Validation	48
4.2	Feature Extraction by Model Fitting	49
4.2.1	Model fit on VEP results.....	49
4.2.2	Fitting the model to SSVEP data in steady state	50
4.3	Model Fitting to the Initial SSVEP Responses	52
4.4	Assessing Classifier Performance.....	59
4.4.1	Cross-validation: SVM in-sample classification performance	59
4.4.2	Cross-validation: SVM out-of-sample classification performance	62
4.4.3	Statistical comparison of classifiers	66
4.5	Classifier Performance in the Context of BCIs	70
4.5.1	Probability of a true positive decoding	70
4.5.2	McNemar testing	71
4.5.3	Hypothetical ITRs	74
5	Discussion	76
5.1	Using Kremláček's Model for fitting SSVEPs	76
5.1.1	Initial model parameters.....	77
5.1.2	Downhill simplex method of parameter estimation.....	77
5.1.3	NRMSE as a goodness-of-fit function	78
5.1.4	Fitting to different portions of the SSVEP	78
5.1.5	Window period	79
5.1.6	Selection of stimulus frequencies	79
5.2	SVM Classifier Performance	80
5.2.1	Cross-validation and overfitting	80
5.2.2	Separation of parameters	81
5.2.3	Initial parameter variability	83
5.2.4	Inter-subject variability	84
5.3	Future work.....	85
5.3.1	SVM training dataset size.....	85
5.3.2	SVM training method.....	85

5.3.3 Hybrid classification schemes	86
6 Conclusion	87
7 References.....	89

List of figures

Figure 1-1: The major components of a BCI system. User intent is determined by interpreting EEG activity induced by the subject's attending a known stimulus. The EEG activity is then processed, classified and used as an input to control an external device (Hong et al., n.d.)	1
Figure 1-2: Examples of stimuli used to elicit VEPs and SSVEPs. A similar effect to that of a flashing light source such as an LED can be achieved by displaying various types of images that alternate in light intensity, colour and contrast (C_0). The most basic of these is a flashing stimulus (C_1), which consists of a single colour that reverses. C_2 is an alternating checkerboard in which the colour within each square alternate. This principle can be extended to the presentation of complex images (C_3) (Wolpaw et al., 2002).	2
Figure 1-3: Generalised transient VEP (top) and SSVEP responses (bottom) are shown here. The corresponding signal is viewed in the frequency domains in F1 and F2 . The stimulus that evoked the response is shown below each plot – adapted from Vialatte et al. (2010)	3
Figure 1-4: Generalised SSVEP BCI. A subject is presented with multiple stimuli (A and B), each with a pre-existing mapping to an action. The subject chooses an action to perform and attends the required stimulus. A VEP is then generated, which is recorded by electrodes on the scalp. A classification scheme decodes this electrical signal and classifies it in terms of the presented stimulus. The BCI executes a command based on the subject's intent and button A is pressed.	4
Figure 1-5: A comparison of classification accuracies achieved using CCA and PSDA for five different window periods T (Bin et al., 2009). N_h denotes the number harmonics of a given reference signal used.	5
Figure 1-6: Comparison of information transfer rates (ITRs) for various SSVEP classification methods for different data window periods. Note that the ITR improves as the window period decreases, even though the classification accuracy may be reduced (Zhang et al., 2014).	6
Figure 1-7: Plot showing how classification accuracies of CCA, PSDA and variants of these methods decrease rapidly when the window period is reduced below one second (Zhang et al., 2012).	6
Figure 1-8: Pattern-reversal VEP fit using Kremláček's model. The contributions of the various components to the model (OSC1, OSC2 and OSC3) are shown by dashed and dotted lines, while the solid thick line represents the final summed model output – a synthetic VEP. The thin solid line represents an actual recorded VEP (Kremláček et al., 2002). OSC = oscillator.	7
Figure 2-1: A pattern-reversal VEP elicited by a single reversal of a checkerboard stimulus. The main characteristics of the response are two negative peaks at 75 ms (N_{75}) and 135 ms (N_{135}) and a positive peak at 100 ms (P_{100}) (Odom et al., 2010).	12
Figure 2-2: The SSVEP response to a 15 Hz stimulus seen in the time-domain (a). Note the initial VEP from 0 to 0,25 s before the signal becomes steady and repeated. The major peaks, P_{100} and N_{135} , can be seen, although they are slightly delayed. In addition, the steady state of the 15 Hz stimulus become evident just after 0.2 s, adding to the response. (b) S shows the same SSVEP response in the frequency domain. Note the characteristic spike at frequency 15 Hz and its harmonics at 30 and 45 Hz (Zhu et al., 2010)	13
Figure 2-3: (A1) Recorded SSVEP waveforms and (A2) synthetic SSVEPs at various frequencies of stimulation. The characteristic peaks at N_{75} , P_{100} and N_{135} are indicated. The black triangle indicates stimulus onset. (B1) and (B2) indicate the corresponding frequency responses as assessed by the PSDA method (Capilla et al., 2011)	14
Figure 2-4: Luminance contrast vs comfort relative to performance measured in ITR. It is seen that higher-contrast systems perform better with spectral methods, owing to the increased SSVEP amplitude. Yet they are rated as more uncomfortable to the end user (Bieger & Molina, 2010).	15
Figure 2-5: The classic 10–20 electrode placement system is seen in A and B above. C shows the extended 10–20 system with 70 electrode sites (BCI2000, 2012).	16
Figure 2-6: Normalised amplitudes of frequency components are shown for a single subject's SSVEP for to different window periods. Window durations greater than 2 s are required for classification (Wang et al., 2006).	18
Figure 2-7: An illustration of CCA processing. EEG signals $x_1, x_2 \dots x_8$ are compared to reference signals $y_1, y_2 \dots y_8$	18

$y_2 \dots y_6$, which have the same frequency as the presented stimuli. Linear combinations of each are generated such that $X = (W_{x1} + \dots + W_{x8})$ and $Y = (W_{y1} + \dots + W_{y8})$. X and Y are compared until they have a high correlation. At this point the largest coefficient in the W_{y_n} set indicates the frequency of the SSVEP (Lin, et al., 2007).	19
Figure 2-8: Individual VEP traces for repeated window periods are shown in blue and their averages in red. Each stimulus frequency will produce a characteristic response in the time domain for a distinct repeat window period. For example, in the case of a 10 Hz stimulus, data averaged across multiple 100 ms windows will produce a characteristic response. Since the subject was attending a 20 Hz stimulus in this instance, averaging data over multiple 50 ms window periods produces a characteristic response, compared to the relatively flat responses at 15 Hz and 12 Hz, and the 20 Hz signal evident on the 10 Hz plot (Manyakov et al., 2010)	20
Figure 2-9: A four-class SLIC BCI schema able to distinguish between four different stimuli by employing four LDAs trained on the subject's response (Luo & Sullivan, 2010a)	21
Figure 2-10: Accuracy of SLIC LDA classification for three subjects, given various lengths of recorded SSVEPs (Manyakov et al., 2010)	22
Figure 2-11: Representation of a binary SVM. Two classes of data exist: Class 1 and Class 2, each a function of the parameters P_1 and P_2 . When the SVM is trained, it fits a hyperplane (H) between the two classes such that they are maximally separable. Given an unseen data point D_1 , the SVM will classify it in relation to the hyperplane that separates the classes. In this instance, D_1 would be identified as belonging to Class 2.	23
Figure 2-12: (A) Kremláček et al.'s (1999) model, as originally published, for visually evoked potentials. $y(t)$ represents the model VEP output while $x(t)$ is the initial input stimulus. Delays are indicated by T_1 , T_2 and T_3 ; weighting factors by K_1 , K_2 and K_3 . (B) A single-oscillator schema, these represent blocks OSC1, OSC2 and OSC3 in (A). Coefficients $-a$ and $-b$ in (B) control the damping factor and the frequency of oscillation, while $v(t)$ represents the output to the next stages. c is a constant of integration.	27
Figure 3-1: VEP and SSVEP checkerboard stimulus. The white block at the bottom left of the stimulus flashed at the stimulus frequency. It was covered at all times by a light-dependent resistor (LDR), which detected changes in light. The white fixation mark is visible in the centre of the screen.	31
Figure 3-2: SSVEP response amplitude as a function of stimulation frequency (Wang et al., 2006).	32
Figure 3-3: (A) Checkerboard stimuli flashing at various frequencies were presented to the subject one at a time. (B) Each stimulus frequency induces a unique SSVEP response in the subject viewing the stimulus. The SSVEP response is recorded and saved (C)	33
Figure 3-4: A single SSVEP trial. Subjects first viewed a fixation cross for 5 s, followed by 10 s of stimulus presentation, and 15 s again viewing a static fixation cross. The fixation cross was kept constant throughout the experiment so as not to induce any VEPs and to give the subject a reference point to focus on.	34
Figure 3-5: Comparison of an SSVEP signal before (left) and after (right) DC offset correction (note the difference in the y-axis values).	35
Figure 3-6: MATLAB Simulink schematic of a single oscillator, OSC2.	36
Figure 3-7: Schematic representation of the model implemented in MATLAB Simulink.	36
Figure 3-8: MATLAB GUI designed to run the model.	37
Figure 3-9: Model outputs for each of the three IP sets. The model output for IP1 is a cosine wave at 12 Hz; IP2 generates a low-frequency 1.6 Hz cosine wave; and IP3 is a set of parameter estimates that produced a successful fit to a 12 Hz flashing checkerboard SSVEP response during model testing.	39
Figure 3-10: Feature extraction by model fitting and parameter estimation. The model starts with one of three sets of initial parameters (IP1, IP2 and IP3). The estimation process adjusts the model parameters (P_j) iteratively using a pattern-search method to obtain a better fit (PG), each time comparing the output to the measured SSVEP response. This process is repeated for 500 iterations. Once the iterations are complete or the normalised root mean square error (NRMSE) has been reduced to zero, the parameter estimation stops and the final estimated parameters (EP) are output.	42
Figure 3-11: Feature extraction by model fitting and parameter estimation for a single SSVEP response: (A) SSVEP data are entered into the model. (B) The model fits the measured data through a process of parameter	

estimation. (C) A set of parameters (P_n) is generated that describes the measured SSVEP signal as a function of the model.	43
Figure 3-12: Code snippet of MATLAB code used to train the SVM's with the defined template <code>Data.SVMData.t</code>	45
Figure 4-1: Output from our replicated model when using the parameters specified by Kremláček et al. (2002) for pattern reversal stimuli. Not	48
Figure 4-2: Model fit (red dashed line) to a stimulus-locked average of 20 occurrences of a single subject's VEP response (solid blue line) when initial starting parameters suggested by Kremláček et al. (2002) for pattern reversal VEPs were used. The model fit produced an NRMSE of 0.26 (goodness of fit = 0.74).	49
Figure 4-3: Model fit to the same stimulus-locked average of 20 occurrences of a single subject's VEP response (solid blue line) when initial starting parameters were set to zero. The NRMSE is 0.86, yielding a goodness of fit of 0.14.	50
Figure 4-4: Comparison of an SSVEP response to a 4 Hz flashing checkerboard stimulus (solid blue line) and the output from the model after 500 iterations (dashed red line). The NRMSE is 21% (goodness of fit = 0.79). Although the initial model fit closely follows the recorded data, it deviates over time due to signal drift.	51
Figure 4-5: Model fit to 1 s of recorded SSVEP data. The black line shows the SSVEP response recorded in subject 4 following onset of a 7.5 Hz flashing checkerboard stimulus at 0 s. The other colours (blue, red, green) show model fits for each of three different initial parameter sets. The goodness of fit for each plot are shown in the legend on the top right.	52
Figure 4-6: Model fit to 0.5 s of recorded SSVEP data. The black line shows the SSVEP response recorded in subject 3 following onset of a 6.5 Hz flashing checkerboard stimulus at 0 s. The other colours (blue, red, green) show model fits for each of three different initial parameter sets. The goodness of fit for each plot are shown in the legend on the top right.	53
Figure 4-7: Model fit to 0.25 s of recorded SSVEP data. The black line shows the SSVEP response recorded in subject 2 following onset of a 6 Hz flashing checkerboard stimulus at 0 s. The other colours (blue, red, green) show model fits for each of three different initial parameter sets. The goodness of fit for each plot are shown in the legend on the top right.	53
Figure 4-8: Box-and-whisker plots showing the distribution of goodness of fit values achieved in all subjects for different frequencies and window periods. The notches in the boxes represent 95% confidence intervals (CI) around the median, defined as $M \pm 1.57 * (IQR / \sqrt{n})$, where M is the median, IQR is the interquartile range and n is the number of measurements. Red crosses denote outliers, that is points more than $1.5 * IQR$ below or above the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data values that are not outliers.	55
Figure 4-9: Box-and-whisker plots showing goodness of fit, averaged over all subjects, stimulus frequencies and initial parameter sets, as a function of data window period. The fact that the notches for 0.25 s, 0.5 s and 1 s do not overlap provides strong evidence at a significance level of 5% that the median the goodness of fit improves when fitting to longer periods of the initial SSVEP response.	56
Figure 4-10: Post-hocs revealed that, when averaged over all window periods, subjects and initial parameter sets, the model achieved significantly better fits to SSVEP responses from 9.3 Hz stimuli than SSVEP responses from 6 Hz stimuli. No other stimulus frequencies showed significant differences.	56
Figure 4-11: Comparison of goodness of fits achieved for each of the three initial parameter (IP) sets, averaged across all subjects, stimulus frequencies and window periods. The fact that the notches for IP2 do not overlap with those of IP1 and IP3, provides strong evidence that the median goodness of fit for IP2 is lower than that of IP1 and IP3 at a significance level of 5%.	57
Figure 4-12: Post-hocs revealed that, when averaged over all window periods, subjects and stimulus frequencies, the model achieved significantly better fits for IP1 and IP3 compared to IP2.	57
Figure 4-13: Post-hocs revealed that, when averaged over all window periods, stimulus frequencies and initial parameter sets, the model achieved significantly better fits to SSVEP responses recorded in subjects 1 and 3	

compared to subject 2.	58
Figure 4-14: Each subject's in-sample resubstitution loss for different SVMs for different window periods of data.	59
Figure 4-15: In-sample confusion matrices of SVMAll for Subject 4 for each window period.	61
Figure 4-16: In-sample class-specific accuracy, precision and recall of SVMAll for subject 4 and a window period of 0.25 s. Accuracy reflects the fraction of all classifications that are true; precision, the fraction of positive classifications that are true; and recall, the fraction of signals belonging to a class that are correctly assigned to that class.	62
Figure 4-17: Each subject's out-of-sample resubstitution loss for different SVMs for different window periods of data.	63
Figure 4-18: Out-of-sample confusion matrices of SVMAll summed across subjects for data window periods 0.25 s, 0.5 s, and 1 s, respectively.	64
Figure 4-19: Out-of-sample confusion matrices of CCA summed across subjects for data window periods 0.25 s, 0.5 s, and 1 s, respectively.	64
Figure 4-20: Class-specific accuracy, precision and recall for the SVMAll classifier across all subjects for window periods 0.25 s, 0.5 s, and 1 s, respectively.	65
Figure 4-21: Class-specific accuracy, precision and recall for the CCA method across all subjects for window periods 0.25 s, 0.5 s, and 1 s, respectively.	65
Figure 4-22: Post-hocs revealed that the SVM classifiers achieved greater accuracies for 8.2 Hz, 9.3 Hz and 10 Hz stimuli compared to 6 Hz stimuli.	66
Figure 4-23: Post-hocs revealed that SVMAll achieved better precision (left) and recall (right) than the other SVM-based classifiers.	66
Figure 4-24: Box-and-whisker plots showing for SVMAll the distribution of class-specific accuracy (left), precision (middle) and recall (right) measures as a function of window period.	68
Figure 4-25: Box-and-whisker plots showing for CCA the distribution of class-specific accuracy (left), precision (middle) and recall (right) measures as a function of window period. Window periods showing significant differences are indicated by brackets above the plot.	68
Figure 4-26: Box-and-whisker plots showing the distribution of class-specific accuracy (left), precision (middle) and recall (right) measures for the different SVM-based classifiers and CCA when data from all the window periods are combined. Significant differences are indicated by brackets above the plot.	69
Figure 4-27: Comparison of the probabilities of a true positive decoding (Prob(TP)) for all subjects combined, for different classification methods as a function of window period.	70
Figure 4-28: Comparison in each subject of Prob(TP) for different classification methods as a function of data window period.	71
Figure 4-29: Summary of the results from McNemar tests performed for each subject at each window period. For each window period, each entry in the matrix shows the number of instances in which the null hypothesis was rejected in any of the 4 subjects. As such, values reflect the number of subjects in whom the classifier in the row performed better than the classifier in the column.	73
Figure 4-30: Average theoretical ITRs for different classification methods as a function of window period.	74
Figure 4-31: Individual subject theoretical ITRs for different classification methods as a function of window period. Discontinuities in subject 4 arise from data points with zero accuracy.	75
Figure 5-1: In the two-dimensional parameter space P_1, P_2 we have two classes (1 and 2). We assume these represent perfect fits onto data and the classes are neatly separable by the hyperplane H . If an error is introduced to the parameters in each fit ($P_1 + \text{Error}$, $P_2 + \text{Error}$), the class groupings are extended into regions R_1 and R_2 . The hyperplane is no longer able to distinguish classes in the region $(R_1 \cap R_2)$	82
Figure 5-2: Prob(TP) of SVMs trained on only three frequencies (6, 8.2 and 12 Hz) for subject 1. An improvement around the 0.5 s point is seen as compared with Figure 4-28. This could, however, be due in part to the decrease in classes used, which resulted in a random chance accuracy of 33%. No statistical analysis was performed on	

these results.	82
Figure 5-3: Classification accuracy ($1 - \text{resubstitution losses}$) of subject 4's SVMs classifying EPs which it had been trained on as well as those it had not been trained on. The plot on the left shows SVM1(EP1), SVM1(EP2) and SVM1(EP3). The middle and right plots show similar results for SVM2 and SVM3 respectively.	83
Figure 5-4: Classification accuracy ($1 - \text{resubstitution losses}$) of subject 1's SVMall classifier using the estimated parameter sets from all subjects. The top trace (green line) shows $1 - \text{in-sample loss}$ as the SVM is classifying the training data. The lower traces show $1 - \text{out-of-sample loss}$, which are within the same range given by the out-of-sample classification rate for subject 1's SVMs (Figure 4-17).	84
Figure 5-5: Examples of (A) a hyperplane implemented with an 8th-degree polynomial function and (B) a hyperplane generated by an RBF kernel.	85

List of Equations

<i>Equation 1: Formula for Shannon's ITR</i>	16
<i>Equation 2: Model based measurement</i>	22
<i>Equation 3: Class accuracy of classification</i>	25
<i>Equation 4: Class precision of classification</i>	25
<i>Equation 5: Recall of classification</i>	25
<i>Equation 6: The model parameter space</i>	28

List of Tables

Table 1: <i>Generic confusion matrix for a three class classifier.</i>	24
Table 3-1: <i>Participant demographics.</i>	34
Table 3-2: <i>Frequency randomisation order.</i>	34
Table 3-3: <i>Input parameters for checkerboard stimuli (Kremláček et al., 2002).</i>	37
Table 3-4: <i>Initial parameters used during model fitting</i>	39
Table 3-5: <i>Initial parameter sets.</i>	40
Table 3-6: <i>A single SVM training set for one subject</i>	44
Table 3-7: <i>SVMs trained per subject.</i>	44
Table 4-1: <i>McNemar test results: p-values for subject 2, window period = 0.25s. Green shaded regions indicate rejection of the null hypothesis which implies that the classifier in the row is more accurate than that in the column.</i>	72
Table 4-2: <i>McNemar test results: p-values for subject 2, window period = 1s. Green shaded regions indicate rejection of the null hypothesis which implies that the classifier in the row is more accurate than that in the column.</i>	72
Table 4-3: <i>McNemar test results: p-values for subject 4, window period = 1s. Green shaded regions indicate rejection of the null hypothesis which implies that the classifier in the row is more accurate than that in the column.</i>	72

GLOSSARY

Stimulus – Stimulus in the context of this research refers to a flashing light source, which is alternating at a known frequency. Specifically Checkerboard pattern SSVEP stimuli are used.

Frequency of stimulation – This refers directly the frequency at which the presented stimulus is flashing

Window-Period – Window period refers to the period from the initial stimulus presentation (0 s) to the specified window period. It is always considered to be from the 0 s point unless explicitly indicated.

Model Parameters – These constitute a set of values of various parameters defined by the model in questions. Such that the any signal generated by a set of given parameters can be linked directly to those parameters as a function of the model.

Model Fit – The process of creating a synthetic signal, in this study an SSVEP response; which matches closely to a recorded signal.

Parameter Estimation – The process whereby the model parameters are varied by some function to create a model fit.

Goodness of fit – This is a metric which describes the error associated with a given model fit.

Initial Parameters – Refer to the parameters which are used to initialise the model before a model fit by parameter estimation is conducted.

Estimated Parameters – Refer to the parameter set generated when a model fit has been successfully completed by a parameter estimation process. These relate, along with some error, to the measured signal as a function of the model

Abbreviations

ANN – Artificial neural Network

ANOVA – Analysis of variance

BCI – Brain Computer Interface

CCA – Canonical correlation analysis

EEG – Electroencephalography

EP – Estimated parameters

fMRI – functional magnetic resonance imaging

GUI – Graphical user interface

ICA – Independent component analysis

IP – Initial Parameters

ITR – Information transfer rate

LDA – Linear discriminant analysis

LDR – Light Dependant Resistor

MEG – Magnetoencephalography

NRMSE – Normalised Root Mean Square Error

PCA – Principle Component analysis

PSDA – Power spectral density analysis

RBF – Radial basis function

SLIC – stimulus-locked inter-trace correlation

SSVEP – Steady State Visually Evoked Potential

SVM – Support Vector Machine

VEP – Visually Evoked Potential

1

Introduction

This study aims to examine the use of a classification scheme for a time-domain model for use with brain computer interfaces (BCIs). Its purpose is to reduce the window period of data required to accurately classify a steady-state visually evoked potential (SSVEP) response, to increase the information transfer rate (ITR) of the BCI. Ideally one would like to classify an SSVEP from the first instance when the visually evoked potential (VEP) occurs under conditions that are practical for everyday use.

Instead of using the normal pathways between the brain and the muscles to actuate a command on a computer, keyboard or similar electronic device, BCIs translate measurements of brain activity into commands (**Figure 1-1**). BCIs are valuable to people with severe motor disabilities that prevent them from communicating through other physical means. Examples include amyotrophic lateral sclerosis (ALS), brainstem stroke, brain or spinal cord injury, cerebral palsy, muscular dystrophies, multiple sclerosis and various other diseases which either impair the neural pathways that control muscles or the muscles themselves (Wolpaw et al., 2008).

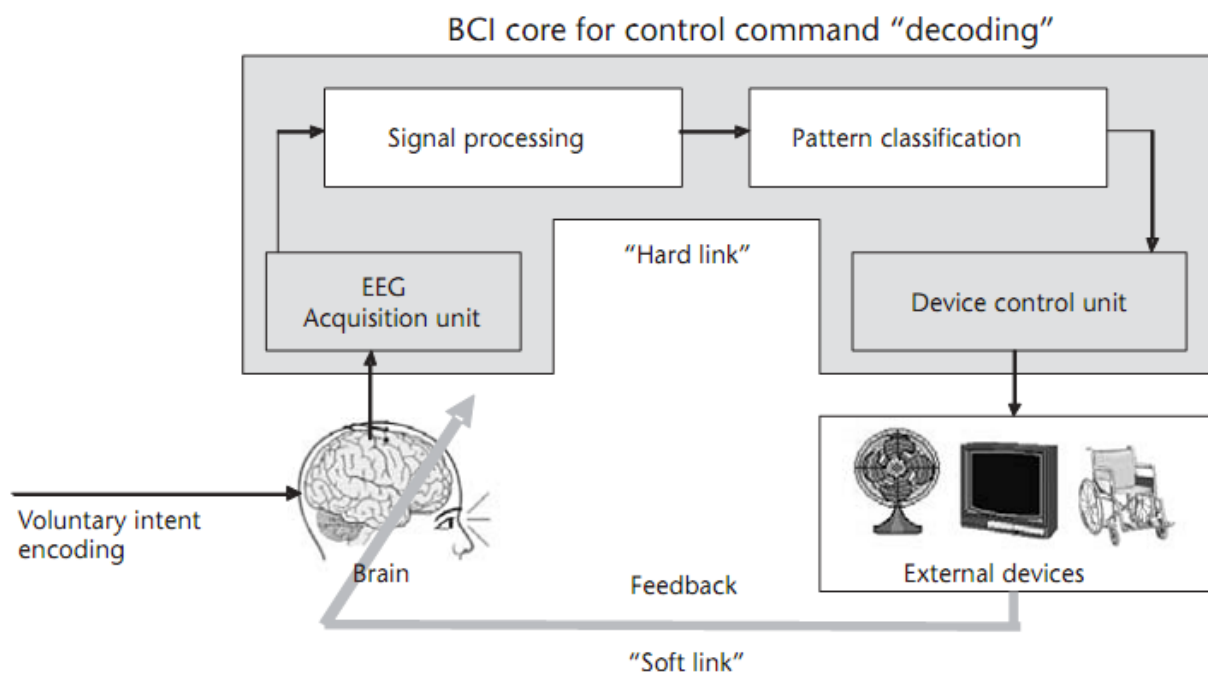


Figure 1-1: The major components of a BCI system. User intent is determined by interpreting EEG activity induced by the subject's attending a known stimulus. The EEG activity is then processed, classified and used as an input to control an external device (Hong et al., n.d.)

BCIs can use different types of brain activity measurement. These can include electroencephalography (EEG), functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS), magnetoencephalography (MEG) or invasive EEG methods such as electrodes inserted directly into the brain (Allison et al., 2010). Of these, EEG recordings are most commonly used owing to their low cost, ease of portability and non-invasive nature (Allison et al., 2008).

The EEG electrical activity that is chosen to drive a given BCI needs to be robust and repeatable in order to be classified. A person's EEG signal offers many possibilities for encoding intent. Primarily, EEG activity can be correlated with actual or imagined movements or the stimulation of a person's auditory or visual senses. The resulting event-related potentials (ERP) can be uniquely classified and used as input to a BCI (Wolpaw et al., 2002).

Visually evoked potentials (VEP) – the electrical response to a visual stimulus – is one way of encoding user intent. Neurons in the visual cortex are highly sensitive to patterns and high-contrast images, which generate clear and distinct EEG waveforms when presented to subjects (Yoshimura & Itakura, 2011). VEPs can be modulated by varying the stimulus pattern, contrast or rate of display; they are present in all members of the population who have a functional visual nervous system. By presenting a specific visual stimulus to a user, intent can be encoded in the EEG waveform, which, once decoded, can be used to control a device. **Figure 1-2** shows examples of stimuli that can be used to elicit VEP and SSVEP responses.

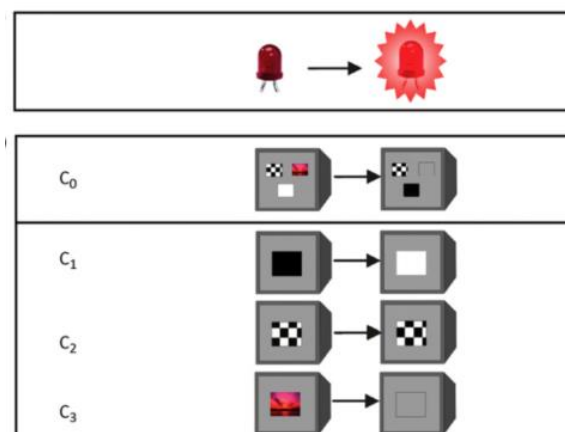


Figure 1-2: Examples of stimuli used to elicit VEPs and SSVEPs. A similar effect to that of a flashing light source such as an LED can be achieved by displaying various types of images that alternate in light intensity, colour and contrast (C_0). The most basic of these is a flashing stimulus (C_1), which consists of a single colour that reverses. C_2 is an alternating checkerboard in which the colour within each square alternate. This principle can be extended to the presentation of complex images (C_3) (Wolpaw et al., 2002).

VEPs are broadly categorised into transient and steady-state. Transient VEPs occur as an initial perturbation in the EEG waveform from 0 to 250 ms after the initial presentation of a light stimulus (Odom et al., 2010). When a flashing stimulus – that is, one that alternates between two distinct stimuli at a constant frequency – is presented, the response is characterised by an initial VEP, followed by an SSVEP at the stimulus frequency (**Figure 1-3**; Manyakov et al., 2010). Owing to the frequency information encoded in the SSVEP signal, SSVEPs are easily identified in both the time and the frequency domains (F2 in **Figure 1-3**), whereas the transient VEP is not distinctive in the frequency domain owing to the absence of major repetitive components (F1 in **Figure 1-3**).

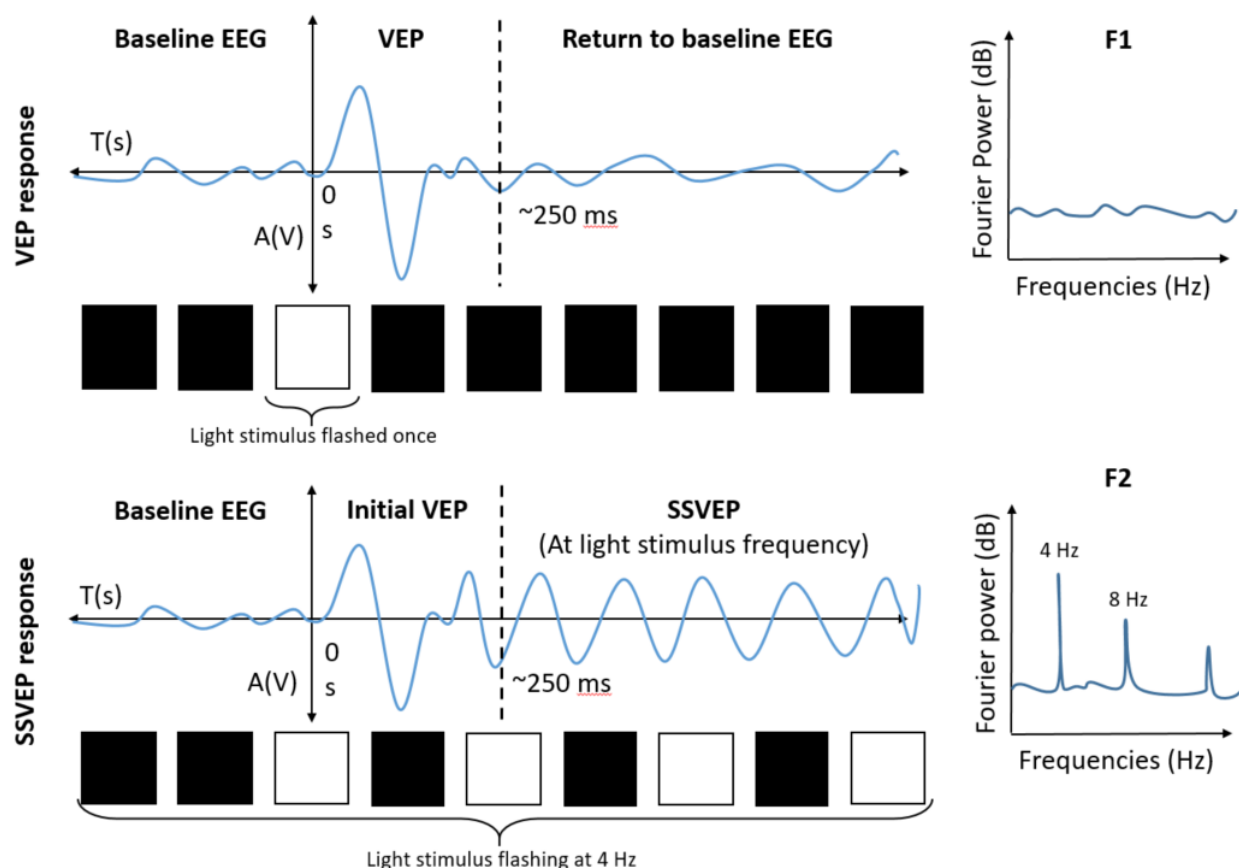


Figure 1-3: Generalised transient VEP (**top**) and SSVEP responses (**bottom**) are shown here. The corresponding signal is viewed in the frequency domains in **F1** and **F2**. The stimulus that evoked the response is shown below each plot – adapted from Vialatte et al. (2010)

More than 80% of publications about BCIs describe the use of EEG to measure brain activity (Mason et al., 2007). Owing to its high temporal resolution, of the order of 512 Hz and above, EEG is well suited to measuring SSVEPs (Allison et al., 2010). Furthermore, the signal can be simply classified by extracting the frequency of the specific flashing visual stimulus directly from the SSVEP (Nicolas-Alonso & Gomez-Gil, 2012), thereby enabling the BCI to initiate specific commands (Zhu et al., 2010). Unlike other EEG signals, SSVEPs do not require the averaging of multiple recordings to be classified accurately. Instead, they can be measured in a single recording from stimulus onset (Wolpaw et al., 2002). Owing to the size and repetitive nature of SSVEP responses they are also more robust to signal noise inherent to the real world settings in which BCIs are used (Volosyak et al., 2009).

BCIs can present multiple stimuli to a subject simultaneously, which enable a variety of user intents to be captured (**Figure 1-4**). The goal of BCI research is to improve the number of stimuli that can be presented, the accuracy and speed of classification and the ease of use. Improvements in these areas allow for a faster, more seamless experience for the user.

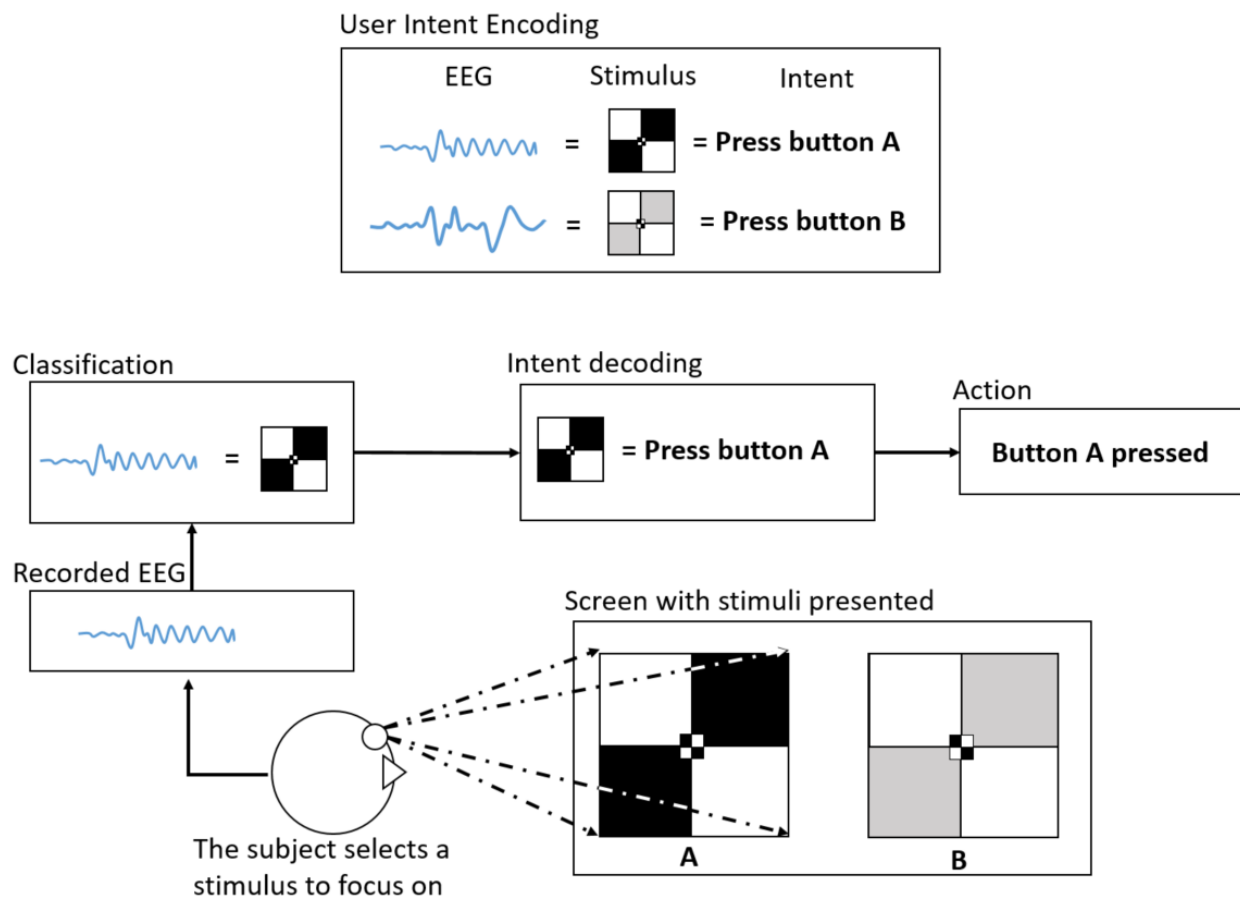


Figure 1-4: Generalised SSVEP BCI. A subject is presented with multiple stimuli (A and B), each with a pre-existing mapping to an action. The subject chooses an action to perform and attends the required stimulus. A VEP is then generated, which is recorded by electrodes on the scalp. A classification scheme decodes this electrical signal and classifies it in terms of the presented stimulus. The BCI executes a command based on the subject's intent and button A is pressed.

The two popular frequency domain classification methods used as a benchmark in this study are (i) power spectral density analysis (PSDA) as described by Marple Jr (1987), and (ii) canonical correlation analysis (CCA), first implemented in SSVEP classification by Lin et al. (2007). Both of these methods achieve classification accuracy rates of around 90% when given a sufficiently long 5 s window period of measured SSVEP data (Kompatsiaris et al., 2016). **Figure 1-5** shows, for each method, how the classification accuracy decreases as the window period of recorded EEG data (T) is reduced. For all window periods, the classification accuracy of CCA is about 10% greater than that achieved using the PSDA method (Hakvoort et al., 2011).

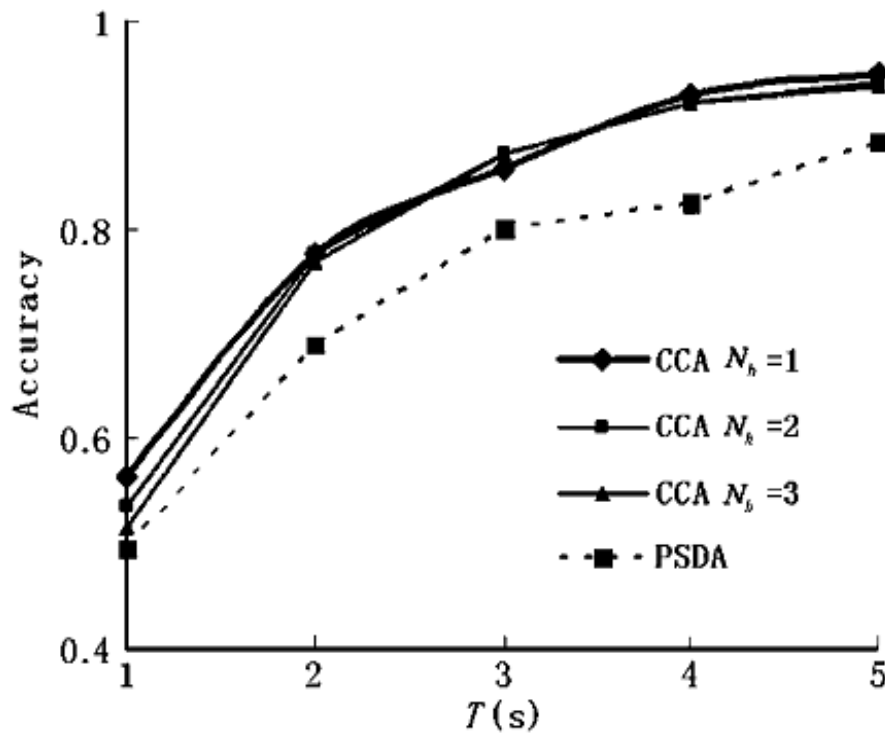


Figure 1-5: A comparison of classification accuracies achieved using CCA and PSDA for five different window periods T (Bin et al., 2009). N_h denotes the number harmonics of a given reference signal used.

