Department of Computer Science and Electronic Engineering, University of Essex

Sound-production Related Cognitive Tasks for Onset Detection in Self-Paced Brain-Computer Interfaces

Author: Youngjae Song

Supervisor: Dr. Francisco Sepulveda

A thesis submitted for the degree of Doctor of Philosophy

Department of Computer Science and Electronic Engineering

University of Essex

July, 2017

Abstract

Objective. The main goal of this research is proposing a novel method of onset detection for Self-Paced (SP) Brain-Computer Interfaces (BCIs) to increase usability and practicality of BCIs towards real-world uses from laboratory research settings.

Approach. To achieve this goal, various Sound-Production Related Cognitive Tasks (SPRCTs) were tested against idle state in offline and simulated-online experiments. An online experiment was then conducted that turned a messenger dialogue on when a new message arrived by executing the Sound Imagery (SI) onset detection task in real-life scenarios (e.g. watching video, reading text). The SI task was chosen as an onset task because of its advantages over other tasks: 1) Intuitiveness. 2) Beneficial for people with motor disabilities. 3) No significant overlap with other common, spontaneous cognitive states becoming easier to use in daily-life situations. 4) No dependence on user's mother language.

Main results. The final online experimental results showed the new SI onset task had significantly better performance than the Motor Imagery (MI) approach. 84.04% (SI) vs 66.79% (MI) TFP score for sliding image scenario, 80.84% vs 61.07% for watching video task. Furthermore, the onset response speed showed the SI task being significantly faster than MI. In terms of usability, 75% of subjects answered SI was easier to use.

Significance. The new SPRCT outperforms typical MI for SP onset detection BCIs (significantly better performance, faster onset response and easier usability), therefore

it would be more easily used in daily-life situations. Another contribution of this thesis is a novel EMG artefact-contaminated EEG channel selection and handling method that showed significant class separation improvement against typical blind source separation techniques. A new performance evaluation metric for SP BCIs, called true-false positive score was also proposed as a standardised performance assessment method that considers idle period length, which was not considered in other typical metrics.

Keywords: Artefact removal; Brain-Computer Interface (BCI); Onset detection; Self-paced BCI; Asynchronous BCI; Covert speech; Sound-production

Acknowledgements

I am glad to have a chance to express my gratitude to all those who supported me to be able to complete this PhD research. First and foremost, I would like to show my deepest appreciation and gratitude to my supervisor Dr. Francisco Sepulveda. I thank you for all your kind advice (from the scientific questions to the PhD researcher life) and enthusiasm for BCIs which made me focus on my research and complete my thesis. I also sincerely appreciate your kind interest on me at every stage of my research since the beginning of my Master's degree.

I would also like to thank all my BCI-Neural engineering lab colleagues: Prof. John Q. Gan, Dr. Ian Daly, Dr. Luca Citi, Prof. Riccardo Poli, Dr Caterina Cinel, Ana Matran-Fernandez, Davide Valeriani, Amir Jahangiri, Andrei lacob, Luz M. Alonso Valerdi and Javier Asensio. Our regular BCI meeting helped me to extend my general background knowledge about BCIs.

Last but not least, I would like to express my love and gratitude to my beloved parents, sister Minju and brother-in-law Kangwing. Your endless love and support helped me a lot to settle down in the UK and complete my PhD. Even though I do not normally express myself on how much I love you and care about you, in my deepest mind, I always thank you and love you above all.

I also would like to thank my PhD viva reviewers and George Kyritsis for proofreading this thesis.

Contents

List of Figures

List of Tables

List of Equations

List of Acronyms

ADC	Analogue to Digital Converter
AR Model	AutoRegressive Model
BCI	Brain-Computer Interface
BMI	Brain-Machine Interface
BP	Band Power
BSS	Blind Source Separation
CCA	Canonical Correlation Analysis
CSP	Common Spatial Pattern
CT	Computerized Tomography
CWT	Continuous Wavelet Transform
DBI	Davis-Bouldin Index
DWT	Discrete Wavelet Transform
EEG	Electroencephalography
EMG	Electromyography
EMG-CCh	EMG Contaminated EEG Channels
ERP	Event-Related Potential
FFT	Fast Fourier Transform
fMRI	Functional Magnetic Resonance Imaging

fNIRS	Functional Near Infrared Spectroscopy
HMI	Human-Machine Interaction
ICA	Independent Component Analysis
ITR	Information Transfer Rate
LDA	Linear Discriminant Analysis
MEG	Magnetoencephalography
MI	Motor Imagery
PCA	Principal Component Analysis
PET	Positron Emission Tomography
ROC	Receiver Operating Characteristic
SCP	Slow Cortical Potential
SSVEP	Steady-State Visual Evoked Potential
SVM	Support Vector Machine
SWT	Stationary Wavelet Transform
WT	Wavelet Transform

List of Publications

Journal Articles

[1] Y. Song and F. Sepulveda, "A novel onset detection technique for brain–computer interfaces using sound-production related cognitive tasks in simulated-online system," *Journal of Neural Engineering*, vol. 14, no. 1, p. 016019, 2017.

Journal Articles Submitted / Under Review

[2] Y. Song and F. Sepulveda, " A Novel Technique for Selecting EMG Contaminated EEG Channels in Self-Paced Brain-Computer Interface Task Onset," IEEE Transaction on Neural Systems & Rehabilitation Engineering, 2017.

Conference Proceedings

[3] Y. Song and F. Sepulveda, "An online self-paced brain-computer interface onset detection based on sound-production imagery applied to real-life scenarios," in *2017 5th International Winter Conference on Brain-Computer Interface (BCI)*, 2017 IEEE, pp. 46-49.

[4] Y. Song and F. Sepulveda, "Classifying Siren-sound Mental Rehearsal and Covert Production vs. Idle State Towards Onset Detection in Brain-Computer Interfaces," presented at the The 3rd International Winter Conference on Brain-Computer Interface, 2015 IEEE, 2015.

[5] Y. Song and F. Sepulveda, "Classifying speech related vs. idle state towards onset detection in brain-computer interfaces overt, inhibited overt, and covert speech sound production vs. idle state," in *Biomedical Circuits and Systems Conference (BioCAS), 2014 IEEE*, 2014, pp. 568-571: IEEE.

1 Introduction

1.1 The Necessity of the Research

Throughout the years, technologies have been developed at a great rate (e.g., smartphones, wearable devices, smart health care systems). Compared to regular PC uses, these smart devices require better Human-Machine Interaction (HMI) such as touch screen and voice command features. These HMI technologies are reducing the barrier between humans and machines. As can be seen from Figure 1.1, in the past, the keyboard and mouse input control was the main option for communicating with machines. However, it is quite a restricted method compared to human natural communication in the real world. Nowadays, this barrier is becoming smaller with many HMI technologies. As a consequence, the question that arises is what the next stage would be beyond this level. The answer could be the use Brain-Computer Interfaces (BCIs). In the near future, the interaction barrier between humans and machines will disappear and users will be able to link their thought directly to a machine command.

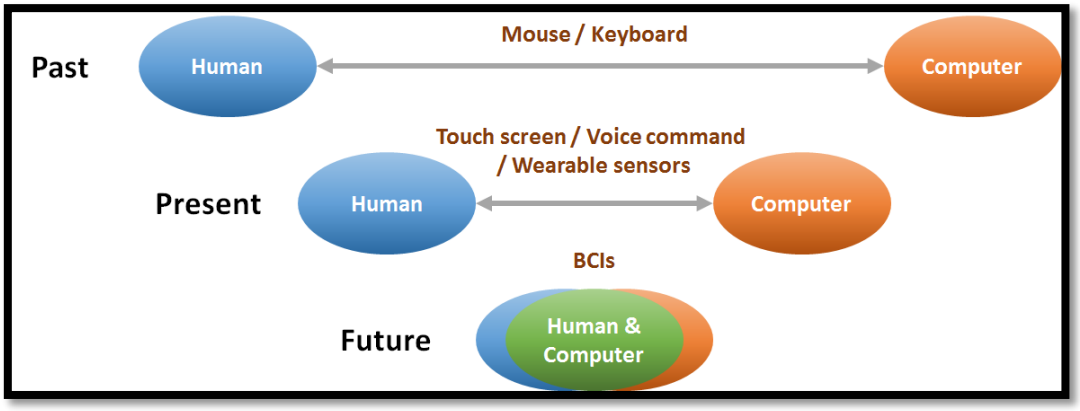


Figure 1.1. Human computer interaction from the past to the future.

The use of BCIs, which are also known as Brain-Machine Interfaces (BMIs – recently used mostly for implanted brain-interfaces) is one of the most phenomenal fields, which has a lot of potential in the human machine interaction area. It enables direct communication between humans and computers (or machines) by interpreting human brain signal activities and it controls and interacts with the users' surrounding environments without any physical involvement [6]. Owing to these huge advantages, it is not only an opportunity for people with motor disabilities but it also gives various interesting options to able-bodied people. It could therefore be one of the most popular communication systems between humans and machines in the near future.

Even though the BCI field has quickly developed in the last few years, it is mainly investigated as a research area due to shortages of practicality and usability. The majority of current BCIs are **cue-based** (**synchronous**) systems, which are controlled by the machine's predefined timing protocol. This means that it forces users to follow the computer's timing commands (locked to the machine). It also requires from the users to keep their mental focus and/or gaze at the computer interface (i.e., P300, SSVEP), which not only is very unnatural to them, but also leads to loss of both user

autonomy and the ability to have a rich interaction with their environment [1, 5, 7, 8]. These are the main issues when BCIs are used outside laboratory settings.

On the other hand, **self-paced** (**asynchronous**) systems analyse the user's brain activities continuously without any specific computer-controlled stimulus [9]. The users control the timing of the BCI system by intentionally performing a specific cognitive task when it suits them [7], thus providing increased autonomy, flexibility, and interaction with the environment (including the people therein, of course). For this reason, self-paced BCIs are more suitable than cue-based BCIs for the ultimate aim of transferring BCIs from laboratory settings towards real-world use [1].

However, there are great challenges in self-paced BCIs [7]. Self-paced BCIs usually require a more difficult system design, worse classification rate and harder analysis than cue-based systems due to the systems' lack of knowledge about the precise time location of the user command. Self-paced BCIs need to continuously analyse the ongoing brain activity in order to distinguish between Intentional Command (IC) and Non-Control (NC) states (also called non-specific or null states). NC states can be any states besides IC states (e.g., idle, daydreaming, other mental activities, irrelevant evoked responses) [8]. One way to distinguish between IC and NC is to use a classifier that treats NC and IC simply as different states in the same classification task. For example, a five-output classifier can include NC states as one of the five output classes. However, given the brain's constant multitasking, this approach—herein called a 'lumped' approach—will lead to a high false-positive (FP) rate for the IC states (and thus a high false-negative rate for the NC states) and to large timing errors in the IC detection. Hence, an alternative approach is needed, namely, separating the 'when' classification task (herein called IC **onset detection**) from the 'what/which' classification stage. This simplifies the problem and leads to reduced

timing errors and lower NC misclassification rates. This onset detection problem should be solved first in order to move self-paced BCIs to real-world uses, apart from the inside laboratory research [1].

In this thesis, sound-production related cognitive tasks (sound imagery) have been proposed for the onset detection method. Based on our thorough literature review (up to 2017), none of the previous works on onset detection or self-paced BCIs systems used speech or sound-production related cognitive tasks. They mostly used instead motor imagery (e.g., [7, 8, 10, 11]). In addition, all the speech related EEG-based BCI studies using different syllables (or syllables / vowels) that were retrieved, focused on the discrimination between various tasks and not on onset detection (i.e., idle versus intentional state). They were also cue-based approaches and not self-paced (e.g., [12-15]), while some of them were ECoG studies [16, 17]. The main novelty in this study is the discrimination between sound-production related cognitive tasks and idle (or non-specific) states for onset detection, which leads to competitive results compared to systems based on typical motor-imagery tasks [10, 18]. A novel score system for the evaluation of a self-paced BCI performance is also introduced [1].

Motor imagery self-paced onset detection BCIs have a crucial issue when they are used outside laboratory settings. The mental procedure is largely overlapping with other common, spontaneous cognitive states. For example, a classifier would not be possible to identify whether the onset detection was from the actual command or from other daily-life gestures such as shaking hands, grabbing an object or cycling.