When assessing the performance of BCIs, the accuracy of a command classification (i.e. how often a command or user intent is correctly interpreted) is not the only metric to consider, however. BCI performance is expressed as an ITR, which is measured in bits/min (Schreiber, 2000). ITR is a function of the accuracy of command classification, the speed of classification and the number of available command options that can simultaneously be presented to a user (Wolpaw et al., 2002). As shown in **Figure 1-6**, the speed of classification, and therefore the ITR, is heavily dependent on the window period of EEG data required for accurate classification, regardless of classification method used. Since the classifier needs to wait for the physiological response of the subject to present sufficient features for classification, the maximum ITR achievable is limited by the window period of SSVEP data that needs to be acquired after the onset of a stimulus. In this work we assume that zero computational time is required for the classifier to output a result.

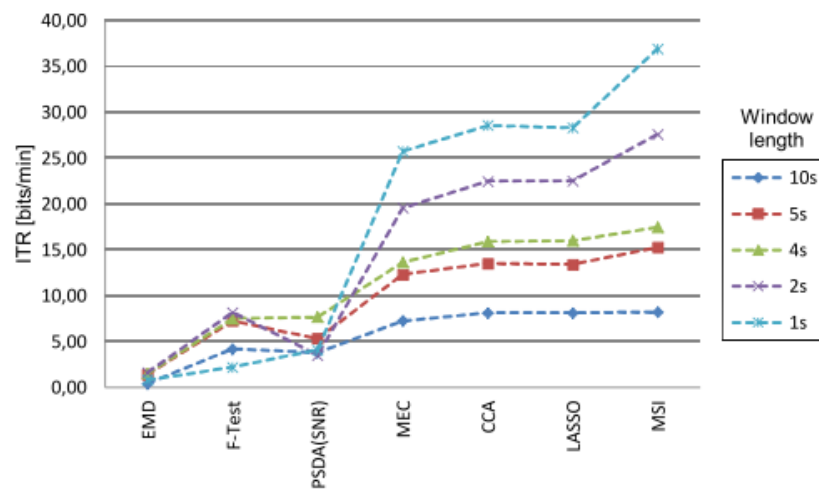


Figure 1-6: Comparison of information transfer rates (ITRs) for various SSVEP classification methods for different data window periods. Note that the ITR improves as the window period decreases, even though the classification accuracy may be reduced (Zhang et al., 2014).

In practice, however, reducing the sampling window below 1 s causes classification accuracies for both PSDA and CCA methods to decrease rapidly, with a window period of 200 ms yielding accuracies of only 30% or less (**Figure 1-7**) (Hakvoort et al., 2011; Lin et al., 2007; Vialatte et al., 2010; Zhang et al., 2014) – thus limiting achievable ITRs.

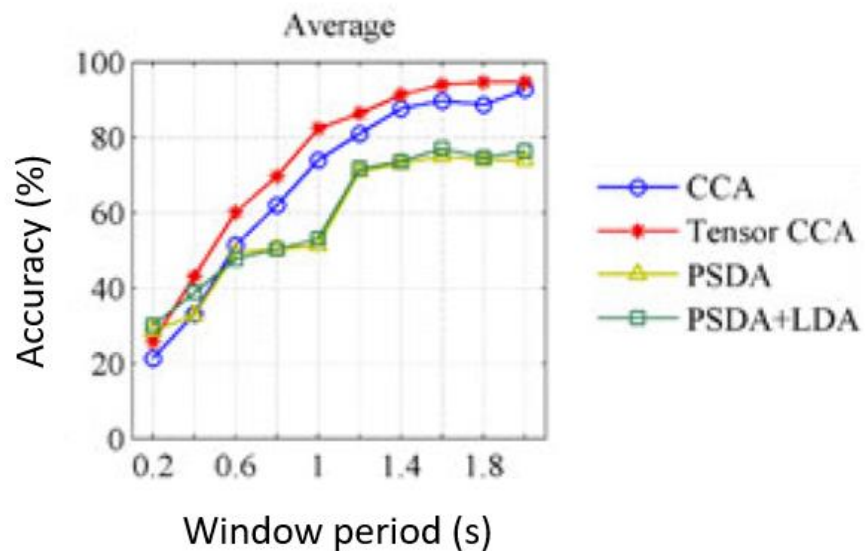


Figure 1-7: Plot showing how classification accuracies of CCA, PSDA and variants of these methods decrease rapidly when the window period is reduced below one second (Zhang et al., 2012).

An alternative approach is to model SSVEPs directly in the time domain as the superposition of continued transient VEP responses to repeated stimuli onsets, with each reversal of the stimulus eliciting a distinct

VEP (Capilla et al., 2011). Time domain classification is more robust to spectral interference from multiple stimuli and underlying EEG dynamics, and has been shown to provide higher accuracies for smaller window periods of SSVEP data when compared to spectral approaches (Abbasi, 2015). One such method – stimulus-locked inter-trace correlation (SLIC) – achieves similar accuracies to those of spectral methods for a window period of 1 s. SLIC requires that multiple VEPs are extracted from the 1 s window period by extracting a VEP for each instance of stimulus reversal. This set of VEPs is then averaged before being processed by linear discriminant analysis (LDA) classifiers (Luo & Sullivan, 2010a). The idea of using time-domain information from the repeated VEPs contained in an SSVEP and classifying the result using LDA inspired the classification approach used in this research.

In this work we used a time-domain VEP model designed by Kremláček et al. (2002) that has been shown previously to successfully create synthetic VEPs. **Figure 1-8** shows a comparison of a synthetic VEP output by the model and a recorded VEP.

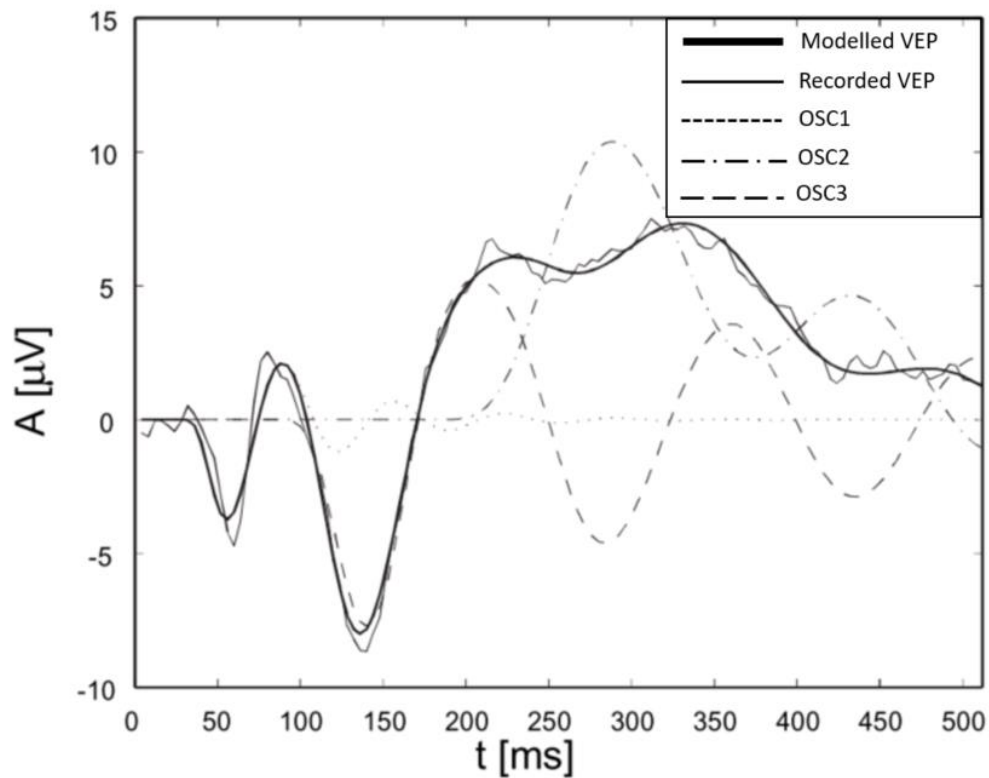


Figure 1-8: Pattern-reversal VEP fit using Kremláček’s model. The contributions of the various components to the model (OSC1, OSC2 and OSC3) are shown by dashed and dotted lines, while the solid thick line represents the final summed model output – a synthetic VEP. The thin solid line represents an actual recorded VEP (Kremláček et al., 2002). OSC = oscillator.

It is hypothesised that this model could be used to extract unique features from the initial portions of the SSVEP response through a process of model fitting and parameter estimation. The parameters could then be classified by a supervised machine-learning model to identify the corresponding stimulus frequency. By determining which of the presented stimuli the user is attending to, user intent can be decoded from less than 250 ms of recorded data, outperforming traditional PSDA and CCA methods.

1.1 Research Rationale and Hypotheses

Since the SSVEP response contains an initial VEP followed by a chain of repeated VEPs at the stimulus frequency, and Kremláček's model (1999) can be adapted to various frequency responses through the use of chained oscillators, it follows that it should be possible to fit the model to the initial portion of an SSVEP.

The initial VEP component of the SSVEP that is to be fit by the model occurs within the first 0.25 s, directly after stimulus presentation. This initial VEP contains frequency-encoded information as the stimulus is alternating, even though the visual pathways have not yet settled into a steady state. If the model fit could generate features that uniquely classify this portion of the recorded SSVEP, it would be possible to determine intent after recording only 250 ms of data. This would reduce the time required by a BCI to generate commands, as compared to traditional spectral techniques that require 1–5 s of recorded EEG data to make a decision with 80% accuracy.

The first question that must be asked is whether it is possible to fit the model with sufficient accuracy to the initial portion of a recorded SSVEP response. The first hypothesis, considered in this thesis, can therefore be stated as:

Hypothesis 1: Kremláček's VEP model can be fit onto the initial VEP portion of an SSVEP response.

A successful model fit to the initial portion of the SSVEP, which is generated as the stimulus is presented, could provide a set of unique features. If the features identified by the model fit process were unique to the frequency of the presented stimulus, the feature sets could be used to train a multiclass support vector machine (SVM) to classify "unseen" or untrained SSVEPs for a given subject. The subject's SSVEPs could then be classified by the frequency of the presented stimulus. This would in turn allow the predictive classification method for the model to generate user commands on a BCI for different stimuli, which would link the frequency of the stimulus that the user is looking at to a command input. The second hypothesis comprises two parts:

Hypothesis 2a: The model fit onto the initial portions of SSVEP signals can generate unique feature descriptive parameters that relate to the frequency of the stimulus presented.

Hypothesis 2b: These unique features enable the SSVEP signal to be classified relative to the stimulus frequency using a multiclass SVM approach.

The ITR of a BCI is a measure of the number of commands that can be successfully executed in a given time. To achieve an accurate classification, this measure is dependent on the time taken to process the data and the duration of EEG data that must be gathered before processing. We were interested in comparing the performance of our model-based approach to widely used spectral methods when the window period of input data is reduced to such an extent that spectral techniques have sub-50% accuracies. This is typically true for window periods of 1 s or less (Bin et al., 2009).

Hypothesis 3: The classification approach based on the time-domain model proposed here outperforms traditional spectral classification methods (PSDA and CCA) when the input to the classifiers comprises a 1 s or less window period of EEG data as recorded from stimulus onset; the shorter the window period, the more pronounced the effect.

1.2 Objectives

The overall objective of this research is to develop a classification system that can more rapidly identify the control input that a subject is attending to on a BCI. Flashing stimuli are used to induce SSVEP waveforms which are then analysed, and a classification is made that attempts to use recorded EEG data to predict the frequency at which they flash.

We aim to classify the SSVEP response in a single occurrence, for a short window period (0.25 s) of data, by examining the initial portions of the SSVEP when it has not yet reached steady state. This is done through a process of parameter estimation using an established mathematical model which creates synthetic VEP responses. The parameter estimation is repeated until the error between the measured SSVEP and synthetic SSVEP is reduced, yielding a set of parameter estimates that describe the SSVEP as a function of the model. This process is repeated for a range of subjects, frequencies and initial model conditions. We then attempt to train a supervised machine learning scheme, in the form of SVMs, to classify the frequency of stimulation from a set of parameter estimates.

The performance of the SVMs is compared against two other commonly used methods of SSVEP classification (PSDA and CCA).

The specific objectives of this study are to:

1. Replicate the synthetic VEP model created by Kremláček et al. (2002).
2. Assess if the VEP model can fit the initial VEP portions of SSVEP waveforms.
3. Examine if there is a significant difference in the goodness of fit when using the VEP model on SSVEP waveforms.
4. Fit the model to a data set of SSVEPs and generate sets of estimated parameters which describe the data. These estimated parameters will then be used to:
 - a. Explore the effect of different window periods (0.25, 0.5, and 1s), from stimulus onset of SSVEPs, on the goodness of fit.
 - b. Investigate the effect of different initial model conditions on the goodness of fit.
5. Use the sets of estimated parameters to train support vector machines, such that they will be able to classify any new estimated parameters by stimulus frequency.
6. Compare the accuracy of classification of the model based SVM approach to the PSDA and CCA methods so as to examine the effect of:
 - a. Window period on classification accuracy.
 - b. Initial model parameters on classification accuracy.
7. Use ITR to compare the performance of the three methods over the window periods used.

1.3 Thesis Outline

This chapter has introduced the scope and objectives of the study. Along with a description of Kremláček's model for transient VEPs, a brief background to BCIs and VEPs plus feature extraction and classification methods are presented in the next chapter. This sets the foundation for the experimental methodology outlined in chapter 3. That chapter comprises three sections, each of which addresses one of the three hypotheses introduced earlier. The first section investigates whether the model can be used to model a VEP response successfully, as described by Kremláček (2002). In the second section we examine whether the method can extract unique features from recorded SSVEP responses which can be used to train SVMs to identify the frequency of stimulation. The accuracy of the model-based method proposed here is compared to that of traditional SSVEP classification methods (PSDA and CCA) in the third section of the chapter.

Testing involved the construction and fit of the model to SSVEP data by estimating parameters. This was done using initial portions of recorded SSVEP responses induced by stimuli flashing at different frequencies. The aim was to extract from the recorded SSVEP unique features based on a time-domain model related to the stimulus frequency. Model fitting yields a set of estimated model parameters that were used together with the model to generate a synthetic SSVEP response that could be compared to the recorded SSVEP.

The SSVEP features described by the model parameters and the frequency of stimulation were then used to train a series of multiclass SVM classifiers. The SVM classifiers, once trained, assesses the parameter set found by the model fit process and makes a classification decision about what the frequency of stimulation viewed by the subject was.

Because this has been an exploratory study examining the feasibility of using the proposed model-based classification method, the effects of several variables on classification accuracy were examined, namely, the initial starting conditions of the model, inter-subject variability, changing the window period of SSVEP data used in the fit, and classifying as well as altering the sets of data used to train the SVMs.

The classification accuracy of the model-based method was finally compared to that of traditional CCA and PSDA methods for various window periods of measured SSVEP data. Each method is assessed by the metrics of accuracy, precision and recall. Potential BCI performance was expressed as relative theoretical ITRs.

The results are presented in chapter 4 and a final discussion and conclusion are to be found in chapters 5 and 6 respectively.

2

Background

As outlined in the introductory chapter, the objective of this study is to investigate the feasibility of an alternative signal classification method that could potentially improve the performance of BCIs. This chapter contextualises and conceptualises the research question while providing a literature review of the applicable BCI and SSVEP classification research.

This chapter briefly introduces VEPs and the classification methods used. The classification methods are generally divided according to feature extraction and feature classification. The CCA and the PSDA methods are introduced along with a time-domain method called stimulus-locked intertrace correlation (SLIC). SLIC is introduced as it gives a context to exploring the classification of SSVEPs, using window periods, in the time-domain.

The idea of model-based feature extraction is outlined and the chosen method of SVM classification is explained in the context of classifying sets of features. The features in this study are made up of various model parameters that are output from the model of VEPs used. The model and its parameters are then introduced. More in-depth background information can be found in the various appendices. A brief introduction to the metrics of accuracy, precision and recall are covered as they form the basis for the comparative assessment.

Following this background chapter, chapter 3: Experimental Methodology describes implementations of the concepts covered, and details about how the hypotheses were tested.

2.1 Visually evoked potentials

2.1.1 VEP and SSVEP responses

VEPs are classified by the stimulus that elicits them. This study uses reversing checkerboard stimuli which, if presented for a single reversal, generate a pattern-reversal VEP. This response is clinically distinguishable by its negative N75 and N135 peaks, with a prominent P100 peak (**Figure 2-1**). Since the specific locations of these peaks vary in time and amplitude from subject to subject, these notations merely serve as a guide for identification. In single trials the peaks may not be very clear and an average of multiple VEPs is recommended to clearly view these features (Odom et al., 2010).

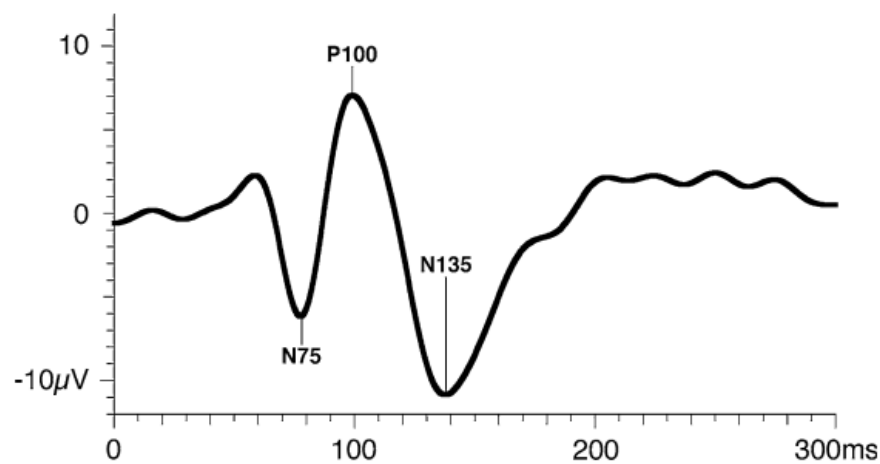


Figure 2-1: A pattern-reversal VEP elicited by a single reversal of a checkerboard stimulus. The main characteristics of the response are two negative peaks at 75 ms (N75) and 135 ms (N135) and a positive peak at 100 ms (P100) (Odom et al., 2010).

The SSVEP response is characterised by an initial VEP, followed by the signal reaching a steady state and oscillating at a fixed frequency. It is therefore best viewed in the frequency domain or in the time-frequency domain, owing to the repetitive nature of the signal. In addition, other brain signals can appear as noise in the measured signal as a result of the additive effect at the electrode sites, noise that can be hard to distinguish from the signal of interest as EEG noise contains multiple frequency components (Vialatte et al., 2010).

The SSVEP response is modulated by the stimulus frequency and luminosity. A change in stimulus frequency will elicit a change in the SSVEP to the frequency of stimulation, while an increase in contrast or luminosity will increase the amplitude of the SSVEP. A stimulus reversing with a higher contrast will be easier to identify because the steady state will be more distinct from background EEG activity (Nishifuji et al., 2009). Increasing the amplitude will result in the peaks of the steady state portion becoming more obvious to spectral classification approaches. It is because of this that SSVEP based BCIs often use high contrast imagery – presented at very bright levels. Presentation in this way is not always

possible in real world situations, where the light levels vary and the presentation of bright-high contrast images can cause user discomfort (Bieger & Molina, 2010).

The initial VEP of an SSVEP is not a traditional VEP: during the period in which the initial VEP appears, the stimulus is still flashing, so it appears as a noticeably large VEP, but it contains information about the frequency of the stimulus as each stimulus reversal induces a new VEP additional to the initial VEP (Zhu et al., 2010). **Figure 2-2** shows a typical EEG recording of an SSVEP response, as well as its associated frequency domain signal.

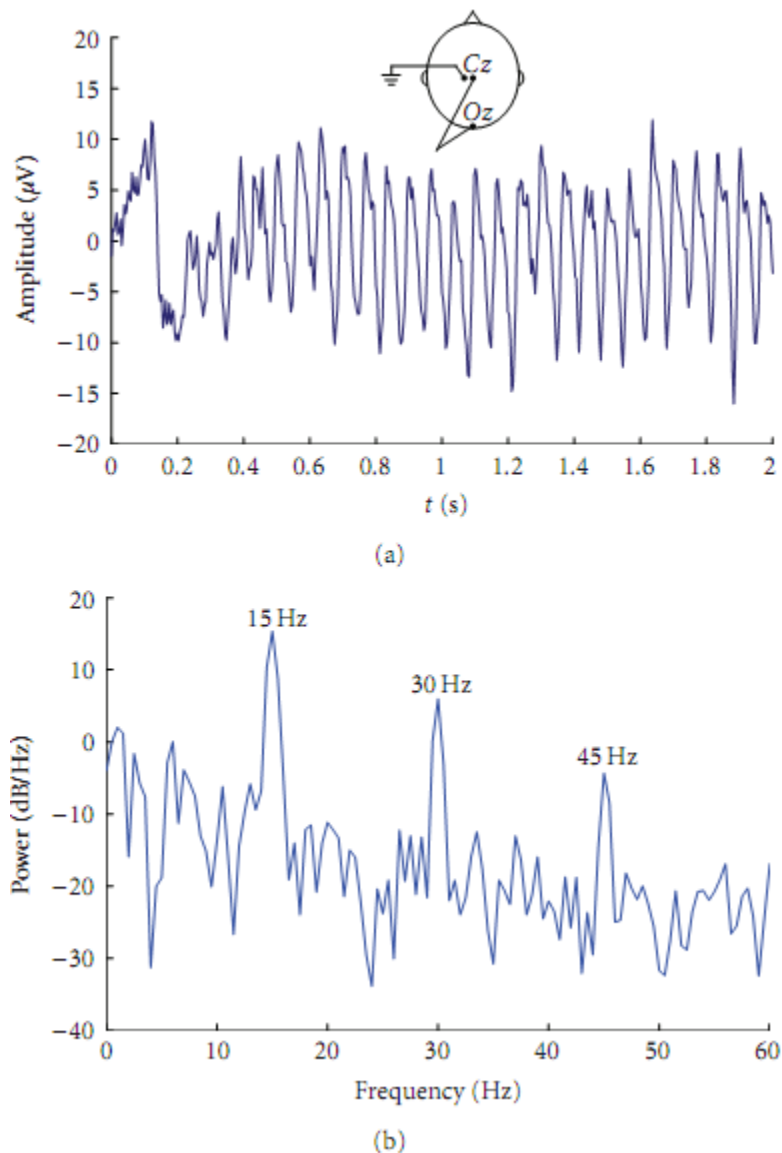


Figure 2-2: The SSVEP response to a 15 Hz stimulus seen in the time-domain (a). Note the initial VEP from 0 to 0,25 s before the signal becomes steady and repeated. The major peaks, P100 and N135, can be seen, although they are slightly delayed. In addition, the steady state of the 15 Hz stimulus becomes evident just after 0.2 s, adding to the response. (b) Shows the same SSVEP response in the frequency domain. Note the characteristic spike at frequency 15 Hz and its harmonics at 30 and 45 Hz (Zhu et al., 2010)

2.1.2 SSVEPs as repeated VEPs

SSVEPs can be explained by the super-position of continued transient VEP responses to stimuli onsets. At every repeat of a flashing stimulus a new VEP is generated. This was explored by Capilla et al. (2011), who constructed a model of SSVEPs using repeated VEPs. The repeated VEPs were attenuated to account for various non-linearity and adaptation phenomena associated with continued steady-state response generation.

Capilla et al. (2011) derived a base VEP response for their subjects from the N75 and P100 peaks presented during a single flash, which was averaged from multiple trials. Following this, the measured VEP was chained together and a synthetic SSVEP response was created (**Figure 2-3**). They found that the synthetic response mirrored the recorded SSVEPs in both the time and the frequency domains (**Figure 2-3**):

“... as we have shown in this study, synthetic steady state responses generated from the linear superposition of transient responses occurring periodically show the same waveform and spectral pattern that characterize the driving phenomenon.” (Capilla et al., 2011: page 11)

Their model operates on time-domain data, as does the model-based approach that is the focus of the present study.

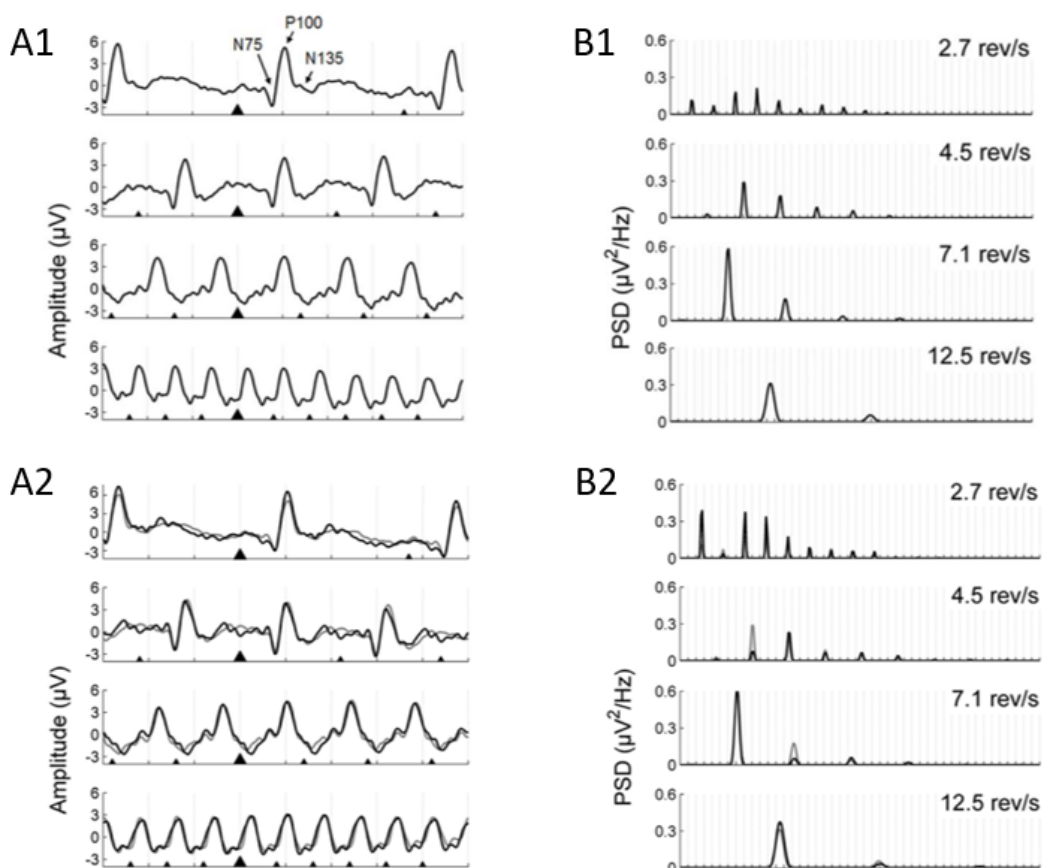


Figure 2-3: (A1) Recorded SSVEP waveforms and (A2) synthetic SSVEPs at various frequencies of stimulation. The characteristic peaks at N75, P100 and N135 are indicated. The black triangle indicates stimulus onset. (B1) and (B2) indicate the corresponding frequency responses as assessed by the PSDA method (Capilla et al., 2011)

2.1.3 SSVEP-based BCIs

The term “BCI” was coined by Vidal (1973), who described the possibility of using the classification of EEG to decode user intent. Vidal (1977) would go on to provide evidence that it was possible to use VEPs as an input to BCI. From there the SSVEP BCI evolved.

The original reasoning behind the use of VEPs remains at the forefront of their use in current-day BCIs. SSVEPs offer comparatively reliable and repeatable signals which are not dependent on the unique higher-order processing of a subject. Furthermore, they do not require much training, if any at all, to obtain accuracies above 70% (Guger et al., 2012). SSVEP BCIs can also achieve high ITRs (Above 70 bits/min) with little training (Zhu et al., 2010).

The main paradigm of SSVEP classification in BCIs is to use the steady-state portion of the SSVEP to inform a classification choice (Vialatte et al., 2010). This is generally induced by presenting the subject with either a single stimulus or multiple stimuli flashing at a known frequency.

Using various signal processing methods the recorded EEG is identified in relation to the stimulus type, generally frequency is used. It is possible to modulate SSVEPs with different configurations of shapes and colours, these do not have a very high modulatory effect of the amplitude and are hard to detect, and unreliable for classification (Vialatte et al., 2010) due to this most BCIs use some form of frequency encoded stimulation. SSVEPs can be induced by multiple stimuli at once, as the additive effect of the stimuli contrast, brightness colour and frequency can create distinct SSVEPs this phenomenon can be used to encode information for use in BCIs (Srihari et al., 2006). This research however only explores SSVEP classification induced by a single flashing stimulus.

While it is possible to increase the amplitude, and thus classification performance of an SSVEP BCI. Increasing the contrast and luminosity can add a certain level of user discomfort (**Figure 2-4**). It is also not always possible to achieve these levels in everyday situations, and most SSVEP research is conducted in controlled laboratory conditions, because of this a method that is able to classify SSVEPs using different features would add to the performance of BCIs in general (Bieger & Molina, 2010).

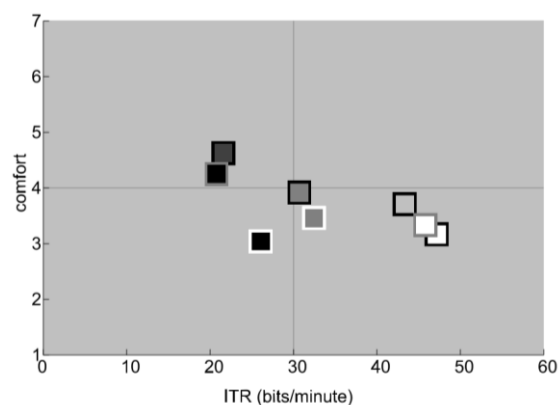


Figure 2-4: Luminance contrast vs comfort relative to performance measured in ITR. It is seen that higher-contrast systems perform better with spectral methods, owing to the increased SSVEP amplitude. Yet they are rated as more uncomfortable to the end user (Bieger & Molina, 2010).

2.1.4 10–20 system of electrode placement

Multiple standard systems exist for placing and naming electrode sites on the human scalp. The most commonly used system is the international 10–20 system that derives its name from the fact that it uses 10% and 20% spacing between electrode sites. This allows the system to accommodate various head sizes. It also uses anatomical landmarks for a constant reference point to the spacing of the electrodes. These include the nasion, inion and preauricular points (Herwig et al., 2003).

In this study, the Electrical Geodesics Inc. net was used to acquire data. The net uses an electrode reference system that is built on the 10–20 naming standard but accommodates configurations of 256- and 128-channel sensor nets (**Figure 2-5**).

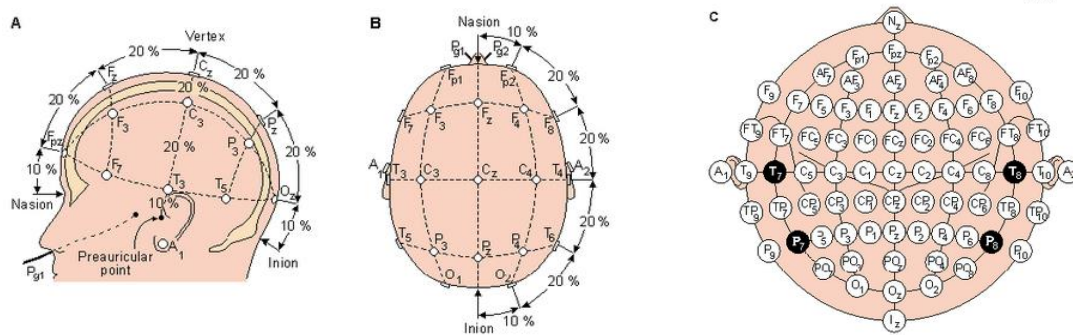


Figure 2-5: The classic 10–20 electrode placement system is seen in A and B above. C shows the extended 10–20 system with 70 electrode sites (BCI2000, 2012).

2.1.5 Information transfer rate (ITR)

Shannon’s ITR can be used to assess the performance of a BCI system, and it extends in general to any communication system (Schreiber, 2000). It is measured in bits/min and this measurement depends on three factors: speed, accuracy of classification and the number of targets (stimuli) simultaneously presented (Volosyak, 2011). It quantifies the number of commands the BCI can successfully execute, where a command is a single event that the BCI can trigger. In the case of a BCI that presents a keyboard of 26 buttons, it will have 26 commands; in the case of a wheelchair, it may have five commands: forward, back, left, right and stop.

For a BCI system with N equally probable commands, s commands performed per minute and a probability p of a true positive decoding of each command, the ITR is given by Equation 1. Given a fixed p , the ITR increases for an increasing number of commands N (Vialatte et al., 2010).

Equation 1: Formula for Shannon’s ITR

$$ITR = s \left[\log_2(N) + p \log_2(p) + (1 - p) \log_2 \left(\frac{1 - p}{N - 1} \right) \right] \quad (1)$$

2.2 SSVEP Classification Methods

A large number of SSVEP classification methods have been developed for use in BCIs. The purpose of an SSVEP classifier is to identify which stimulus was viewed by the subject – this is identified by the frequency at which the stimulus was presented. There are three stages to classification: (i) signal pre-processing, (ii) feature extraction and (iii) feature classification. In some methods (such as PSDA), feature extraction and classification are combined (Mason et al., 2007).

Once SSVEP data have been pre-processed, features can more easily be extracted and classified. Three methods of feature extraction are considered in this study:

1. Power spectral density analysis (PSDA)
2. Canonical correlation analysis (CCA)
3. Model-based classification.

Each method is outlined below. In addition, the SLIC time-domain classification method is outlined because it serves as the motivation for exploring a time-domain model approach.

After features have been extracted, they are classified in a feature-classification stage. Since we used support vector machines (SMV) to classify the model fit output, a brief introduction is provided.

2.2.1 Pre-processing

The purpose of pre-processing is to reduce the signal-to-noise ratio and remove unwanted artefacts. Most pre-processing uses frequency filtering, such as removing the 50 Hz band of electrical noise or unwanted higher- and lower-frequency EEG activity. SSVEPs are also generally filtered by a bandpass filter that is set to the range of stimulus frequencies presented. Spatial filtering can also be applied in which a linear combination of electrodes is used to reduce unwanted noise (Liu et al., 2013). Filtering is also used to remove distinct artefacts in the signal, such as those resulting from eyeblinks or movement of the subject.

2.2.2 Feature classification of power spectral density analysis (PSDA)

The PSDA method involves taking the power spectral density of a segment of EEG data. This can be done from a single electrode, or from multiple electrodes (Ortner et al., 2011). It is the most common approach to SSVEP classification because it is computationally cheap and easy to implement. In the study of BCIs it offers a very quick solution to the challenge of online signal classification (Liu et al., 2013).

The SSVEP data are broken down into their frequency components using fast Fourier transforms. The power of each frequency present in the SSVEP can be plotted as an amplitude vs a frequency plot (**Figure 2-6**). A threshold is applied to the data and an SSVEP is classified as occurring when a frequency has power above the selected threshold (Xia et al., 2013).

The longer the window period used, the more distinct the frequency components become. When the window period is decreased below 2 s, the frequency components of the presented stimulus start to decrease, to the point where background EEG data can add unwanted noise producing mis-classifications (**Figure 2-6**). This method is also susceptible to noise from other stimuli that may be simultaneously presented to the subject: if more than one stimulus is presented simultaneously, multiple SSVEP responses can be induced, which may either mask the frequency component in question or trigger the classifier's threshold (Zhang et al., 2013; Xu et al. 2014).

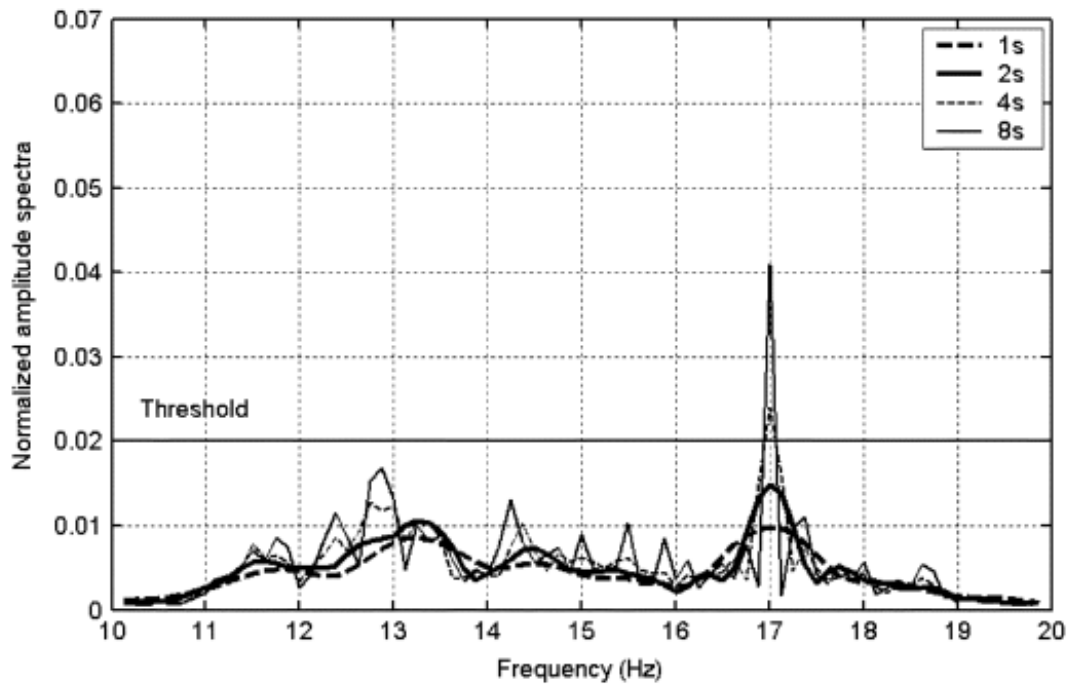


Figure 2-6: Normalised amplitudes of frequency components are shown for a single subject's SSVEP for to different window periods. Window durations greater than 2 s are required for classification (Wang et al., 2006).

Fourier methods are limited by the window used to measure the data and therefore require a window period of more than a second to achieve reasonable results. The window period selected also adds to the delay between command inputs achievable in a BCI context (Hansson-Sandsten, 2010). **Figure 2-6** shows that a window period of at least 4 s is required before user intent can be detected, which implies a 4 s delay before a command can be actioned.

2.2.3 Feature classification of canonical correlation analysis (CCA)

Canonical correlation analysis (CCA) is another method for classifying SSVEPs with accuracies that outperform PSDA methods given the same window period of SSVEP data (Lin et al., 2007). CCA exploits the fact that the frequency of the presented stimulus is known. It requires that reference signals, normally sinewaves, be generated at the same frequency at which the stimuli are presented (Lin et al., 2007). CCA draws a statistical conclusion by comparing the reference frequency signal to that of measured SSVEP data.

Measured EEG data are linearly combined with a weighting coefficient for each measured channel. The resulting signal is then correlated with a linear combination of the reference signals, each with their own weighting coefficients. When a strong correlation between these two linear combinations is found, the weighting coefficients are examined. The reference signal with the largest coefficient is assigned as the frequency of the measured SSVEP. Essentially, the reference signal that is closest to the recorded data is selected (**Figure 2-7**).

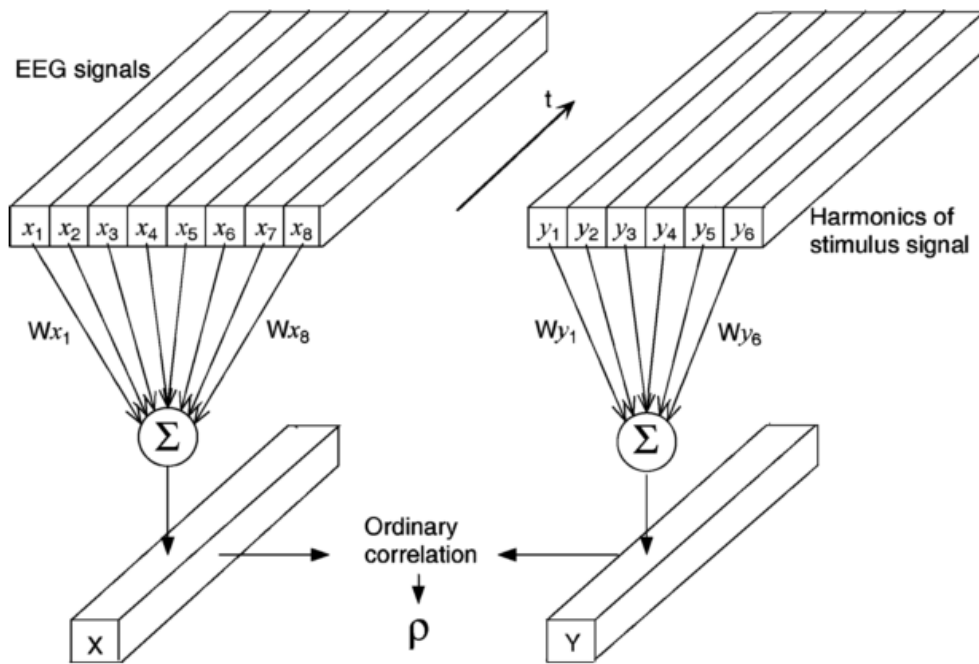


Figure 2-7: An illustration of CCA processing. EEG signals $x_1, x_2 \dots x_8$ are compared to reference signals $y_1, y_2 \dots y_6$, which have the same frequency as the presented stimuli. Linear combinations of each are generated such that $X = (W_{x_1} + \dots + W_{x_8})$ and $Y = (W_{y_1} + \dots + W_{y_6})$. X and Y are compared until they have a high correlation. At this point the largest coefficient in the W_{y_n} set indicates the frequency of the SSVEP (Lin, et al., 2007).

As with PSDA, the accuracy of the CCA method is affected by the window period of SSVEP data used in the classification. CCA outperforms PSDA classification methods, given the same window period. A window period of 4 s is required to achieve classification accuracies above 90% (**Figure 1-7**).

There are many variants of the CCA method, these could involve using multiple numbers of electrodes or various filtering methods before the EEG is analysed or constraining the coefficients or weighting the components differently (Bin et al., 2009). CCA can also be used in combination with PSDA for an improved classifier selection (Liu et al., 2013). This study uses the CCA method proposed by Lin et al., (2007) as a baseline for comparison, this is the standard method described above.

2.2.4 Extraction and classification of time-domain features

Time-domain classification of SSVEP responses has been successfully implemented in SSVEP BCIs (Luo & Sullivan, 2010). Furthermore, time-domain classification has been shown to outperform spectral approaches (Manyakov et al., 2010; Abbasi et al., 2015; Müller-Putz et al., 2008).

This method of classification relies on detecting features in the time-domain signal (Müller-Putz et al., 2008). One such method is SLIC (Luo & Sullivan, 2010a). This method exploits the fact that when a flashing stimulus is attended by a subject, initial and repeated VEP components can be found.

SLIC takes a recorded EEG segment containing SSVEP data and divides it into repeated windows, starting at time 0 s (stimulus onset). The same EEG segment is assessed with varying repeated window periods. These are then averaged to create a VEP response. The window periods used are dependent on the frequency of the stimulus being presented (Manyakov et al., 2010).

The window periods are chosen such that a full period of oscillation induced by the stimulus can occur. This averaged response contains information about the stimulus being presented, and is used as input a linear discrimination analysis (LDA) classifier **Figure 2-8** shows an example of a four-class BCI with stimulus frequencies at 10, 12, 15 and 20 Hz. The subject in this instance is attending the 20 Hz stimulus.

This method looks for the repeated VEPs of which the SSVEP is composed by assessing a single channel using different stimulus-locked windows.

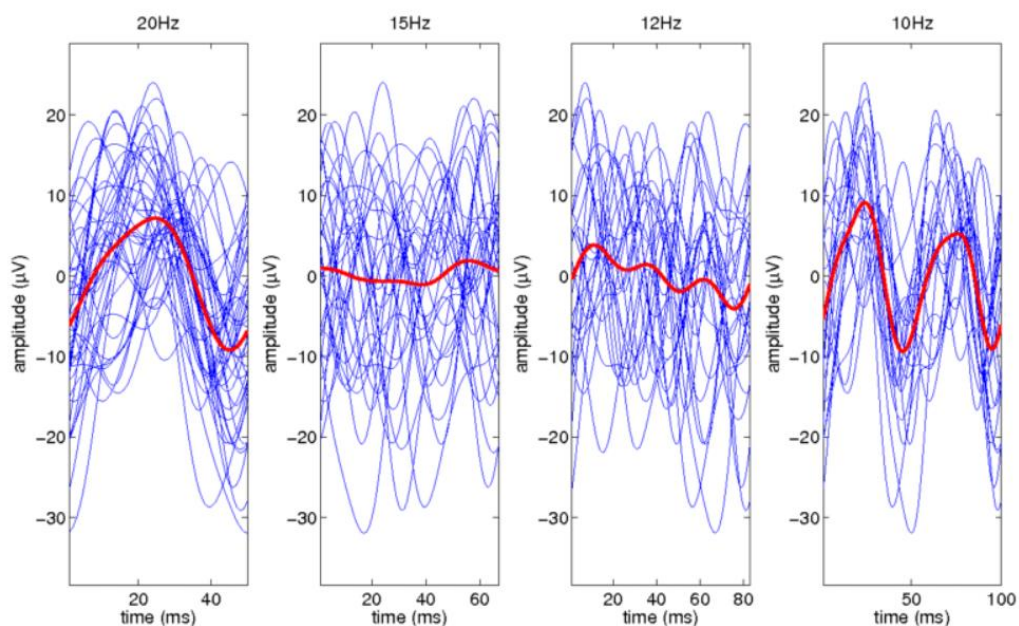


Figure 2-8: Individual VEP traces for repeated window periods are shown in blue and their averages in red. Each stimulus frequency will produce a characteristic response in the time domain for a distinct repeat window period. For example, in the case of a 10 Hz stimulus, data averaged across multiple 100 ms windows will produce a characteristic response. Since the subject was attending a 20 Hz stimulus in this instance, averaging data over multiple 50 ms window periods produces a characteristic response, compared to the relatively flat responses at 15 Hz and 12 Hz, and the 20 Hz signal evident on the 10 Hz plot (Manyakov et al., 2010)

This set of averaged VEPs is then input into a trained classifier scheme using multiple LDAs, with each LDA trained on the features that present at the various stimuli frequencies. The coefficients of the LDAs indicate the probability of a stimulus being attended (**Figure 2-9**).

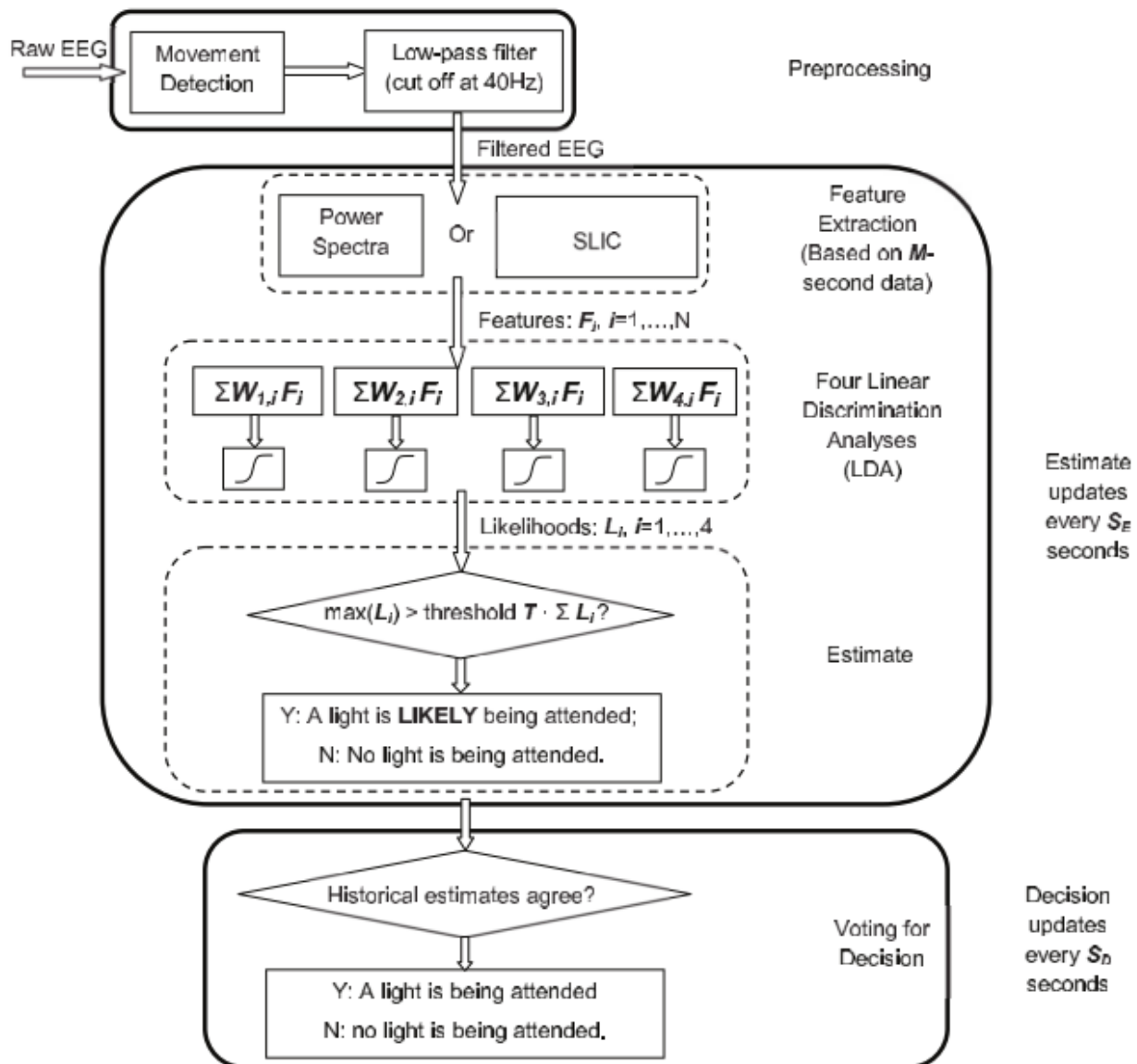


Figure 2-9: A four-class SLIC BCI schema able to distinguish between four different stimuli by employing four LDAs trained on the subject's response (Luo & Sullivan, 2010a)

Although sufficient EEG data is required to generate a useful average of VEP responses, SLIC has been shown to achieve good accuracy with a data segment length of about 1 s (**Figure 2-10**), compared to roughly 5 s of EEG data required for CCA and PSDA methods to achieve a similar degree of accuracy. This enables SLIC to improve the BCI's ITR because fewer SSVEP data are required for classification decisions, which increases ITR. In addition, because this method uses time-domain features and not frequency-domain features, it is less susceptible to spectral noise or to interference induced by multiple stimuli being displayed at once, as well as the refresh-rate limitations imposed by monitors (Abbasi et al., 2015).

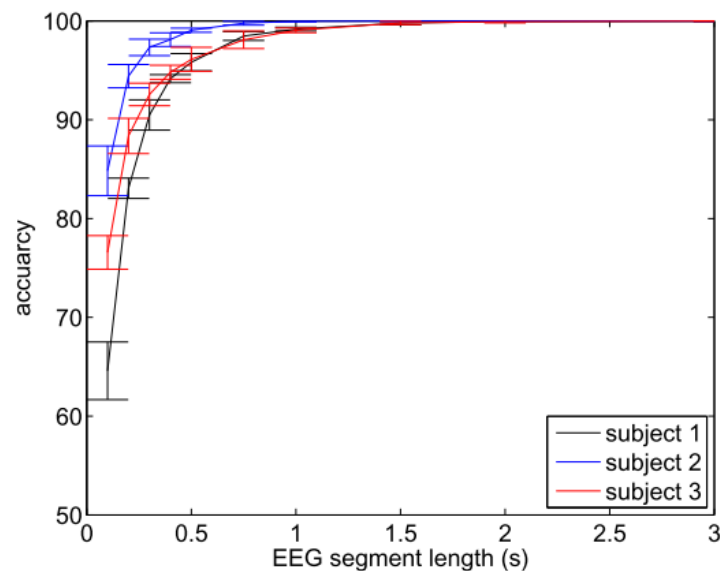


Figure 2-10: Accuracy of SLIC LDA classification for three subjects, given various lengths of recorded SSVEPs (Manyakov et al., 2010)

2.3 Model-based Feature Extraction

Model-based feature extraction involves constructing a mathematical model of a signal in question, in this instance the human SSVEP response. In the ideal case, as a function of a given set of model parameters (P_n), the model is able to output a synthetically generated signal that matches the measured signal exactly. In general there may be some degree in error between the measured and synthetic signals, Equation 2 describes this.

Equation 2: Model based measurement

$$Model_SSVEP(P_n) + Error = Measured_SSVEP \quad (2)$$

From this we are able to associate a given set of model parameters (P_n) with a measured SSVEP. The task of classifying which “user intent” (stimulus) a certain parameter set belongs to is then performed by a multiclass support vector machine. If the parameter space P is separable into uniquely identifiable regions such that a combination of parameters can be linked to unique responses, a classification is able to be made.

2.4 Feature Classification of Support Vector Machine (SVM)

SVM-based classification has been directly applied to SSVEP classification problem in several studies (Jian & Tang, 2014; Singla, 2014; Ioannis Kompatsiaris et al., 2016) as well as to transient VEPs (Yoshimura & Itakura, 2011).

SVMs are supervised machine-learning classifiers; they are binary classification systems able to distinguish between two class types. SVMs function by describing the classes as a function of the parameters that constitute each class. In the two-dimensional example (**Figure 2-11**) there are two parameters, P1 and P2. A hyperplane is fit onto separate the classes; in this instance, Class 1 is all data points that are above the line H [Class 1 = $(P1, P2) > H$]. Class 2 is all points below line H [Class 2 = $(P1, P2) < H$]. This principle can be extended into a multi-dimensional parameter space of n dimensions. If we use all data presented for classes 1 and 2 as training data for the SVM and we are given a new data point D1, the SVM will classify D1 as Class 2, as it satisfies $D1 < H$. SVMs can be extended to multiclass classification systems by chaining together multiple binary SVMs (Hsu & Lin, 2002).

The hyperplane (H) used to separate classes can be mapped to various shapes, from linear to n^{th} order polynomials. This is referred to as the kernel function and it dictates the shape of the hyperplanes. Various kernels have been used in SSVEP identification: quadratic (Singla, 2014), radial biased functions (Yoshimura & Itakura, 2011) and polynomials (Singla et al., 2014b). An 8th-degree polynomial was chosen, because the SVMs were being trained on parameters output from the model and not directly on the recorded SSVEP data. A high-order polynomial is able to accommodate the non-linearity of parameters (Ben-Hur & Weston, 2010), as is expected from the non-linear behaviour of the model's oscillators.

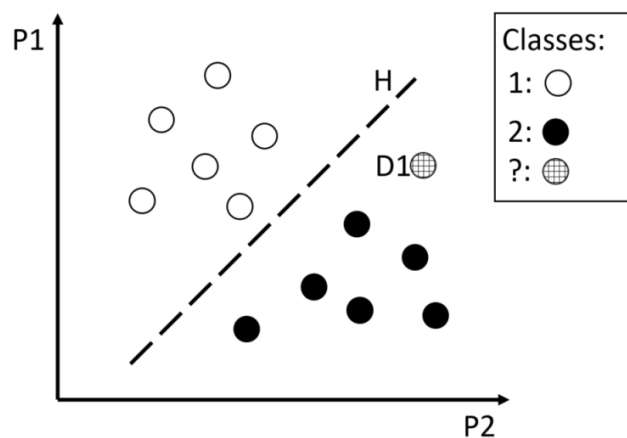


Figure 2-11: Representation of a binary SVM. Two classes of data exist: Class 1 and Class 2, each a function of the parameters P1 and P2. When the SVM is trained, it fits a hyperplane (H) between the two classes such that they are maximally separable. Given an unseen data point D1, the SVM will classify it in relation to the hyperplane that separates the classes. In this instance, D1 would be identified as belonging to Class 2.

2.5 Assessing Multiclass Classifier Performance

In this study, the output of each of the classification methods is constrained, since they are able to classify the input data only as belonging to one of a predefined set of classes. Given all the frequencies under assessment, the classifiers will label an unseen SSVEP response as one of the following: 6 Hz, 6.5 Hz, 7 Hz, 7.5 Hz, 8.2 Hz, 9.3 Hz, 10 Hz.

Three measures of performance are used to describe the overall ability of a multiclass classifier to classify data. These measures, accuracy, precision and recall, are derived using a confusion matrix.

A confusion matrix for a hypothetical three class classifier is shown in Table 1. Each element in the confusion matrix is an integer number representing how many times a particular classification (or misclassification) occurred. The rows represent the actual class value, while the columns represent the predicted class value. The values TP(A), TP(B) and TP(C) show true positives, i.e. TP(A) shows the number of times class A was correctly classified as being class A, this is the True Positive occurrence.

The number of false negatives (FN) for a class is given by the sum of all other values (excluding the TP value) in the corresponding row of the matrix. The total number of false positives (FP) for a given class is the sum of all other values (excluding the TP value) in the column corresponding to the class. The true negative for a given class is given by the sum of all values in the matrix, excluding the row and column to which that class belongs (i.e. for A: $TN = TP(B) + TP(C) + E(BC) + E(CB)$).

Table 1: Generic confusion matrix for a three class classifier.

	Predicted class			
		A	B	C
	A	TP(A)	E(AB)	E(AC)
	B	E(BA)	TP(B)	E(BC)
	C	E(CA)	E(CB)	TP(C)

When assessing a multiclass classifier it is important to assess it in the context three measures, accuracy, precision and recall rate. The singular metric of accuracy can over represent the performance of the classifier. In the context of BCIs this is important as the aim is to have the highest possible accuracy, precision and recall because this leads to correct predictions of user intent and higher ITRs.

2.5.1 Accuracy of classification

For each class, accuracy is defined as the ratio of the sum of the True Positives (TP) and True Negatives (TN) to the sum of all positive (TP+FP) and all negative cases (TN+FN):

Equation 3: Class accuracy of classification

$$\text{Class accuracy} = (TP+TN) / (TP+FP+TN+FN) \quad (3)$$

An overall accuracy of the classifier, across all classes, is calculated as the sum of the diagonal over the sum of all occurrences in the matrix. This accuracy in the context of BCIs reflects the number of TP as a ratio to all other attempts. It is referred to in this research as *Prob(TP)* – as it is effectively the probability of a TP occurring.

2.5.2 Precision of classification

Precision represents the proportion of classifications assigned to a class that are correct. For each class in the confusion matrix this is the ratio of TPs to the total number of times the class was predicted, i.e. the sum of all values in the column. Precision quantifies how probable it is that the classification was correct.

Equation 4: Class precision of classification

$$\text{Class precision} = TP / (TP + FP) \quad (4)$$

2.5.3 Recall of classification

Recall is a counterpart of precision and is also referred to as sensitivity or true-positive rate. It represents the number of times a given class was classified correctly, and is defined by equation 5. Recall quantifies the likelihood that a class was misclassified.

Equation 5: Recall of classification

$$\text{Class recall} = TP / (TP + FN) \quad (5)$$

2.6 A Model of Visually Evoked Potentials (VEPs) as a Feature Extractor

VEPs can be modelled as an oscillation with perturbations such as changes of amplitude, frequency and phase (Kremláček et al., 2002). Kremláček et al. (1999) proposed a mathematical model of transient VEPs structured around three chained damped harmonic oscillators, a model that is replicated and explored in this research. An outline of this model is given below. Models constructed using oscillators have previously been used with to successfully model VEPs (Wang et al., 2013). For example, Wang et al (2013) used forced Van der Pol oscillators to describe a VEP waveform.

2.6.1 Underlying principles of the model

The Kremláček et al. model is based on the principle that different neuronal groups along the visual pathway can be activated sequentially and that each contributes jointly to a final VEP measured over the occipital cortex (Kremláček et al., 1999).

Kremláček conducted both principal component analysis (PCA) and independent component analysis (ICA) on VEP data. The results indicated that three components, each representative of neuronal groups along the visual pathway, described the data optimally (Kremláček et al., 1999). Based on the results of the PCA and ICA, each component was modelled as an oscillator. The oscillators were connected in series, with delays introduced between consecutive oscillators, thereby mimicking the sequential activation of the visual pathway.

2.6.2 Model implementation

The model was built by Kremláček in MATLAB Simulink. A schematic representation is given in **Figure 2-12(A)**. Three distinct areas of neuronal groups within the visual system were modelled using the same oscillator scheme (configured individually and labelled OSC1, OSC2 and OSC3, respectively, in **Figure 2-12(B)**).

Each oscillator represented one of the following: primary visual cortex, secondary visual cortex and, finally, higher-order processes, possibly associated with perception (Kremláček et al., 2002). T_1 , T_2 and T_3 describe the delay in activation as the signal propagates through the visual cortex. For each neuronal group a delay time T_n transpires before the following oscillator is activated. Finally, each oscillator's contribution to the final signal is weighted by coefficients K_1 , K_2 and K_3 . Since the SSVEP recorded on the scalp represents an additive combination of all electrical activity below the electrode site, the model computes the weighted sum of the components.

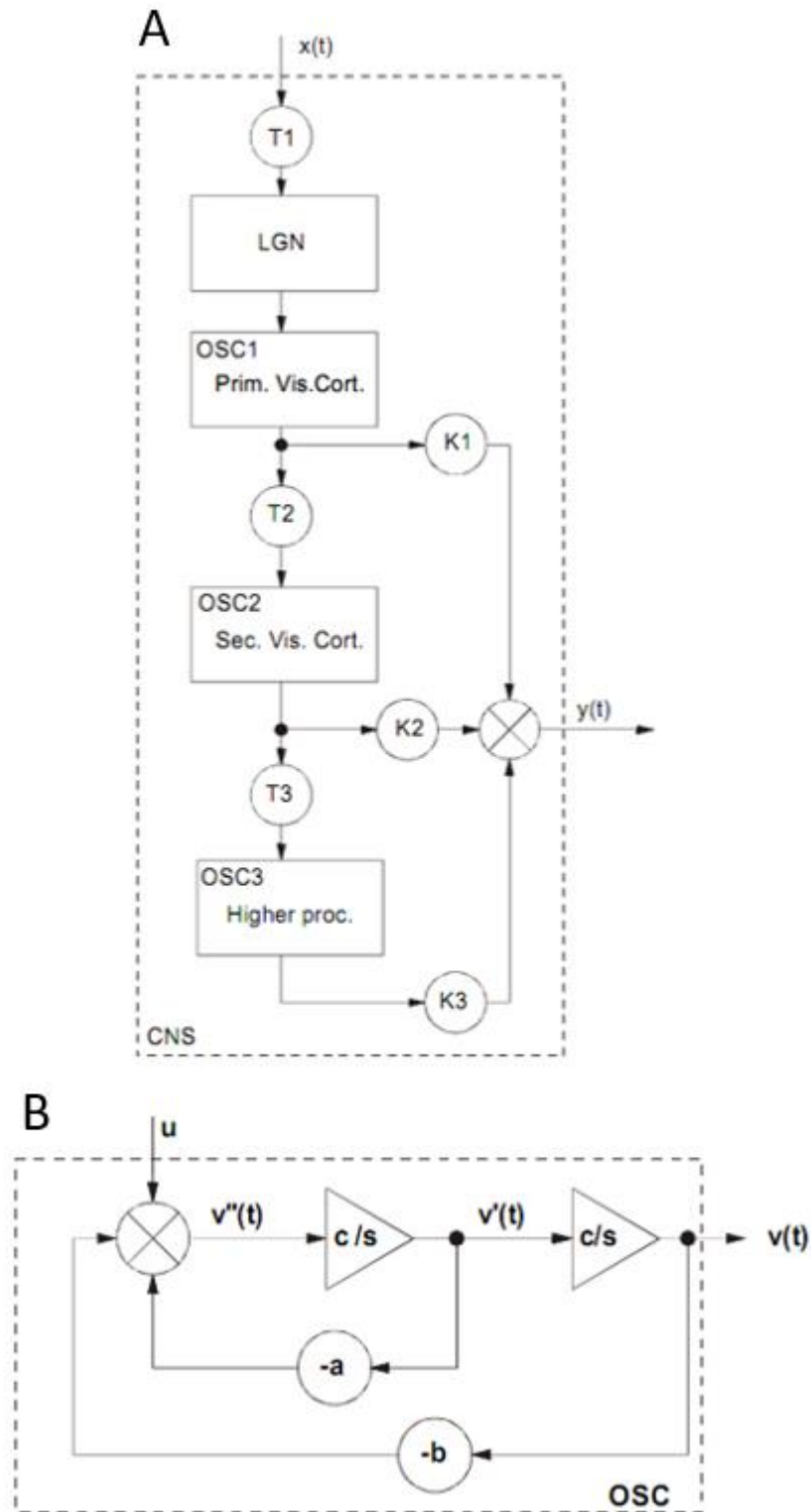


Figure 2-12: (A) Kremláček et al.'s (1999) model, as originally published, for visually evoked potentials. $y(t)$ represents the model VEP output while $x(t)$ is the initial input stimulus. Delays are indicated by T1, T2 and T3; weighting factors by K1, K2 and K3. **(B)** A single-oscillator schema, these represent blocks OSC1, OSC2 and OSC3 in (A). Coefficients $-a$ and $-b$ in (B) control the damping factor and the frequency of oscillation, while $v(t)$ represents the output to the next stages. c is a constant of integration.

2.6.3 Model parameters (feature descriptors)

The model has 12 unique parameters that make up its parameter space. These parameters are categorised into three different types according to their position and function in the model. The three parameter types are:

1. **Frequency of oscillation:** There are three sets of (a , b) parameters, one for each oscillator. They describe the frequency of relaxed oscillations and the damping factor of OSC1, OSC2 and OSC3.
2. **Time delays:** Onset delay before each oscillator is activated by the previous oscillator; indicated by parameters T1, T2 and T3 in **Figure 2-12(A)**.
3. **Weighting coefficients:** These amplify or attenuate the oscillator outputs before they are jointly summed; indicated by parameters K1, K2 and K3 in **Figure 2-12(A)**.

For a given simulation of the model, the parameters are grouped as a set, such that the parameter set (P_n) represents the model's state at a given time:

Equation 6: The model parameter space

$$P_n = [T1, T2, T3, K1, K2, K3, \text{Osc1a Osc1b, Osc2a, Osc2b Osc3a Osc3b}] \quad (6)$$

a, b parameters and the frequency of oscillation

The parameters a and b together describe the frequency of oscillation for each oscillator as well as the damping factor that attenuates it. Each parameter name includes the oscillator with which it is associated: (Osc1a Osc1b, Osc2a, Osc2b Osc3a Osc3b).

T parameters and the time delay

Each oscillator models a separate system along the visual pathway. The time delays (T1, T2, T3) account for the delay in signal propagation between these systems. In effect, this models the time taken for a neuronal group to propagate an EEG signal.

K parameters as weightings

The scaling factor K , applied to each oscillator's output, represents the contribution of the oscillator as a cortical source to the measured scalp potential. The gains are linear scalar values.

2.6.4 Model output

Figure 1-8 shows the output of Kremláček's model, after it has been fit to averaged measured VEP data. Each individual oscillator's contribution is also, along with the final modelled VEP.

2.6.5 Model validation

Kremláček et al. (2002) gathered data from four subjects using three VEP stimulation paradigms, namely, fast- and slow-motion onset stimuli, as well as pattern-reversal stimuli. Twenty responses were recorded per subject. The current investigation focuses on steady-state pattern-reversal stimulation; as a result, only the P-VEP (pattern-reversal VEP) data from Kremláček et al. (2002) is discussed.

Kremláček et al. (2002) took a grand average of each subject's VEPs on which the model was fit. It was found that the initial model parameter values had a large influence on the model's performance. As a result, initial starting points were selected via a direct search.

The goodness of fit was reported as a normalised root mean square error (NRMSE) value. In all cases the NRMSE was below 13% (Kremláček et al., 2002). It should be noted that in this study a goodness-of-fit indicator was used which was equal to 1 minus NRMSE – this is explained in the methodology section.

3

Experimental methodology

The overall objective of this study was to establish the feasibility of using a time-domain based classification scheme of SSVEP data, required to encode intent in a BCI system. The proposed method was compared to two widely used frequency-domain based classification methods with particular interest in the length of window period required to facilitate successful classifications.

All the experiments carried out in this study were approved by the Human Research Ethics Committee (HREC) of the Faculty of Health Sciences of the University of Cape Town (UCT) (Ref. number 073/2012).

3.1 EEG Data Acquisition

3.1.1 Recording equipment and electrode placement

EEG data were recorded using an EGI 300 Geodesic EEG System (GES 300) with a 128-channel HydroCel Geodesic Sensor Net (HCGSN) and a sampling rate of 512 Hz. Electrodes were spaced in accordance with the 10–20 system (**Figure 2-5**). Data captured with electrodes O1, Oz and O2 were used during analysis, with a reference at electrode Fz. All the other channels were recorded but not processed.

Electrode impedance was kept below 6 kOhm. The impedance of the sensor net was checked at the start and end of the recording using the calibration and impedance recording functionality of the GES 300. If a recording channel had increased to above 6 kOhm of characteristic impedance by the end of the recording session, the affected datasets were discarded. The net was then adjusted, additional conductive saline solution was applied, and the measurement session was repeated.

3.1.2 Stimulus design and presentation

Three different types of stimuli were designed, namely a VEP-inducing checkerboard stimulus and two different SSVEP stimuli – pattern-reversal checkerboards and two white flashing LED lights.

The subject was positioned 0.5 m from the CRT monitor on which the checkerboard VEP and SSVEP stimuli were presented. A CRT monitor with a refresh rate of 100 Hz was used because a periodic flashing stimulus requires a refresh rate of at least double the stimulus frequency to avoid distortion as per the Nyquist sampling criterion. With a 100 Hz refresh rate, the maximum presentable frequency is therefore 50 Hz.

The recordings were performed in a closed room with the curtains drawn to control the amount of ambient light entering the recording area. A Lux meter was used to ensure that a level of 250 to 300 Lux was maintained during the recording session. This range is similar to that of a normally lit room and just below the brightness of office lighting, which is between 350 and 500 Lux. The light level was controlled to reduce the impact on the recordings and also to represent a configuration similar to that in which a BCI would typically be used.

Stimuli displayed on monitors and point light sources induce different EEG waveforms. An LED light source induces the largest-amplitude SSVEP response, while a checkerboard stimulus is more robust to noise and

stimulates more of the visual pathway. Checkerboard stimuli are therefore more widely used in SSVEP-based BCIs (Vialatte et al. 2010).

Each frame of the alternating achromatic checkerboard stimulus was designed in MATLAB and then compiled as a compressed .avi file. The checkerboard stimulus was presented at a resolution of 1024 x 768 px. In this configuration the stimulus had a field of view of 15 cycles per degree. The CRT monitor was connected to a PC separate from the EEG acquisition system to ensure that the stimulus would be displayed without stutters or discontinuities caused by the processing load experienced by a PC when recording raw EEG data. Anomalies in stimulus presentation would manifest as artefacts in the EEG data (Vialatte et al., 2010).

The checkerboard stimulus had a small centred fixation target, as can be seen in **Figure 3-1** that remained on the screen at all times, providing the subject with a reference point that they were instructed to focus on. This prevented the subject's eyes from drifting when no stimulus was being presented.

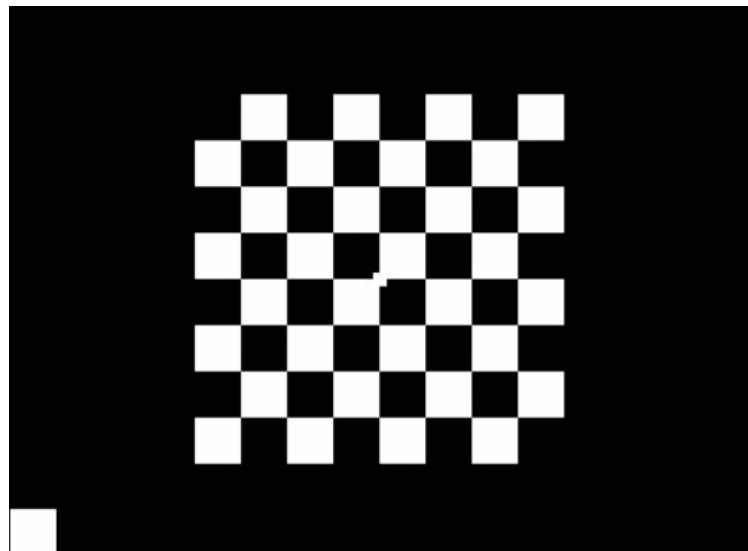


Figure 3-1: VEP and SSVEP checkerboard stimulus. The white block at the bottom left of the stimulus flashed at the stimulus frequency. It was covered at all times by a light-dependent resistor (LDR), which detected changes in light. The white fixation mark is visible in the centre of the screen.

Two white LEDs spaced 10 cm apart and driven by a programmable function generator were used as the flashing light stimulus. During the LED stimulus recordings the two white LEDs were positioned at the centre of the CRT monitor screen, ensuring the same distance and field of view to the subject as the checkerboard stimuli.

The SSVEP stimuli were designed to oscillate in the low-frequency region where the amplitude of the SSVEP response is greatest (**Figure 3-2**). Notably, a frequency around 12 Hz produces the greatest amplitude SSVEP response. This increased amplitude response was motivation for the chosen frequencies of stimulation.

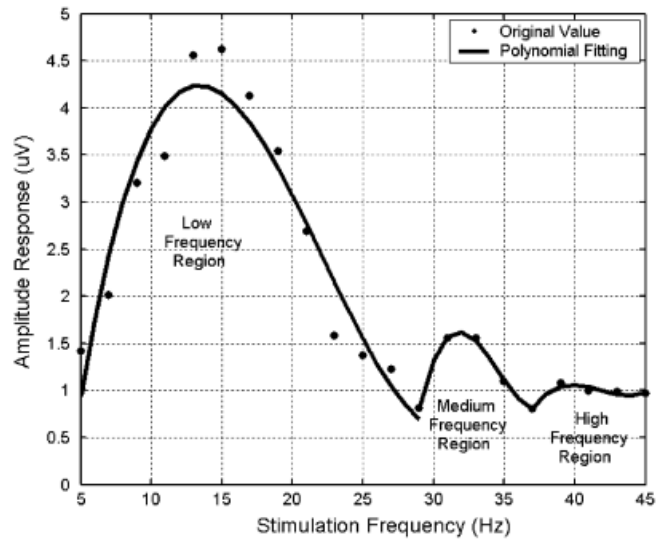


Figure 3-2: SSVEP response amplitude as a function of stimulation frequency (Wang et al., 2006).

3.1.3 Synchronisation of stimulus display and EEG recording

A light-dependent resistor (LDR), as well as one of the LED stimuli, were connected both to the CRT monitor and to the EEG recording device. At every presentation of the stimulus, the LDR was triggered and the EEG data tagged. In this way any delay or variation in stimulus frequency could be detected and accounted for in the analysis. The white block used for this synchronisation (**Figure 3-1**) was completely covered by a 3D printed adapter that held the LDR in place on the screen; this ensured that it would not be visible to the subject and thus not add additional SSVEP stimulation.

3.1.4 Exploratory data set

The implementation and feasibility of the proposed model was initially assessed using data recorded in a single subject (male, 24 years) at the occipital lobe while observing the stimuli described above, designed to elicit either a distinct VEP or SSVEP response. The VEP response was evoked by the single presentation of a flashing checkerboard stimulus. Each VEP stimulus was presented to the subject 20 times, with a rest period between each exposure of 15 s. SSVEP responses were recorded separately for both single-frequency flashing light and pattern-reversal checkerboard stimuli. Stimuli were presented at various frequencies around 12 Hz, specifically 1, 2, 4, 9, 10, 11, 13, 16 and 19 Hz (**Figure 3-3**). Each frequency was presented 5 times, for 5 s duration, with a 15 s break between presentations, for a total of 45 recordings per SSVEP stimulus.

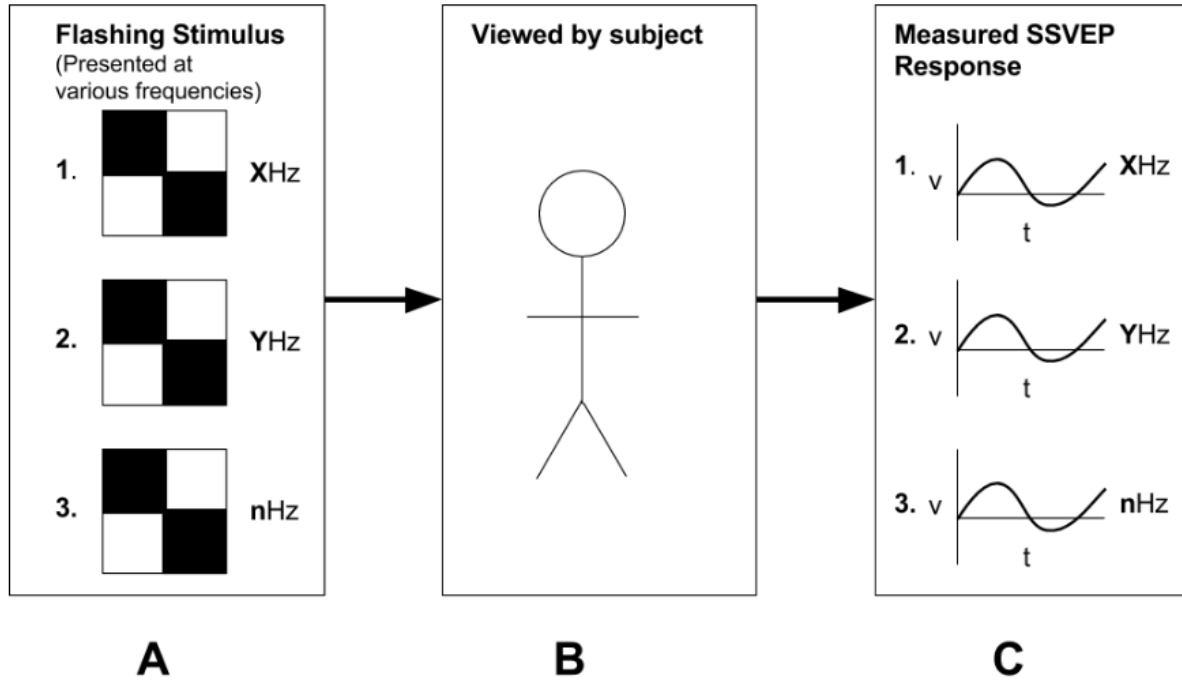


Figure 3-3: (A) Checkerboard stimuli flashing at various frequencies were presented to the subject one at a time. (B) Each stimulus frequency induces a unique SSVEP response in the subject viewing the stimulus. The SSVEP response is recorded and saved (C).