On the other hand, a sound-production related cognitive task is also needed to reduce the chances of IC false positives but this can be addressed by choosing cognitive tasks that do not significantly overlap with other common, spontaneous and frequent

cognitive states [4]. Using specific words/syllables/letters for the onset detection would likely increase both the onset false positives as well as the task-related false negatives due to the large overlap with the continuous internal speech in normal thought processes. For this reason, we have chosen the thought of high tones or siren-like sound production tasks as onset switches, which are both unlikely to overlap with normal thought processes. The chosen tasks are easy to produce and control voluntarily and there is no dependence on the subjects' mother-language. In addition, sound imagery tasks are very intuitive for the vast majority of people as we almost constantly 'speak' internally or think many words in normal life. This is also a big advantage for people with severe motor disabilities, an important target population for BCIs [1].

Furthermore, for the self-paced onset detection system, artefact handling is an important issue as large EMG artefacts are likely emerging by executing onset tasks from the idle state. However, previous BCI and EEG studies applied artefact handling techniques to all EEG channels. As such, e.g., in the case of artefact removal using a blind source separation, a common approach in BCIs, there may be significant loss of useful EMG-free EEG information [19-22]. Thus, a new artefact handling method is required to investigate the improvement of system performance and reliability.

In summary,

- BCI is the next generation of human machine interaction technology.
- However, current BCIs are mostly investigated in research due to shortages of practicality and usability.
- The Cue-based approach is the main reason of the practicality problems.
- Therefore, self-paced BCIs are more suitable for taking BCIs from laboratory settings and put them to real-world uses.

- However, there are great challenges in self-paced BCIs, where the onset detection problem is the main issue.
- In order to resolve this, our new sound-production related cognitive tasks were investigated as a potential onset detection solution in this thesis.
- Motor imagery tasks have a significant issue in case they are used in outdoor laboratory settings (largely overlap with other common, spontaneous cognitive states).
- In addition, a new artefact handling method needs to be investigated for the performance and reliability of the system.

For all the above reasons, this sound imagery onset detection research is necessary to move forward BCIs from their use in laboratory experiments to real-world uses.

1.2 Goal of the Research

Main Goal

The goal of the research is to propose a novel method of onset detection in self-paced Brain-Computer Interfaces (BCIs). This research presents intuitive sound-production related cognitive tasks for detecting the onset for the system in order to establish full control of self-paced BCIs. It will increase the usability and the practical uses of BCI systems towards real-world settings and not only in laboratory research.

In order to achieve this goal, various sound imagery tasks were tested against the idle state in offline and simulated-online experiments. An online experiment was then

carried out where a messenger dialogue was turned on when a message arrives by executing an onset task in real-life scenarios such as watching video, reading text. The results were compared to typical motor imagery tasks in order to showcase the advantage of the proposed sound imagery task over the motor imagery onset detection. As a further addition, the message onset detection experiment was also conducted outside the usual laboratory settings.

Research Objectives and Scopes

This PhD research investigated a novel method of onset detection towards pure self-paced BCIs in real-life uses. There are two general objectives: (1) Background studies of sound imagery vs. idle state. (2) A self-paced onset detection system in daily-life task scenarios.

Background Studies and Sound Imagery vs. the Idle State

- Signal processing and EMG & EOG artefacts handling studies:

  Regarding the first objective, various signal processing, feature extraction and classification methods need to be studied. Furthermore, the EMG and EOG artefacts handling procedure has to be investigated and tested in order to ensure that the sound imagery onset detection is purely based on the cognitive tasks and not the artefacts. In the onset detection system, users stay calm and relaxed during the idle period and they then execute task for the onset. During this procedure, it is likely that the participants might generate some artefacts. Thus, it is important to handle these artefacts in advance.

- Investigation of high-tone sound production cognitive tasks versus the idle state in an off-line cue-based system:

In order to investigate our hypothesis whether the sound-production related cognitive tasks can be distinguished from the idle state, an offline cue-based study needed to be carried out in advance. The classification of high tone sound production tasks vs. the idle state was done in three different speech modes (overt, inhibited overt and covert). These three different speech modes were tested in order to examine their differences and to find the optimal mode for the onset detection in BCIs.

- Investigation of siren-like sound mental rehearsal versus the idle state in an off-line self-cue-based system:

  In addition to the above second objective, a siren-like sound imagery task (in covert speech) was also investigated in order to realise whether it can be used for an onset detection system.

- Development of a system that can simulate an online situation and classification of various sound imagery tasks from the idle state:

  Following the investigation of the above objectives (offline cue-based), high tone and siren-like sound production related cognitive tasks were tested with the use of covert and inhibited overt speech modes in a simulated-online situation. In order to move our hypothesis from offline cue-based settings towards an online self-paced system, this simulated-online system was explored in advance.

Self-paced Onset Detection System in Daily-Life Task Scenarios

- Proposal of a new metric for the self-paced BCI systems' performance evaluation assessment score:

In order to compare our sound imagery onset detection approach with other typical motor imagery systems, a common performance assessment metric for self-paced systems was necessary. However, current existing evaluations did not take into account altogether the true-positive, false-positive and idle period length aspects. For this reason, we proposed a new performance evaluation metric called the true-false-positive score for self-paced BCIs.

- Comparison of sound imagery and motor imagery onset detection systems in online daily-life task scenarios:

  In order to demonstrate the advantages of our sound imagery task over the motor imagery task for the self-paced onset detection system, the two cognitive tasks were compared in an online setting. The users were trying to open a message dialogue when a new message arrived by executing the onset detection task during daily-life task scenarios such as watching videos and reading text.

- Testing the sound imagery onset detection system outside the laboratory settings:

  In order to investigate the problem of applying BCIs in a real-life setting, a message opening an online self-paced onset detection system was tested outdoors at a cafeteria. The results were compared to an indoor laboratory experiment.

1.3 Scientific Contributions & Expected Effects of the Research

Much research has been carried out so far in the field of BICs. However, there are still many areas to be investigated. Even though cue-based (synchronous) systems

do indeed have higher accuracies and simpler designs than self-paced (asynchronous) BCIs, they prevent BCIs from being applied to real-world situations due to their lack of practicality and usability. Therefore, a self-paced system is required. however, there is a big challenge in this case called the **'onset detection problem'**, which allows the user to control the machine freely whenever they desire. Once the onset detection problem is solved, the practicality of BCI systems will be greatly increased.

- **A novel onset detection system was suggested and developed with sound-production related cognitive tasks (sound imagery).**
    - **The sound imagery onset detection system showed significantly better performance results than the motor imagery onset task.**
    - **It also showed a faster onset response and better usability.**
    - **The system was tested at an outdoor laboratory environment with daily-life task scenarios.**

In this PhD research, **a novel onset detection system was suggested** with a sound-production related cognitive task (sound imagery). Based on our thorough literature review, none of the works on onset detection or self-paced BCIs used a speech or sound-production related approach. They instead mostly used the motor imagery (e.g., [7, 8, 10, 11]). However, there are a couple of clear advantages of our sound imagery method over the typical motor imagery. Firstly, **it is very intuitive and easy to apply** as the majority of people almost constantly 'speak' internally or think many words in normal life. In addition, it is advantageous for people with motor disabilities for whom motor imagery may not be suitable. Secondly, the chosen **sound imagery task** (high tone, siren-like sound production) **does not significantly overlap with other common, spontaneous cognitive states in order to make it feasible to be used**

in a daily-life situation. On the other hand, motor imagery or specific words/syllables/letters production tasks would likely have a large overlap with daily-life situations such as body movements, gestures or conversations. Thirdly, **the sound imagery onset detection system showed significantly better performance results and a faster onset response speed than motor imagery** (details on Chapter 7). Finally, **the proposed onset detection method has been tested at an outdoor laboratory environment with daily-life scenarios.** This research showed the current problems of BCIs when used in an outdoor laboratory setting, which is an essential investigation at this stage in order to apply BCIs in real-world scenarios and not only in indoor laboratory experiments.

- **A novel EMG contaminated EEG channel selection and handling procedure was proposed**
    - **Our method showed a significantly better class separation (reduced useful information loss) than the typical blind source separation techniques (i.e., ICA, PCA and BSS-CCA) on both our data set and the BCI competition data set.**

This work also **proposed a new technique of selecting an EMG contaminated EEG channel and handling procedure** in chapter 4. Artefact handling is an essential procedure in EEG based studies. However, commonly used blind source separation methods could cause some useful information loss (cross-talk of brain and muscle artefacts were observed in [19-22]). In this research, the information loss of useful EEG sources was minimised by applying the artefact handling techniques only to the selected EMG contaminated EEG channels. This new technique had gone through statistical reliability tests in order to ensure it is strict enough to be used. The result

showed that our EMG contaminated channel selection and handling procedure significantly improved class separability on both our data set and the BCI competition IV data set 2a [23].

- **A new performance evaluation metric for self-paced BCIs was proposed.**
  - **This new true-false-positive score considered all the important aspects of a self-paced system (e.g., idle period) and it can be therefore used as a common evaluation method for other studies.**

In addition, **a new performance evaluation metric for self-paced BCIs, called the true-false-positive score was proposed** in this study. Currently, there is no common and standardised performance assessment method of self-paced BCI systems. The results would therefore all vary according to different papers and experiment settings. For example, some papers show performance results with a hit rate that can only be applied to their own experimental settings (e.g., [8, 10, 24, 25]), which makes it difficult to compare the system performance. In addition, there is no metric that considers the idle period length, which is a very important aspect of self-paced systems. For example, if two systems have the same number of true-positives and false-positives but one system has a longer idle period, then it is a certainly better system as it has a smaller false-positive ratio during the experiment. For this reason, we proposed a new performance evaluation metric that takes all these matters into account and can be used from all self-paced BCI systems as a standard evaluation metric.

Because of all the above reasons, this research has a potential effect on the BCI field and it can be divided into a short-term and long-term effect:

- **Short-term effect:** From a short-term point of view, this work provides background knowledge (including its processing algorithms) about a novel sound imagery onset detection system, which has clear advantages over the typical motor imagery method in self-paced BCIs. In addition, a new EMG artefact contaminated EEG channel selection and handling method was proposed to be used in EEG based studies and not only in BCIs, but also in brain mapping and clinical areas. Finally, a new performance evaluation metric for self-paced BCIs was suggested to be used as a standardised assessment method in the BCI field.
- **Long-term effect:** From a long-term point of view, once this sound imagery onset detection method is further developed by improving its practicality, then this system can be implemented to all kinds of cue-based BCI systems so that they can be used as a self-paced approach by providing the onset command to the machine. Thus, the current prototype onset detection system, which was tested in real-life task scenarios at outdoor laboratory settings, illustrates the direction of future BCIs in terms of practical uses. In addition, the new true-false-positive score metric can be used as a standardised evaluation method in self-paced BCIs so that it makes it easier to compare the system performances in BCI studies.

1.4 Limitations of the Research

As this PhD research was not funded by any companies or organisations, there were a couple of limitations that concerned it. Firstly, people with motor disabilities

are an important target population for BCIs in general but the participants who took part in the experiments were all able-bodied subjects. This would be a huge limitation in motor imagery BCIs as there could be some performance difference between able-bodied and motor disability subjects. However, in our sound imagery task case, ideally, there would not be any difference. Thus, it will not be a huge problem in this thesis.

Secondly, in the final online self-paced onset detection experiment, the simulation of a message arriving / opening application was used without the actual text messenger. The reason was that a specially designed message arriving motion (details in Chapter 7) had to be used in order to minimise any event related potentials (visual and audio) in order to ensure this onset detection was done with the sound imagery mental task and not the event related potentials. In addition, artefacts related to mobile phone signals had to be avoided by the study at this investigation stage.

Thirdly, there was a limitation with regards to the order of the outdoor experiments. The first indoor laboratory experiment was conducted continuously to all the subjects and afterwards, the outdoor experiment was tested. This biased recording order could impact the final results and therefore had to be randomised. However, the installation of BCI equipment at an outdoor cafeteria area and then move it back to the indoor environment takes more time and costs more human resources than if the opposite took place. In addition, the main aim of Chapter 7 was to investigate and compare the motor imagery and our sound imagery task in terms of onset detection at indoor laboratory settings. Therefore, this limitation was ignored. However, in order to minimise the biased recording order effect, there was a small tea/coffee break in-between the two experiments and the total experiment time did not exceed 1 and a half hours, which helped maintain the participants' concentration level similar to that of an indoor lab experiment.

Lastly, there are possible issues with the experimental protocol. Even though we made a concentrated effort to minimise VEPs by having a circular progress bar, which is considered a visual angle that avoids VEPs, there could still some VEPs be present during the online real-life scenarios experiment (e.g., message arriving notification, watching a video task). Furthermore, there could be possible mental tasks that overlap even though the participants were prevented from counting the timer during the onset activation. Moreover, although the participants were instructed to refer to the circular progress bar in order to answer the onset response time, it was difficult to clearly avoid timer count.

1.5 Structure of Thesis

This thesis consists of nine chapters and its structure as follows. Chapter 2 describes a short overview of EEG and brain physiology & structure in order to provide the readers with some background knowledge. It gives brief information about the visual system, auditory system, somatic sensory system, motor control and language processing & sound production related brain physiology.

Chapter 3 contains the literature review about BCIs and their general applications. It also does a comparison review of self-paced and cue-based systems followed by speech related BCIs and the onset detection in BCIs. In addition, various signal processing, feature extraction and classification methods were examined.

The experimental part of the thesis is discussed in Chapter 4. It suggests a new EMG artefact contaminated EEG channel selection and handling technique. It

describes the new method, which minimises the information loss compared to the typical blind source separation techniques. It also shows the statistical reliability of the new method which can be used in BCI fields and other EEG-based studies.

Chapter 5 analyses the classification of various speech related cognitive tasks against the idle state in an offline setting in order to explore the potential possibility of a speech related onset detection system. It consists of two different experiments, where the first experiment is about high tone sound-production in covert, inhibited overt and overt speech modes and the second one is a siren-like sound imagery task.

Chapter 6 explores a more in-depth sound-production related onset detection system in a simulated-online situation in order to investigate whether the proposed method can be used in online settings with a cue-based approach to a self-paced system. This chapter also suggests the new performance evaluation metric (true-false-positive score) for self-paced BCIs.

Chapter 7 contains two experiments. The first one shows the comparison result between our new sound imagery onset detection task and the motor imagery task in an online situation with daily-life task scenarios such as watching video and reading text. The second experiment of this chapter is about testing the new onset detection system at outdoor laboratory settings in order to demonstrate real-life uses and show its limitations.

Chapter 8 provides an overall in-depth discussion of the main findings and the scientific contribution to the BCI field followed by a summary of conclusions along with proposals for future work in Chapter 9.

1.6 Ethical Matters

In this research, all the hardware resources were provided by the BCIs-Neural Engineering laboratory at the University of Essex and all the experiments were done in accordance with the University of Essex's Ethics Committee guidelines. The brain signal recording method was non-invasive and no thoughts were monitored, apart from very specific and deliberate sound imagery and motor imagery. The data were also anonymised as per ethical guidelines.

2 Neuro Electric Physiology Background

2.1 Electroencephalography

Electroencephalography (EEG) is a brain signal recording method that is widely used in clinical areas and BCI research. It normally refers to a non-invasive method that detects oscillations of the brain electrical potential (electric currents during synaptic excitations) from the surface of the scalp [26].

Because of its ease of uses, inexpensiveness and safety (non-invasive manner), EEG is the most widespread method for brain signal recording by far [6]. However, there are some disadvantages. The electrical voltage that is generated from a single neuron is not high enough to be measured. Thus, EEG measures hundreds of thousands of synaptic excitations and their synchronous activation, which increases the electrical amplitude [27]. Even though they have a synchronous electrical activation, the signals still have to cross the scalp, skull and many other layers. It makes the signal weaker and of poor quality. Moreover, this recording method is easily affected by artefacts [6].

EEG Rhythms

The brain signal has an individual variation and it changes easily according to the conscious level and mental activities even of the same person [27]. However, EEG rhythms have specific patterns and they can be categorised according to their frequency

range: Delta (δ), Theta (θ), Alpha (α), Beta (β), Gamma (γ) and Mu (μ). These frequency ranges were defined in accordance with their distribution over the scalp or biological significance [6].

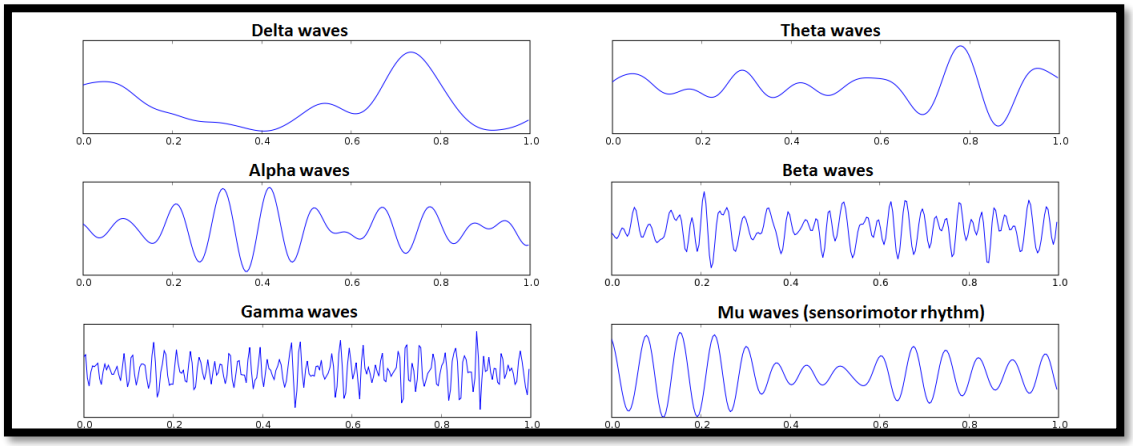


Figure 2.1. Various EEG frequency ranges for a 1 second sample (acquired in the Oz position) [28].

Figure 2.1 shows 1 second samples of various EEG rhythms, which are categorised by their frequency range. Their characteristics are shown in Table 2.1.

Table 2.1. Characteristics of various EEG bands.

EEG (Freq Range)	Properties
Delta (<4 Hz)	- Deep sleep stage [27]. - Brain disorders (if a large amount of delta activity is found in awake adults) [29]. - Due to its low frequency, the delta wave is easy to confuse with artefacts such as muscles of the neck or jaw [6].
Theta (4-7 Hz)	- During sleep states (not deep) or meditative concentration [27]. - The theta wave can be seen from some cognitive processes (e.g., calculation) [30].
Alpha (8-12 Hz)	- Quiet and relaxed state and the amplitude increases when the eyes are closed [27]. - The alpha rhythm can be found from frontal, temporal, parietal and occipital regions and it can be categorised according to its functional roles (e.g., mu, occipital and tau) [31]. - The Alpha rhythm from the occipital region is related to visual processing and metal effort [32].
Beta (12-30 Hz)	- Busy or active concentration state [27]. - This rhythm is recorded in the frontal and central area of the brain and it is related to motor behaviour [33]. - The frontal beta rhythm can be recorded for a few milliseconds post-stimulus and it is related to stimulus assessment and decision making [34].

Gamma (30-100 Hz)	- The gamma rhythm can easily be affected by EMG or EOG artefacts [35]. - It is related to perception and consciousness or the REM sleep stage [27]. - This rhythm appears during the linguistic processing of meaningful words at around 30 Hz [36, 37].
Mu (8-13 Hz)	- The mu rhythm can be found at a similar frequency range with the alpha wave. However, it is more associated with motor activities then the alpha rhythm [38]. - It appears from the sensory-motor cortex during either performing or imagining motor actions [27].

EEG Recording

There is a number of different brain signal recording systems such as Functional magnetic resonance imaging (fMRI), Positron Emission Tomography (PET), Computerised Tomography (CT), Magnetoencephalography (MEG), Functional Near-Infrared Spectroscopy (fNIRS) and EEG. Even though EEG has some disadvantages such as poor spatial resolution due to the volume conduction problem, it is by far the most common system and is widely used in BCIs because of its advantages, such as inexpensiveness, portability, safeness and excellent time resolution [9].

An EEG recorder amplifies (between around 1k and 100k times) the electrical brain signal due to its negligible amplitude (between 1 and 100 μV) and then filters it with an analogue signal processer followed by an analogue to digital converter (ADC) process [39].

The placement of electrode sensors usually follows the conventional 10-20 system (Figure 2.2 – A and B), which has been standardised by the International Federation of Societies for Electroencephalography and Clinical Neurophysiology. This system places electrodes at a 10% or 20% interval distance on the scalp. Another option is the extension of a 10-10 system (Figure 2.2 – C) which can be used in order to improve spatial resolution [40].

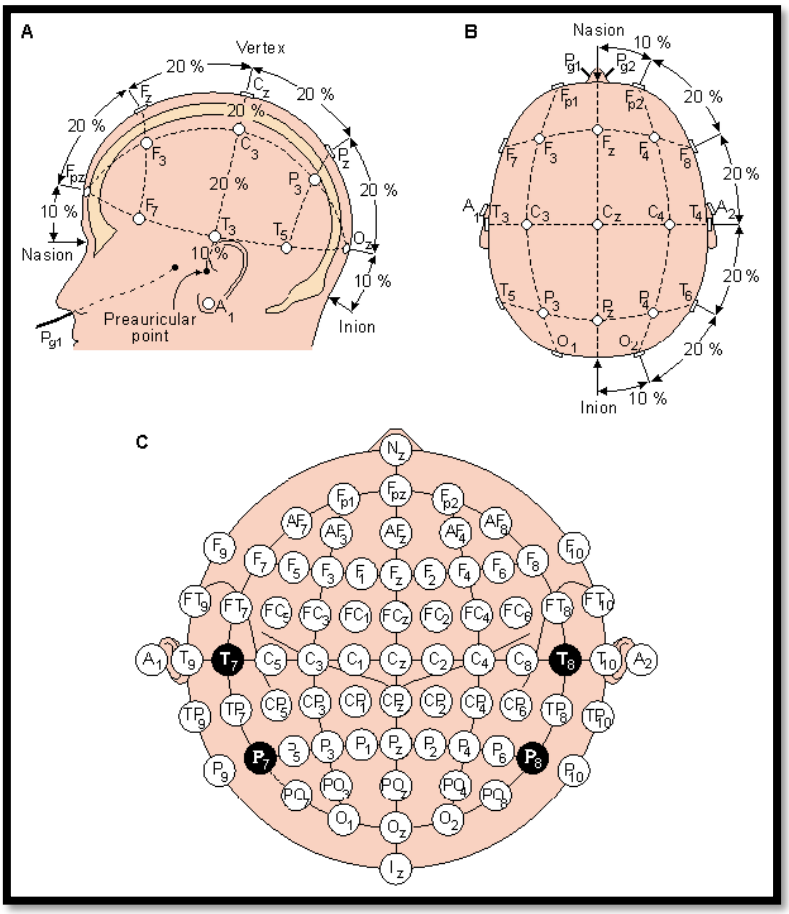


Figure 2.2. Conventional 10-20 EEG electrode placement over the scalp (A and B). Extended 10-10 system (C) [41].

One of the main issues of EEG recording is the volume conduction problem. Volume conduction is the transmission of electric current sources from the biological

tissue to the EEG sensors. As it was previously mentioned, the voltage from the scalp is much smaller than the membrane potential. Moreover, each tissue has different conductivities and impedances. For these reasons, the EEG signal which is measured from the sensor is very different from the signal that is measured from inside the scalp with the invasive-measurement method [26].

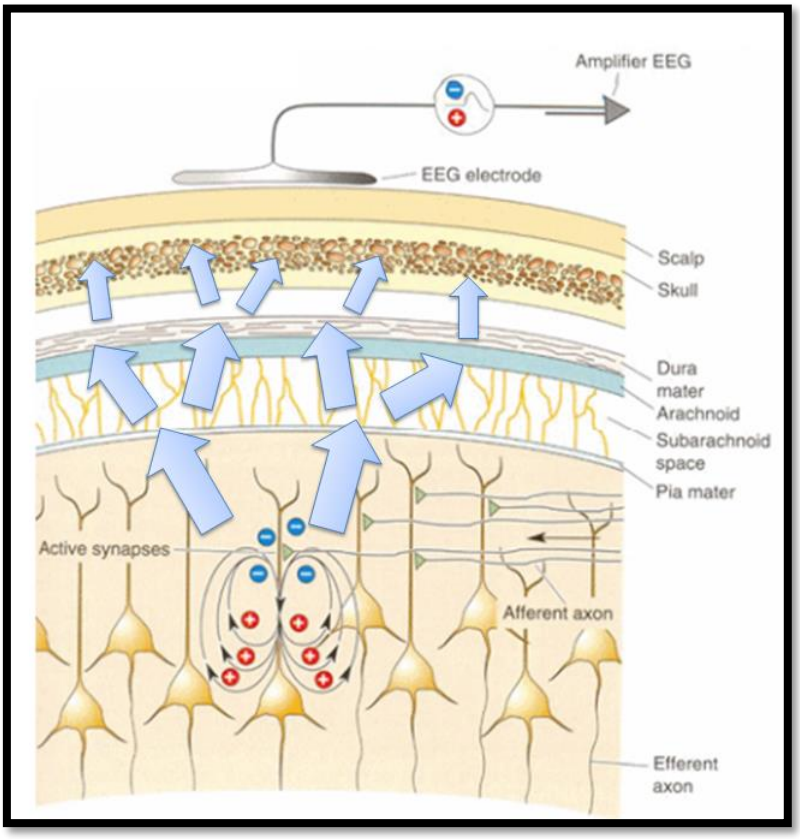


Figure 2.3. The volume conduction problem (modified from [27]).