3.1.5 Main data set used for hypothesis testing

The performance of the proposed time-domain model for feature extraction and classification was investigated using published SSVEP data (Vilic, 2015). This data set comprised SSVEP data from four subjects in response to single-frequency flashing checkerboard stimuli presented at 6, 6.5, 7, 7.5, 8.2, 9.3 and 10 Hz, respectively. A single SSVEP trial consisted of 5 s of EEG data recorded while the subject viewed a fixation cross at the centre of a blank screen, followed by 10 s of stimulus presentation at a particular frequency, and 15 s rest while again viewing the fixation cross (**Figure 3-4**). This sequence was repeated four times for each stimulus frequency. As such, 28 SSVEP recordings were available per subject. The order of stimulus frequencies was randomised. EEG recording was continuous. Table 3.1 provides demographic details of the subjects and Table 3.2 the order in which the different frequencies were presented to individual subjects. Electrode Oz was with reference to Fz with ground placed on Fpz as per the 10-20 system (**Figure 2-5**).

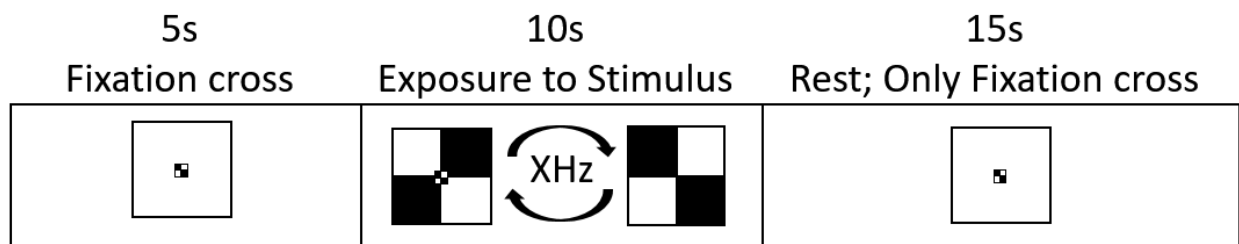


Figure 3-4: A single SSVEP trial. Subjects first viewed a fixation cross for 5 s, followed by 10 s of stimulus presentation, and 15 s again viewing a static fixation cross. The fixation cross was kept constant throughout the experiment so as not to induce any VEPs and to give the subject a reference point to focus on.

Table 3-1: Participant demographics.

	Age	Gender	Handedness
Subject 1	32	Male	Right
Subject 2	27	Male	Right
Subject 3	27	Male	Right
Subject 4	31	Female	Right

Table 3-2: Frequency randomisation order.

	Order of stimulus frequencies (Hz)
Subject 1	6, 6.5, 7, 7.5, 8.2, 9.3, 10
Subject 2	7, 6, 10, 6.5, 8.2, 7.5, 9.3
Subject 3	8.2, 6, 7.5, 9.3, 10, 6.5, 7
Subject 4	10, 7, 6, 7.5, 6.5, 9.3, 8.2

3.2 Data Pre-processing

3.2.1 SSVEP window period selection

Recordings for each trial were initially truncated to 512 samples from stimulus onset, which corresponds to 1 s of recorded SSVEP data. Subsequently, data for three window periods were generated, namely 0.25 s, 0.5 s and 1 s corresponding to 128, 256 and 512 samples, respectively.

3.2.2 Removal of DC offset

Impedance of the electrodes or the choice of reference could cause a DC offset in the recorded data. Although a reference electrode was used to remove the bulk of the DC offset, drifts in the baseline EEG reference voltage and fluctuations in the offset of the SSVEP signal could cause additional DC offset. Since the model does not have a parameter that accounts for DC offset, additional offset correction was necessary.

A filter-based approach was deemed unsuitable in this instance because it would add a time-domain shift to the recorded data. Such a shift would not have been accounted for by our time-domain-based model and could therefore potentially have affected the accuracy of the classification.

The DC offset was removed by subtracting the mean of the signal over the window period from each point in the signal (**Figure 3-5**). Although this method does not remove all the DC components, it yielded data that were clean enough and centred about the 0 V point while not inducing phase shifts or non-linearity.

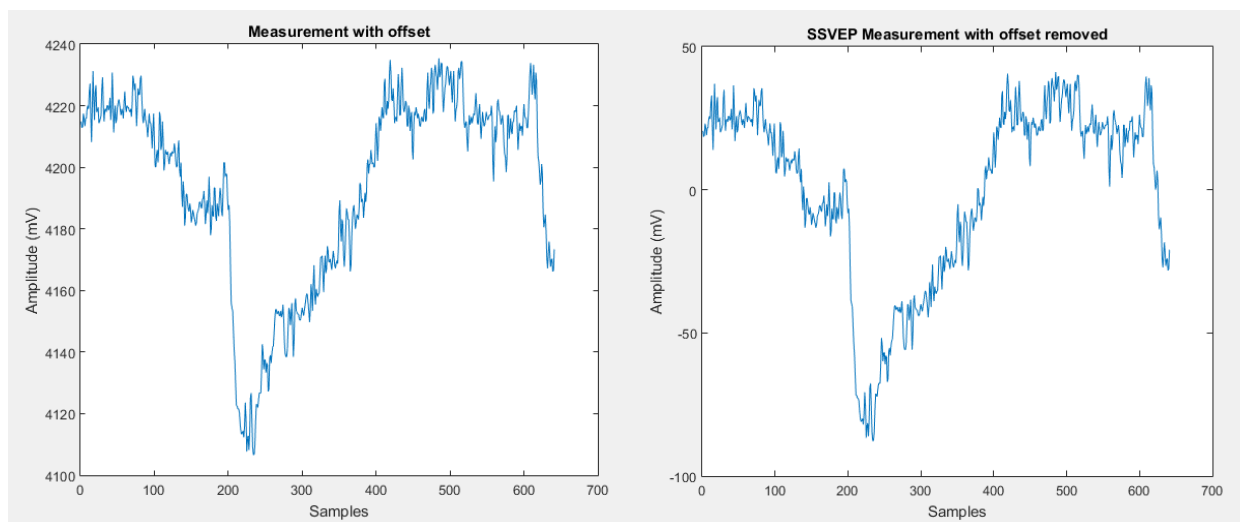


Figure 3-5: Comparison of an SSVEP signal before (left) and after (right) DC offset correction (note the difference in the y-axis values).

3.2.3 Noise reduction

The data were filtered to remove any noise arising from the electrical mains, because such noise can add a large frequency component to the signal at 50 Hz. In this instance, a 50 Hz notch filter was applied (Vilic, 2015).

3.2.4 Eyeblink removal

Eyeblinks were identified by using a plugin for EEGLAB called *corrmap*. This plugin identifies ocular artefacts based on templates. As accounting for eyeblinks is beyond the scope of this project, any trial in which an eyeblink was detected in the first second after stimulus onset was discarded. After exclusions, at least three trials per stimulus frequency remained for each subject.

3.3 Model Construction

The model described in Chapter 2 was constructed using MATLAB and Simulink in accordance with the original literature (Kremláček et al., 2002). Here, three serially connected forced oscillators were created so that the model parameters and model dynamics could be replicated. Each oscillator was modelled as a separate block titled OSC1, OSC2 and OSC3, respectively. **Figure 3-6** is a schematic representation of an oscillator.

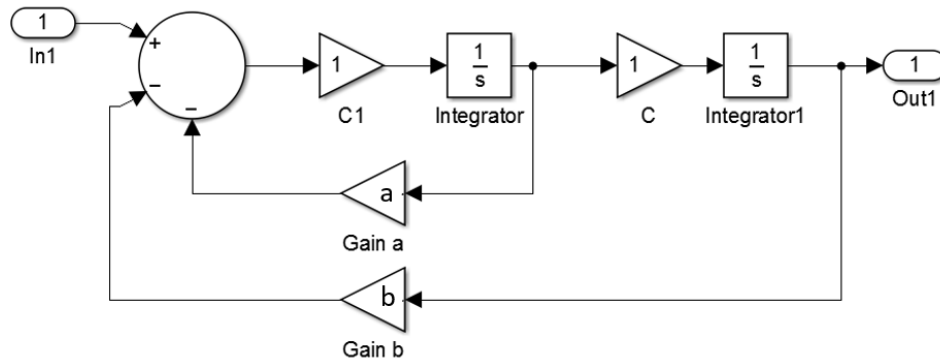


Figure 3-6: MATLAB Simulink schematic of a single oscillator, OSC2.

The three oscillators were connected in series, including the delay and weighting coefficients (**Figure 3-7**). Output probes were added at every major junction in the system, as these points were of interest in the tuning and performance-evaluation of the system. This enabled the individual components of the modelled data to be accessible.

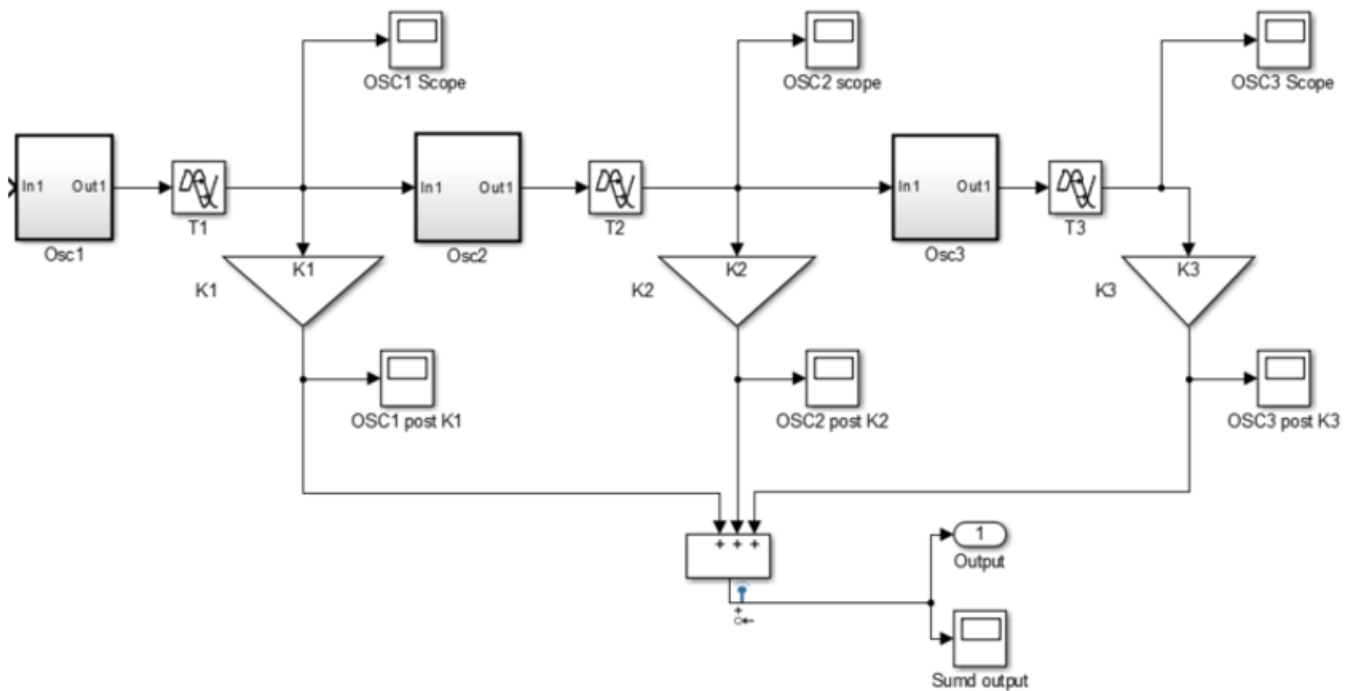


Figure 3-7: Schematic representation of the model implemented in MATLAB Simulink.

A graphical user interface (GUI; **Figure 3-8**) was designed to run the model and display the parameter set and resultant output.

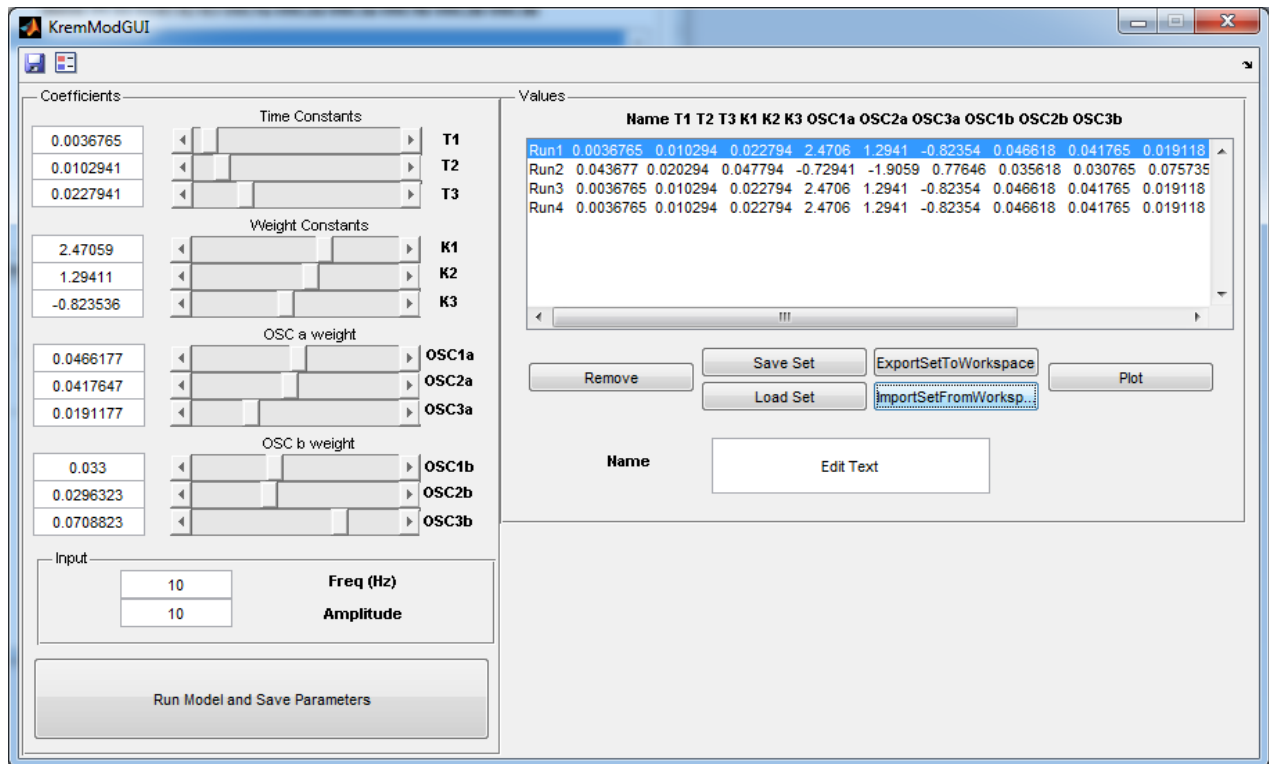


Figure 3-8: MATLAB GUI designed to run the model.

The sampling time was set to 1.953125 ms, which corresponds to a sampling rate of 512 Hz to match that of the recorded SSVEP responses. The simulation time would therefore be determined by the number of samples (i.e. 512 for 1 s, 256 for 0.5 s, and 128 for 0.25 s).

The model generates output data by means of a 4th-order ordinary differential equation (ODE) solver that uses the Runge-Kutta method (Cheever, 2015). This method performs well when a small fixed-sample time step is chosen. Given that 3rd-order ODEs describe each individual oscillator block, lower-order solvers would not have been appropriate.

3.4 Model Validation

The model output was initially visually compared to the plots published by Kremláček et al (2002) when using the input parameters recommended by the authors for pattern reversal checkerboard stimuli (Table 3-3).

Table 3-3: Input parameters for checkerboard stimuli (Kremláček et al., 2002).

K1	K2	K3	OSC1a	OSC1b	OSC2a	OSC2b	OSC3a	OSC3b	T1	T2	T3
-5.2293	-2.4574	0.2404	0.0335	0.0091	0.0058	0.0017	0.0142	0.0001	24.6984	53.1513	6.4955

3.5 Feature Extraction by Model Fitting

Model fitting refers to the process whereby the model output is computed after initialisation by an initial parameter (IP) set, the difference (error) between the model output and the signal being fit is evaluated, and the model parameters are subsequently adjusted programmatically using a pattern search algorithm to reduce the error. This process is repeated iteratively until either a suitable fit is found (i.e. the error is reduced to zero) or the maximum number of iterations has been completed. Model fitting therefore produces a set of estimated parameters (EPs), also known as a feature set (P_n), that when substituted back into the model produces a synthesised version of the signal to which the model had been fit, and some error.

In the original study from which the model was adapted, the normalised root mean square error (NRMSE), which represents the difference between the model output for a given P_n and the measured data, was used to assess performance (Kremláček & Holčík, 1999). In this study goodness of fit was calculated using the goodness-of-fit function (`goodnessOfFit`) available in MATLAB using the default MATLAB configurations. Goodness of fit provides a measure of the similarity between the model output and signal being fit. The function uses NRMSE as a cost function when determining error in the fit and returns a value of (1-NRMSE). Thus a goodness of fit value of 1 (100%) and a NRMSE of zero indicate a perfect fit. The model seeks to maximise the goodness of fit by iteratively estimating parameters that reduce the NRMSE from that of the previous iteration.

Model fitting was set to perform 500 iterations, regardless of whether the error was reduced, with a maximum of two restarts. The iteration and error threshold values were based on the work of the developers of the model (Kremláček & Holčík, 1999). The impact of IPs on fitting performance was examined in the stimulus-locked average of the VEP recordings from the exploratory dataset by initialising the model with either the input parameters recommended by Kremláček et al. (2002) for pattern reversal stimuli or zeroes.

The feasibility of fitting the model to SSVEPs was examined using SSVEP data from the exploratory data set once it has reached steady state.

Based on these initial experiments, initial parameters and parameter constraints were determined.

3.5.1 Initial parameter selection

In this work, three sets of initial parameters (IPs) were defined, each of which will generate a unique set of estimated parameters during model fitting (Table 3-4).

Although the IP's impact fit performance and a direct search method can be used to find good starting points (Kremláček, et. al 2002), IP selection and optimisation was not a focus of this study. The three possible IP sets (IP1, IP2 and IP3) were defined in a way that each would relate to different characteristics of the SSVEP response being fit. It should be noted, however, that restricting the possible IPs to three fixed sets may limit the performance of the model across multiple subjects.

Table 3-4: Initial parameters used during model fitting.

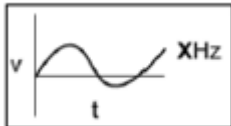
SSVEP response:	Model starting point	Estimated parameters	Satisfying:
	IP1	EP1	Model(EP1) + NRMSE1 = Measured SSVEP
	IP2	EP2	Model(EP2) + NRMSE2 = Measured SSVEP
	IP3	EP3	Model(EP3) + NRMSE3 = Measured SSVEP

Figure 3-9 shows the model output for each of the three initial parameter sets for 256 samples (i.e. 0.5 s). The sampled points, which correspond to the measurements of the recorded EEG signal, are shown on the x-axis, and the amplitude of the model output on the y-axis. The amplitude is scaled during parameter estimation by the K1, K2 and K3 parameters to match that of the recorded SSVEP signal. Table 3-5 gives the parameter values for each of the three IP sets.

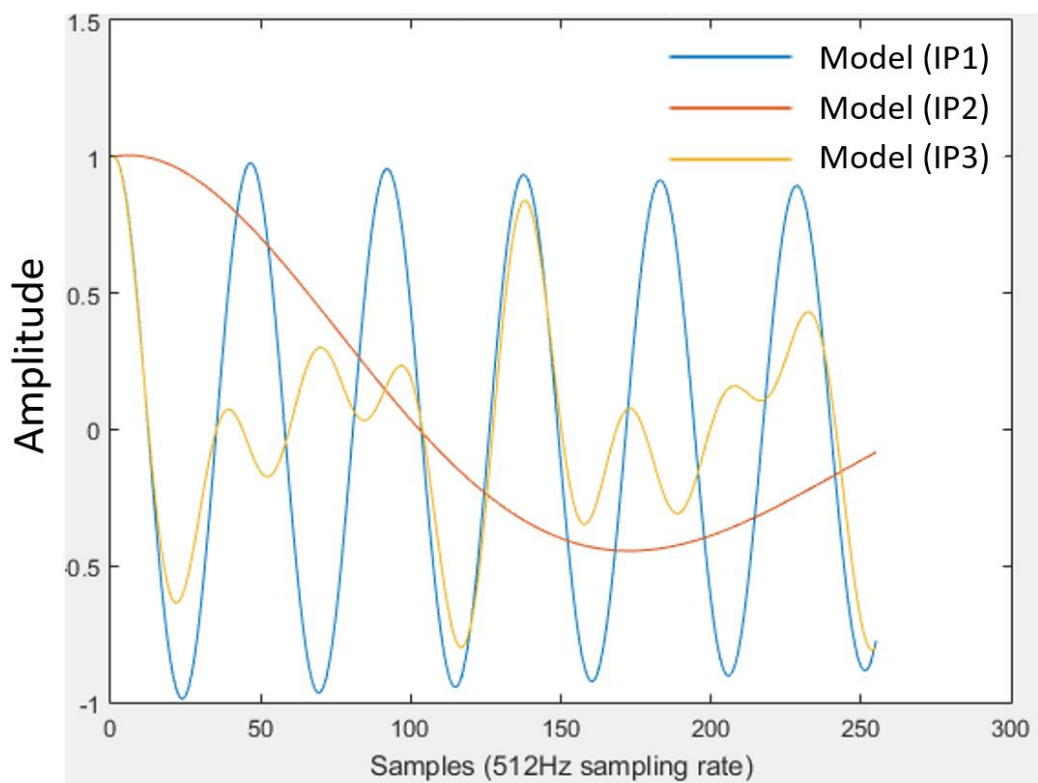


Figure 3-9: Model outputs for each of the three IP sets. The model output for IP1 is a cosine wave at 12 Hz; IP2 generates a low-frequency 1.6 Hz cosine wave; and IP3 is a set of parameter estimates that produced a successful fit to a 12 Hz flashing checkerboard SSVEP response during model testing.

IP1 was constructed so that when it is input into the model, a 12 Hz cosine wave with unit amplitude would be generated. This frequency was selected as it produces the largest amplitude SSVEP response (Wang et al., 2013; **Figure 3-2**).

IP2 corresponds to a low-frequency 1.6 Hz cosine wave. This frequency was selected so that the negative gradient increases in the region where the P100 peak (sample 51) of the VEP response transitions to its N135 peak (sample 69). The largest change (steepest gradient) in the VEP response (**Figure 2-1**) occurs when the signal amplitude decreases from P100 to the N135 inflection point. The first low point at around 330 ms (sample 170) would be optimised during fitting to approach the N135 point. If a VEP response is heavily smoothed a similar shape would emerge, which would move towards the shape of a 1.6Hz cosine wave. Since the pattern simplex search examines the gradient of the function being fit and attempts to match the model output to distinct gradients, it was hoped that the gradient of the 1.6Hz wave would be matched to distinctive VEP features – specifically the N135 point.

IP3 was based on a parameter set that produced a good fit to a 12 Hz flashing checkerboard SSVEP response during model testing.

Table 3-5: Initial parameter sets.

	K1	K2	K3	OSC1a	OSC1b	OSC2a	OSC2b	OSC3a	OSC3b	T1	T2	T3
IP1	1	1	1	0.5	5 000	0.5	5 000	0.5	5 000	0	0	0
IP2	1	1	1	5	100	5	100	5	100	0	0	0
IP3	1	1	1	0.5	2 000	0.75	5 000	0.9	9 000	0	0	0

3.5.2 Parameter constraints

To improve the accuracy of the model fit and parameter estimation, scaling factors were applied to parameter estimates and upper and lower bounds were imposed. Scaling factors were used to weight each parameter value relative to the other values because the downhill-simplex method otherwise assumes that all parameters change at the same rate. This allows the relative magnitudes of parameters to be taken into account when making adjustments. The scaling factors and adjustments were kept constant for every iteration performed.

K1, K2 and K3, which scale the amplitudes of each oscillator in the model, were assigned weightings of 1 relative to the other parameters.

Time parameters were assigned weightings of 0.1, reducing the sensitivity of the model output to these parameters by a factor of ten compared to the gain parameters K1, K2 and K3. The delays T1, T2 and T3 were restricted to positive values as the SSVEP signal starts at 0 s, and the step size of the solver was set to a minimum of 1.9 ms, which corresponds to a 512 Hz sampling rate. The latter reduces calculation

complexity and erroneous pattern matching caused by trying to fit data generated by the model between recorded samples. Finally, the sum of all the time delays was constrained to be less than the period in which the initial VEP occurs (i.e. $T_1+T_2+T_3 < 0.5$ s). This prevents the model from maximising the delays or setting delays greater than the window period of the acquired data, ensuring that each oscillator is able to contribute to the synthetic SSVEP.

OSC1a, OSC1b, OSC2a, OSC2b, OSC3a and OSC3b, which describe the frequency and damping of each oscillator, were scaled to match the characteristics of a damped oscillator. To have a critically damped system, the damping factor ($OSCa/OSCb$) must be less than 1.

3.6 Model Fitting to the Initial SSVEP Responses

The performance of model fitting to initial portions of the SSVEP signal was examined in the published data described previously. 28 SSVEP recordings were available in each of 4 subjects – 4 SSVEP trials for each of 7 stimulus frequencies. Since at least 3 trials per stimulus frequency remained for each subject after exclusions due to eyeblinks, the trial with the worst signal-to-noise ratio was excluded for stimulus frequencies with 4 surviving trials, leaving 3 SSVEP trials per stimulus frequency for each subject (i.e. 21 SSVEP trials in each of 4 subjects).

As described previously, the SSVEP trials were truncated to 512, 256 and 128 samples, producing three data sets corresponding to window periods of 1 s, 0.5 s, and 0.25 s, respectively (i.e. 21 SSVEP trials for each of 3 window periods in each subject; 84 SSVEP trials per window period; 252 SSVEP trials in total).

Figure 3-10 illustrates the model fitting and parameter estimation process for a single SSVEP trial. During fitting the model is initialised by one of the three IP sets (IP1, IP2 or IP3). Each model fit produces for a specific SSVEP trial a set of estimated parameters (EPs) (**Figure 3-11**).

For each SSVEP trial, the model fitting process was repeated three times, once for each IP set (IP1, IP2 or IP3), yielding three sets of EPs (EP_i , where $i = 1$ to 3). As such, a total of 756 EP sets were generated. Goodness of fit was calculated for each of the estimated parameter sets, the distribution of which was investigated using notched box-and-whisker plots. The notches in the boxes represent 95% confidence intervals (CI) around the median, defined as $M \pm 1.57 * (IQR / \sqrt{n})$, where M is the median, IQR is the interquartile range and n is the number of measurements. Data points more than $1.5 * IQR$ below or above the 25th and 75th percentiles, respectively, are considered to be outliers. The whiskers extend to the most extreme data points that are not classified as outliers.

Multi-way ANOVA was used to examine the effect of window period, stimulus frequency, IP set, and subject on goodness of fit.

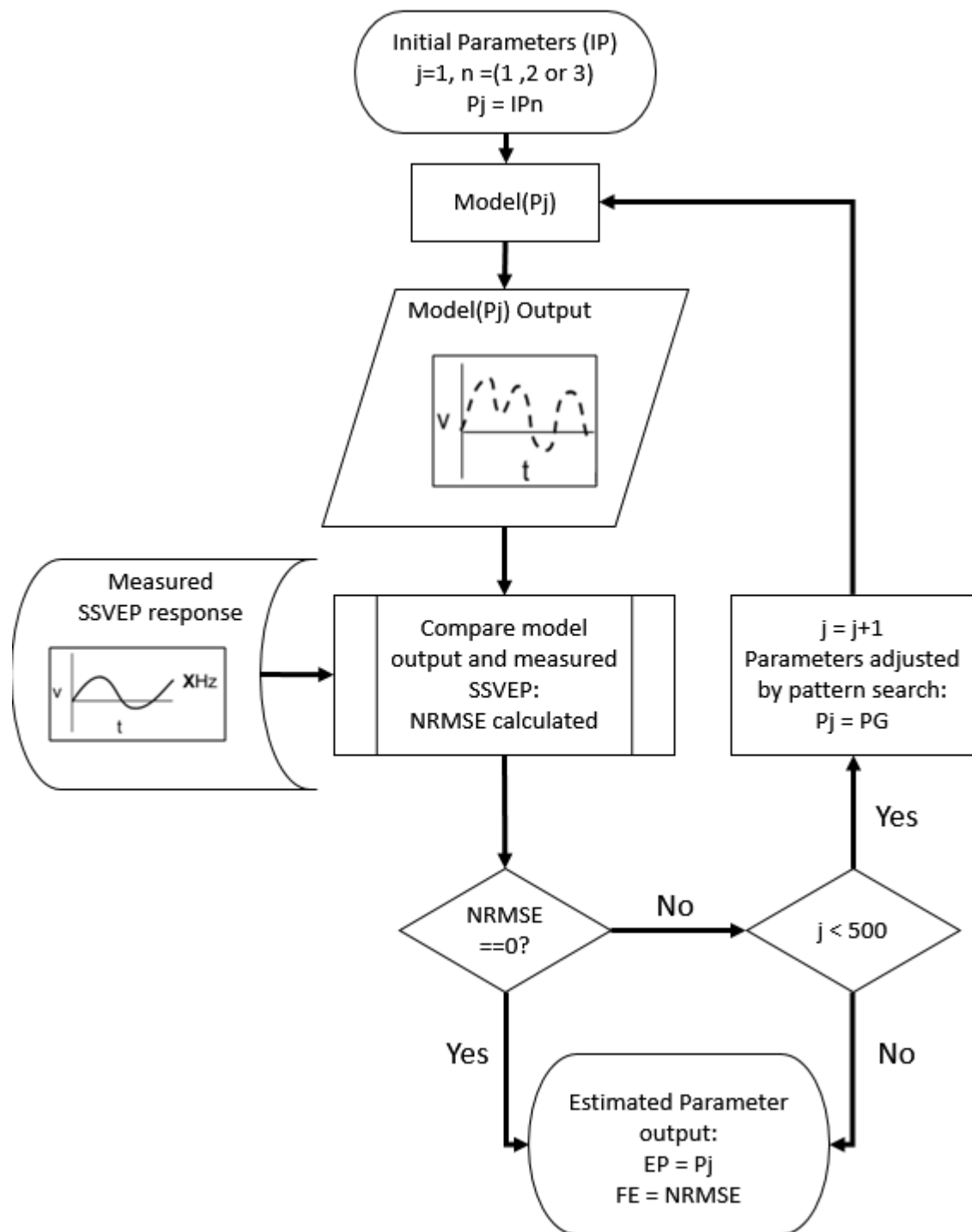


Figure 3-10: Feature extraction by model fitting and parameter estimation. The model starts with one of three sets of initial parameters (IP1, IP2 and IP3). The estimation process adjusts the model parameters (P_j) iteratively using a pattern-search method to obtain a better fit (PG), each time comparing the output to the measured SSVEP response. This process is repeated for 500 iterations. Once the iterations are complete or the normalised root mean square error (NRMSE) has been reduced to zero, the parameter estimation stops and the final estimated parameters (EP) are output.

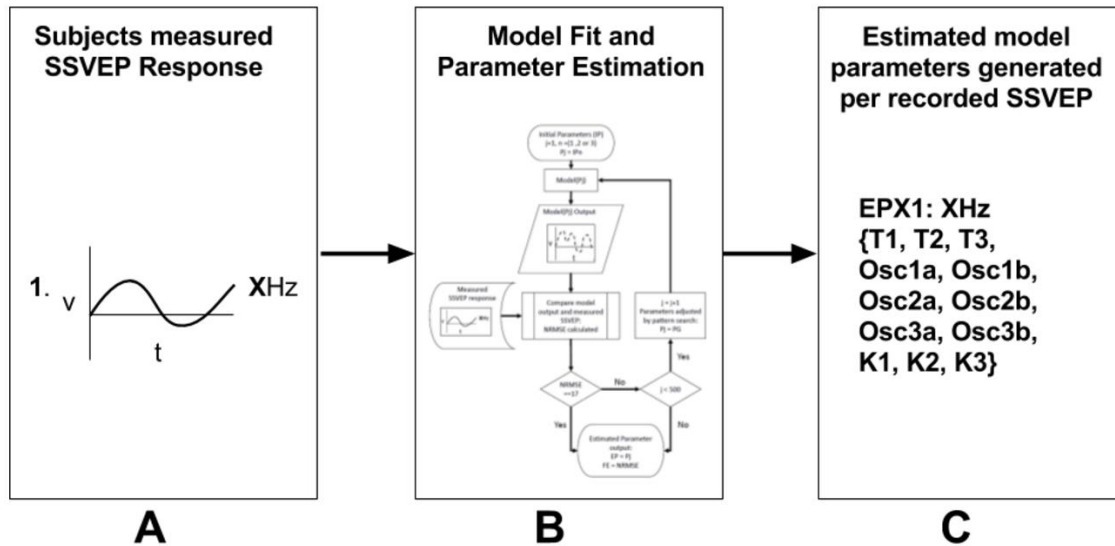


Figure 3-11: Feature extraction by model fitting and parameter estimation for a single SSVEP response: **(A)** SSVEP data are entered into the model. **(B)** The model fits the measured data through a process of parameter estimation. **(C)** A set of parameters (P_n) is generated that describes the measured SSVEP signal as a function of the model.

3.7 Feature Classification

In this study three feature-classification methods were compared: a model-based approach with a multiclass SVM scheme and the commonly used spectral-domain CCA and PSDA methods. While the CCA and PSDA methods combine feature extraction and classification into one closed process, the proposed model-based approach has distinct feature-extraction (by model fit) and classification (by SVM) stages. Every set of parameters output by the model fit is considered a descriptive feature set.

3.7.1 Multiclass SVMs

The objective of the multiclass SVM classifiers in this context is to determine from only the estimated parameter sets (EP_n) generated by the model for each SSVEP response which stimulus frequency the subject was attending at a given time.

All SVMs were trained in the same way and with the same configuration in order to ensure that the SVM scheme would be an independent variable when SVM classifiers are later compared to other classification methods. Each of the SVMs consisted of a set of chained binary-class SVMs to allow for a multiclass SVM classifier configuration. The chaining of SVMs was handled by the MATLAB `fitcecoc` function and is not discussed further here.

For each subject, SVMs were trained for a specific window period and IP set using the associated estimated parameter sets that were generated by the model. The frequency of the stimulus represents the class. Table 3-6 shows a single SVM training set for one subject for a window period of 0.25 s and IP1. X, Y and Z are generic place holders for the frequencies of 6, 6.5, 7, 7.5, 8.2, 9.3 and 10 Hz.

Table 3-6: A single SVM training set for one subject

Subject n , SVM1 ($t = 0.25$ s):	
Frequency of stimulation (class)	Estimated parameters (EPs) from model fit, with $t = 0.25$ s and IP1
X Hz	EPX1: {T1, T2, T3, Osc1a, Osc1b, Osc2a, Osc2b, Osc3a, Osc3b, K1, K2, K3}
Y Hz	EPY1: {T1, T2, T3, Osc1a, Osc1b, Osc2a, Osc2b, Osc3a, Osc3b, K1, K2, K3}
Z Hz	EPZ1: {T1, T2, T3, Osc1a, Osc1b, Osc2a, Osc2b, Osc3a, Osc3b, K1, K2, K3}

A total of 12 SVMs were trained per subject. The first nine SVMs include estimated parameters generated from fitting, for each SSVEP trial, three different window periods of data for each of three IP sets. Since each fit produces a unique set of EPs, nine unique SVM training sets are generated per subject. The final three SVM training sets, one for each of the three window periods, includes estimated parameters from all three initial parameter sets. Table 3-7 summarises the 12 SVM training sets used for each subject.

These twelve SVMs enabled classifier performance to be assessed as a function of initial parameters and window period. Each SVM is named after the set of EPs (linked to the IPs) on which it was trained. As such, SVM n ($n = 1,2,3,\text{all}$) denotes the SVM trained using the sets of EPs generated from model fitting using IP set n .

Table 3-7: SVMs trained per subject.

SVMs trained per subject	Window period of SSVEP data used in the fit process	Model initialisation parameters used in the fit	Estimated parameters
SVM1($t = 0.25$ s)	0.25 s	IP1	EP1(0.25)
SVM1($t = 0.5$ s)	0.5 s	IP1	EP1(0.5)
SVM1($t = 1$ s)	1.0 s	IP1	EP1(1)
SVM2($t = 0.25$ s)	0.25 s	IP2	EP2(0.25)
SVM2($t = 0.5$ s)	0.5 s	IP2	EP2(0.5)
SVM2($t = 1$ s)	1.0 s	IP2	EP2(1)

SVM3(t = 0.25 s)	0.25 s	IP3	EP3(0.25)
SVM3(t = 0.5 s)	0.5 s	IP3	EP3(0.5)
SVM3(t = 1 s)	1.0 s	IP3	EP3(1)
SVMAll(t = 0.25 s)	0.25 s	IP1, IP2 and IP3	EP1(0.25), EP2(0.25) and EP3(0.25)
SVMAll(t = 0.5 s)	0.5 s	IP1, IP2 and IP3	EP1(0.5), EP2(0.5) and EP3(0.5)
SVMAll(t = 1 s)	1.0 s	IP1, IP2 and IP3	EP1(1), EP2(1) and EP3(1)

3.7.2 SVM construction

Each SVM was built using the same configuration. A polynomial kernel function of degree 8 was used as the SVM kernel since polynomial kernel functions are better suited to learned models in which the features have a non-linear relationship. When assessing the position of a single value in the parameter set, a polynomial kernel function takes into account not only the relative spacing of a single variable, but also the combinations of all the features (Rai, 2011).

The predictor data of the SVM were standardised such that each column of input data is weighted by a column mean and a standard deviation. This reduces the effect of outlying data points. This functionality, and that of the polynomial kernel, are dealt with entirely by MATLAB's Machine Learning toolbox.

The MATLAB function, `fitcecoc`, was used to train the multiclass SVM classifiers. MATLAB's random variable generation was reset before each SVM was trained using the command `rng(1)`. This enables the training outcome to be repeatable (**Figure 3-12**):

```
% SVM template
Data.SVMData.t = templateSVM('Standardize',1,'KernelFunction' ...
,'polynomial', 'PolynomialOrder',8,'SaveSupportVectors',true);
% Train Models:
rng(1)
Data.SVMData.SVM1.Mdl = fitcecoc(X1,Y,'Learners',Data.SVMData.t);
```

Figure 3-12: Code snippet of MATLAB code used to train the SVM's with the defined template `Data.SVMData.t`.

3.7.3 CCA feature extraction and classification

CCA was introduced and first used by Lin et al. (2007). The MATLAB functions to perform CCA classification of SSVEPs have been made available by the authors of a study that compared CCA to another classification method (Zhang, 2014). The only change made to their implementation was to replace the reference sine waves they employed with sine waves that corresponded to the stimulus frequencies presented in this research. Essentially CCA classification compares unknown signals to a set of known patterns.

3.7.4 PSDA feature extraction and classification

PSDA feature classification is available in the MATLAB EEG-processing-toolbox. The toolbox was developed specifically to compare different SSVEP classification methods in a study published by Kompatsiaris et al. (2016). These validated functions were used in the present study (Ioannis Kompatsiaris et al., 2016) to build the PSDA classification system. Since PSDA uses the frequency responses in the data as features for classification, the time-series SSVEP data needed to be transformed into the frequency domain prior to classification.

The classifier returns the largest primary frequency component, which is assumed to represent the stimulus frequency.

3.8 Assessing Classifier Performance

3.8.1 Cross-validation

A cross-validation model was constructed for each SVM that was trained. This was generated using MATLAB's *crossval* function. For each subject, the cross-validation models for SVM1, SVM2, and SVM3 each contained 21 sets of EPs (3 per class), while that of SVMAll consisted of 63 sets (9 per class). For each subject, both in-sample and out-of-sample resubstitution losses were evaluated for each SVM for each window period using the cross-validation model.