Figure 2.3 illustrates the volume conduction problem. The power of the signals (the size of the arrow) is gradually decreasing when it passes through many layers and the path of the signals (the direction of the arrow) is changing variably. The volume

conduction problem is therefore one of the main disadvantages of the EEG recording system.

In addition, there are some statistical EEG problems. Firstly, EEG recording displays a **stochastic behaviour** and it requires statistical analysis as single values are not reliable (future values are not predictable). Secondly, EEG is a **nonstationary and nonlinear (time varying)** signal. Lastly, it is **noisy (large signal to noise ratio)** due to the volume conduction problem discussed above. For these reasons, various signal processing and feature extraction techniques must be used in order to increase the reliability of EEG measured BCIs [42].

2.2 Brief Review of Brain Physiology and Structure

The human brain can be divided in three major parts: Brainstem, Cerebellum, and Cerebrum. The following explanation was based on [26, 27].

- Brainstem: The brainstem relays signals (action potential) from the body (nerve fibres) to the brain centres and vice versa. It controls involuntary functions (e.g. heart regulation, hormone sections and biorhythms)
- Thalamus: The thalamus relays all the sensory inputs and motor signals to the cerebral cortex except the smell.
- Cerebellum: The cerebellum is related to motor control (voluntary movements) and some cognition.
- Cerebrum: The cerebrum contains around 10^{10} neurons, which are strongly interconnected. It plays an important role in the main brain

function such as in movement, sensory processing, olfaction, language and communication, and learning and memory.

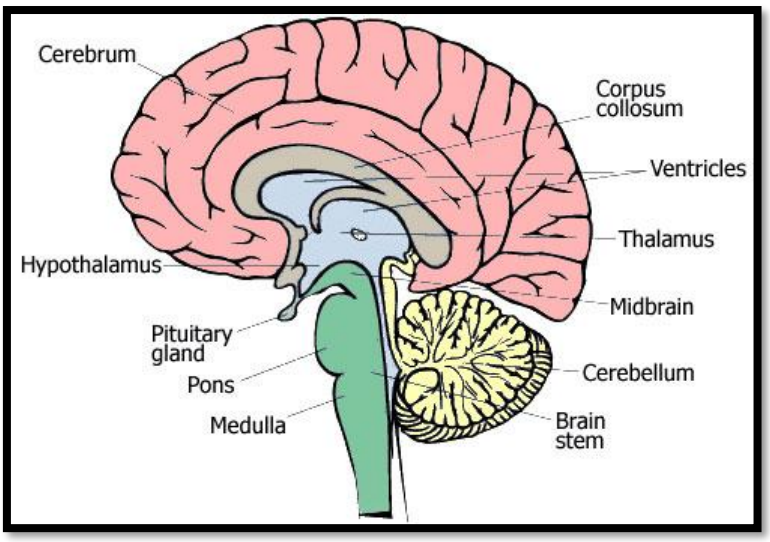


Figure 2.4. The human brain structure [43].

In the following sections, a detailed brain physiology of language processing and the sound production system will be discussed as it provides important background information for our sound production cognitive onset task. In addition, some brief explanation about visual, auditory, somatosensory, motor systems will be related to brain functions simultaneously. For example, on the sound imagery onset task, the visual and auditory, somatosensory and motor related signals can be found as artefacts.

Visual System

The procedure of visual information in the brain is through the ***eye → thalamus → visual cortex*** process. There are three major parts in this procedure, called the Optic Nerve, Optic Chiasm, and Optic Tract. The visual information, which originates from

the retina, passes through optic nerve and these optic nerves are combined at the optic chiasm. The information from the eyes is combined and divided according to the visual field at the optic chiasm [44].

Figure 2.5 shows how the visual pathway looks like. The information from the right side of the visual field goes to the left optic tract (to the left thalamus) and vice versa. From the thalamus, the visual information travels to the visual cortex, which is located in the backside of the brain (occipital lobe) and it is the largest part in the human brain system. The main functions of this part are the processing of information about objects (static or moving) and pattern recognition [44].

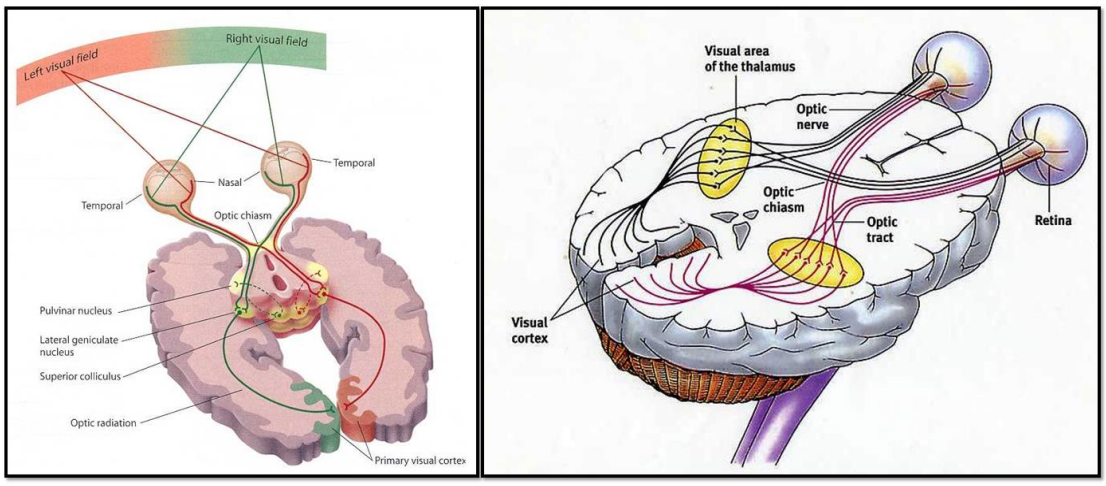


Figure 2.5. The visual processing pathway [45].

Auditory System

Ears can detect the variations of air pressure and translate them through the neural activity to the brain [27]. The procedure of auditory pathway is as follows:

1. The sound wave beats the eardrum (tympanic membrane).

2. The eardrum moves the auditory bones.
3. These moves cause the fluid in the cochlea to move.
4. The cochlea's movement stimulates the sensory neurons.

The neural signals from the cochlea, travel to the auditory cortex, which processes the sound information in the cerebral cortex. The sound information passes through many pathways so it is more complicated than the visual information [27]. Figure 2.6 shows how the auditory pathway looks like.

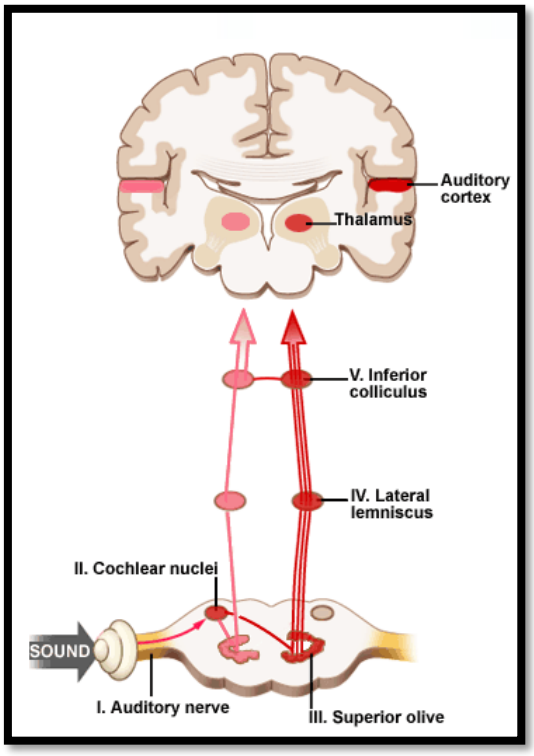


Figure 2.6. The auditory pathway [45].

As can be seen from the figure, the sound information from the auditory nerve goes to the superior olive via the cochlear nuclei in the brainstem [27]. This information

goes through the lateral lemniscus and the inferior colliculus, which are located in the midbrain. Finally, the sound information arrives to the auditory cortex through the medial geniculate nucleus (MGN) of the thalamus. The primary auditory cortex is located on the superior temporal gyrus in the temporal lobe [46].

Somatic Sensory System

The sensory system refers to the vision, auditory (hearing), somatic sensation (touch), gustatory (taste), olfaction (smell) and vestibular (balance/movement) senses. In this section, the somatic sensory system, which contains touch, pain, temperature and pressure, will be discussed. This system has receptors, which receive signals from the outside environment, throughout the whole body. The information from the skin touch or vibration is different from how the pain and temperature works. In this chapter, we will only discuss about the touch information [27].

The skin is the largest part of the sensory system. It describes the procedure of how the touch information is translated into neural signals and is detected by the brain. The information, from the receptors, goes through the sensory nerves and the spinal cord to the brain, which is in the parietal lobe of the cerebral cortex [47].

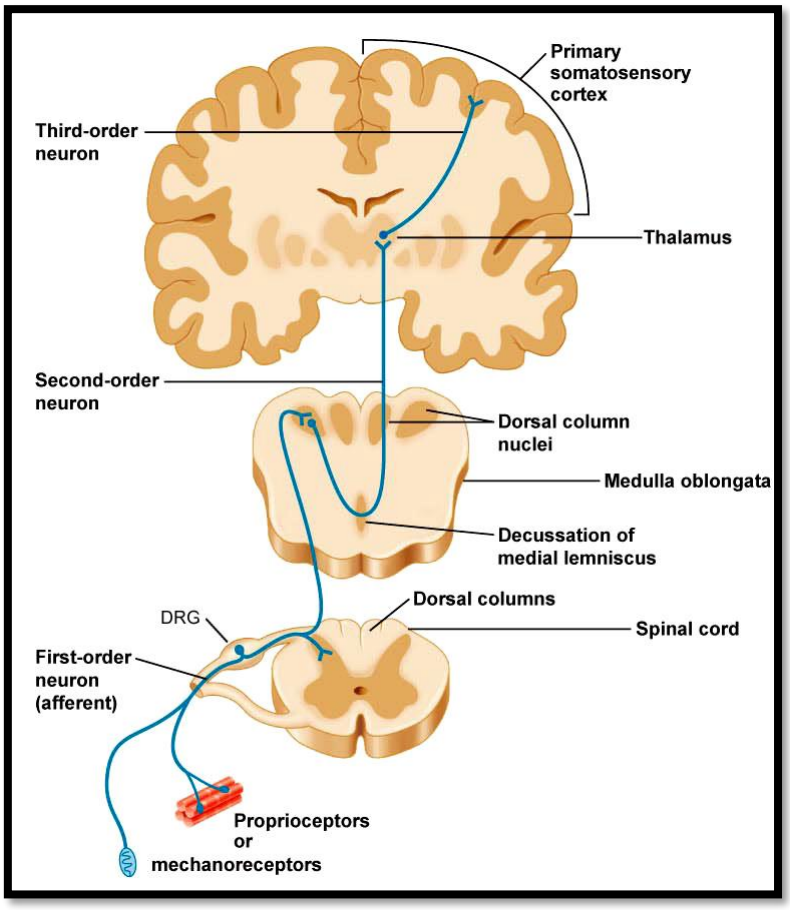


Figure 2.7. The touch and proprioceptive information pathway [48].

The pathway in Figure 2.7 is called the dorsal column-medial lemniscal pathway [27]. From the receptors, the touch information ascends to the thalamus and arrives at the primary somatosensory cortex the sensory information is processed. Figure 2.8 shows the location of the somatosensory cortex in the cerebral and the functions in each area. From a microscopic point of view, this location map may vary depending on the person.

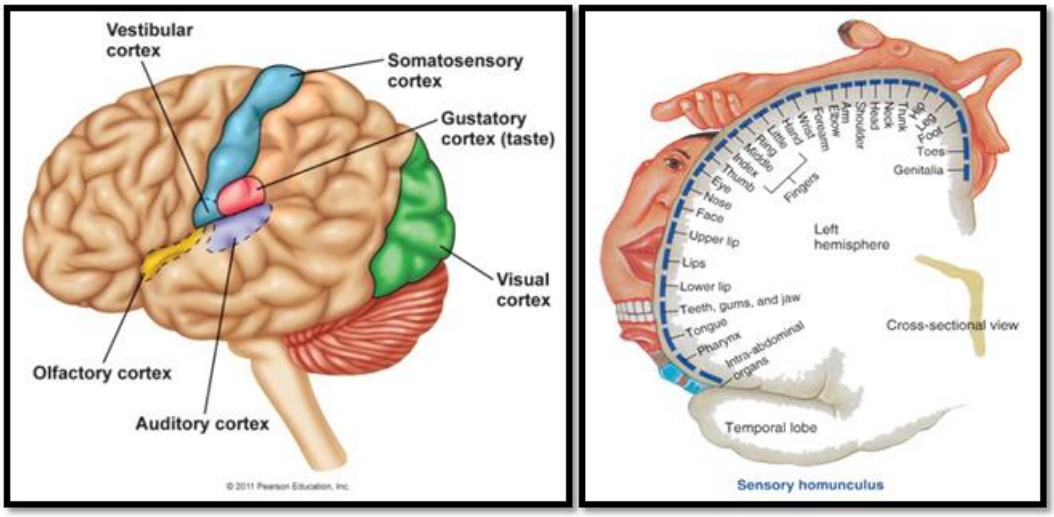


Figure 2.8. The somatosensory cortex (left) and somatotopic map (right) [48, 49].

Motor Control

The first step of human motor movement starts from the neocortex and basal ganglia of the forebrain. The basal ganglia (group of nuclei) are believed to be associated with action selection. It decides on how a person should behave and what the best movement is according to each case. Secondly, the motor cortex, which is located in the cerebral cortex, plans and controls voluntary movement. Lastly, the brain stem and the spinal cord execute the movements [27].

As can be seen from Figure 2.9, the motor cortex can be divided into five different parts; the Primary motor cortex, the Premotor cortex, the Supplementary the motor area, the Posterior parietal cortex, and the Primary somatosensory cortex. The primary motor cortex is the main part which sends neural signals to the spinal cord. Axons are descending to the spinal cord. There are two groups of pathways (lateral pathways, ventromedial pathways). The lateral pathways are related to voluntary movement. In contrast, the ventromedial pathways control the pose of the body and its balance [27].

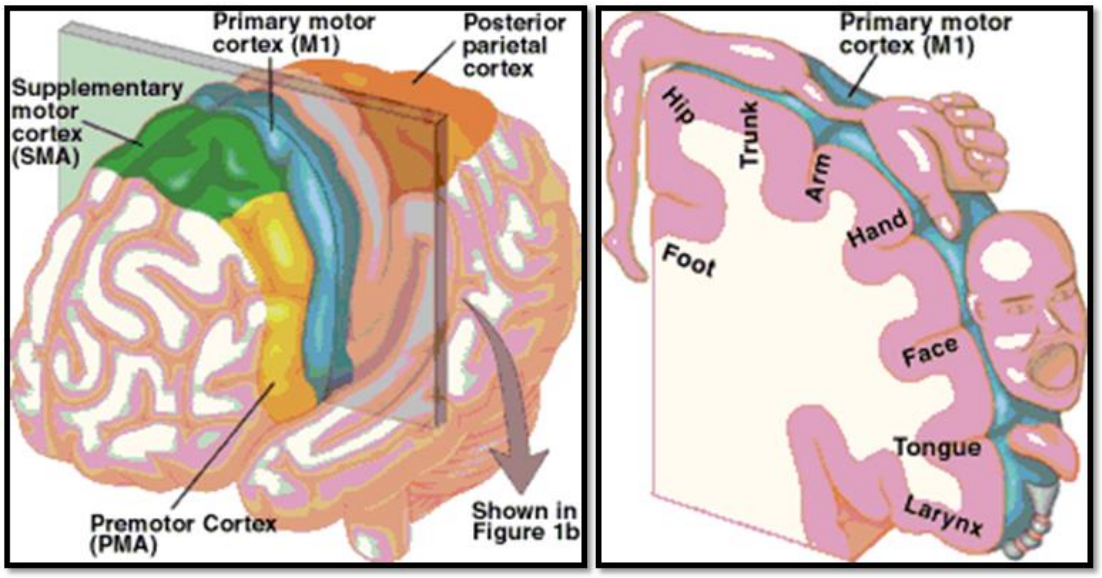


Figure 2.9. The motor cortex [50].

Language Processing & Sound Production

Language is the most natural form of human-to-human communication. It could propose a better way of future human-machine interaction. There are two important brain areas for language processing: Broca's area and Wernicke's area.

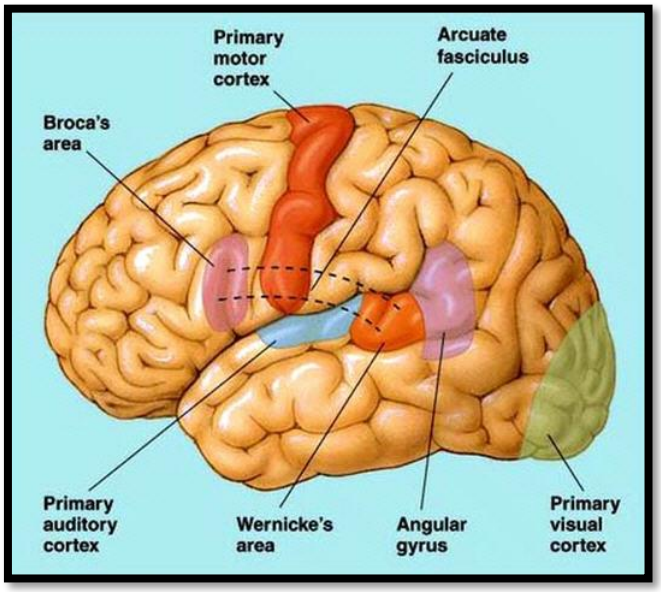


Figure 2.10. Cortical representation for language processing (left hemisphere) [51].

Broca's area is located near the motor cortex area, which controls the mouth and lips. It deals with fluent pronunciation by passing signals to the motor cortex. Wernicke's area is located between the auditory cortex and the angular gyrus. This area is better known for language comprehension [27, 52]. For this reason, people with Broca's aphasia showcase good comprehension but have a non-fluent and ungrammatical speech. On the other hand, Wernicke's aphasia shows poor language comprehension but fluent and grammatical speech, which are meaningless [27]. As can be seen from Figure 2.10, the left hemisphere is an important region for language processing. The key components of the language system are located in a small area.

Wernicke proposed a model for language processing and Norman Geschwind extended it, now called the Wernicke-Geschwind model, for sound/speech production in the human brain. Figure 2.11 shows the model for the two different tasks. Task (A) is repeating a spoken word and task (B) is repeating a written word. Stimulus from the outside environment is firstly processed by the auditory (for model A) or visual (for model B) cortex and it then comes to Wernicke's area in order to be understood as meaningful words. After this stage, in order to repeat the words (sound production), signals are moved to Broca's area from Wernicke's area via the arcuate fasciculus. In Broca's area, words are firstly converted to a code for muscular movements and then pass to the nearby motor cortex in order to control the movements of the lips, tongue, larynx, etc. [27].

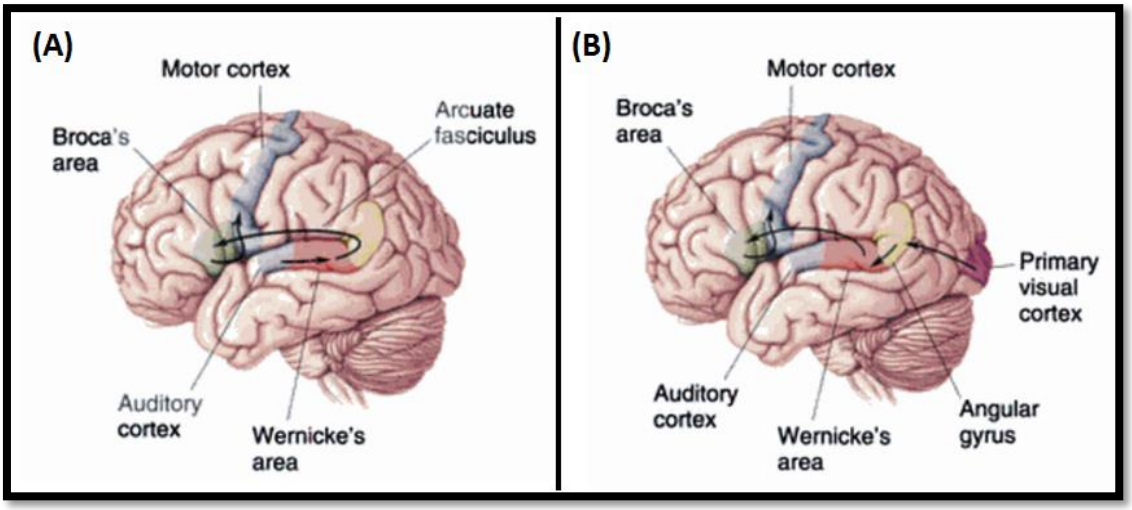


Figure 2.11. The Wernicke-Geschwind model. (A): repeating a spoken word, (B): repeating a written word [27].

However, this model has several errors and is oversimplified [27]. In a real sound production behaviour, it would not need stimuli from the outside environment as suggested in the above two tasks. In addition, one dangerous simplification is the overstatement of the significance of the given cortical areas for the certain functions. Researchers have recently found that aphasia is influenced by damage to subcortical structures (e.g., thalamus, caudate nucleus), which are not in the model. It is also often found that other cortical areas can sometimes compensate the language function after a stroke and actually recover it. Therefore, it is difficult to define sharp functional distinctions between regions, which are suggested by the model. In spite of these problems, the Wernicke-Geschwind model is continuously being used in clinical areas due to its simplicity and approximate validity [27].

In [53], possible new models for the language processing pathway were reviewed (Dorsal and Ventral streams). The major fibre tract which is anchoring the dorsal stream is the Superior Longitudinal Fasciculus (SLF) / Arcuate Fasciculus (AF). This

SLF/AF component has received the most attention for language and there exist several models of SLF/AF connectivity, such as the 'three-segment' model presented by Catani et al [54] and the 'two-segment' model by Friederici et al [55]. In terms of functionality of SLF/AF, it has been suggested that it can transfer information between Wernicke's and Broca's areas. However, it may be an important pathway in language learning [56]. It may also be involved in processing complex syntactic structures for language comprehension [57]. However, precise anatomical characterisation and functionality of these pathways still remains under investigation [53]. In terms of Ventral streams, the Uncinate Fasciculus (UF), Extreme Capsule (EmC), Middle Longitudinal Fasciculus (MdLF) and the Inferior Longitudinal Fasciculus and Inferior Fronto-Occipital Fasciculus (ILF and IFOF, respectively), have been claimed as new pathways as part of the ventral language stream but their role in language remains unclear and controversial [53].

Many studies repeatedly showed the important roles of Broca's and Wernicke's areas in the speech production and comprehension in different methodologies (lesion and behavioural studies: [58, 59], PET or fMRI studies: [60-62]). Some studies ([60, 63]) have also showed that language processing involves a widely distributed network of different cortical areas (but its neurological pathway remains unclear). Recently, ECoG studies have started investigating the neural correlation of speech tasks (e.g., [64-66]) [67]. They mostly showed a high gamma response from the left temporal gyrus. However, the task design was repeating words after auditory stimuli. It is therefore hard to characterise spectral and spatiotemporal dynamics in particular, for covert self-paced language processing.

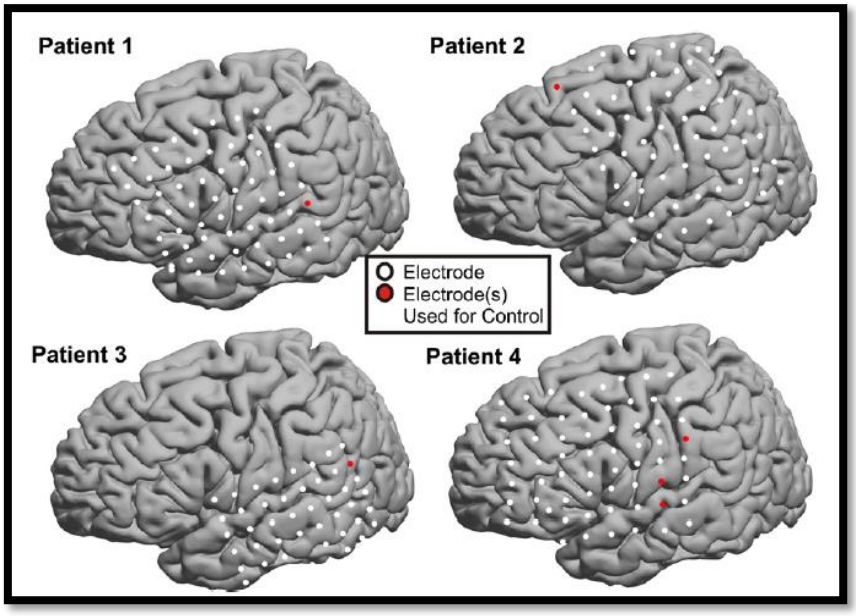


Figure 2.12. The physiological location for cursor control (red dot) in an ECoG speech production study [68] (figure from [68]).