The in-sample resubstitution loss gives an indication of the ability of the SVM to classify data from within the training set. All data are included in the training set and the SVM attempts to classify a data set from the training set. In the ideal case this would return no errors, however a perfect in-sample loss does not imply that the SVM is perfect at classification, as it may be over fit to the training data (i.e. it may only be good in the context of the training data).

The out-of-sample resubstitution accuracy was calculated using a leave-one-out cross-validation approach. This method trains the classifier with the given data, leaving out one data set (a class and its associated parameters). The ability of the trained SVM to then correctly classify the left-out data set is assessed. This process is repeated for every data set in the training set.

For each subject, in-sample confusion matrices were generated for SVMAll and window period. Class-specific accuracy, precision and recall were calculated from the confusion matrices for SVMAll and window period **for each subject**.

Out-of-sample confusion matrices for all subjects combined were generated for SVMAll and window period by adding the out-of-sample confusion matrices of individual subjects. Class-specific accuracy, precision and recall across subjects were then computed for each class and window period.

For comparison, out-of-sample confusion matrices were constructed for CCA classification for each window period, across all subjects, and class-specific accuracy, precision and recall computed.

3.8.2 Statistical comparison of classifiers

For the four SVM-based classifiers, three-way ANOVAs were used to examine the effect of window period, classifier, and frequency on class-specific accuracy, precision and recall. In addition, the effect of window period on accuracy, precision, and recall was evaluated separately in the SVMall and CCA classifiers using one-way ANOVAs. Finally, three one-way ANOVAs were performed to compare accuracy, precision and recall between the four SVM-based classifiers and CCA.

3.9 Classifier Performance in the Context of BCIs

3.9.1 Probability of a true positive decoding

The performance of the classifiers in a multi-class BCI application were evaluated using the probability of a true positive decoding, $Prob(TP)$, to quantify the fraction of all classified signals which are assigned to the correct class. It is defined as the sum of the diagonals (TPs) over the sum of all entries in the confusion matrix:

$$Prob(TP) = TP/[TP+FP+TN+FN] \quad (3-1)$$

$Prob(TP)$ values were computed for each classifier for each window period, both for subjects individually and for all subjects combined.

3.9.2 McNemar testing

Asymptotic McNemar testing was performed for each subject for each window period to examine whether $Prob(TP)$ differed significantly between the various SVM-based classifiers and CCA. McNemar testing is a categorical statistical test designed for comparing multi-class classification methods. McNemar testing was implemented here using the MATLAB function `testcholdout`. The null hypothesis was that classifier A is *at most as accurate as* classifier B. Rejection of the null hypothesis therefore indicates with 95% certainty that classifier A is more accurate than classifier B. Since each classifier was compared with every other classifier (20 comparisons) for every window period (3) and each subject (4), a total of 240 tests were performed. The resulting p -values were tabulated in 5x5 matrices - one for each subject for every window period.

The McNemar results were summarised across subjects by producing for every window period a 5x5 matrix where every entry shows the number of subjects for which the null hypothesis was rejected at that position in the individual subject matrices.

3.9.3 Hypothetical ITRs

Hypothetical ITR values were then computed for each classifier for each window period, both for subjects individually and for all subjects combined. During computation of the ITRs, some limits were placed on the variables. For instance, the number of possible commands, N , was fixed at 7, which corresponds to the number of unique frequency classes presented. Further, processing time was assumed to be instantaneous, so that the number of commands possible per second was fixed:

$$\text{Number of commands per second} = 1/(\text{data window period}) \quad (3-2)$$

4

Results

4.1 Model Validation

Figure 4-1 shows the output from our replicated model when using model parameters specified by Kremláček et al. (2002). This output is similar to that produced by the authors using the same model parameters (see **Figure 1-8**).

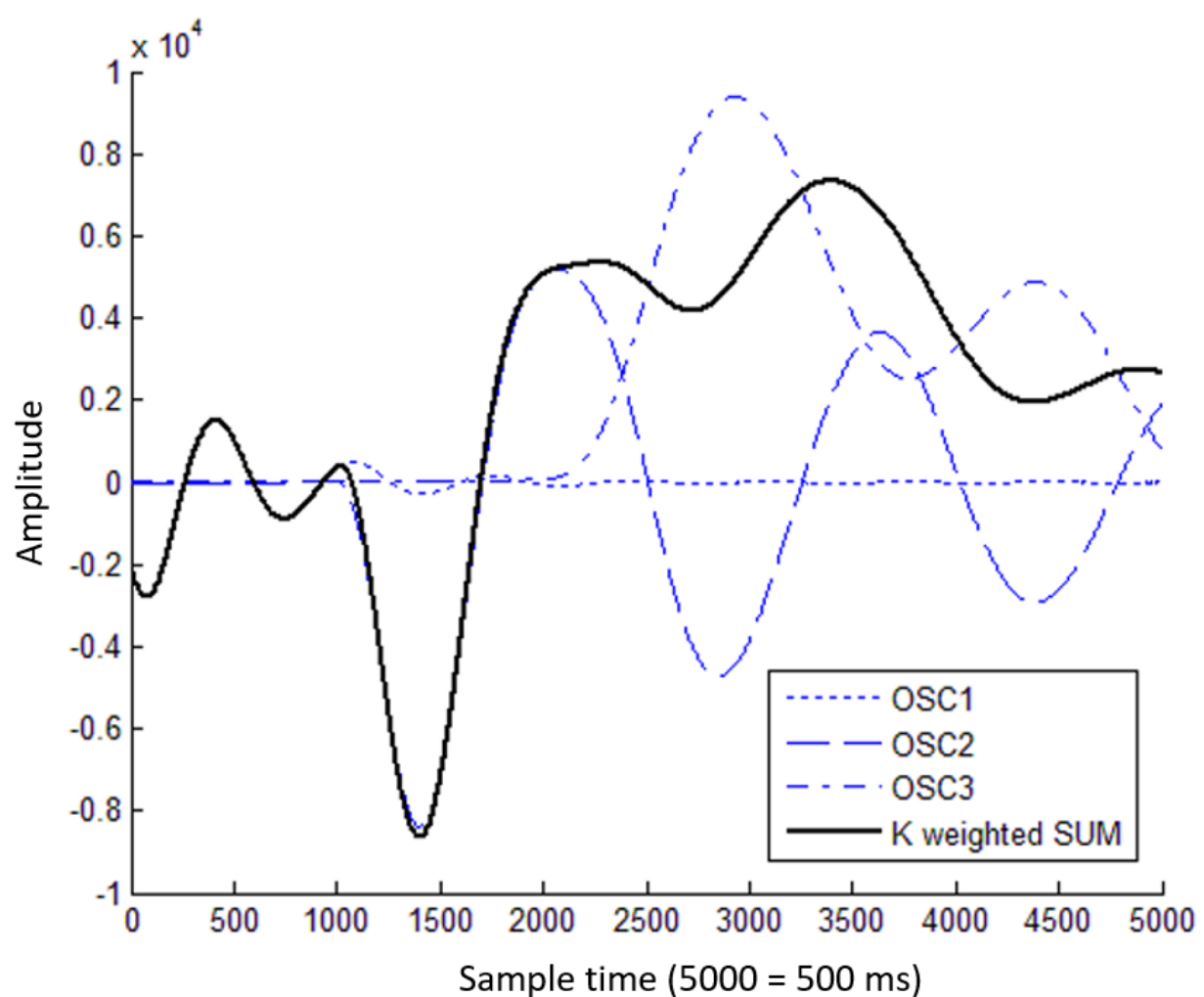


Figure 4-1: Output from our replicated model when using the parameters specified by Kremláček et al. (2002) for pattern reversal stimuli. Not

4.2 Feature Extraction by Model Fitting

4.2.1 Model fit on VEP results

The dashed red lines in **Figure 4-2** and **Figure 4-3** show the outputs when fitting the model to a stimulus-locked average of 20 occurrences of a single subject's VEP response for two different initial parameter sets. For the fit in **Figure 4-2**, the initial starting parameters suggested by Kremláček et al. (2002) for pattern reversal VEPs were used, producing a fit with an NRMSE of 0.26. Although this is higher than the 0.13 NRMSE reported by Kremláček et al. (2002) for the same starting parameters, the characteristic features of the VEP response (N70, P100, N135) are preserved. In contrast, setting initial parameters to zero produced a poor fit with an NRMSE of 0.86 (**Figure 4-3**). Notably, the model, which increases or decreases from a starting value of zero, is not able to account for the DC offset of $\sim 2.5 \mu\text{V}$ at 0 s (present in the grand averaged VEP data).

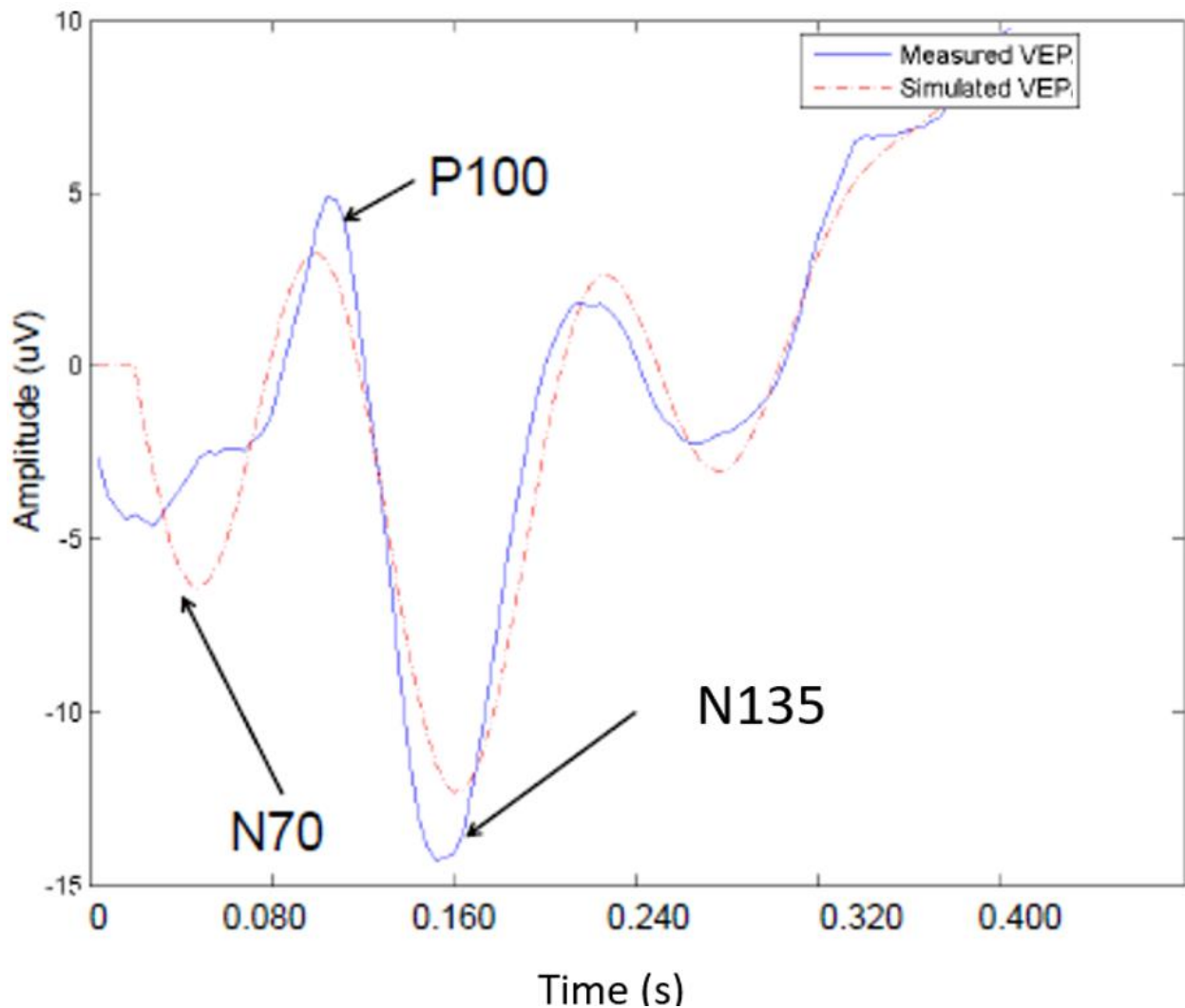


Figure 4-2: Model fit (red dashed line) to a stimulus-locked average of 20 occurrences of a single subject's VEP response (solid blue line) when initial starting parameters suggested by Kremláček et al. (2002) for pattern reversal VEPs were used. The model fit produced an NRMSE of 0.26 (goodness of fit = 0.74).

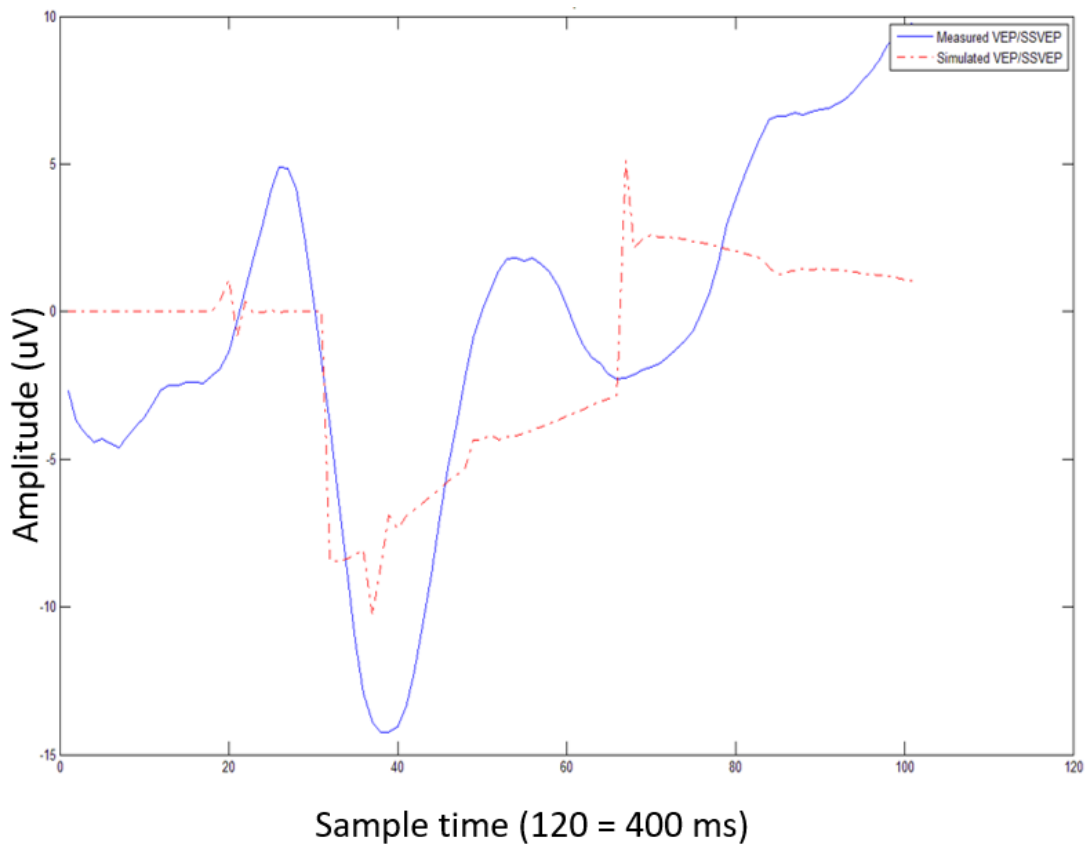


Figure 4-3: Model fit to the same stimulus-locked average of 20 occurrences of a single subject's VEP response (solid blue line) when initial starting parameters were set to zero. The NRMSE is 0.86, yielding a goodness of fit of 0.14.

4.2.2 Fitting the model to SSVEP data in steady state

Fitting the model to SSVEP waveforms, excluding the initial 1 s portion of the SSVEP response before the signal reaches steady state, consistently yielded NRMSEs below 25% (goodness of fit > 0.75). **Figure 4-4** shows an example of a recorded SSVEP response to a 4 Hz flashing checkerboard stimulus (solid blue line) together with the synthetic SSVEP generated after 500 iterations of parameter estimation by the model (dashed red line). Note that the x-axis is the sample length and does not indicate the stimulus onset. The NRMSE for this particular fit is 21% (goodness of fit = 0.79). Although the shape of the synthetic SSVEP closely matches the data, it can be seen to deviate from the recorded SSVEP over time due to signal drift.

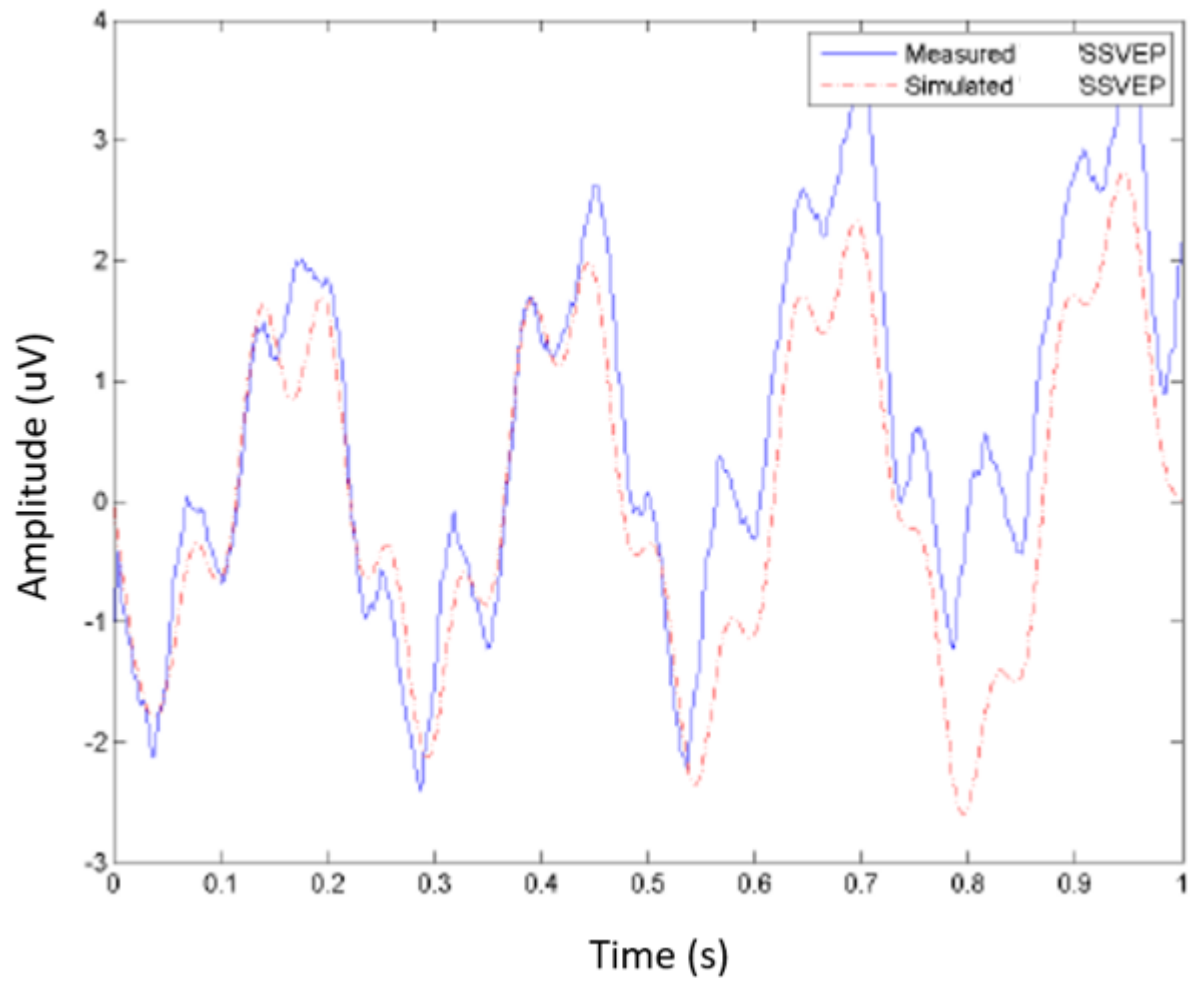


Figure 4-4: Comparison of an SSVEP response to a 4 Hz flashing checkerboard stimulus (solid blue line) and the output from the model after 500 iterations (dashed red line). The NRMSE is 21% (goodness of fit = 0.79). Although the initial model fit closely follows the recorded data, it deviates over time due to signal drift.

4.3 Model Fitting to the Initial SSVEP Responses

Figure 4-5, **Figure 4-6** and **Figure 4-7** show examples of model fits to the initial 1 s, 0.5 s and 0.25 s of SSVEP responses recorded for different subjects for different stimulus frequencies. In each plot the solid black line indicates the recorded SSVEP response following stimulus onset at 0 s. The other lines show the synthetic SSVEPs generated by the model for each of three different initial parameter sets. The fits shown are part of the dataset used to train the SVMs. The goodness of fit values for each fit are shown in the legends found in the top right hand corner of the figures.

In general the plots demonstrate that the model can be fit to the initial SSVEP response. Although the goodness of fits are lower than that achieved when fitting to VEPs, the main features of the SSVEPs are generally preserved. Small perturbations in the SSVEP response are, however, not fitted well (see, for example, features at 0.1 s and 0.4 s in **Figure 4-6**). It is also evident from the red curve in **Figure 4-7** that it is possible to generate a fit with a reasonable goodness of fit (0.41) that in no way mirrors the physiological SSVEP response. Notably, since each SSVEP represents a single trial and not an average of trials, the features of the initial SSVEP response are highly variable and the characteristic features of a VEP response (N75, P100, N135) are generally not evident.

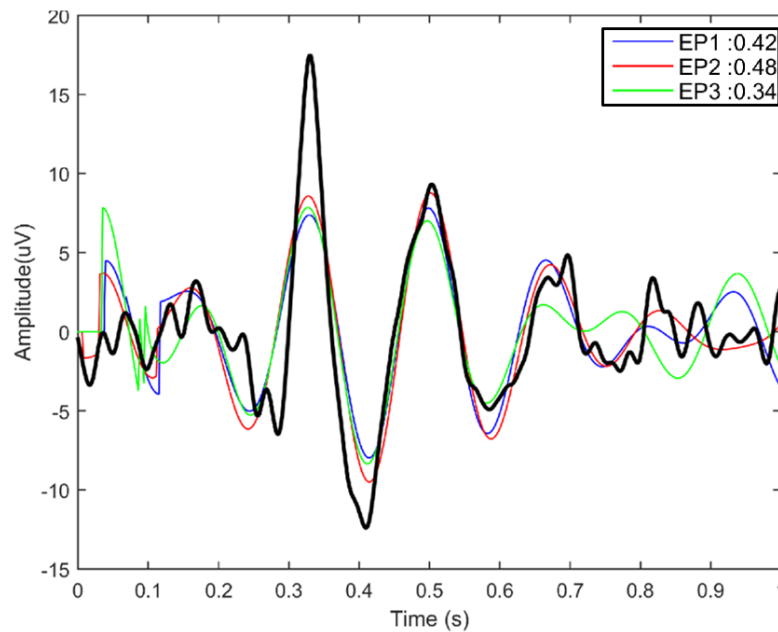


Figure 4-5: Model fit to 1 s of recorded SSVEP data. The black line shows the SSVEP response recorded in subject 4 following onset of a 7.5 Hz flashing checkerboard stimulus at 0 s. The other colours (blue, red, green) show model fits for each of three different initial parameter sets. The goodness of fit for each plot are shown in the legend on the top right.

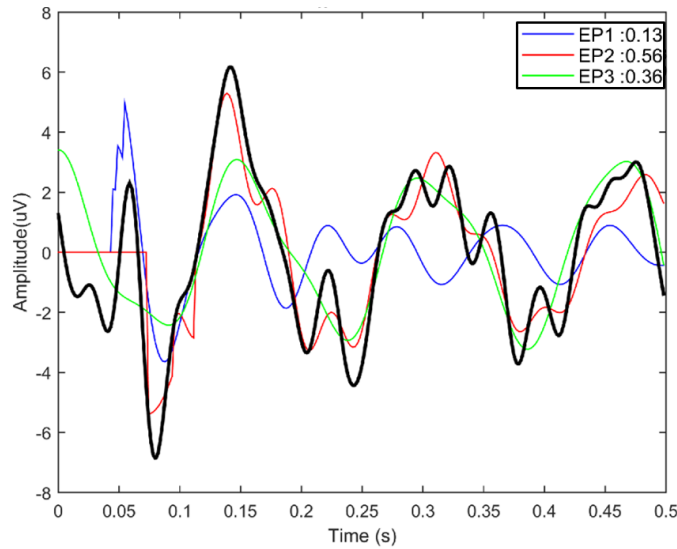


Figure 4-6: Model fit to 0.5 s of recorded SSVEP data. The black line shows the SSVEP response recorded in subject 3 following onset of a 6.5 Hz flashing checkerboard stimulus at 0 s. The other colours (blue, red, green) show model fits for each of three different initial parameter sets. The goodness of fit for each plot are shown in the legend on the top right.

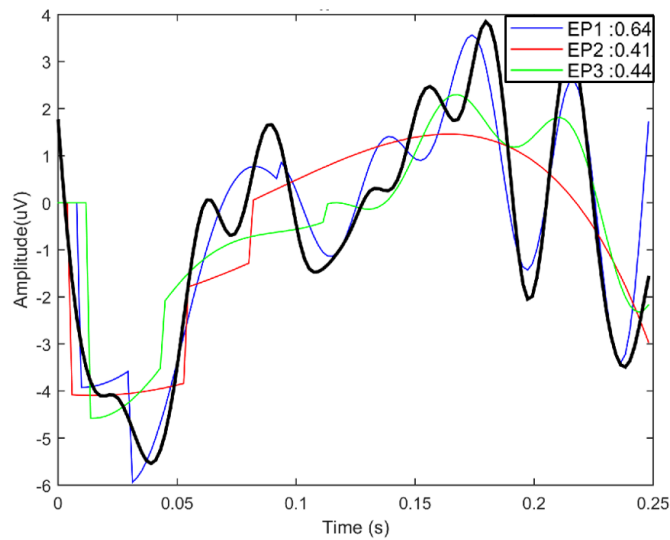


Figure 4-7: Model fit to 0.25 s of recorded SSVEP data. The black line shows the SSVEP response recorded in subject 2 following onset of a 6 Hz flashing checkerboard stimulus at 0 s. The other colours (blue, red, green) show model fits for each of three different initial parameter sets. The goodness of fit for each plot are shown in the legend on the top right.

Figure 4-8 shows the distribution of the goodness of fits achieved, for all subjects and initial parameter sets, when fitting the model to recorded SSVEPs from different stimulus frequencies and for different window periods of data. **Figure 4-9** shows the same data collapsed across stimulus frequencies. Figures **4-8** and **4-9** demonstrate for each stimulus frequency separately and all frequencies combined that the goodness of fit increases significantly when the model is fit to longer window periods of data ($F = 203.4$, $p < 0.0001$, $df = 2$). This is supported by the fact that the notches for all three window periods do not

overlap (Figure 4-9), which provides strong evidence that the median goodness of fit for the three window periods is different from each other at a significance level of 5%.

Although visual inspection of the goodness of fits, for each window period of data, appear similar for different stimulus frequencies, multi-way ANOVA showed a significant effect of frequency ($F = 2.52$, $p = 0.02$, $df = 6$). Post-hocs revealed that (when averaged over all window periods, subjects and initial parameter sets) the model produced better fits to SSVEPs recorded from 9.3 Hz stimuli than 6 Hz stimuli (**Figure 4-10**); goodness of fits did not differ significantly between any other stimulus frequencies. Notably, when fitting to 1 s of data, the goodness of fits achieved for different stimulus frequencies show greater variability than when fitting to shorter window periods of data (**Figure 4-8**).

The choice of initial parameter set also significantly impacted fitting performance ($F = 21.67$, $p < 0.0001$, $df = 2$). **Figure 4-11** shows that the median goodness of fit averaged across all subjects, stimulus frequencies and window periods. The range of goodness of fit values found for all three IPs is comparable, although post-hocs do reveal that the median value for IP2 is lower at the 5% significance level as compared to IP1 and IP3 (**Figure 4-12**). It should be noted however that they differed by only 0.07 in goodness of fit value.

Finally, goodness of fits (when averaged over all window periods, stimulus frequencies and initial parameter sets) differed significantly between subjects ($F = 5.4$, $p = 0.0011$, $df = 3$), with SSVEP responses from subject 2 producing poorer fits than those recorded in subjects 1 and 3 (**Figure 4-13**).

None of the factors showed significant interaction effects (all p 's > 0.1).

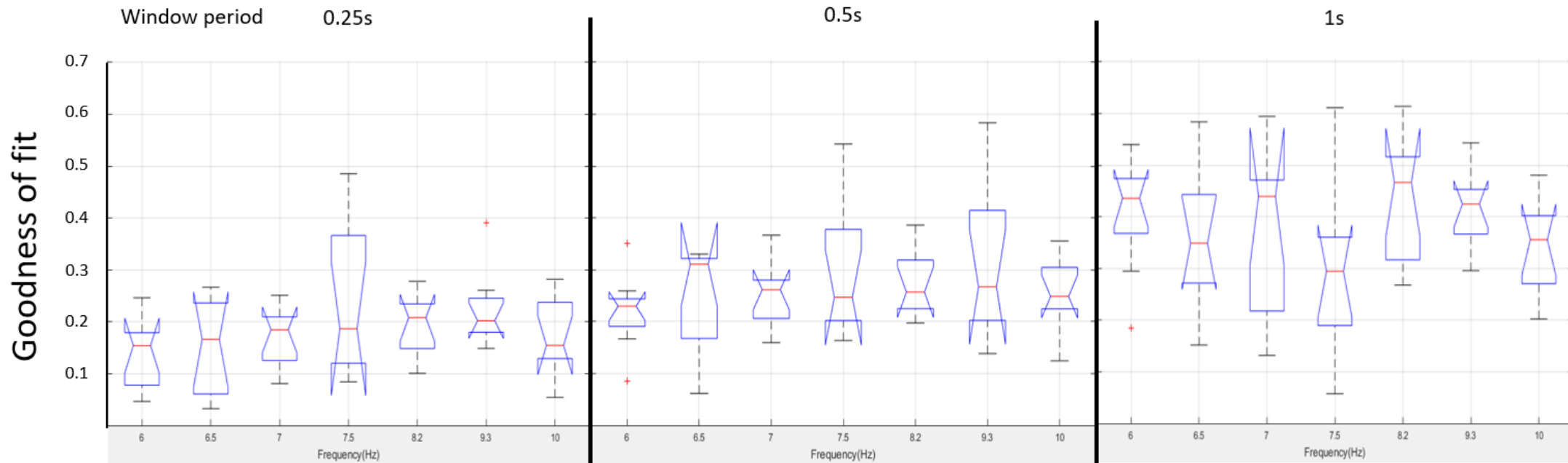


Figure 4-8: Box-and-whisker plots showing the distribution of goodness of fit values achieved in all subjects for different frequencies and window periods. The notches in the boxes represent 95% confidence intervals (CI) around the median, defined as $M \pm 1.57 \cdot (IQR / \sqrt{n})$, where M is the median, IQR is the interquartile range and n is the number of measurements. Red crosses denote outliers, that is points more than $1.5 \cdot IQR$ below or above the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data values that are not outliers.

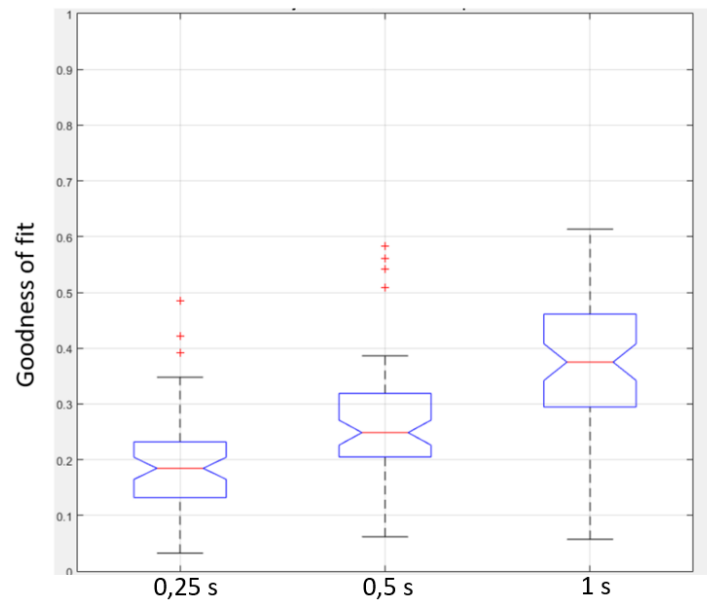


Figure 4-9: Box-and-whisker plots showing goodness of fit, averaged over all subjects, stimulus frequencies and initial parameter sets, as a function of data window period. The fact that the notches for 0.25 s, 0.5 s and 1 s do not overlap provides strong evidence at a significance level of 5% that the median the goodness of fit improves when fitting to longer periods of the initial SSVEP response.

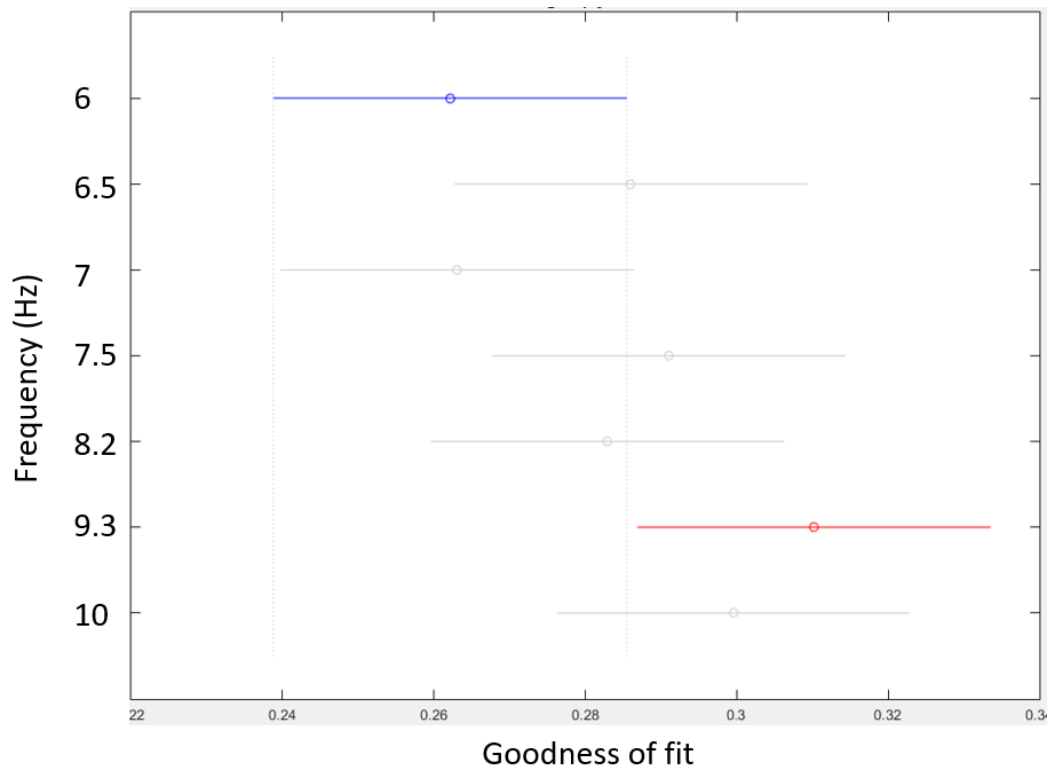


Figure 4-10: Post-hocs revealed that, when averaged over all window periods, subjects and initial parameter sets, the model achieved significantly better fits to SSVEP responses from 9.3 Hz stimuli than SSVEP responses from 6 Hz stimuli. No other stimulus frequencies showed significant differences.

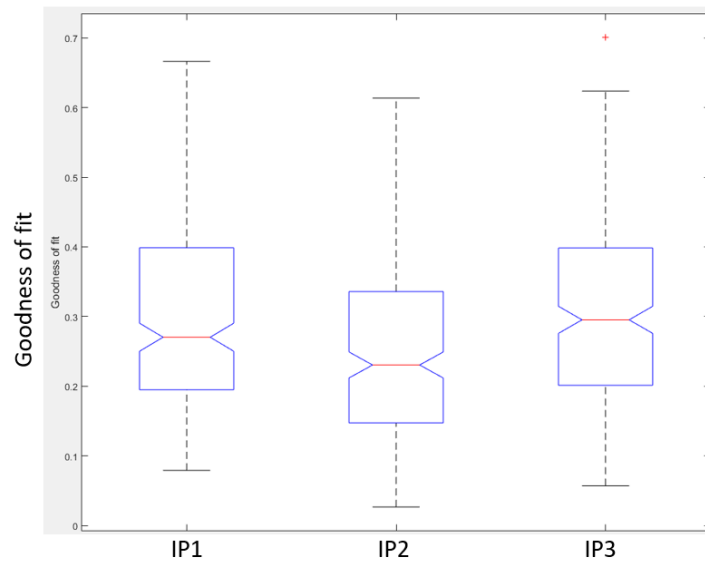


Figure 4-11: Comparison of goodness of fits achieved for each of the three initial parameter (IP) sets, averaged across all subjects, stimulus frequencies and window periods. The fact that the notches for IP2 do not overlap with those of IP1 and IP3, provides strong evidence that the median goodness of fit for IP2 is lower than that of IP1 and IP3 at a significance level of 5%.

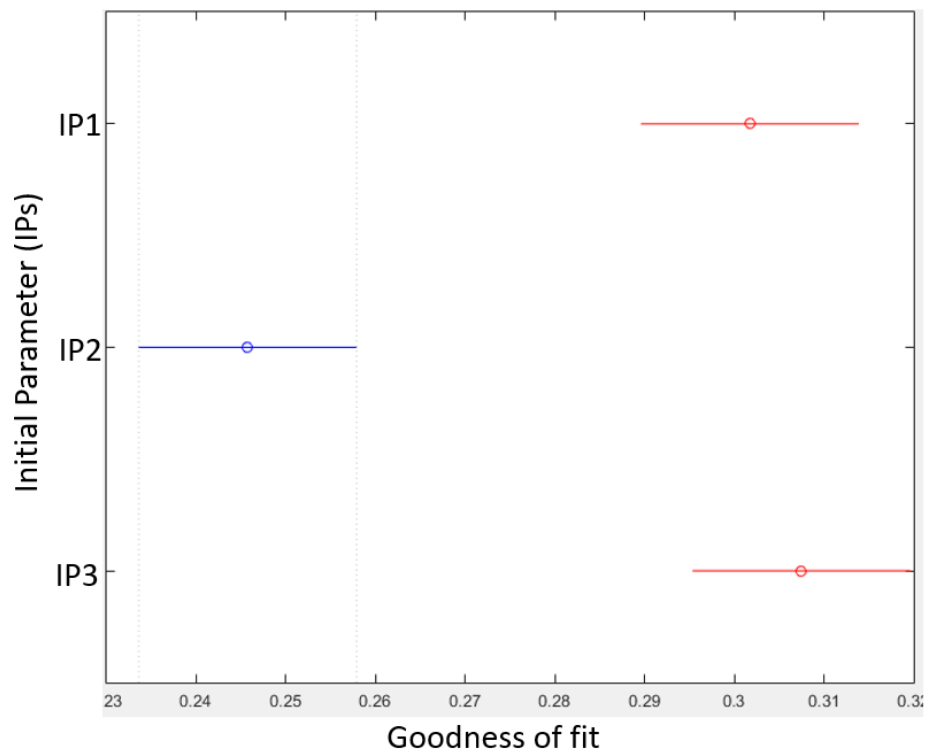


Figure 4-12: Post-hocs revealed that, when averaged over all window periods, subjects and stimulus frequencies, the model achieved significantly better fits for IP1 and IP3 compared to IP2.

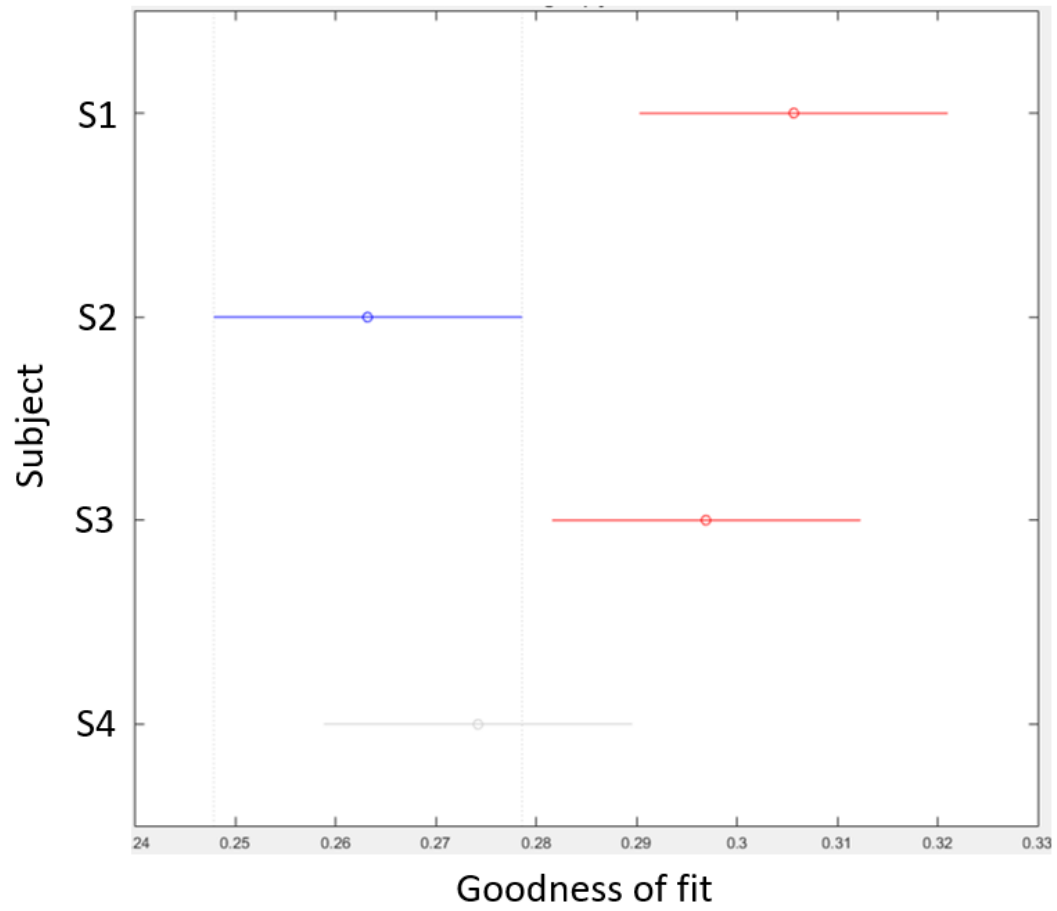


Figure 4-13: Post-hocs revealed that, when averaged over all window periods, stimulus frequencies and initial parameter sets, the model achieved significantly better fits to SSVEP responses recorded in subjects 1 and 3 compared to subject 2.

4.4 Assessing Classifier Performance

4.4.1 Cross-validation: SVM in-sample classification performance

Figure 4-14 shows the in-sample resubstitution losses of each SVM for each window period and subject. Since the in-sample loss represents the proportion of data used in the training set that is classified incorrectly by the SVM, low losses are typical. Although in-sample loss values were generally between 10 and 20%, the loss reached as high as 30% in some instances. In all subjects, SVMAll demonstrated good classification performance with losses of 20% or less for all three data window periods.

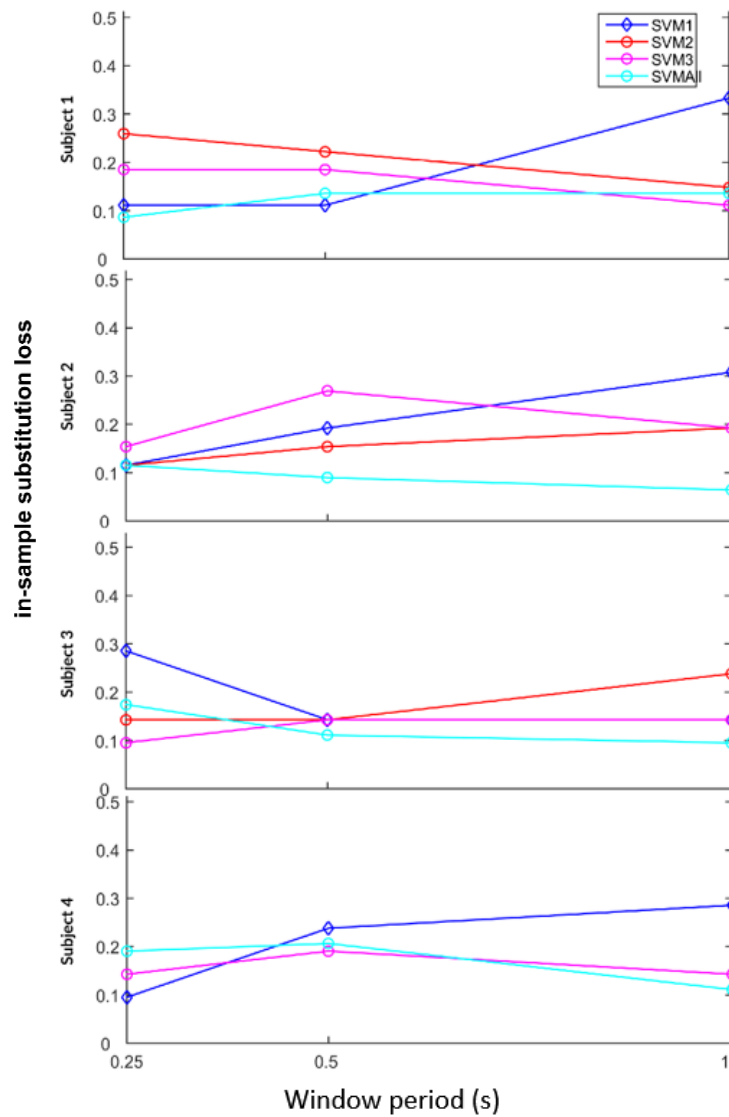


Figure 4-14: Each subject's in-sample resubstitution loss for different SVMs for different window periods of data.

Figure 4-15 shows for subject 4 the in-sample confusion matrices of SVMAll for each window period. The confusion matrix indicates true classes on the y-axis vs predicted classes on the x-axis. As such, the diagonal represents true positives (TP).

For example, in the confusion matrix for the 0.25 s window period instance, 6 Hz was classified correctly nine times (TP = 9), and never incorrectly (FN = 0), as seen by the horizontal 6 Hz row. The vertical 6 Hz column, tells us that one 10 Hz signal was incorrectly classified as belonging to the 6 Hz class (FP = 1). All other signals (matrix entries not found in the 6Hz row or column) were correctly classified as not belonging to the 6 Hz class (TN = 53).

In contrast, 10 Hz was classified correctly six times (TP = 6) and incorrectly three times (as 6 Hz, 7 Hz and 7.5 Hz; FN = 3). Two signals, one 7.5 Hz and one 9.3 Hz signal, were incorrectly classified as belonging to the 10 Hz class (FP = 2), and all other signals were correctly classified as not belonging to the 10 Hz class (TN = 52).

From the confusion matrices of multiclass classifiers the accuracy, precision and recall, defined in section 2.5, can be calculated for each class. Accuracy reflects the fraction of all classifications that are true (i.e. $(TP+TN)/(TP+TN+FP+FN)$); precision, the fraction of positive classifications that are true (i.e. $TP/(TP+FP)$); and recall, the fraction of signals belonging to a class that are correctly assigned to that class (i.e. $TP/(TP+FN)$). **Figure 4-16** shows for subject 4 the in-sample class-specific accuracy, precision and recall of SVMAll for the 0.25 s window period.

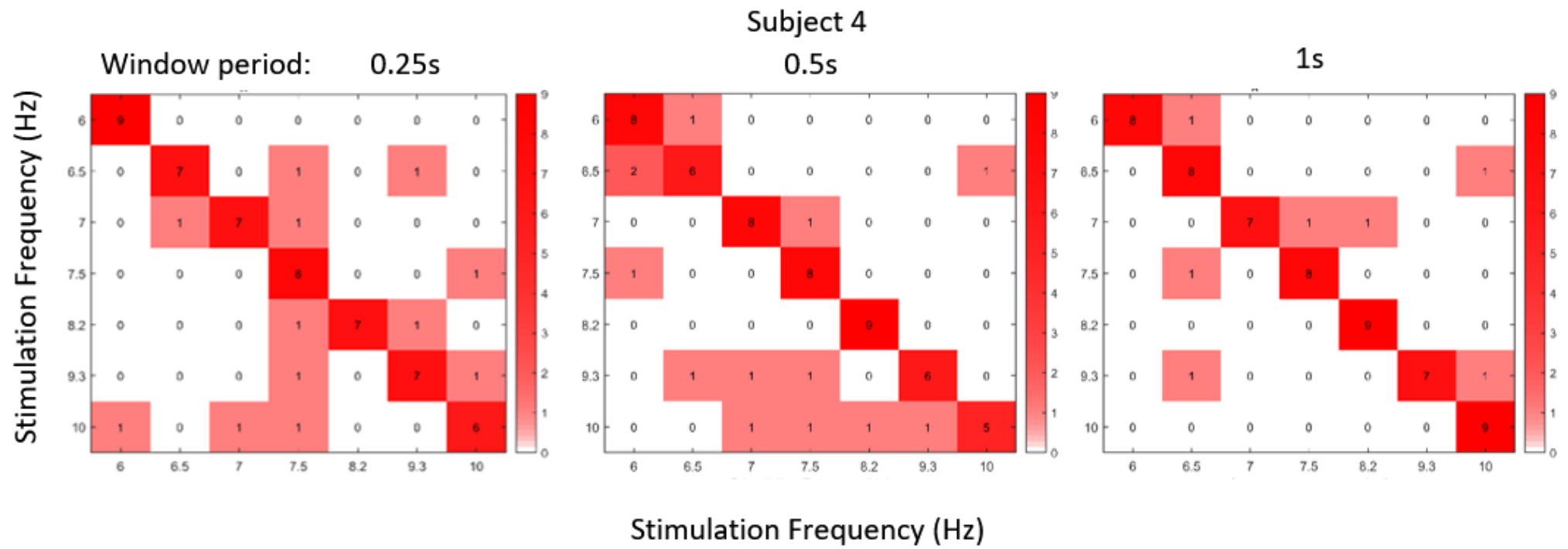


Figure 4-15: In-sample confusion matrices of SVMAll for Subject 4 for each window period.

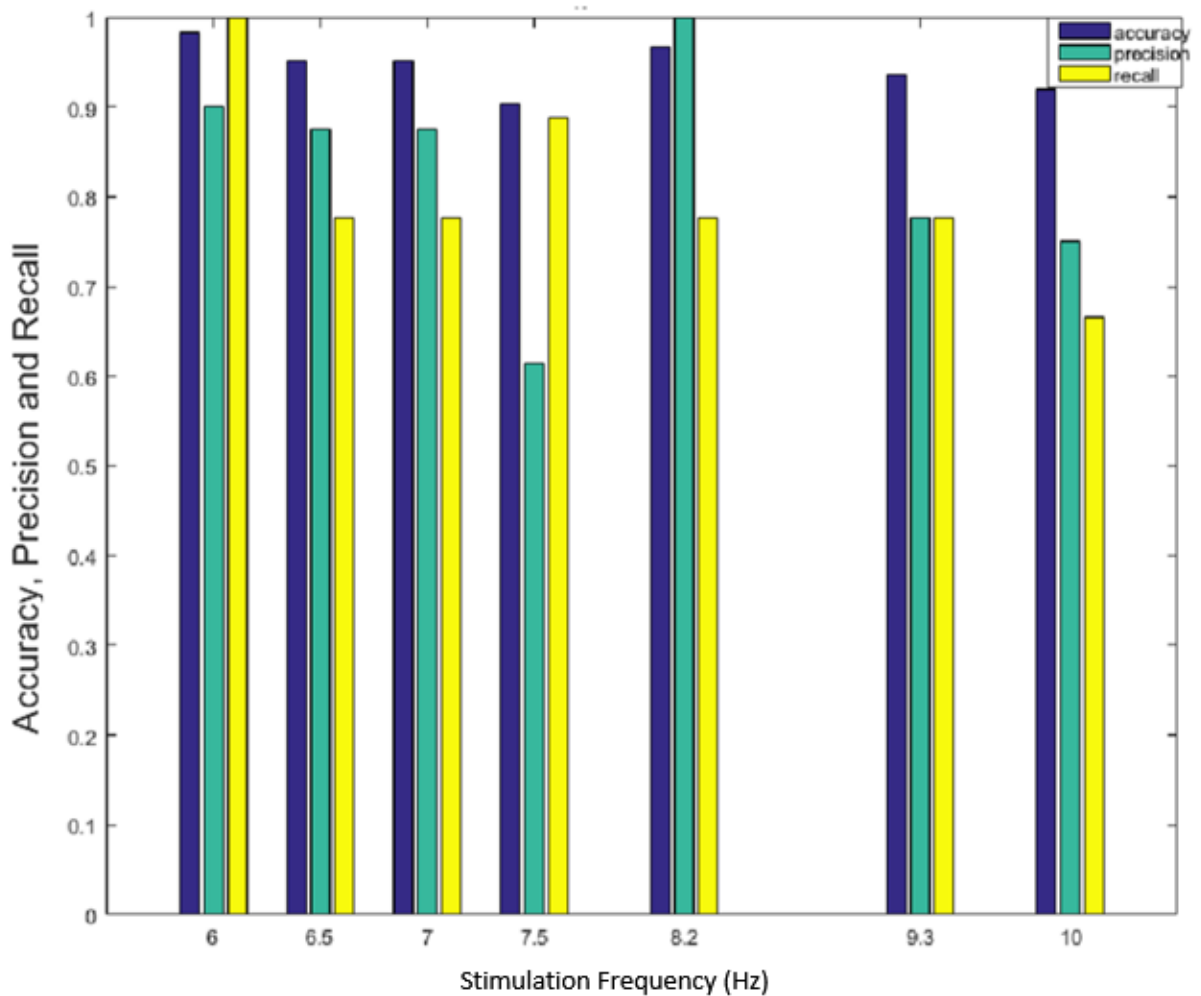


Figure 4-16: In-sample class-specific accuracy, precision and recall of SVMAll for subject 4 and a window period of 0.25 s. Accuracy reflects the fraction of all classifications that are true; precision, the fraction of positive classifications that are true; and recall, the fraction of signals belonging to a class that are correctly assigned to that class.

4.4.2 Cross-validation: SVM out-of-sample classification performance

Figure 4-17 shows the out-of-sample resubstitution losses of each SVM for each window period and subject. The out-of-sample losses were typically high (between 80 and 100%), corresponding to values one would expect by chance. Since there were seven classes to classify, a random guess would have a 1 in 7 chance of correctly classifying the data, corresponding to a loss of 85.8%. A loss of 100% indicates that none of the data excluded from the training set were correctly classified by the SVM. As with the in-sample case, SVMAll was least affected by the data window period.

Out-of-sample confusion matrices for each window period are presented for the SVMAll and CCA classifiers in **Figure 4-18** and **Figure 4-19**, respectively, for all subjects combined. These were constructed by adding the out-of-sample confusion matrices of individual subjects. As such, for SVMAll, which includes fits from all three initial parameter sets in the training set, entries on the diagonal represent the number from a total of 36 signals (3 per subject for each of 4 subjects and 3 IPs) per class

that were classified correctly (TP). For CCA, diagonal entries represent the number from a total of 12 signals per class that were classified correctly. The colour coding serves to highlight the performance of the classifier. A large number of true positives, compared to FNs and FPs, would result in hot colours on the diagonal and cool colours elsewhere. As such, the distribution of hot colours throughout the confusion matrix of SVMAll for the 0.25 s window period reflects poor performance, compared to relatively more hot colours along the diagonal and cool colours off the diagonal in the CCA confusion matrix for the 1 s window period signifying better performance. **Figure 4-20** and **Figure 4-21** show the corresponding class-specific accuracy, precision and recall for each window period for the SVMAll and CCA classifiers, respectively.

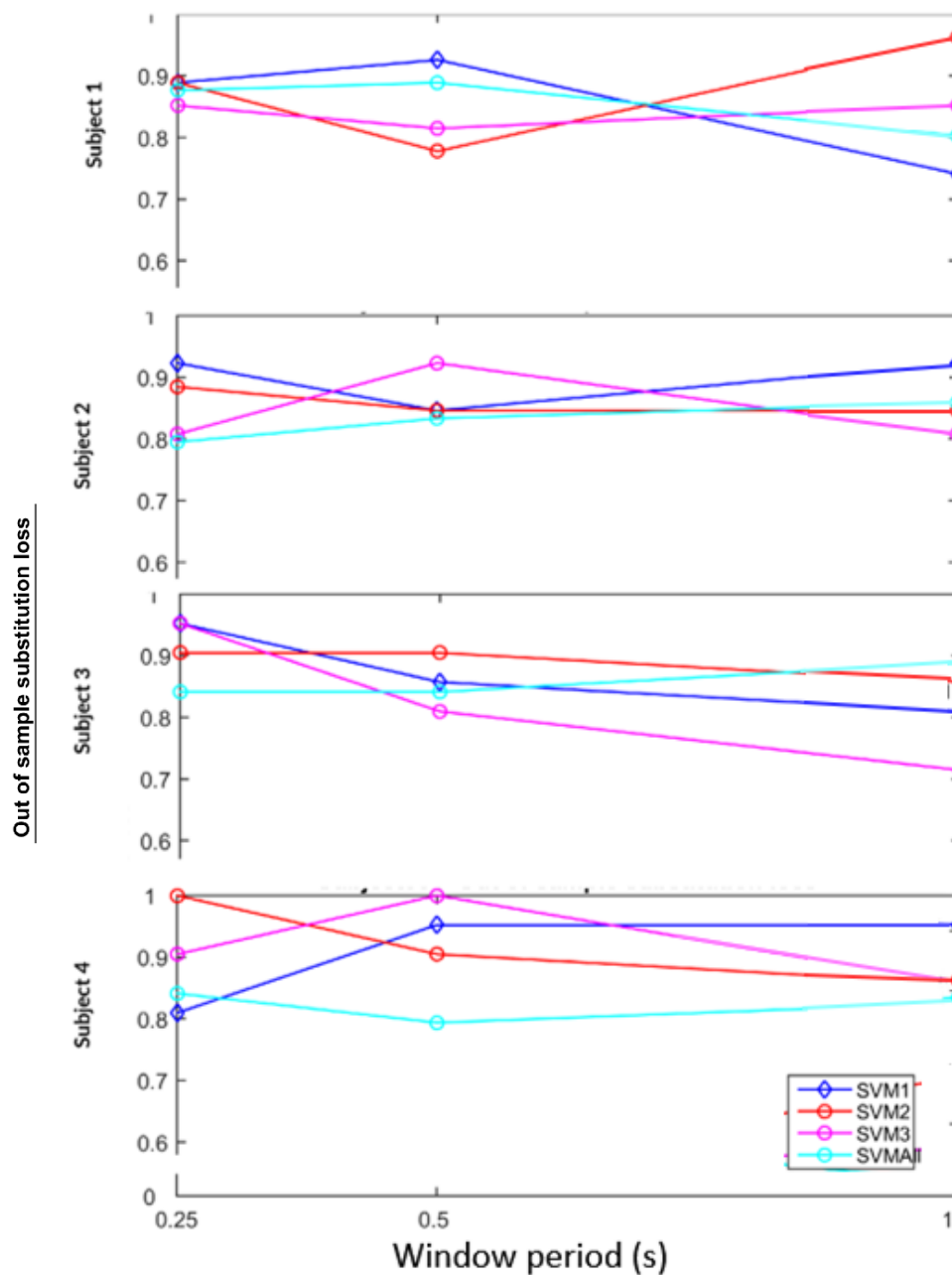


Figure 4-17: Each subject's out-of-sample resubstitution loss for different SVMs for different window periods of data.

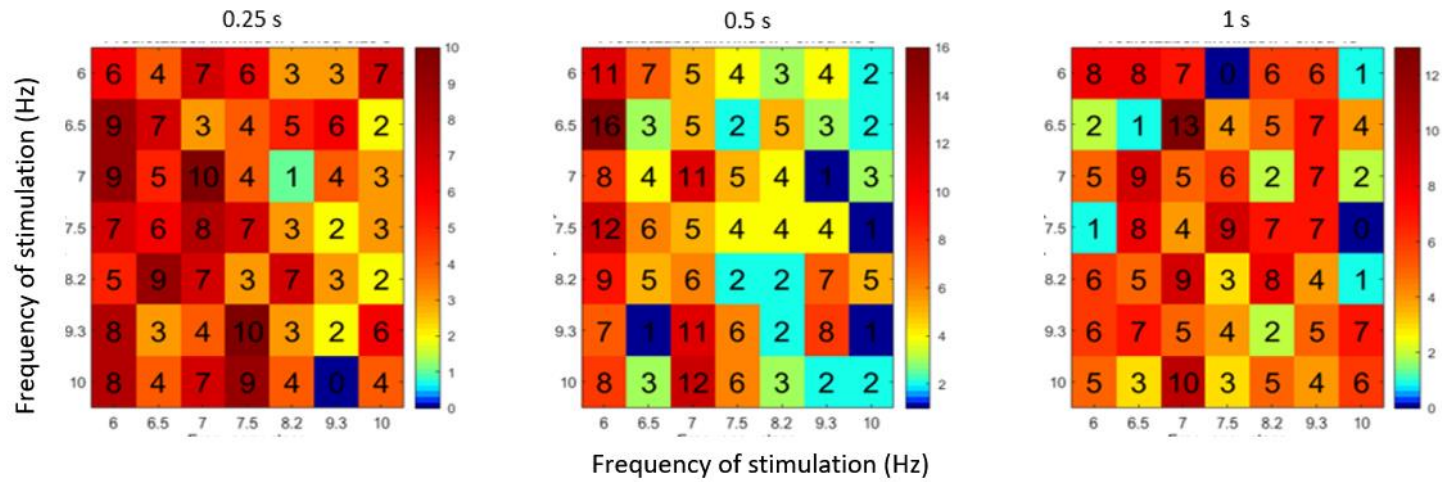


Figure 4-18: Out-of-sample confusion matrices of SVMAll summed across subjects for data window periods 0.25 s, 0.5 s, and 1 s, respectively.

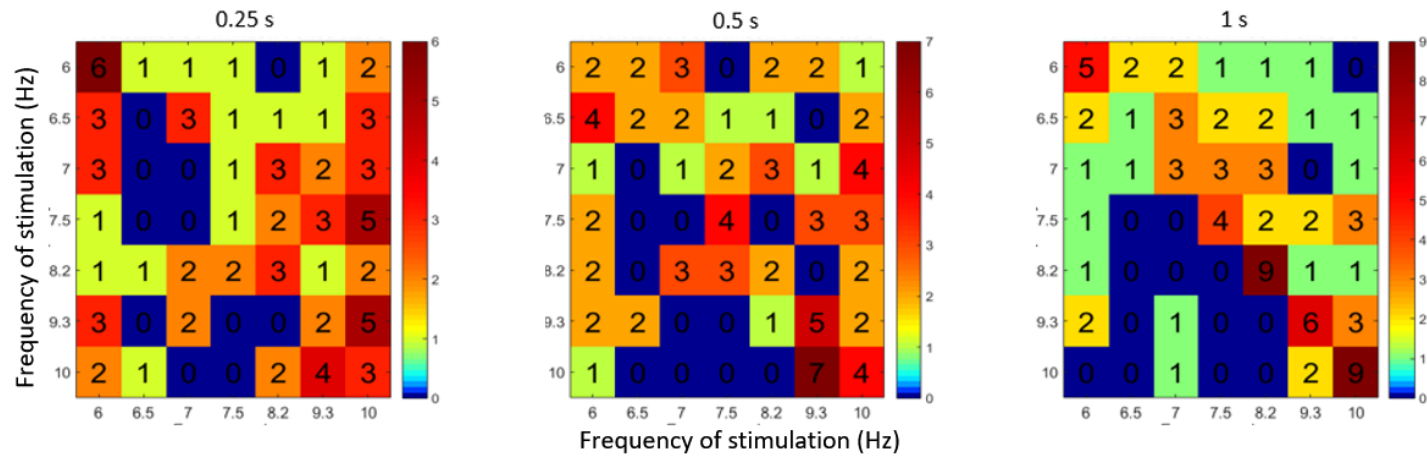


Figure 4-19: Out-of-sample confusion matrices of CCA summed across subjects for data window periods 0.25 s, 0.5 s, and 1 s, respectively.

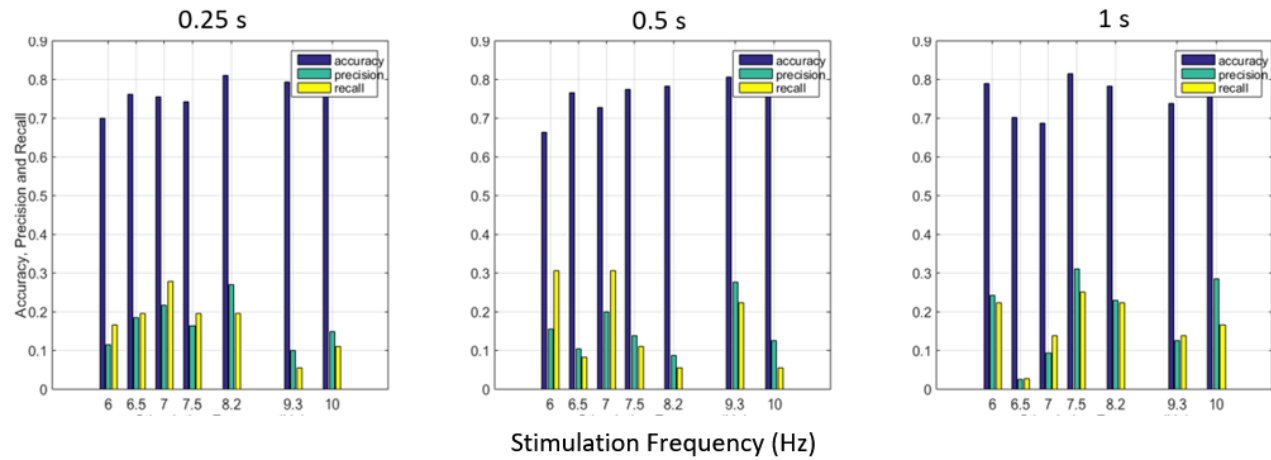


Figure 4-20: Class-specific accuracy, precision and recall for the SVMAll classifier across all subjects for window periods 0.25 s, 0.5 s, and 1 s, respectively.

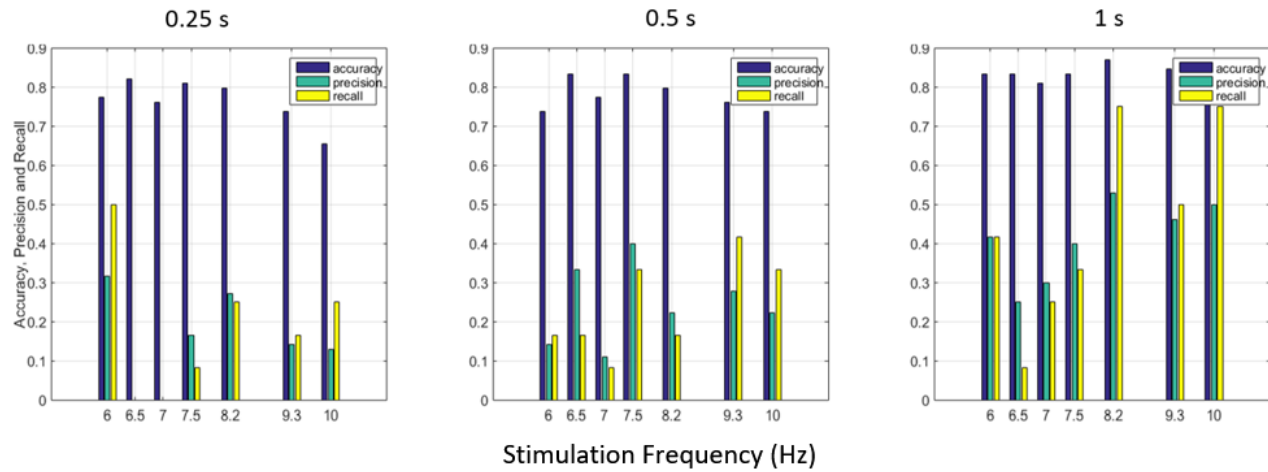


Figure 4-21: Class-specific accuracy, precision and recall for the CCA method across all subjects for window periods 0.25 s, 0.5 s, and 1 s, respectively.

4.4.3 Statistical comparison of classifiers

For the SVM classifiers, the effects of window period, classifier and frequency class on accuracy, precision, and recall were examined using three-way ANOVAs. Accuracy only showed a significant effect of frequency ($F = 5.05$, $p = 0.0008$, $df = 6$), with 8.2 Hz, 9.3 Hz and 10 Hz stimuli achieving better accuracies than 6 Hz stimuli (**Figure 4-22**). In contrast, precision and recall both showed significant effects of classifier ($F = 6.01$, $p = 0.002$, $df = 3$ and $F = 4.61$, $p = 0.0079$, $df = 3$, respectively), with SVMAll performing better than SVM1, SVM2, and SVM3 on precision and better than SVM1 on recall (**Figure 4-23**). Neither accuracy, precision nor recall showed any effect of window period (all p 's > 0.40).

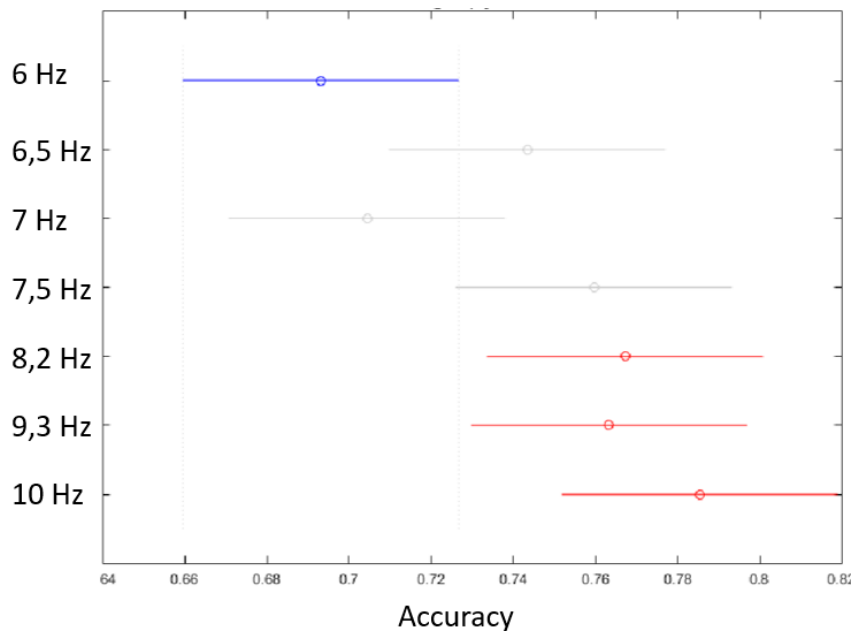


Figure 4-22: Post-hocs revealed that the SVM classifiers achieved greater accuracies for 8.2 Hz, 9.3 Hz and 10 Hz stimuli compared to 6 Hz stimuli.

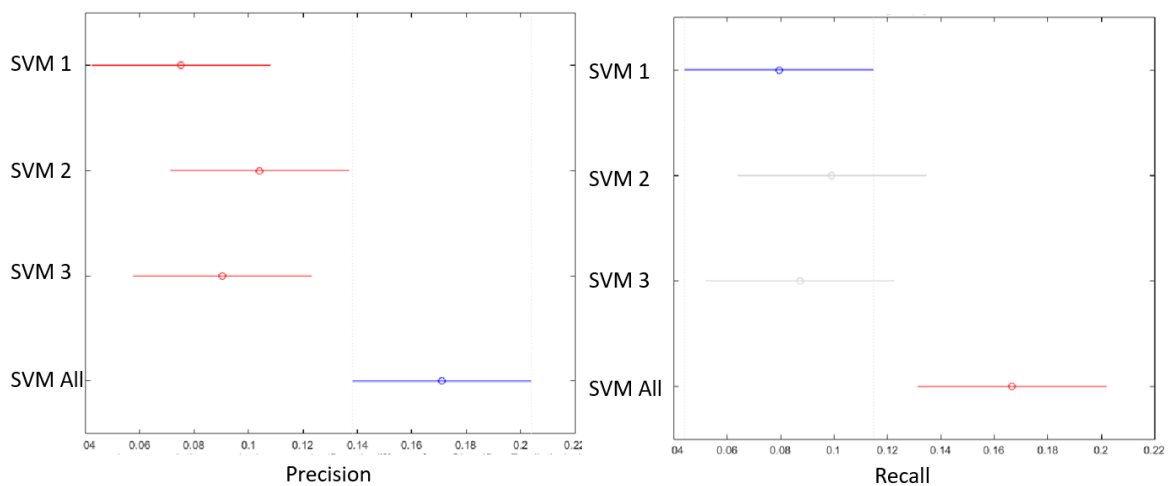


Figure 4-23: Post-hocs revealed that SVMAll achieved better precision (left) and recall (right) than the other SVM-based classifiers.

In contrast to SVMAll that did not show an effect of window period on accuracy, precision or recall (all p 's > 0.75; **Figure 4-24**), for CCA all three measures showed a significant effect of window period (accuracy: $F = 6.2$, $p = 0.0089$, $df = 2$; precision: $F = 10.3$, $p = 0.001$, $df = 2$; and recall: $F = 3.67$, $p = 0.046$, $df = 2$) (**Figure 4-25**). Post hocs revealed that accuracy and precision improved for the 1 s window period as compared with the two shorter window periods. In the case of recall, the 1 s window period produced significantly better results than the 0.25 s window period.

Figure 4-26 shows the distribution of class-specific accuracy, precision and recall measures for the four SVM-based classifiers and CCA when data from all the window periods were combined. All three measures showed significant effects of classifier (accuracy: $F = 4.11$, $p = 0.004$, $df = 4$; precision: $F = 12.63$, $p < 0.0001$, $df = 4$; and recall: $F = 10.67$, $p < 0.0001$, $df = 4$). Post-hocs revealed that CCA yielded better precision and recall than all the SVM-based methods, as well as better accuracy than SVM1, SVM2, and SVM3. Only on precision did SVMAll perform better than SVM1 and SVM2.

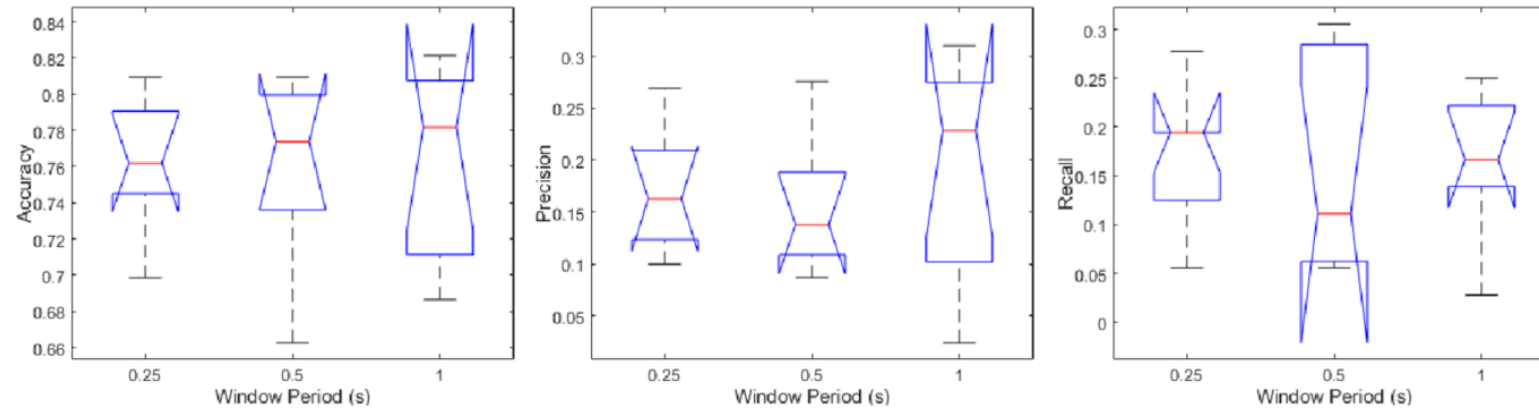


Figure 4-24: Box-and-whisker plots showing for SVMall the distribution of class-specific accuracy (left), precision (middle) and recall (right) measures as a function of window period.

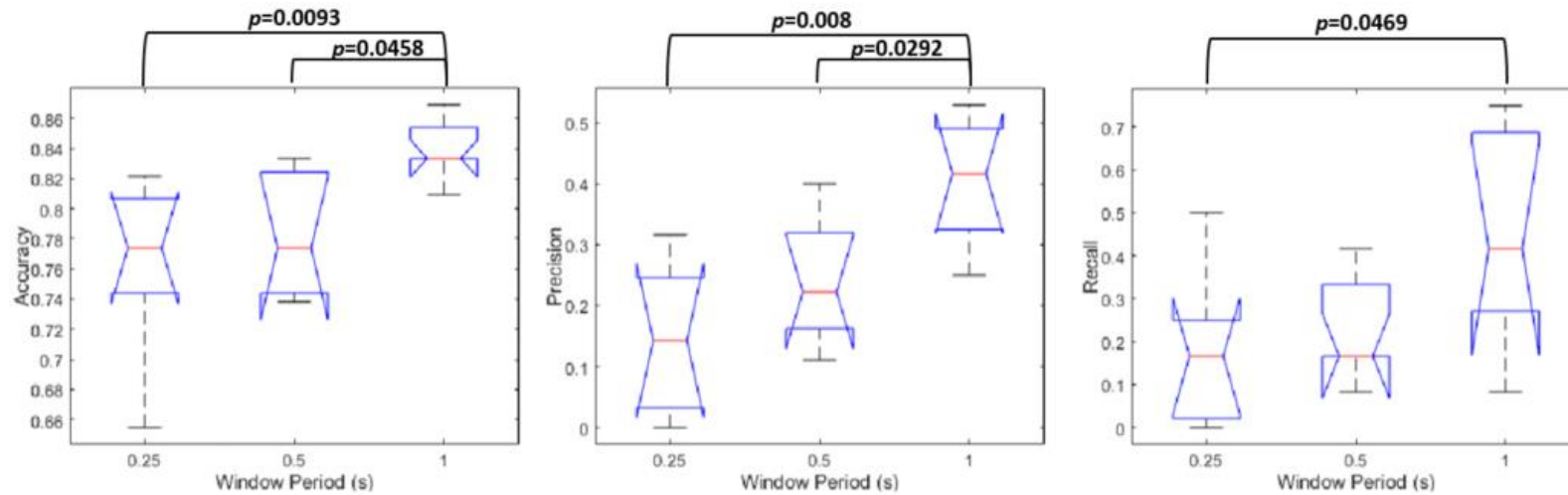


Figure 4-25: Box-and-whisker plots showing for CCA the distribution of class-specific accuracy (left), precision (middle) and recall (right) measures as a function of window period. Window periods showing significant differences are indicated by brackets above the plot.