In a ECoG study [68], a controlled computer cursor (one-dimension) was using sound-production tasks. There were five different tasks that were repeatedly dictating overtly or covertly expresses of ‘oo’, ‘ah’, ‘eh’ and ‘ee’, along with a rest period. Participant 1 (P1) used ‘ee - right’ vs ‘oo – left’, participant 2 (P2): ‘oo’ vs ‘ah’, participant 3 (P3): ‘ah’ vs ‘rest’ and participant 4 (P4): ‘ee’ vs ‘rest’. The optimal frequency bands for each participant would all vary (92.5-97.5 Hz, 410-420 Hz and 75-100Hz for P1, P2 and P3, respectively and 40Hz, 560Hz, 550Hz for P4). The physiological locations that were used for the classification would also all vary depending on their subjects (shown in Figure 2.12 - Brodmann area 42, 6 and 40 for P1, P2 and P3, respectively and 3, 22 and 43 for P4). Furthermore, another ECoG study [69] tested sound production tasks (i.e., ‘ah’ and ‘ee’ with six consonants ‘p’, ‘b’, ‘t’, ‘d’, ‘k’ and ‘g’) and the result showed that in addition to high gamma frequencies, lower frequencies (0-40Hz) are useful for speech activity detection [69]. These results

indicate that the actual optimal brain physiological area and the frequency bands for language processing / sound-production related tasks could all vary depending on their subject or they could appear widely in a distributed cortical area.

Moreover, the work in [70] reviewed facial EOG / EMG artefacts and showed correlation coefficient results between the facial artefacts and the EEG signal. Subjects were asked to stick their tongue out three times (which we think it may overlap with the speech task and give us similar facial EMG artefacts) and found the highest correlation coefficient value of 0.9675 between the artefacts signal and EEG channels T8 and T7, followed by channels FC6 and FC5 with a value of 0.8693. The F4 and F3 channels also reported a high correlation value of 0.8105 and 0.8095 for channels AF4 and AF3, respectively [70]. These results indicate that an artefacts handling procedure for speech related and sound production tasks (e.g., involuntary mouth and vocal cord movement, facial EMG) would be necessary.

2.3 Summary

In this chapter, the neuro electric physiology background was covered. The characteristics of EEG and its recorded issues were explained. A brain physiology and structure was also described. A brief explanation about visual, auditory, somatosensory, motor systems was provided, followed by a detailed background explanation of language processing & the sound production system. Having described this neuro electric physiology background, the next chapter will discuss in more detail the BCI related literature review. The general BCI applications, self-paced onset detection BCIs,

speech related BCIs and their EEG signal processing, artefact handling, feature extraction, classification and performance assessment will be covered.

3 Literature Review

3.1 Brain-Computer Interface

A Brain-Computer Interface (BCI), also known as Brain-Machine Interface (BMI), is a direct communication system, which includes hardware and software aspects, between humans and computers (or machines). It interprets human brain signal activities (e.g., EEG) and controls and interacts with their surroundings without any physical muscle involvement such as a touch screen or a voice control. Thus, it is not only an opportunity for people with severe motor disabilities but also offers various interesting applications for able-bodied people such as BCI controlled home devices or games [6].

In general, a BCI consists of five consecutive stages: data acquisition, pre-processing or signal enhancement, feature extraction, classification and control, and feedback [71]. Figure 3.1 shows the general proceeding structure of a BCI system.

At the signal acquisition stage (stage A in Figure 3.1), the system captures the brain signals. At this stage, brain activity can vaguely be categorised into two different ways: 1) the system monitors ongoing brain activities while the user performs specific mental tasks such as in-brain simulation of hand movement. 2) the system provides stimuli (e.g., flashing objects or beep sounds) and captures the user's brain involuntary

responses. From the signal processing (stage B), a noise reduction, artefact processing and further processing (e.g., frequency band filters, spatial filters) are applied to the recorded brain signals in order to increase the signal-to-noise ratio. The feature extraction (stage C) reduces the dimensionality of signals without any relevant information loss in order to minimise the complexity of the data. It increases the reliability of classification (stage D), which can discriminate the user's mental tasks. In addition, some BCIs give feedback to the users (stage E) [9].

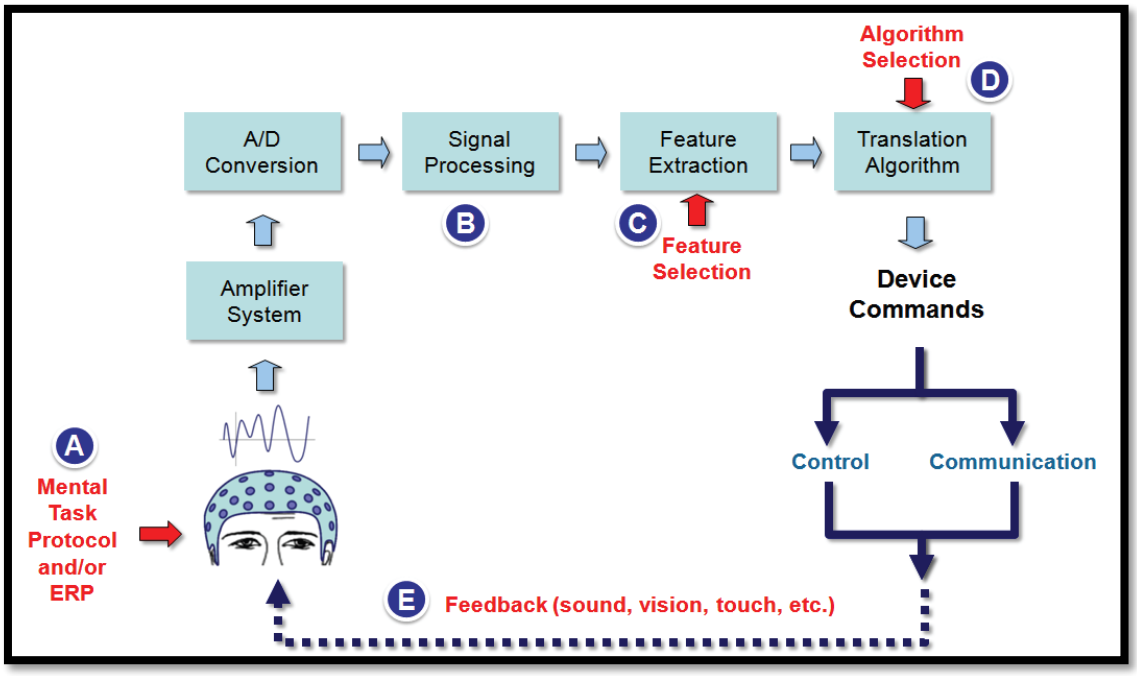


Figure 3.1. Schematic model of a BCI system [9].

3.2 General BCI Application Paradigms

P300 BCIs

Event-Related Potentials (ERPs) are electrical brain signals which are generated in response to visual (flashing letters), auditory (sound), tactile (sensory provocations) and mental counting stimuli. These are time-locked to the stimuli and it is clearer when the stimuli are unpredictable or the subject concentrates on the task [9, 72].

Figure 3.2 shows an ERP wave, which consists of many different components such as P300, N100 and N400. As can be seen, the P300 wave, which has a positive peak, appears between around 300 and 600ms post stimuli and has the most significant peak and clear latency. For this reason, P300 has been widely used in many BCI fields and it has many applications.

The most common P300 BCI application is a word speller. In the BCI competition Ⅲ the data set Ⅱ from [73], had 36 characters (6 rows by 6 columns). The rows or columns of the matrix were flashed in a random order and user would gaze at the target symbol and count how many times the chosen character was flashed. P300 is elicited only when the target character's row or column of the matrix is flashed and BCI uses this effect to determine which the target character was. The first rank of the competition achieved a 96.5% and 73.5% accuracy when each row and column blinked 15 and 5 times, respectively [73].

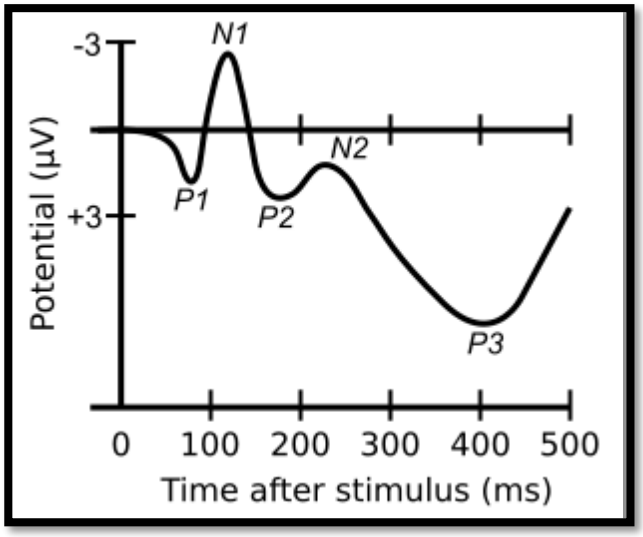


Figure 3.2. ERP components [73].

SSVEP BCIs

Steady-State Visual Evoked Potentials (SSVEPs) are similar to the P300 approach. However, there are individual flashing objects at different frequencies, which are usually between 6 Hz and about 35 Hz [74]. Users fix their gaze on a target object and then the object flashes at a specific frequency. It will generate a strong target frequency-domain element on the visual cortex, thus the BCI can determine which the target object was. This approach is usually faster than the P300, however, it is stricter for the user to gaze at the target object [9].

Motor Imagery BCIs

A Motor Imagery (MI) task is the imagination or mental rehearsal of limb movements such as hands, feet or the tongue. It does not need to have external stimuli (e.g., visual or auditory stimuli for P300 or SSVEP), therefore the users do not need to entirely focus on the BCI monitor in order to execute a command. In addition, one advantage of MI BCIs is that the movement-related brain activity is well localised as it was illustrated in Figure 2.9 [9]. Thus, MI widely used BCI techniques and it was

extensively investigated in many well-known BCI groups (such as Wadsworth [75], Berlin [76], or Graz [77]) [6].

In the latest BCI competition IV [73], the data set 2a [23] used 4-class motor imagery tasks (left hand, right hand, foot and tongue) with 22 EEG channels. The winner of the competition showed that the average (9 subjects) of the kappa value for the predicted class-labels of the evaluation set was 0.57 [73].

Other Tasks

There are many other examples of BCI applications. **Slow Cortical Potentials (SCPs)** are slow voltage changes that can be found over the vertex at low frequencies (mainly less than 1 Hz which can be extended up to 4 Hz). They require a long-term recording with many trials in order to be able to monitor the overall EEG activity trends [78].

Other cognitive tasks, which generate different neural activities in different areas can be used for BCIs. In [18, 79], a variety of different mental tasks were tested in order to be able to identify them from an idle (neutral) state. The auditory recall, navigation imagery, sensorimotor attention, calculation and many other cognitive states were investigated.

3.3 Self-paced vs. Cue-based BCIs

Among other various definitions, BCIs can be categorised as cue-based (synchronous) or self-paced (asynchronous) systems. Figure 3.3 shows different

examples between Self-Paced BCIs (**SP-BCIs**) and Cue-Based BCIs (**CB-BCIs**). As can be observed, CB-BCI systems have to inform the user when to start and stop thinking the command to the machine. It means the user must follow the computer's own timing commands [5]. The majority of the current stage of EEG-based BCI systems are CB-BCIs, where the analysis and classification of brain signals are locked to the machine's predefined timing protocol [7]. The advantage of CB-BCIs is that they provide a better classification rate and an easier analysis than SP-BCIs as the machine knows the precise time location of relevant events by providing specific cues or triggers to the users. For example, very common cue-based approaches are the P300 and SSVEP BCI systems, which were discussed earlier. However, this approach forces the users to keep concentrating on the computer's command (e.g., looking at blinking objects) which is a very unnatural interaction approach. In addition, CB-BCI systems always need an external stimulus from the computer so that the computer can make decisions for the users [8].