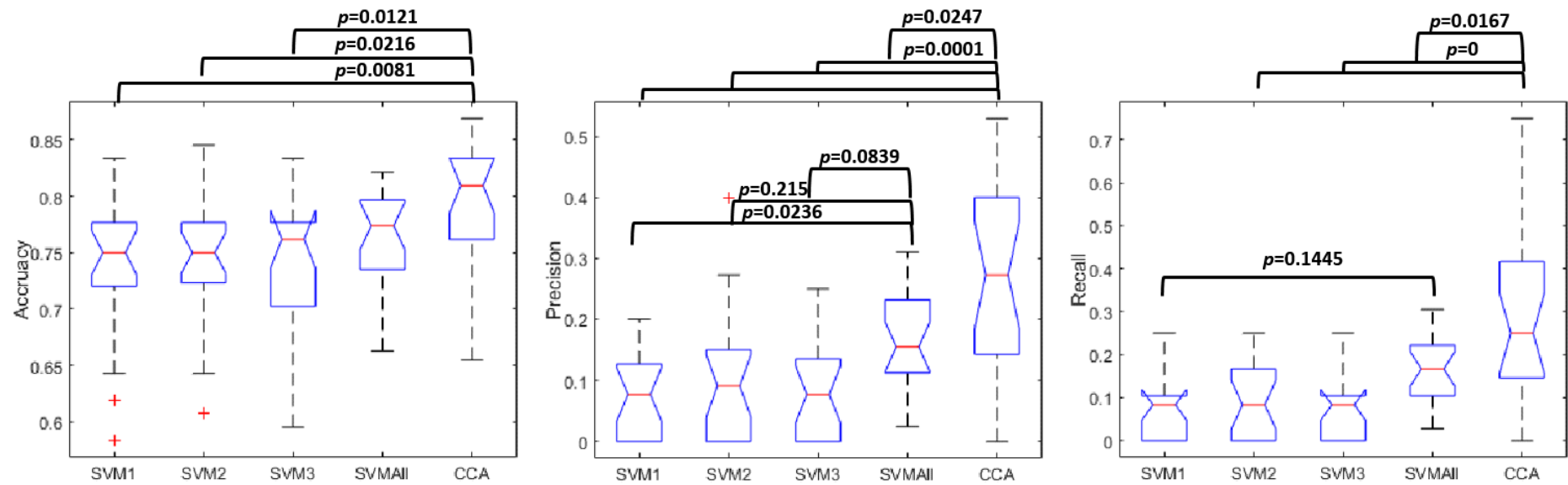


Figure 4-26: Box-and-whisker plots showing the distribution of class-specific accuracy (left), precision (middle) and recall (right) measures for the different SVM-based classifiers and CCA when data from all the window periods are combined. Significant differences are indicated by brackets above the plot.

4.5 Classifier Performance in the Context of BCIs

4.5.1 Probability of a true positive decoding

To evaluate the performance of the classifiers in a multi-class BCI application, we calculated the probability of a true positive decoding ($Prob(TP) = TP/[TP+FP+TN+FN]$). **Figure 4-27** shows $Prob(TP)$ for all subjects combined as a function of window periods for each of the classification methods. These values reflect the fraction of all signals classified that are assigned to the correct class. Consistent with the results shown previously for accuracy, precision and recall (Figures 4-24 and 4-25), $Prob(TP)$ is unaffected by window period for the SVM-based classifiers while it increases with increasing window period for CCA and PSDA. **Figure 4-28** shows $Prob(TP)$ as a function of window period for each subject individually. Similar trends are seen as those observed in Figure 4-27 for the case of all subjects combined.

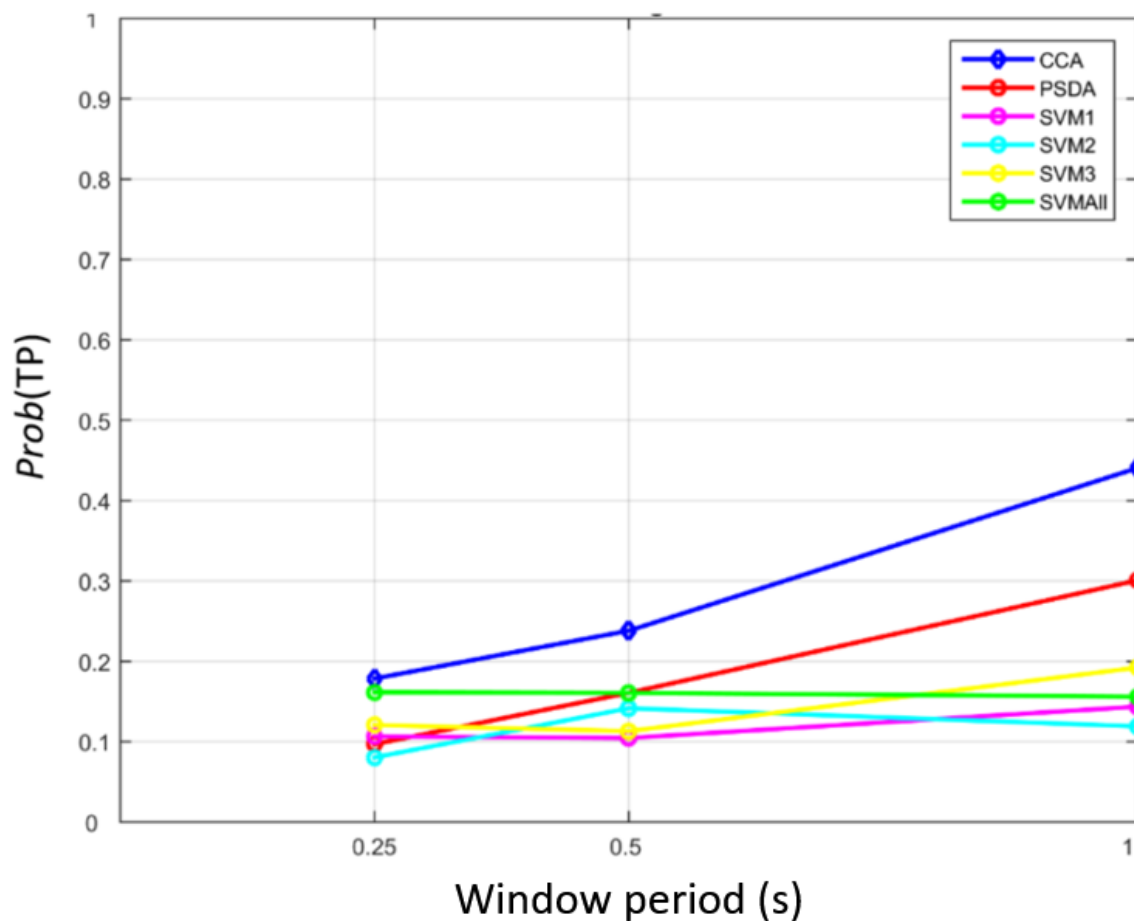


Figure 4-27: Comparison of the probabilities of a true positive decoding ($Prob(TP)$) for all subjects combined, for different classification methods as a function of window period.

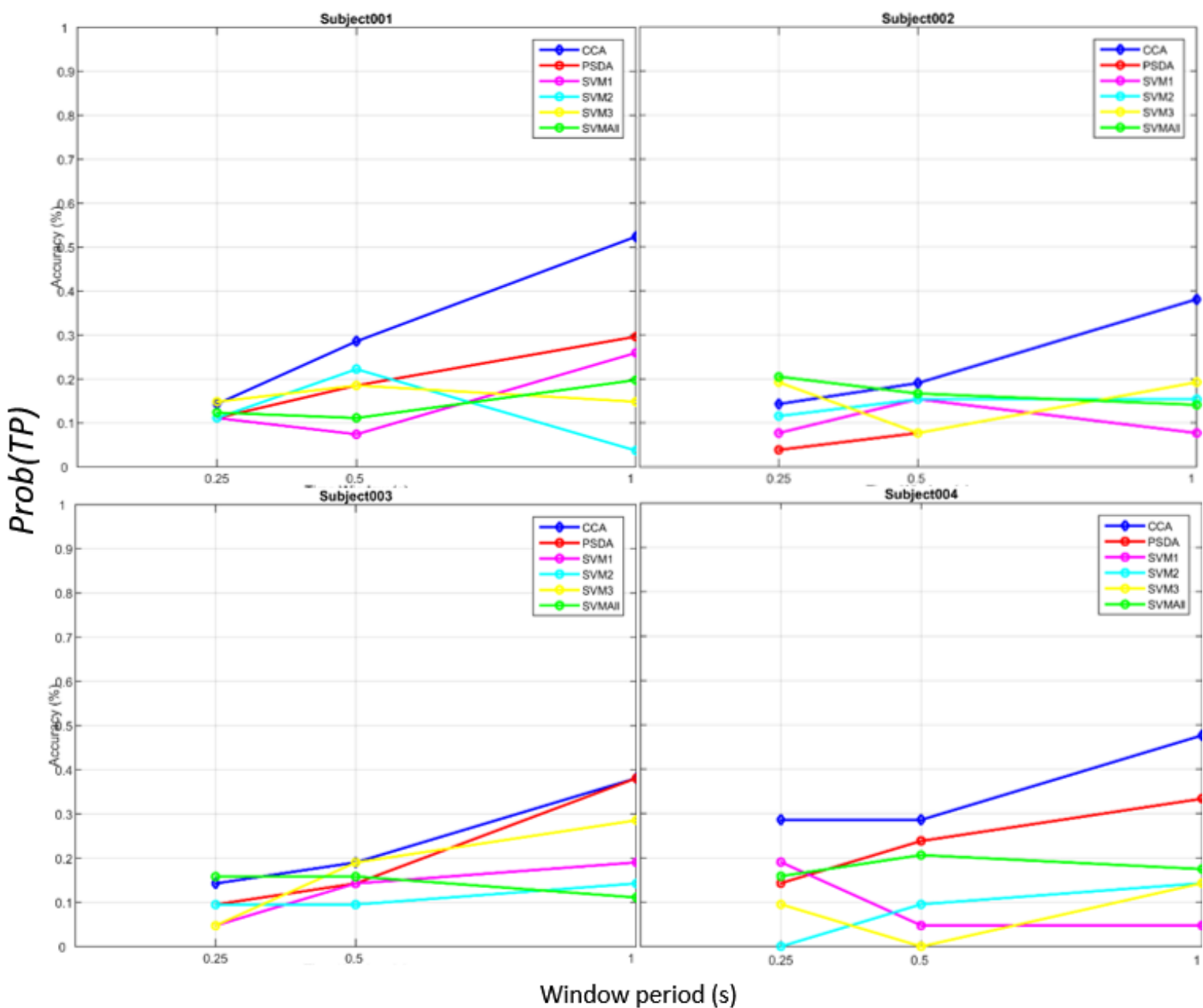


Figure 4-28: Comparison in each subject of $Prob(TP)$ for different classification methods as a function of data window period.

4.5.2 McNemar testing

Mc Nemar testing was used at each window period to examine whether $Prob(TP)$ was significantly better for certain classification methods as compared to others. Tables 4-1 to 4-3 show examples for three different subjects and different window periods of results from Mc Nemar tests. The null hypothesis was that the classifier in each row is *at most as accurate as* the classifier in each column. Green shaded regions indicate rejection of the null hypothesis with 95% certainty, i.e. that the classifier in the row is more accurate than the classifier in the column. Each table entry shows the p-value of the statistical comparison. McNemar tests show, for example, that no classifier outperforms any other classifier in subject 2 for 0.25 s of data (Table 4-1, all p 's > 0.28). In contrast,

for the same subject but a 1 s window period of data, the CCA method is significantly more accurate than SVM1, SVM2, and SVMAll (Table 4-2). Similarly, CCA outperforms all the SVM-based methods in subject 4 for a 1 s window period of data (Table 4-4).

The results of the McNemar tests performed in all subjects are summarised in **Figure 4-29**. For each window period, each entry in the matrix shows the number of instances in which the null hypothesis was rejected in any of the 4 subjects, notably, for window period 1 s, CCA outperforms the SVM-based methods in at least 3 of 4 subjects. At shorter window periods, CCA and SVMAll outperforms some of the other SVM-based methods in only 1 or 2 subjects.

Table 4-1: McNemar test results: p-values for subject 2, window period = 0.25s. Green shaded regions indicate rejection of the null hypothesis which implies that the classifier in the row is more accurate than that in the column.

	SVM1	SVM2	SVM3	SVMAll	CCA
SVM1	1.0000	0.3274	0.5000	0.2819	0.5000
SVM2	0.6726	1.0000	0.6726	0.5000	0.6726
SVM3	0.5000	0.3274	1.0000	0.2819	0.5000
SVMAll	0.7181	0.5000	0.7181	1.0000	0.6726
CCA	0.5000	0.3274	0.5000	0.3274	1.0000

Table 4-2: McNemar test results: p-values for subject 2, window period = 1s. Green shaded regions indicate rejection of the null hypothesis which implies that the classifier in the row is more accurate than that in the column.

	SVM1	SVM2	SVM3	SVMAll	CCA
SVM1	1.0000	0.7181	0.9488	0.1587	0.9902
SVM2	0.2819	1.0000	0.9584	0.0786	0.9711
SVM3	0.0512	0.0416	1.0000	0.0127	0.8716
SVMAll	0.8413	0.9214	0.9873	1.0000	0.9977
CCA	0.0098	0.0289	0.1284	0.0023	1.0000

Table 4-3: McNemar test results: p-values for subject 4, window period = 1s. Green shaded regions indicate rejection of the null hypothesis which implies that the classifier in the row is more accurate than that in the column.

	SVM1	SVM2	SVM3	SVMAll	CCA
SVM1	1.0000	0.1284	0.0899	0.2071	0.9873
SVM2	0.8716	1.0000	0.5000	0.7181	0.9943
SVM3	0.9101	0.5000	1.0000	0.6726	0.9943
SVMAll	0.7929	0.2819	0.3274	1.0000	0.9959
CCA	0.0127	0.0057	0.0057	0.0041	1.0000

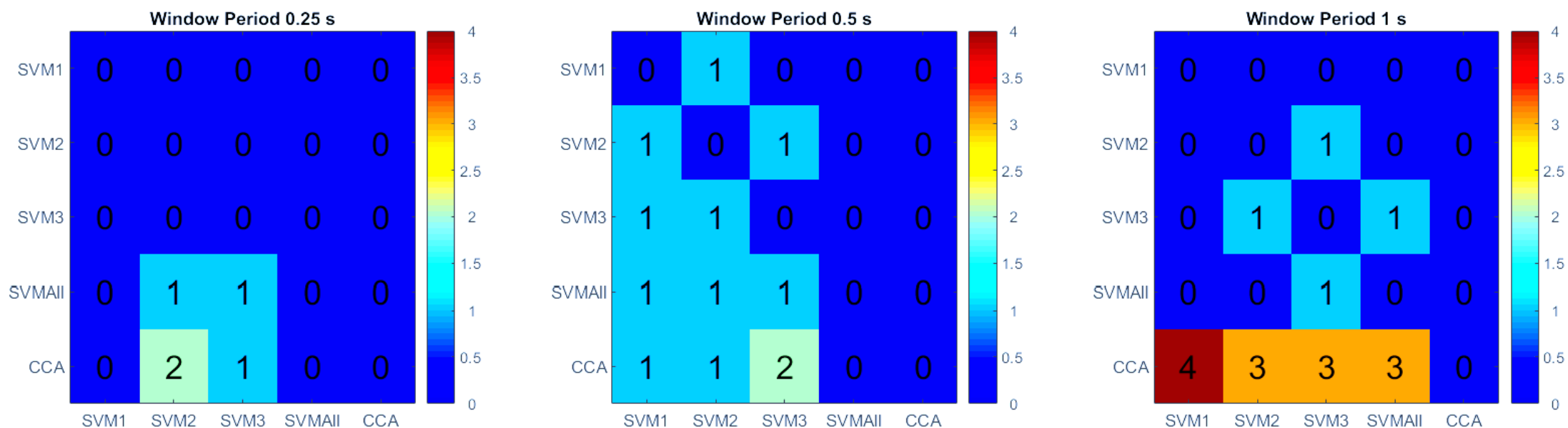


Figure 4-29: Summary of the results from McNemar tests performed for each subject at each window period. For each window period, each entry in the matrix shows the number of instances in which the null hypothesis was rejected in any of the 4 subjects. As such, values reflect the number of subjects in whom the classifier in the row performed better than the classifier in the column.

4.5.3 Hypothetical ITRs

Figure 4-30 shows average hypothetical ITRs corresponding to the above $Prob(TP)$ values for the SVM, CCA and PSDA classification methods as a function of window period. ITRs for each subject individually are shown for the different classification methods as a function of window period in **Figure 4-31**. The hypothetical ITR generally follows a trend of decreasing with increased window period.

The ITR generally decreases, as expected, with an increase in window period, which demonstrates a strong need to reduce the window period to improve BCI performance. Whereas the SVM scheme was less accurate, it allowed for a higher theoretical ITR than that of the PSDA method. The CCA method still outperformed the SVM.

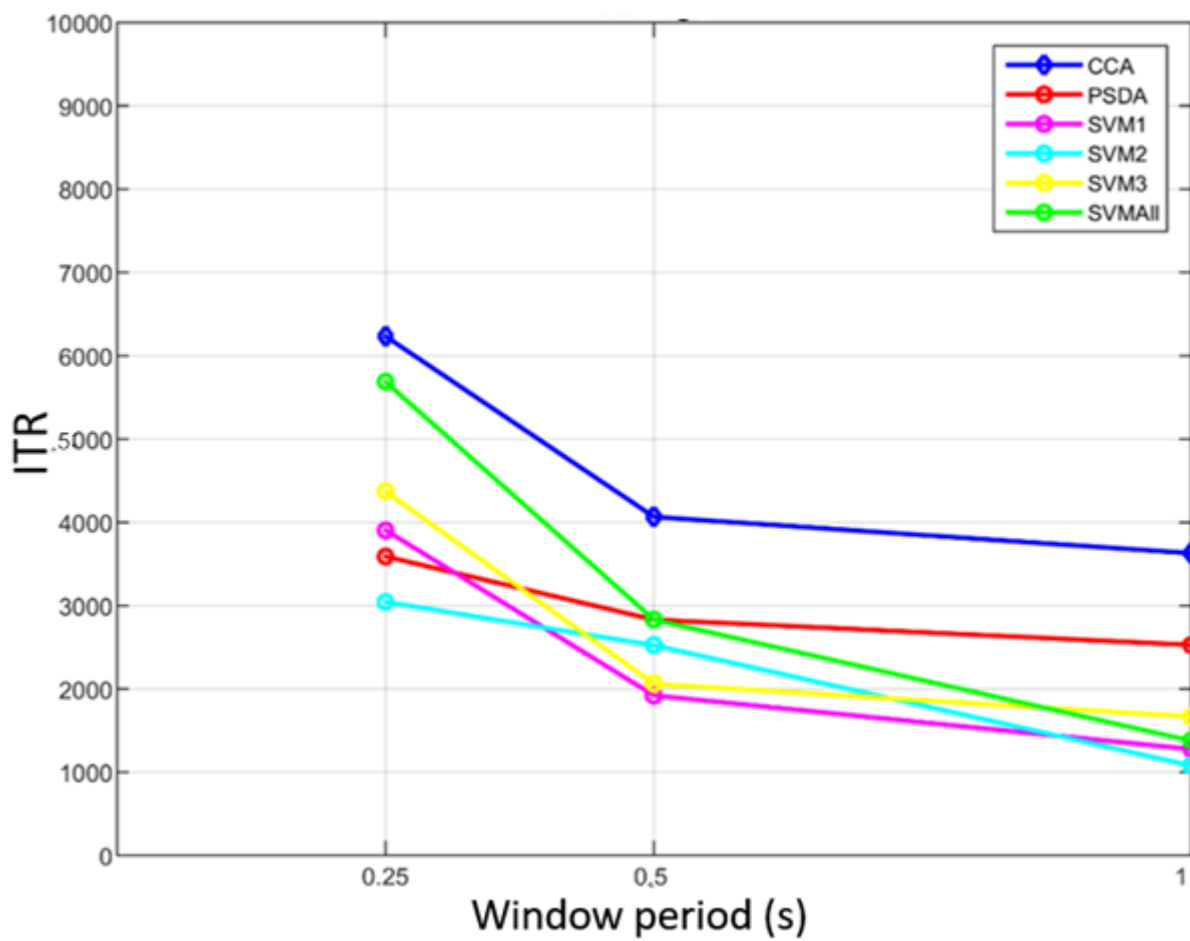


Figure 4-30: Average theoretical ITRs for different classification methods as a function of window period. ITR values are artificially inflated due to the assumption of instantaneous processing speeds. The hypothetical ITR is thus orders of magnitude larger than normal ITR values.

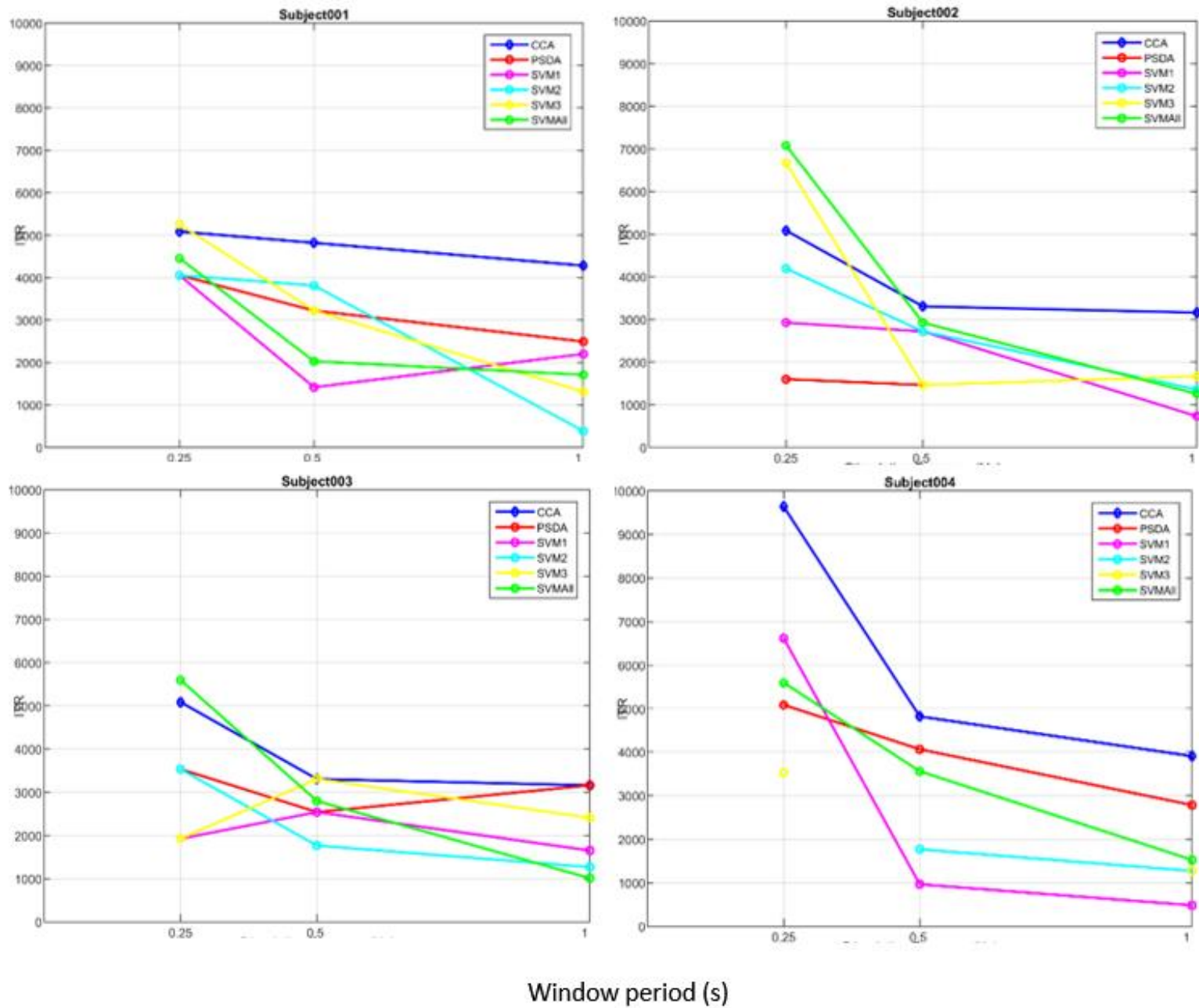


Figure 4-31: Individual subject theoretical ITRs for different classification methods as a function of window period. ITR values are artificially inflated due to the assumption of instantaneous processing speeds. Discontinuities in subject 4 arise from data points with zero accuracy.

5

Discussion

The findings of this study show that it is possible to fit Kremláček's model of VEPs to SSVEP responses, with a goodness of fit within the range of approximately 0.05 to 0.6, depending on the window period and the stimulus frequency (Figure 4-8). This result validates **Hypothesis 1: Kremláček's VEP model can be fit onto the initial VEP portion of an SSVEP response.**

However, the classification performance of the SVM classifiers, informed by the estimated model parameters, was found to be poor. On average, the probability of a true positive decoding ($Prob(TP)$; **Figure 4-27**) was found to be roughly the same as that of a random choice ($\sim 14\%$). On a subject-specific basis, only a few SVMs performed better than random choice, the highest $Prob(TP)$ value achieved being 28% (subject 1, SVM1 at 1 s; subject 3, SVM3 at 1 s; **Figure 4-28**). Overall, it was not possible to determine with any reliability which stimulus was attended by a subject; this may be due to the small dataset (4 subjects) used to train the SVMs. We therefore reject both **Hypothesis 2a: The model fit onto the initial portions of SSVEP signals can generate unique feature-descriptive parameters that relate to the frequency of the stimulus presented** and **Hypothesis 2b: These unique features enable the SSVEP signal to be classified relative to the stimulus frequency, using a multiclass SVM approach.**

The CCA method at any of the three window periods (0.25 s, 0.5 s or 1 s; **Figure 4-29**) outperformed the SVM-based classifiers in contradiction of **Hypothesis 3: The classification approach based on the time-domain model proposed here outperforms traditional spectral classification methods (PSDA and CCA) when the input to the classifiers comprises a 1 s or less window period of EEG data as recorded from stimulus onset; the shorter the window period, the more pronounced the effect.**

A few observations could be made regarding the SVM classifiers. The SVM classification performance at the shortest window period (0.25 s) is on average comparable to the CCA and PSDA approaches (**Figure 4-28** and **4-30**). At larger window periods the CCA and PSDA methods improved in performance as expected; however, the SVM-based approach did not improve noticeably. On a subject-specific basis, the SVM performance varied without any discernible patterns being evident (**Figure 4-28** and **4-31**).

What follows is a discussion of the various factors that contributed to these results in the context of this and future research.

5.1 Using Kremláček's Model for fitting SSVEPs

When fitting VEP data, Kremláček et al. (2002) averaged data to remove spontaneous EEG activity and to account for noise. No such averaging was performed in this study, although an attempt was made to train and classify the data using a single stimulus exposure which resulted in much more spurious EEG activity being present in the data. The dataset used was also very small, consisting of only 4 subjects. Future work would need to address these limitations by using a larger dataset that might show improvements in the model's potential as a classifier when applied to SSVEPs.

Various factors influence the model fit and classification process, namely: (i) the choice of initial model parameters; (ii) the downhill-simplex method of parameter estimation; (iii) the use of NRMSE as a goodness-of-fit function; (iv) the portion of SSVEP being fit to; (v) the window period used in the fit, and (vi) the selection of the stimulus frequencies used.

5.1.1 Initial model parameters

The ability of the model to fit to the VEPs and SSVEPs was dependent on the choice of initial parameters (IPs). This was clearly evident during the initial exploration, where a set of initial parameters all set to zero (**Figure 4-3**) produced far worse fits than those obtained using the input parameters recommended by Kremláček et al. (2002) (**Figure 4-2**). Furthermore, the choice of IP was seen to produce a significant difference in the goodness-of-fit values (**Figure 4-12**), even though the size of this difference was small (0.07). An improved selection of the IPs may therefore yield improved fit results, which Kremláček et al. (2002) also postulate.

A possible method to identify better IPs would be to assess a larger range of subjects and stimulus frequencies, the resultant estimated parameters found in this process could then be used as initial parameters for future fits. Furthermore, EPs found in this way should be obtained from fits on grand averages of SSVEP data instead of the single-stimulus exposure used in this study. It is expected that averaging out spurious EEG activity would result in better fits and therefore improved IPs.

5.1.2 Downhill simplex method of parameter estimation

The specific pattern search method used in the fit process, namely the downhill simplex method, may have played a significant role in the model's performance. The model fit was limited to a set number of 500 iterations and two restarts; it is possible that increasing the number of iterations could improve the fit, simply because more possible solutions could be tested. However, the disadvantage of that would be increased processing time. The downhill simplex method is computationally expensive, with a single SSVEP fit of 500 iterations taking approximately 30 minutes to complete. Such long fit times may be acceptable for training a BCI system on a user offline; it is, however, a limitation of using a time-domain model in a BCI context, because SSVEP systems that use spectral classification techniques generally do not require extended periods of training.

Future work should look at exploring the effect of alternative parameter estimation methods on the goodness of fit and subsequent classification performance, as well as on the processing time required. Furthermore, a better selection of initial parameters would reduce the fit time required as the IPs would be closer to the expected EPs.

Information transfer rate (ITR)

The theoretical ITR used in this study makes an assumption regarding the window period required to classify, namely that the computational time was zero. If the actual computational time were taken into account, the CCA would perform by an order of magnitude better than the SVM classifiers, as a result of the large difference in processing times. This highlights a limitation of the theoretical ITR used in this study due to its extending the ideas behind ITR and fixing certain values.

5.1.3 NRMSE as a goodness-of-fit function

Using NRMSE as the cost function for the goodness-of-fit criteria may not be the best approach for reducing the errors in fit: whereas it gives a mathematical description of fit, there are instances where a perceived ‘good’ fit may not have an actual valid fit.

Figure 4-7 shows three fits for each EP generated (subject 3 stimulated at 6 Hz). The goodness of fit is indicated as 0.41 for EP2 in red and 0.44 for EP3 in green. While there is a very small difference (0.03) in the goodness-of-fit values between the two fits, visual inspection of **Figure 4-7** shows that EP3 is closer to the measured signal in form, whereas the fit from EP2 has sudden, almost discontinuous changes in amplitude that are uncharacteristic of EEG activity.

It is possible that the model may fit onto underlying trends in the data that are visually inconsistent but which are mathematically valid when assessed by the NRMSE method. Mean squared error has the disadvantage of heavily weighting outlying points, so that when an attempt is made globally to minimise such points during the pattern search, the model may have exaggerated responses to the forced linear minimisation of global maxima (Chai & Draxler, 2014). Owing to this weighting, the signal amplitude dominates the fitting process over other aspects such as frequency of oscillation or the selection of the oscillators’ damping factors. This effect is seen in **Figure 4-5**, which shows that the EPs fit better to high-amplitude, low-frequency oscillations (0.35–0.65 s) compared to the lower-amplitude, higher-frequency oscillations (0.7–1 s).

There are many other goodness-of-fit indicators that place emphasis on different aspects of error. These include mean absolute error (Chai & Draxler, 2014) and mean absolute percentage error (Myttenaere, et. al, 2015). In possible future explorations of this work, alternative goodness-of-fit indicators should be examined to find a measure that is more sensitive to matching the fit, based on frequency and not on signal amplitude.

Poor fits have repercussions for machine-learning approaches to data classification, as pointed out by Singla (2014) in the case of neural networks. It may be mathematically possible that a model fit that appears uncharacteristic of the signal being measured can generate features that are unique to the signal and are interpreted as being classifiable.

5.1.4 Fitting to different portions of the SSVEP

An SSVEP consists of two portions: an initial transient response and a steady-state response. The steady-state response emerges ~ 1 s after stimulus onset.

The model was able to produce a synthetic VEP that had the traditional characteristics of a VEP response (N70, P100, N135) and a relatively high goodness of fit (0.74; **Figure 4-2**). However, it was not able to account for the DC voltage offset ($\sim 2.5 \mu\text{V}$ at $t(0)$). The existing model could potentially be improved by adding a new voltage offset parameter to the model in order to shift the model response and more accurately fit the VEP. However, this would be at the cost of increased model complexity. Alternatively, any DC offset present in the SSVEP data could first be eliminated from the signal to ensure that the first SSVEP data point is always at $0 \mu\text{V}$ before presenting the signal to the

model. A more drastic approach would involve using a completely new model, such as those explored by Wang et al. (2013), who used a forced Van der Pol oscillator to model VEP responses, and Zhang et al. (2013), who used their own multivariate synchronisation index (MSI) method to better model the dynamics of the SSVEP response.

The model's ability to fit the transient response (goodness of fit = 0.74; **Figure 4-2**) was found to be less than the fit achieved for the steady-state response of the SSVEPs (goodness of fit = 0.79; **Figure 4-4**). One possible cause of this difference in fit performance is that the attributes of the model – already being oscillatory in nature – have a stronger ability to fit periodic signals, as is typical of SSVEPs.

5.1.5 Window period

This study used shorter window periods than would normally be used with SSVEP-based BCIs and spectral classification methods. The longest window period used (1 s) is generally where the lower bounds of other classification methods start, spectral methods typically not being able to detect SSVEPs reliably below 0.5 s (Bin et al., 2009; see **Figure 1-5**). In this case, three window periods of interest were identified (0.25 s, 0.5 s and 1 s).

Similarly to the spectral methods, the model showed that increasing the window period during a fit produced a corresponding improvement in the goodness of fit (**Figure 4-9**). This may be a combination of an increased window period incorporating more of the steady-state SSVEP, thus enabling the oscillators to fit better onto the signal, and smaller errors in the goodness-of-fit function for the damped higher-frequency oscillations (compare the differences in EPs and SSVEP response for the 1 s window period in **Figure 4-5** against those of the 0.5 s window period in **Figure 4-6**).

5.1.6 Selection of stimulus frequencies

The frequency of the stimulus used did not have a noticeable effect on the fit results (**Figure 4-10**). The goodness-of-fit values were shown to be of similar distribution for all but two (9.3 Hz significantly outperforming 6 Hz) of the stimulation frequencies used. Furthermore, there was no noticeable relationship between frequency of stimulation and goodness of fit across all window periods (**Figure 4-8**). These results are an indication of the robustness of the technique; varying the frequency (as one would do in BCI) yields similar fit results independently of the frequency of stimulation.

Typically, SSVEPs can be seen from 1 Hz to 100 Hz, although the frequency response is not linear and the amplitude decays as the frequency increases (Vialatte, Maurice, Dauwels & Cichocki, 2010). A shortcoming of the selection of frequencies used in this study is that only a narrow band of all possible SSVEP-inducing frequencies was investigated (6–10 Hz). The choice of frequencies investigated was based on a need to obtain the highest possible SSVEP amplitude response (**Figure 3-2**).

Figure 4-22 shows the class-specific accuracy across the SVM-based classifiers. It is interesting to note that 8.2, 9.3 and 10 Hz performed better than 6 Hz, with the accuracy increasing as the stimulus frequency increased. It is possible that this increased accuracy is a result of the larger SSVEP amplitudes (a 10 Hz SSVEP typically has twice the amplitude of a 6 Hz SSVEP; **Figure 3-2**).

Future work could expand the selection of stimulation frequencies to include SSVEPs with reduced signal amplitude. This would be of interest for comparing them against spectral classification methods that rely heavily on amplitudes to distinguish the signals. At higher SSVEP responses, the additive effect of EEG noise makes for a signal that is much more difficult to distinguish (Wang et al., 2006). This may not be the case for the time-based method. Furthermore, the SSVEP amplitude is connected to the stimulus luminosity, with brighter – higher-contrast – images producing more easily classifiable SSVEPs for spectral methods. However, such systems report low usability scores, because they require intense changes in contrast (**Figure 2-4**). If the time-based model proves better suited to lower-amplitude signals, and therefore to the detection of lower-contrasting stimuli, it would be more comfortable for end-users.

Stimulation frequency could potentially be used to improve the performance of the model. It has been shown that SSVEPs can be modelled as compositions of multiple VEPs (Luo & Sullivan, 2010a). As such, the initial oscillator (OSC1) could be excited with a repeated wave oscillating at the stimulus frequency to which a fit is trying to be made. This would add an extra parameter to the model; however, it may force the model to act more in line with that of a repeated VEP – by constantly triggering the model with a forced input. Alternatively, the IPs could be preconfigured to certain generic stimulus frequencies, thereby offering a better initial starting point for the parameter estimation process. This would also reduce the fit time required as the IPs would be closer to the expected EPs.

5.2 SVM Classifier Performance

Overall, the performance of the SVM classifiers was poor, as measured by accuracy, recall, precision and *Prob(TP)*. Possible reasons for this are discussed below, but they include overfitting (5.2.1), separation of parameters (5.2.2), the influence of initial parameter selection (5.2.3) and inter-subject variability (5.2.4).

5.2.1 Cross-validation and overfitting

The SVM classifiers have an average in-sample resubstitution loss of about 20% (**Figure 4-14**). This indicates that when they were based on previously seen data, the SVMs had, on average, an 80% level of accuracy. In-sample resubstitution is an optimistic measure of classifier performance, because good in-sample loss may not be a predictor of classifier performance on unseen data. Yet a bad in-sample loss generally leads to poor classification performance (Anguita et al., 2012). The ability of the SVMs to distinguish accurately (> 80%) between seen data indicates that the data are potentially neatly separable (as described in section 5.2.2).

The out-of-sample losses calculated during the cross-validation process were very high, having an average misclassification rate of about 85% (**Figure 4-17**). Viewed in conjunction with the in-sample losses this misclassification rate indicates that the SVMs may have been overfit (Anguita et al., 2012). Overfitting can be seen in the 0.25 s window period confusion matrix of SVMAll (**Figure 4-18**). Specifically by noticing the increased weights in the 6 Hz column it can be seen that the classifier favours ‘guessing’ that the stimulus is 6 Hz. This could explain the significant difference in class-specific accuracy seen for 6 Hz in **Figure 4-22**. An overfit model tends to perform very well on seen

data and poorly on unseen data, as was found to be the case for the model (see sections 4.4.1 and 4.4.2).

There are two main causes of overfitting: insufficiently large training datasets and the complexity of model parameters. The dataset used in this study was small, each class being observed only three times per subject per window period. The model parameter space was relatively larger, consisting of 12 parameters. If this study were to be repeated, a much larger dataset should be used to reduce the potential for classifier overfitting. A reduction in model parameters may also help to reduce the chance of overfitting.

In an exploratory principal component analysis (PCA), eight optimal parameters were identified. However, the motivation behind the 12 parameters used is that each oscillator is associated with a biological process in the human visual system. The ultimate goal is that the parameters produced by a model fit are able to be used for some form of clinical diagnosis. For example, if there were a much larger than average T1 (initial delay), it might indicate a physiological difference in the subject's primary visual cortex. Whereas the diagnostic aspect can potentially add further clinical value to BCI applications, it is not a requirement for BCI to replicate SSVEP physiology, but only to identify when a subject attends a stimulus.

5.2.2 Separation of parameters

The ability of the classifiers to distinguish between classes correctly is dependent on the distribution of parameters, in the model the parameter space. This is illustrated for a hypothetical two-dimensional parameter space in **Figure 5-1**

Initially the two groups of parameters (or classes) are spaced closely together, but they are still considered to be classifiable because a hyperplane can be fit between them. If each fit has an error introduced to the parameters ($P1 + \text{Error}$, $P2 + \text{Error}$) such that the class grouping is extended into regions R1 and R2, the groupings become less distinct and start to overlap. The hyperplane is then unable to distinguish classes in this region ($R1 \cap R2$). In the context of the SVM classifiers, any variance introduced by the fit error could result in the misclassification of closely spaced EPs.

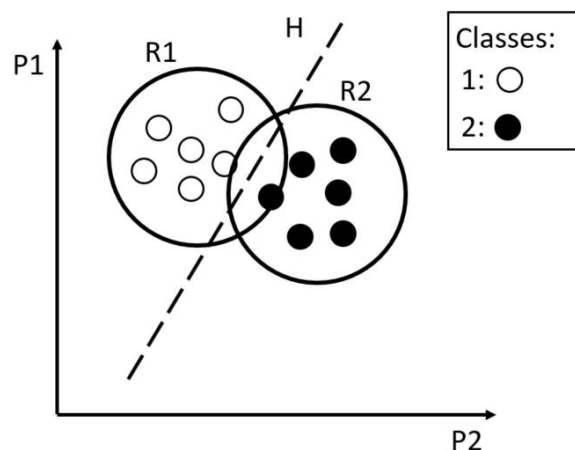


Figure 5-1: In the two-dimensional parameter space P_1, P_2 we have two classes (1 and 2). We assume these represent perfect fits onto data and the classes are neatly separable by the hyperplane H . If an error is introduced to the parameters in each fit ($P_1 + \text{Error}$, $P_2 + \text{Error}$), the class groupings are extended into regions R_1 and R_2 . The hyperplane is no longer able to distinguish classes in the region $(R_1 \cap R_2)$.

Consider, for example, the very closely spaced stimulus frequencies used in this study. The smallest spacing separating the stimuli is only 0.5 Hz. This may have produced EP sets with insufficient distance between the parameters to enable the stimuli to be classified accurately. It is possible that stimulus frequencies with larger frequency differences between the stimuli may yield EPs which themselves have larger distances between the parameter groupings, making them more distinct and thus easier to classify. This theory was explored by training the SVMs on only three frequencies of stimulation (6, 8.2 and 12 Hz). **Figure 5-2** shows the $\text{Prob}(TP)$ results for subject 1. An interesting improvement around the 0.5 s mark is seen for SVM1, SVM2 and SVM3, as compared to the results shown in **Figure 4-28**. It should be noted, however, that this improvement could in part be due to the decrease in the number of classes, resulting in a random chance $\text{Prob}(TP)$ of 33%.

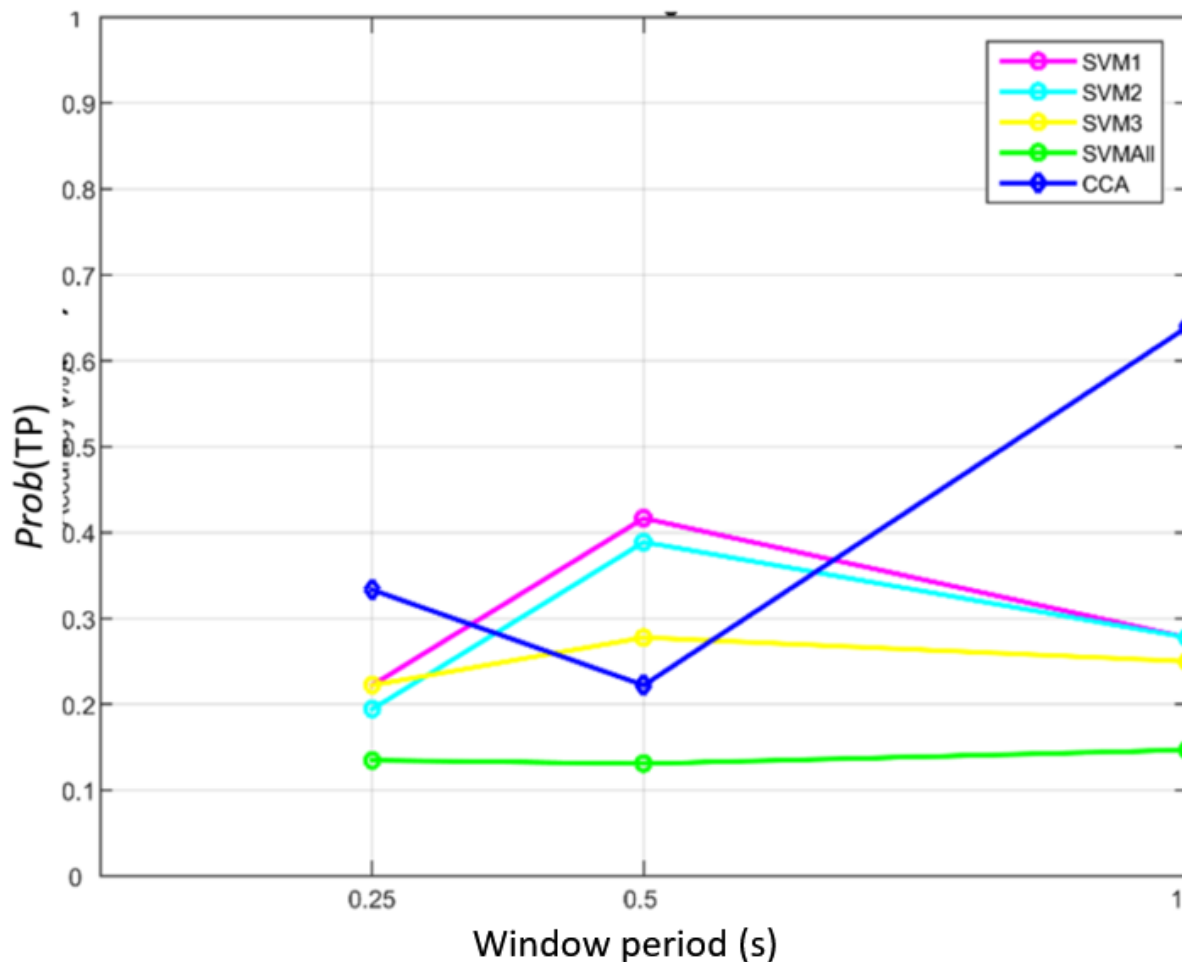


Figure 5-2: $\text{Prob}(TP)$ of SVMs trained on only three frequencies (6, 8.2 and 12 Hz) for subject 1. An improvement around the 0.5 s point is seen as compared with Figure 4-28. This could, however, be due in part to the decrease in classes used, which resulted in a random chance accuracy of 33%. No statistical analysis was performed on these results.

The fit error is due in part to the spurious EEG activity inherent in the single-trial SSVEPs used in this study. By averaging SSVEP signals according to stimulus frequency, a reduction in the spurious activity could also reduce the fit error by training fits on this averaged data. In doing so, estimated parameter sets that are more generally representative of a given subject's SSVEP would be produced. However, the classification scheme would still aim to classify on a single trial because averaging in the BCI system would involve exposing subjects to the same stimulus multiple times before a BCI command could be classified and acted on. This would affect the ITR, because the window period would be multiplied by the number of exposures required to classify the SSVEP.

5.2.3 Initial parameter variability

In an ideal case, the model output would be independent of the initial parameters used, as all fits would converge on the same EP set for a given SSVEP signal, regardless of the starting point used for the model. The extent to which this holds true for the SVM classifiers was investigated by considering the variability seen in the classification accuracy attributed to IP1, IP2 and IP3 by using the SVMs to classify estimated parameters that they had not seen. For instance, SVM1, which was trained on EP1, was used to classify EP1, EP2 and EP3.

Figure 5-3 shows the resultant classification accuracy for subject 4. The high classification accuracy scores in each window are effectively 1 less the in-sample resubstitution loss, using trained data for the prediction. The low classification accuracy scores are effectively 1 less the out-of-sample resubstitution loss, using unseen data for the prediction. These results indicate large initial parameter variability in the SVM-based classifiers used in this study. Improvements in the fitting approach used by the model might reduce this effect.

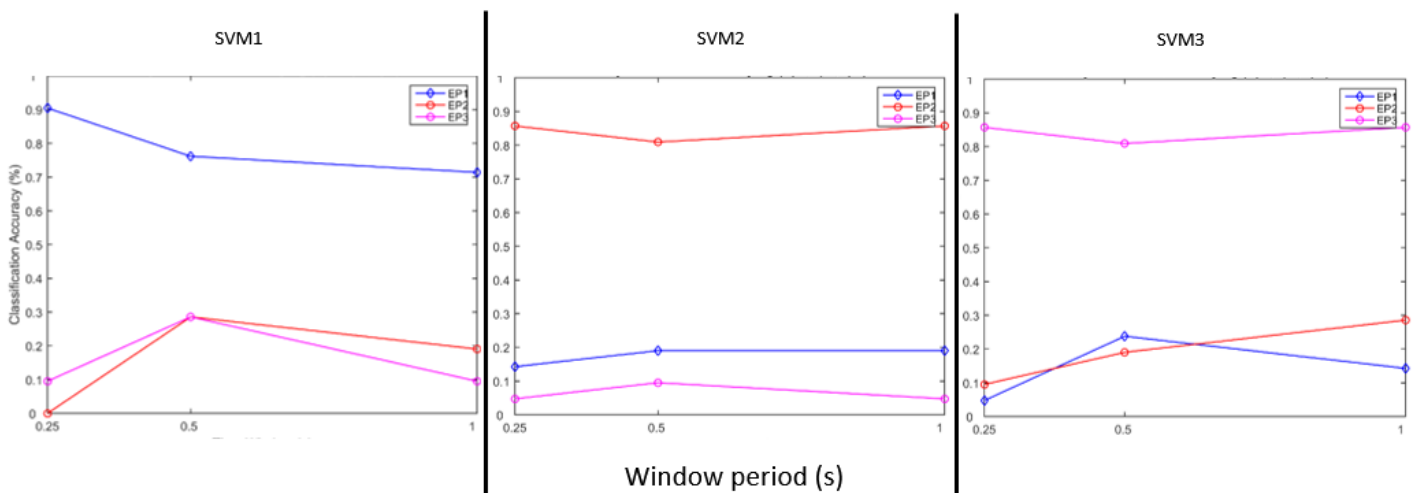


Figure 5-3: Classification accuracy (1 – resubstitution losses) of subject 4's SVMs classifying EPs which it had been trained on as well as those it had not been trained on. The plot on the left shows SVM1(EP1), SVM1(EP2) and SVM1(EP3). The middle and right plots show similar results for SVM2 and SVM3 respectively.

5.2.4 Inter-subject variability

Each subject had at least one SVM classifier that showed increasing $Prob(TP)$ as the window period increased (**Figure 4-28**); however, there was no consistency as to which of the SVMs performed best across subjects. This seems to indicate that there is at least some inter-subject variability inherent in the performance of the SVM classifiers. The extent of this variability was investigated by training SVMall on a single subject's data and then classifying the estimated parameters from all subjects. **Figure 5-4** shows the classification accuracy results for subject 1. The top trace (subject 1) is effectively 1 less the in-sample resubstitution loss, whereas the lower traces (all other subjects) are 1 less the out-of-sample losses. The large difference between the traces indicates a large inter-subject variability inherent in the SVM-based classifiers. For this reason, in a BCI context, it is imperative that the classifiers are trained for each subject individually using subject-specific data.

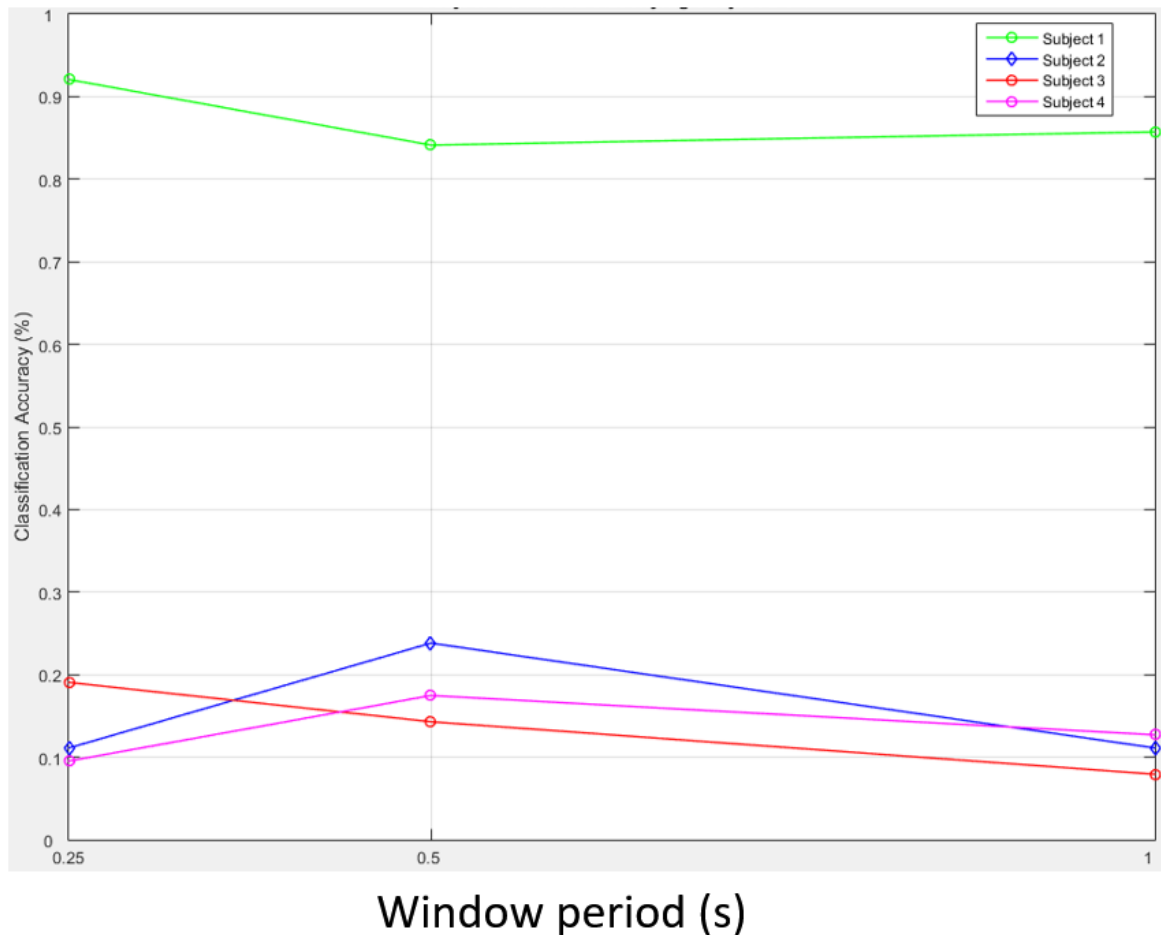


Figure 5-4: Classification accuracy (1 – resubstitution losses) of subject 1's SVMall classifier using the estimated parameter sets from all subjects. The top trace (green line) shows 1 less the in-sample loss as the SVM is classifying the training data. The lower traces show 1 less the out-of-sample loss, which are within the same range given by the out-of-sample classification rate for subject 1's SVMs (Figure 4-17).

5.3 Future work

5.3.1 SVM training dataset size

ANOVAs (**Figure 4-23** and **Figure 4-26**) performed on the accuracy of the SVMs revealed that there was no significant difference in performance between SVM1, SVM2, SVM3 and SVMAll. SVMAll did, however, outperform SVM1, SVM2 and SVM3 with regard to precision, while for recall it outperformed only SVM1. These results are interesting because SVMAll was trained on all the available EPs, and thus had a larger training set than those used for SVM1, SVM2 or SVM3. This suggests the possibility that a larger training would produce better results.

5.3.2 SVM training method

The method of training SVMs was kept constant so as to better assess other factors that influenced classification. There is scope to explore different approaches to SVM training with a view to improving the results. The choice of kernel function plays a large role in identifying support vectors of SVMs that are used as boundaries between classes (Hsu & Lin, 2002).

In this study, an 8th-degree polynomial kernel was used, specifically because it is better suited to fitting parameters with a non-linear relationship. However, the spacing of parameters may lend itself more readily to a radial-basis function kernel (RBF). A two-dimensional example of 8th-degree polynomial and RBF kernel functions is shown in **Figure 5-5**, the RBF kernel being better suited to grouping clusters of disjointed parameters.

RBF has been shown to work well with SVM features (Jian & Tang, 2014). Another option is a quadratic kernel, which has been shown to outperform both the RBF and a 3rd-order polynomial when directly assessing SSVEP data and not model parameters (Singla, 2014). A custom kernel mapping could also be considered. In order to assess which of these kernels would be most appropriate, they would need to be tested on a much larger cross-validated dataset.

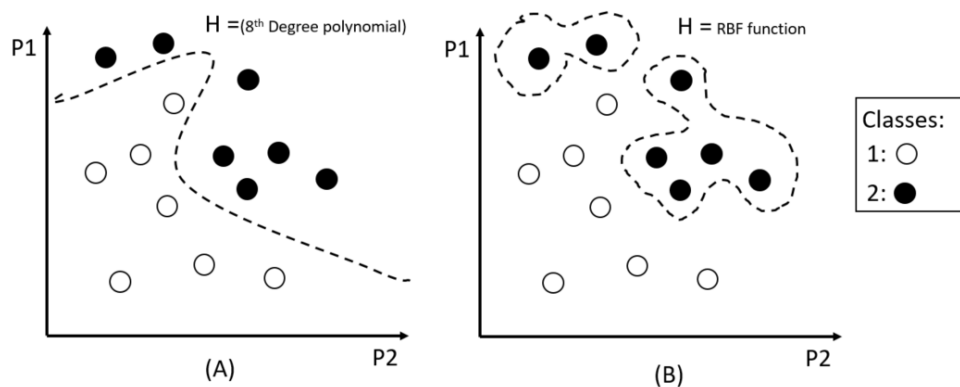


Figure 5-5: Examples of (A) a hyperplane implemented with an 8th-degree polynomial function and (B) a hyperplane generated by an RBF kernel.

5.3.3 Hybrid classification schemes

A method of improving classification accuracy would be to combine various classification methods. Zhang et al. (2012) showed that adding a linear discriminant analysis step to the PSDA method improved performance, and similarly using tensors to enhance the CCA method yielded better results, as did combining CCA with other Fourier methods.

A possible combination of an adapted CCA method with the model-based approach could be examined. The CCA method compares reference sine waves of various frequencies to the measured SSVEP and makes a classification based on the reference sine wave that has the highest correlation to the measured SSVEP. Transferring this idea to the model-based approach, instead of reference sine waves being used and correlated to the measured signal, reference model fits could be used (generated by training on the subject). A high correlation with a synthetic SSVEP would indicate to which class of SSVEP the measured signal belongs. In so doing, the model would attempt to classify the SSVEPs using the weighting of a coefficient attached to each parameter.

In recent years, the field of machine learning has taken several large leaps into the field of artificial neural networks (ANN). A combination SVM and ANN methods recently demonstrated accuracies of 88.5%, with a window period of 1 s (Singla, 2014). This work investigated the SSVEPs recorded directly and focused only on the steady-state portions of the SSVEP, with a rolling average segmenting the data every 0.25 s. Another possible approach to consider would be to use such a combination method, including the model-based features, to find fits on the initial SSVEP in the first 0.25 s.

6

Conclusion

This study assessed the performance of an SSVEP model-based classification method. The ability of the model to fit SSVEP data was examined, with different initial model parameters used in the parameter estimation process and for three different window periods. It was found that the model was able to fit SSVEP data within a reasonable goodness-of-fit range.

Estimated parameters which described the measured SSVEPs, with a degree of error, were used to train a series of SVMs. The performance of the SVM classifiers was then tested, while considering the impact of initial parameters, the window period and the frequency of stimulation. This performance was compared to CCA and PSDA methods, which are traditional methods of SSVEP classification, and it was found that in most cases the CCA method outperformed the SVM method. The SVM performance tended to be erratic and dependent on the subject, initial parameters and window period used.

Specifically, a number of hypotheses were examined in this study. They are repeated here for ease of reference.

Hypothesis 1: *Kremláček's VEP model can be fit onto the initial VEP portion of an SSVEP response.*

Hypothesis 2a: *The model fit onto the initial portions of SSVEP signals can generate unique feature-descriptive parameters that relate to the frequency of the stimulus presented.*

Hypothesis 2b: *These unique features enable the SSVEP signal to be classified relative to the stimulus frequency, using a multiclass SVM approach.*

Hypothesis 3: *The classification approach based on the time-domain model proposed here outperforms traditional spectral classification methods (PSDA and CCA) when the input to the classifiers comprises a 1 s or less window period of EEG data as recorded from stimulus onset; the shorter the window period, the more pronounced the effect.*

We found that Hypothesis 1 was possible, and that the Kremláček et al. (2002) model of VEPs was able to fit onto the initial portion of SSVEPs. The fit was not as good as Kremláček's reported NRMSE; however, some improvements could be made to the model possibly to reduce the error in fit.

The performance of the SVM classifiers was generally poor. The classifiers showed dependencies on numerous factors, including EEG data window period, initial parameters, goodness-of-fit function and choice of SVM configuration used. However, a limitation of the current study that would have had an impact on the performance measures, was the small dataset (four subjects). Nonetheless, in the context of the results presented here, Hypothesis 2 has to be rejected.

The model-based approach was compared to the CCA and PSDA methods. CCA outperformed the SVM method, even though the window period was greatly reduced. This was not in line with Hypothesis 3 – which was therefore rejected.

Overall, the accuracy of this novel system did not perform at expected levels, although there were peaks of performance which, given the construction of the model and the classifiers, points to the fact that this may be a viable method of signal classification. But it is acknowledged that a fair amount of work will have to go into refining the model in order for it to attain performance levels comparable to that of CCA.

7

References

- Abbasi, M. a, Gaume, A., Member, S., & Francis, N. (2015). Fast Calibration of a Thirteen - command BCI by Simulating SSVEPs from Trains of Transient VEPs – Towards Time - domain SSVEP BCI Paradigms, 22–24.
- Allison, B., Lüth, T., Valbuena, D., Teymourian, A., Volosyak, I., & Gräser, A. (2010). BCI demographics: How many (and what kinds of) people can use an SSVEP BCI? *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(2), 107–116. <http://doi.org/10.1109/TNSRE.2009.2039495>
- Allison, B. Z., McFarland, D. J., Schalk, G., Zheng, S. D., Jackson, M. M., & Wolpaw, J. R. (2008). Towards an independent brain-computer interface using steady state visual evoked potentials. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 119(2), 399–408. <http://doi.org/10.1016/j.clinph.2007.09.121>
- Anguita, A. D., Ghio, A., Oneto, L., Ridella, S., & Anguita, D. (2012). Estimation for Support Vector Machines In – sample and Out – of – sample Model Selection and Error Estimation for Support Vector Machines. *Ieee Transactions on Neural Networks and Learning Systems*.
- Arakawa, K., Tobimatsu, S., Tomoda, H., Kira, J., & Kato, M. (1999). The effect of spatial frequency on chromatic and achromatic steady-state visual evoked potentials. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 110, 1959–1964.
- Bashashati, A., Fatourehchi, M., Ward, R. K., & Birch, G. E. (2007). A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. *Journal of Neural Engineering*, 4(2), R32–57. <http://doi.org/10.1088/1741-2560/4/2/R03>
- BCI2000. (2012). The 10-20 International System. Retrieved January 1, 2015, from http://www.bci2000.org/wiki/index.php/User_Tutorial:EEG_Measurement_Setup
- Ben-Hur, A., & Weston, J. (2010). A user’s guide to support vector machines. *Methods in Molecular Biology (Clifton, N.J.)*, 609, 223–239. http://doi.org/10.1007/978-1-60327-241-4_13
- Bieger, J., Garcia-molina, G., & Zhu, D. (2010). Effects of Stimulation Properties in Steady-State Visual Evoked Potential Based Brain-Computer Interfaces. *32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 3345–3348.
- Bieger, J., & Molina, G. G. (2010). Technical note TN-2010-00315 Light Stimulation Properties to Influence Brain Activity. *Electronics*.
- Bin, G., Gao, X., Wang, Y., Hong, B., & Gao, S. (2009). VEP-based brain-computer interfaces: time, frequency, and code modulations [Research Frontier. *IEEE Computational Intelligence Magazine*, 4(4), 22–26. <http://doi.org/10.1109/MCI.2009.934562>
- Bin, G., Gao, X., Yan, Z., Hong, B., & Gao, S. (2009). An online multi-channel SSVEP-based brain-computer interface using a canonical correlation analysis method. *Journal of Neural Engineering*, 6(4), 46002. <http://doi.org/10.1088/1741-2560/6/4/046002>

- Bo Hong, Yijun Wang, X., & Gao, S. (n.d.). Quantitative EEG-Based Brain-Computer Interface.
- Brainloop. (2016). BCI - Brain-computer interfaces.
- Brunner, C., Allison, B. Z., Krusienski, D. J., Kaiser, V., Müller-putz, G. R., Pfurtscheller, G., & Neuper, C. (2010). Improved signal processing approaches in an offline simulation of a hybrid brain – computer interface, *188*, 165–173. <http://doi.org/10.1016/j.jneumeth.2010.02.002>
- Capilla, A., Pazo-alvarez, P., Darriba, A., Campo, P., & Gross, J. (2011). Steady-State Visual Evoked Potentials Can Be Explained by Temporal Superposition of Transient Event-Related Responses, *6*(1). <http://doi.org/10.1371/journal.pone.0014543>
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. <http://doi.org/10.5194/gmd-7-1247-2014>
- Cheever, E. (2015). Fourth Order Runge-Kutta. Retrieved from <http://lpsa.swarthmore.edu/NumInt/NumIntFourth.html>
- Di Russo, F., Pitzalis, S., Spitoni, G., Aprile, T., Patria, F., Spinelli, D., & Hillyard, S. a. (2005). Identification of the neural sources of the pattern-reversal VEP. *NeuroImage*, *24*, 874–886. <http://doi.org/10.1016/j.neuroimage.2004.09.029>
- Diez, P. F., Mut, V. a, Avila Perona, E. M., & Laciár Leber, E. (2011). Asynchronous BCI control using high-frequency SSVEP. *Journal of NeuroEngineering and Rehabilitation*, *8*(1), 39. <http://doi.org/10.1186/1743-0003-8-39>
- Dougherty, D. D., Rauch, S. L., & Rosenbaum, J. F. (2004). Essentials of neuroimaging for clinical practice. *American Psychiatric Pub.*, 117–118.
- Foxe, J. J., Strugstad, E. C., Sehatpour, P., Molholm, S., Pasieka, W., Schroeder, C. E., & McCourt, M. E. (2008). Parvocellular and magnocellular contributions to the initial generators of the visual evoked potential: High-density electrical mapping of the “C1” component. *Brain Topography*, *21*, 11–21. <http://doi.org/10.1007/s10548-008-0063-4>
- Gao, X., Xu, D., Cheng, M., & Gao, S. (2003). A BCI-based environmental controller for the motion-disabled. *IEEE Transactions on Neural Systems and Rehabilitation Engineering : A Publication of the IEEE Engineering in Medicine and Biology Society*, *11*(2), 137–40. <http://doi.org/10.1109/TNSRE.2003.814449>
- Guger, C., Allison, B. Z., Großwindhager, B., Prückl, R., Hintermüller, C., Kapeller, C., ... Edlinger, G. (2012). How Many People Could Use an SSVEP BCI? *Frontiers in Neuroscience*, *6*(November), 169. <http://doi.org/10.3389/fnins.2012.00169>
- Hakvoort, G., Reuderink, B., & Obbink, M. (2011). Comparison of PSDA and CCA detection methods in a SSVEP-based BCI-system. Retrieved from <http://eprints.eemcs.utwente.nl/19680/>
- Hansson-Sandsten, M. (2010). Evaluation of the optimal lengths and number of multiple windows for spectrogram estimation of SSVEP. *Medical Engineering & Physics*, *32*(4), 372–83. <http://doi.org/10.1016/j.medengphy.2010.01.009>
- Heinrich, S. P. (2007). A primer on motion visual evoked potentials. *Documenta Ophthalmologica. Advances in Ophthalmology*, *114*(2), 83–105. <http://doi.org/10.1007/s10633-006-9043-8>

- Herwig, U., Satrapi, P., & Schönfeldt-Lecuona, C. (2003). Using the international 10-20 EEG system for positioning of transcranial magnetic stimulation. *Brain Topography*, 16(2), 95–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/14977202>
- Hsu, C.-W., & Lin, C.-J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Transactions on Neural Networks*, 13(2), 415–425. article.
- Ioannis Kompatsiaris, Spiros Nikolopoulos, Katerina Adam, Elisavet Chatzilari, Kostantinos Georgiadis, Georgios Liaros, & Oikonomou, V. P. (2016). Comparative evaluation of state-of-the-art algorithms for SSVEP-based BCIs. *Arxiv Statistics - Machine Learning*, (January), 1–33. Retrieved from <http://arxiv.org/abs/1602.00904v2>
- Jian, H.-L., & Tang, K.-T. (2014). Improving classification accuracy of SSVEP based BCI using RBF SVM with signal quality evaluation. In *Intelligent Signal Processing and Communication Systems (ISPACS), 2014 International Symposium on* (pp. 302–306). inproceedings.
- Kelly, S. P., Lalor, E. C., Finucane, C., McDarby, G., & Reilly, R. B. (2005). Visual spatial attention control in an independent brain-computer interface. *IEEE Transactions on Bio-Medical Engineering*, 52(9), 1588–96. <http://doi.org/10.1109/TBME.2005.851510>
- Kelly, S. P., Lalor, E., Finucane, C., & Reilly, R. B. (2004). A comparison of covert and overt attention as a control option in a steady-state visual evoked potential-based brain computer interface. *Conference Proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, 7, 4725–4728. <http://doi.org/10.1109/IEMBS.2004.1404308>
- Kim, K. H., Odom, J. V., Bach, M., Barber, C., Brigell, M., Marmor, M. F., ... Holder, G. (2004). Visual evoked potentials standard (2004). *Documenta Ophthalmologica*, 1–10.
- Kremláček, J., & Holcík, J. (1999). Visual Evoked Potentials Model of Magnocellular System.
- Kremláček, J., & Kuba, M. (1999). Global brain dynamics of transient visual evoked potentials. *Physiological Research*.
- Kremláček, J., Kuba, M., & Holcík, J. (2002). Model of visually evoked cortical potentials. *Physiological Research / Academia Scientiarum Bohemoslovaca*, 51(1), 65–71. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12071292>
- Kronegg, J. (2005). Analysis of bit-rate definitions for Brain-Computer Interfaces. *Human-Computer Interaction*.
- Lin, Z., Zhang, C., Wu, W., & Gao, X. (2007). Frequency recognition based on canonical correlation analysis for SSVEP-based BCIs. *IEEE Transactions on Bio-Medical Engineering*, 54(6 Pt 2), 1172–6. <http://doi.org/10.1109/TBME.2006.886577>
- Liu, Q., Chen, K., & Ai, Q. (2013). Recent Development of Signal Processing Algorithms for SSVEP-based Brain Computer Interfaces, 34(4), 299–309. <http://doi.org/10.5405/jmbe.1522>
- Luo, A., & Sullivan, T. J. (2010a). A user-friendly SSVEP-based brain-computer interface using a time-domain classifier. *Journal of Neural Engineering*, 7(2), 26010. <http://doi.org/10.1088/1741-2560/7/2/026010>
- Luo, A., & Sullivan, T. J. (2010b). A user-friendly SSVEP-based brain-computer interface using a time-domain classifier. *Journal of Neural Engineering*, 7(2), 26010. <http://doi.org/10.1088/1741-2560/7/2/026010>

- Manyakov, N. V, Chumerin, N., Combaz, A., Robben, A., & Van Hulle, M. M. (2010). Decoding SSVEP responses using time domain classification. *International Conference on Neural Computation*, 376–380. <http://doi.org/10.5220/0003106103760380>
- Marple Jr., S. ~L. (1987). *Digital spectral analysis with applications*. Englewood Cliffs, NJ, Prentice-Hall, Inc., 1987, 512 p. book.
- Mason, S. G., Bashashati, A., Fatourech, M., Navarro, K. F., & Birch, G. E. (2007). A comprehensive survey of brain interface technology designs. *Annals of Biomedical Engineering* (Vol. 35). <http://doi.org/10.1007/s10439-006-9170-0>
- Mizoguchi, T., Takahashi, G., & Inoue, K. (2012). A Study of Stimulation Methods in Visual Evoked Potential Based Brain-Computer Interface, 2169–2174.
- Müller-Putz, G. R., Eder, E., Wriessnegger, S. C., & Pfurtscheller, G. (2008). Comparison of DFT and lock-in amplifier features and search for optimal electrode positions in SSVEP-based BCI. *Journal of Neuroscience Methods*, 168(1), 174–181. article.
- Myttenaere, A. De, Golden, B., Myttenaere, A. De, & Golden, B. (2015). Using the Mean Absolute Percentage Error for Regression Models To cite this version : *ESANN 2015 Proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 22-24 April 2015, i6doc.com Pub*, (April), 22–24.
- Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors (Basel, Switzerland)*, 12(2), 1211–79. <http://doi.org/10.3390/s120201211>
- Nieuwenhuis, S., Jepma, M., Fors, S. La, & Olivers, C. N. L. (2008). The role of the magnocellular and parvocellular pathways in the attentional blink. *Brain and Cognition*, 68, 42–48. <http://doi.org/10.1016/j.bandc.2008.02.119>
- Nishifuji, S., Shigeyama, T., & Tanaka, S. (2009). Dependence of EEG response to flicker stimuli on stimulus parameters. *Stimulus*, 827–832.
- Nooh, A. A., Yunus, J., & Daud, S. M. (2011). A Review of Asynchronous Electroencephalogram-based Brain Computer Interface Systems. *Engineering and Technology*, 11, 55–59.
- Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottureau, B. R., & Rossion, B. (2015). The steady-state visual evoked potential in vision research : A review. *Journal of Vision*, 15(6), 1–46. <http://doi.org/10.1167/15.6.4>
- Odom, J. V., Bach, M., & Brigell, M. (2010). ISCEV STANDARDS ISCEV standard for clinical visual evoked potentials (2009 update), 111–119. <http://doi.org/10.1007/s10633-009-9195-4>
- Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, 112(4), 713–9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11275545>
- Ortner, R., Allison, B. Z., Korisek, G., Gaggl, H., & Pfurtscheller, G. (2011). An SSVEP BCI to control a hand orthosis for persons with tetraplegia. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, 19(1), 1–5. <http://doi.org/10.1109/TNSRE.2010.2076364>

- Rai, P. (2011). Kernel Methods and Nonlinear Classification. Course notes, CS5350/6350: Machine Learning. Retrieved from <https://www.cs.utah.edu/~piyush/teaching/15-9-print.pdf>
- Schlögl, A., Vidaurre, C., & Müller, K.-R. (2009). Brain–Computer Interfaces: A Gentle Introduction, 1–28. <http://doi.org/10.1007/978-3-642-02091-9>
- Schreiber, T. (2000). Measuring Information Transfer. *Phys. Rev. Lett.*, 85(2), 461–464. article. <http://doi.org/10.1103/PhysRevLett.85.461>
- Sincich, L. C., & Horton, J. C. (2005). The circuitry of V1 and V2: integration of color, form, and motion. *Annual Review of Neuroscience*, 28, 303–326. <http://doi.org/10.1146/annurev.neuro.28.061604.135731>
- Singla, R. (2014). Comparison of SSVEP Signal Classification Techniques Using SVM and ANN Models for BCI Applications. *International Journal of Information and Electronics Engineering*, 4(1), 6–10. <http://doi.org/10.7763/IJIEE.2014.V4.398>
- Singla, R., Khosla, A., & Jha, R. (2014). Influence of stimuli colour in SSVEP-based BCI wheelchair control using support vector machines. *Journal of Medical Engineering & Technology*, 38(3), 125–134. article. <http://doi.org/10.3109/03091902.2014.884179>
- Srihari Mukesh, T. M., Jaganathan, V., & Reddy, M. R. (2006). A novel multiple frequency stimulation method for steady state VEP based brain computer interfaces. *Physiological Measurement*, 27(1), 61–71. <http://doi.org/10.1088/0967-3334/27/1/006>
- Tobimatsu, S., Tomoda, H., & Kato, M. (1995). Parvocellular and magnocellular contributions to visual evoked potentials in humans: stimulation with chromatic and achromatic gratings and apparent motion. *Journal of the Neurological Sciences*, 134, 73–82. [http://doi.org/10.1016/0022-510X\(95\)00222-X](http://doi.org/10.1016/0022-510X(95)00222-X)
- Vangelis P. Oikonomou, Georgios Liaros, Kostantinos Georgiadis, Elisavet Chatzilari, Katerina Adam, S. N. and I. K. (2016). Matlab code for processing EEG signals.
- Vialatte, F.-B., Maurice, M., Dauwels, J., & Cichocki, A. (2010). Steady-state visually evoked potentials: focus on essential paradigms and future perspectives. *Progress in Neurobiology*, 90(4), 418–38. <http://doi.org/10.1016/j.pneurobio.2009.11.005>
- Vidal, J. J. (1973). Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering*, 2(1), 157–180. article.
- Vidal, J. J. (1977). Real-time detection of brain events in EEG. *Proceedings of the IEEE*, 65(5), 633–641. article.
- Vilic, A. (2015). AVI SSVEP Dataset. Retrieved January 1, 2015, from <http://www.setzner.com/avi-ssvep-dataset/>
- Volosyak, I. (2011). SSVEP-based Bremen-BCI interface--boosting information transfer rates. *Journal of Neural Engineering*, 8, 36020. <http://doi.org/10.1088/1741-2560/8/3/036020>
- Volosyak, I., Cecotti, H., Valbuena, D., & Gr, A. (2009). Evaluation of the Bremen SSVEP based BCI in real world conditions, 322–331.
- Wang, Y., Member, S., Wong, N., Wang, Y., Wang, Y., Huang, X., ... Cheng, C. (2013). Study of Visual Stimulus Waveforms via Forced van der Pol Oscillator Model for SSVEP - Based Brain -

Computer Interfaces, 475–479.

- Wang, Y., Wang, R., Gao, X., Hong, B., & Gao, S. (2006). A practical VEP-based brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering : A Publication of the IEEE Engineering in Medicine and Biology Society*, 14(2), 234–9. <http://doi.org/10.1109/TNSRE.2006.875576>
- Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, 113(6), 767–91. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12048038>
- Y. Zhang, G. Zhou, Q. Zhao, A. Onishi, J. J. (2012). Multiway Canonical Correlation Analysis for Frequency Components Recognition in SSVEP-Based BCIs, 7666(November 2011), 356–363. <http://doi.org/10.1007/978-3-642-34478-7>
- Yoshimura, N., & Itakura, N. (2011). Usability of Transient VEPs in BCIs. *Recent Advances in Brain-Computer Interface Systems*, 119–134. <http://doi.org/10.1109/CNE.2007.369683>
- Zhang, Y. (2014). LassoSSVEP_Demo.zip. Retrieved March 20, 2016, from https://www.mathworks.com/matlabcentral/fileexchange/47498-lassossvep-demo-zip/content/LassoSSVEP_Demo/LassoSSVEP.m
- Zhang, Y., Xu, P., Cheng, K., & Yao, D. (2013). Multivariate Synchronization Index for Frequency Recognition of SSVEP-based Brain-computer Interface. *Journal of Neuroscience Methods*, 1–9. <http://doi.org/10.1016/j.jneumeth.2013.07.018>
- Zhang, Y., Xu, P., Cheng, K., & Yao, D. (2014). Multivariate synchronization index for frequency recognition of SSVEP-based brain-computer interface. *Journal of Neuroscience Methods*, 221, 32–40. <http://doi.org/10.1016/j.jneumeth.2013.07.018>
- Zhang, Y., Zhou, G., Jin, J., Wang, X., & Cichocki, A. (2014). SSVEP recognition using common feature analysis in brain-computer interface. *Journal of Neuroscience Methods*. <http://doi.org/10.1016/j.jneumeth.2014.03.012>
- Zhu, D., Bieger, J., Garcia Molina, G., & Aarts, R. M. (2010). A survey of stimulation methods used in SSVEP-based BCIs. *Computational Intelligence and Neuroscience*, 2010, 702357. <http://doi.org/10.1155/2010/702357>