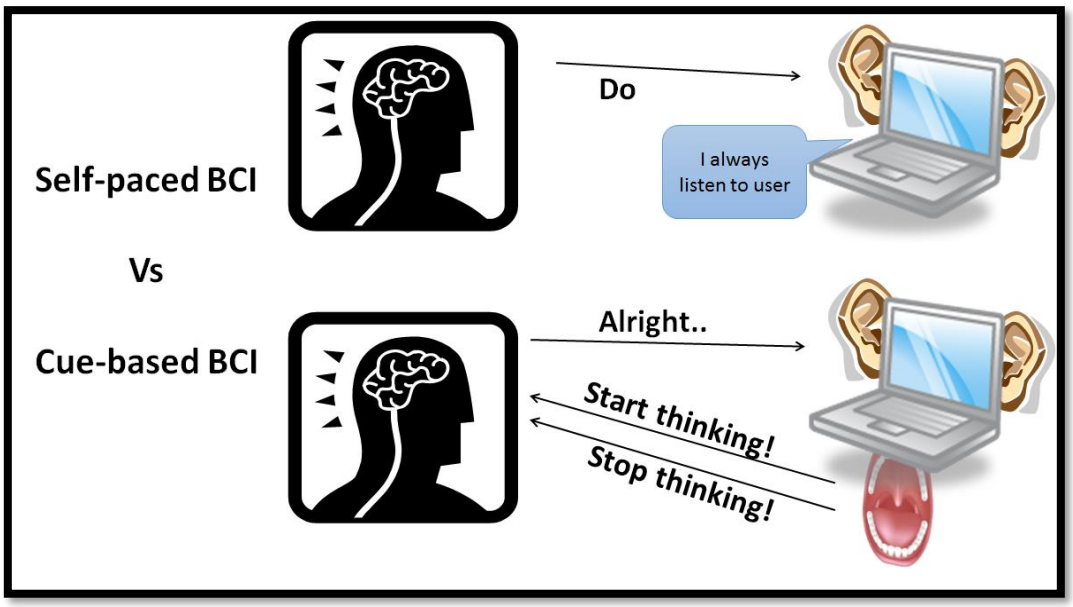


Figure 3.3. Example of human and BCI system interaction (self-paced vs. cue-based BCIs).

On the other hand, SP-BCIs analyse the user's brain signals continuously without a specific computer-controlled stimulus [9]. The users control the BCI system by intentionally performing a specific mental/cognitive task at any point they want [7]. This design is more intuitive to the users. It enables them to control the system in a more natural way according to the user's own timing and speed of communication. It increases the usability and flexibility of the BCI systems [10]. In order to expand the BCIs from the indoor laboratory settings into real-world applications, the machine dependent timing constraints have to be resolved. For this reason, the SP-BCI system is essential for the real-world use of BCIs in the near future.

However, there is still a great challenge for SP-BCIs. It is much more complicated to analyse them than CB-BCIs as they have a lack of knowledge about the precise time location of the user's command. The user's control intention and timing are usually unknown to the machine [5, 7]. Therefore, SP-BCIs should continuously analyse the ongoing brain activity and they should be able to distinguish between Intentional-Control (IC) and Non-Control (NC) states. An IC state describes the process where the intended brain activity is supposed to produce a BCI output, which is quite straightforward. However, the NC state could be any other state besides the IC state. It can be also called a non-specific state (i.e., idle, daydreaming, other mental activities or performing some other actions) [8]. Distinguishing between the IC and NC state is very important in order for SP-BCIs to reduce the false-positive rate. In order to solve this issue, onset detection methods can be introduced (it will be explained in the following section).

In the case of these timing analysis difficulties, the classification performance in SP-BCIs is usually poorer than in CB-BCIs. In addition, the performance assessment

method of SP-BCI systems is not standardised and therefore it may vary depending on the values provided by papers and experimental settings. The following section, recent studies of SP-BCIs will be discussed together with their applications, evaluation methods and performance results.

In [80], a speech related SP-BCI was tested with a functional Near Infrared Spectroscopy (fNIRS) for 5 male subjects. The experiment was carried out as an off-line scenario and there were three types of speech activity: a normal audible speech, a silent speech (moving the articulatory muscles but without sound production), and speech imagery. The experiment consisted of 10 sentences (around 66 characters), which were taken from a news broadcast. The subjects were asked to produce each speech mode followed by pause periods, which were regarded as the idle state [81]. SVM was used for classification. In this paper, precision and recall were used for the evaluation method. The average result of 5 subjects had a 74% accuracy, 61% true-positive rate, 16% false-positive rate, 84% true-negative rate and the precision and recall values were 0.73 and 0.61, respectively [80].

In [82], motor imagery was used for the self-paced system. There were three right-handed subjects who were asked to perform real movements (i.e., extending their right wrist, holding it sill for about 1-2 seconds and then relaxing) on their own pace without any cue from the system. However, they were asked to leave at least a 4 seconds interval between the tasks. They used an electromyogram (EMG) in order to identify correct onset muscle activities. For the feature extraction, the Thomson Multitaper method was used for the Power Spectral Density (PSD) and the Davis-Bouldin Index (DBI) was applied for the selection of the features. The Naïve Bayes classifier was used for classification. In this study, the performance was analysed with a True-False (TF) difference rate and an average time between the correctly detected onset and the real

movement onset. The TF rate was defined as: $TF = (TP/E – FP/(E+FP)) * 100$, where *TP*: True- Positive, *FP*: False-Positive and *E*: total number of events. In this experiment, the authors achieved a TF rate of 95%, 69% and 59% in 3 different subjects and the average time was 325 ms, 788 ms and 688 ms respectively [82].

In [8], a tetraplegic subject, who got injured in the spinal cord, was trained with the cue-based motor imagery system during a 4 months period. After that stage, for the self-paced system study, the authors used two electrodes; Cz (foot representation area) and Fz (ground electrode). A single logarithmic band power feature was applied and a simple threshold was used to distinguish between the imaginary thought of foot movement and the rest (non-control) state. The aim of this experiment was to move a virtual wheelchair to the target area and lay it still there for a couple of seconds. The authors evaluated the subject's performance with the percentage of the accurate stop at the target place. They achieved an around 90% accuracy value. However, for most of the duration of the experiment, the time that the wheelchair stopped moving was too short (between 0.08 and 0.88 seconds), where it was supposed to last for at least a couple of seconds. Therefore, even though, the results reported quite a high performance, there was still some doubt concerning the actual classification accuracy between the active and non-control states.

3.4 Speech Related BCIs

Speech related BCIs can be divided into overt and covert systems. Overt speech takes place when one is speaking naturally so that they can clearly be heard and moves

all the related muscles such as the lips and the tongue. In contrast, covert speech refers to when a subject imagines speaking internally in their mind without any actual speech related muscle movements.

Overt speech experiments have been successfully studied with an EEG measurement so far. However, overt speech exhibits a bad signal to noise ratio (SNR) which is caused by muscle movement artefacts that accompany it [83]. In addition, overt speech does not have a reasonable form for ultimate BCI systems (it can be used speech recognition technology can be used instead of BCI).

In contrast, a covert speech BCI has big advantages even though it is very difficult to classify and it may have different processes compared to the overt speech production [83]. Firstly, it is suitable when the oral speech is undesirable such as in a quite library. Secondly, it can be a solution for people who are not able to speak out overtly [84].

There is a number of papers in published literature, which investigated speech related BCIs. In [13], three subjects were instructed to perform three different tasks: imaginary vocalisation of the ‘a’ and ‘u’ vowels (including imaginary mouth and lip movements) and no action state. These tasks were performed with the appearance of a visual cue. Strong positive Speech Related Potentials (SRPs) of around 300 ms were reported in this experiment, after the visual cue appeared, at channels C3, Cz and C4 (15 Hz cut-off low pass filtered). The results showed that the difference between the ‘u’ task and the idle state had an around 78% accuracy value on average for the three participants. The ‘a’ task vs. the idle state accuracy was around 72%. On the other hand, the ‘a’ task vs. the ‘u’ task showed a lower accuracy result of around 62.6% [13].

In [12, 14, 85], the authors investigated a similar imaginary speech case. The subjects were asked to think either the syllable ‘ba’ or ‘ku’ at a specific rhythm with audio cues. In [85], where the highest classification result of 99.76% was achieved, the authors studied the covert production itself rather than the imaginary speech muscle movements.

In [81], the functional Near Infrared Spectroscopy (fNIRS) signal was recorded with various speech modes (audible, silently uttered and imaginary speech) and no speech production state. These speech modes were compared to the no-speech state and the classification rates were between 69% and 88%. During the speech states, subjects read around 66 characters followed by 10 seconds of a no-speech state.

An ECoG study [16] showed a continuously oral speech decoding, called the brain-to-text system in an offline setting. Seven subjects participated and they were instructed to read text (historical political speeches) aloud which was cut into 21-49 phrases. They then trained the data with 23 phones (e.g., ‘aa’, b’, ‘ch’, ‘eh’, etc.). At this stage, three of the participants were rejected as they showed speech related activations with a lower than normal probability. The example was decode from “/w/ih/aa/r/ /k/aa/m/ih/t/aa/t/ /t/aa/t/eh/” to “we are committed today”. The recoding results showed that the system yielded significantly higher accuracies than the random models. Participants 1, 3, 5 and 6 showed an around 10-15% accuracy value while participants 2 and 7 showed just under 40% and around 50% accuracy values, respectively [16]. Even though the results are very promising and the system itself has great potential, the artefact removal procedure could have been considered as the reading task was overtly conducted, which may influence the classification outcome with some artefacts such as facial EMG.

3.5 Onset Detection in BCIs

Onset detection allows self-paced BCIs to detect when a user wants to send a command. It should be able to classify a specific active task against the idle (no control) state [4]. However, determining the precise onset time is quite difficult and this is called the onset detection problem in BCIs.

There are a couple of reasons why onset detection is important. Firstly, the identification of the idle state is compulsory for self-paced BCI applications. The system must not only have a reliably high enough true-positive accuracy but also an as low as possible false-positive rate in order to be used in realistic applications (especially when safety is an issue). Secondly, for a full function of a self-paced BCI, the user must be able to turn on and off the machine when they desire. Onset detection can therefore be used as an on/off switch [86].

The various onset detection methods (e.g., classifying the idle versus the task states) will be introduced and discussed in this section. The idle state in BCIs can be categorised in two types. The first one is a relaxing state, where the subject tries to think nothing and relaxes. The second one is where the subjects can perform almost any mental tasks except the active mental task, which belongs to the class [87]. However, it is very difficult to classify various non-class mental states as an idle state because of its diversity. For this reason, many papers regarded the resting period as an idle state.

In [88], the authors classified the motor imagery tasks against the idle state. They made two two-class classifiers for three different classes (left hand, right foot and idle). The experiment was performed with three different states: left hand, right hand motor imagery tasks and relax. 32 channels at a 256 Hz sampling rate were recorded and an

Event-Related Desynchronisation (ERD) in mu and beta rhythms was used as a feature. If the feature did not belong to the motor imagery tasks, they assumed it belonged to the idle state and it accordingly achieved true-positive rates of around 40% from the relax state. A paper [89] also suggested a similar method of idle state detection and achieved a 48.73% accuracy value.

In [90], the imagination of motor movements, cube rotation, subtraction calculation, word association and relax (closed eyes) states were tested as an onset of a self-paced BCI. The task was to type 27 symbols in a keyboard and it achieved a 0.22 bit rate (detailed keyboard layout and paradigm can be found in [90]).

In [18], various mental tasks (Movement imagery, Mental arithmetic, Navigation imagery, Auditory imagery and Phone imagery; details of mental tasks can be found in [18]) were compared with the idle state. The idle state was defined as not performing any of the active tasks and the subjects remained focused on the fixation cross in the same manner as the active task. In this experiment, the average accuracy of the five subjects was 74.4% (motor left), 71.8% (motor right), 73.2% (calculation), 66.8% (navigation), 62.4% (auditory recall) and 62% (phone imagery) [18].

Qian et al. [91] developed a motor imagery-based brain-controlled switch. They tried to develop a system that minimises the false positive rate in order to be feasible in real world switch system. They employed event-related potential-based systems, which instructed the subject to repeat 1 s during the right index finger pinch-urging time (intentional control) followed by 2 s of resting time, with the use of sound cues until the switch turned on. The results on average of the four subjects showed an around 0.8% false-positive rate with the response speed being around 37 s (urging time: 12.3s).

3.6 EEG Signal Processing

EEG signals usually have a poor signal-to-noise ratio as electrodes are placed on the subject's scalp, which forbids the direct contact between the brain and the machine. Therefore, the EEG signal is very weak and contains various artefacts. For this reason, a pre-processing process is required in order to achieve an optimal classification accuracy.

Referencing

One of the most common and easiest ways to remove common environmental noise in an EEG recording system is referencing. Here are some common referencing methods in BCIs:

- Common Reference: One or two extra electrodes are placed and the noise data are subtracted sample by sample in the time domain. The reference point should not have any signal information and it must be placed near to the scalp which contains similar noise features (e.g., ear lobes or mastoid) [9].
- Bipolar Reference: The electrodes are grouped in pairs and the noise signal is subtracted between the two electrodes. It describes the difference of the linked electrodes [92].
- Common Average Reference (CAR): The average value of all the channels is removed from each channel. This zero-centered output value can be regarded as a spatial filter as it removes common noise for all the present channels [93].

- Laplacian Filtering: This method is useful because it maximises the spatial difference by the subtraction of one channel from the average of the surrounding channels [9].

Frequency Band Filtering

The raw EEG signal has various frequency components and artefacts. In order to keep informative frequency ranges and remove undesired signals such as mains interference and EMG noises, a band-pass filtering process is necessary.

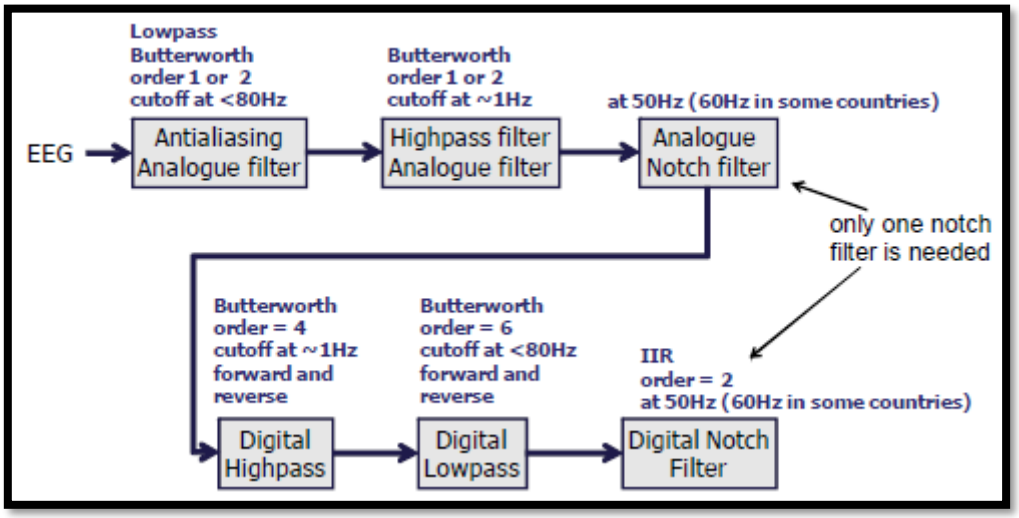


Figure 3.4. Typical frequency band filtering of EEG signals [9].

Figure 3.4 illustrates a basic example of a filtering process for an EEG signal. The high-pass filter with a value of around 0.5 Hz is used to remove motion related artefacts, such as body and electrode cable lines movements. The low-pass filter is used to reduce EMG artefacts. The notch filter is required to remove the mains interference

(50 Hz in the UK, 60 Hz in USA). However, the filter type, cut-off frequencies and order numbers can all be variably chosen depending on the systems.

3.7 Artefact Removal

The EEG artefacts are normally much higher in amplitude than EEG signals and cover a wide frequency range [36]. In order to have clear EEG data, the artefact removal is necessary. There are three main EEG artefacts:

- Physiological variability: It is caused by other electrophysiological effects such as EMG (e.g., jaws, facial muscle twitches and swallowing), EOG (eye movement and blinking) and ECG (cardiac activity) [9, 36].
- Transducer artefacts: These artefacts are related to the movement of cables and the change of the electrodes' contact from the scalp during the long sessions of recording. This could change the baseline level of the EEG recording [36].
- Mains interference: Mains interference is appearing at 50 Hz in the UK and 60 Hz in some countries). This artefact can be removed by referencing or a notch / stop-band filter [9].

In order to remove these EEG artefacts, various techniques can be used. Transducer artefacts and mains interference can be removed by referencing and frequency band filtering. Most motion related artefacts can be eliminated with a high-

pass filter with a cut-off frequency of around 0.5 Hz. However, some EMG or EOG signals overlap with the EEG signal. The study in [70] showed correlation coefficient results between facial artefacts (e.g., blink, left/right wink, raise brow, smile and clench) and the EEG signal. They found that channels F3 and F4 had the highest correlation coefficient values of 0.8579 followed by channels F8 and AF3 (0.8242) for the blink artefacts. For the swallowing artefacts, channels FC6 and FC5 had the highest value of 0.9543 followed by F8 and F7 with 0.8819. These results indicate that specific channels were highly correlated with the facial EMG artefacts [70]. For this reason, it is difficult for the muscle movement artefacts to be fully removed from EEG [9]. In order to reduce these artefacts, some statistical methods can be applied.

Recent surveys on BCI artefacts [94, 95] demonstrated blind source separation techniques that are mostly used in current BCI systems, such as PCA, ICA and BSS-CCA, which have been shown to provide satisfactory artefact removal results. Therefore, the BSS-CCA, ICA and PCA techniques will be briefly described in the following sections in addition to wavelet de-noising.

Canonical Correlation Analysis (CCA)

Canonical Correlation Analysis (CCA) is a statistical analysis that measures the linear relationship between two multidimensional variables. It finds two bases for each variable and the correlation matrix between the variables is diagonal so that the correlations on the diagonal can be maximised [96]. The Blind Source Separation (BSS) problem can be solved with CCA by taking the EEG multi-channel source vector as the first multidimensional variable and the temporally delayed EEG source vector as the second variable [97]. In order to solve the BSS problem, BSS-CCA assumes mutually uncorrelated sources that are maximally auto-correlated. It can therefore be used to

separate the brain signal from the muscle activity sources as the muscle artefacts have relatively low autocorrelation compared to the brain signal [98]. (a detailed calculation procedure with equations for the EEG artefacts handling can be found in [98]).

Independent Component Analysis (ICA)

The Independent Component Analysis (ICA) is another widely used technique which can solve the BSS problem. It divides EEG sources into multi-components, which maximises their mutual independence [99]. By removing the artefact component ICs and reconstructing the signal, it produces artefact-free EEG data (a detailed calculation procedure with equations for the EEG artefacts handling can be found in [100, 101]). The ICA procedure usually requires a manual visual inspection from the experts in order to clarify the artefact component ICs, therefore it is not suitable for an online system. However, the work in [102, 103] demonstrated an automatic artefact IC detection technique by using Kurtosis and Entropy.

Principal Component Analysis (PCA)

The Principal Component Analysis (PCA) is a statistical procedure that identifies patterns in high dimensions of data (multi channels or multi trials) by finding their similarities and differences [104]. It uses a linear transformation and it generates a set of principal components (Eigenvalues). The higher value of the PCs represents a strong correlation between multiple signals (e.g., channels or trials) and the smaller value of PCs can be regarded as a noise component, which normally has less correlation. By

sorting these PCs in descending order and eliminate the last few PCs and reconstruct the signals (reversing the PCA process), noise-removed signals will be generated.

Wavelet De-noising

The Wavelet Transform (WT) decomposes a time domain signal into time-frequency space. Due to this feature, it is one of the superior techniques for processing non-stationary signals such as EEG. It also works well for finding and removing artefacts. In this section, the wavelet de-noising method will be introduced and a detailed WT technique for feature extraction and its background knowledge can be found in the next section.

Eye-blinks and movements generate electrical signals known as Ocular Artefacts (OA), which are the main contaminating recourse for EEG. In [105, 106] the Stationary Wavelet Transform (SWT) was applied in order to detect OAs and correct EEG signals by removing the OAs from the contaminated EEG.

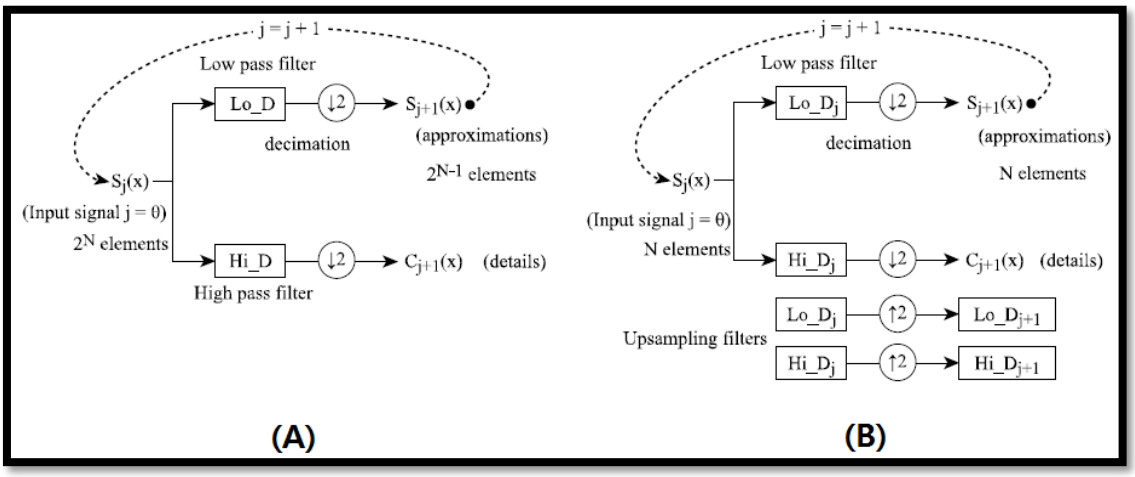


Figure 3.5. Decomposition process of using a DWT (A) and a SWT (B) [107].

The SWT is a modified version of DWT (Figure 3.5 shows the difference). While DWT uses samples downwards at each level, SWT is never sub-sampled but the filters are upwards sampled at each level of decomposition. Thus, each set of coefficients contains the same number of samples as the input [107].

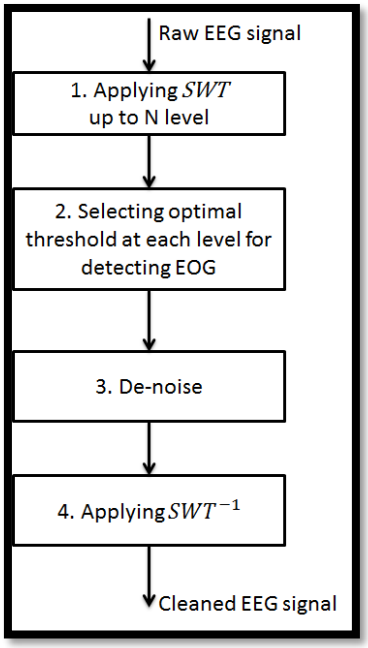


Figure 3.6. Block diagram of the wavelet de-noising method.

Figure 3.6 illustrates the de-noising procedure of using SWT. In [105, 106] the 'Coiflet 3' and 'Symlet 3' wavelet has been chosen respectively at stage 1 (from the above figure) as it resembles the shape of the eye blink artefact. In order to detect the EOG artefact, some threshold techniques such as the adaptive threshold were used in [105, 106] from stage 2. The SWT was found to be working well for the removal of the EOG artefact. In [108], various techniques have been compared for the removal of OAs and it was reported that the combination of SWT and the adaptive filter outperformed other methods.

3.8 Feature Extraction

For BCI systems, the use of raw EEG signals is very difficult to interpret and classify due to their characteristics [109]:

- Noisy: EEG signals are very noisy and have a poor signal-to-noise ratio.
- High dimension: BCI systems usually deal with a high dimension of data (e.g., multi-channel, multi-trials, frequency analysis).
- Non-stationary: EEG signals are non-stationary and time-variant.
- Small training sets: In BCIs, the training sets are usually small.

For these reasons, feature extraction plays an important role in helping to improve the recognition performance prior to the classification stage. In this section, general feature extraction methods in BCIs will be briefly introduced and some popular approaches will be discussed in detail.

Table 3.1. Brief explanation of various feature extraction methods in BCIs [6].

	Method	**Properties**
Dimension Reduction	*PCA*	- A statistical procedure that uses linear transformation - Transforms a set of correlated variables into a set of linearly uncorrelated Principal Components (PCs) - Noise and dimension (channels) reduction methods - Noise should be uncorrelated with the EEG signal in a PCA process
	ICA	- A method that splits a mixed signal into its original sources

		- Good for artefact removal such as eye blinking - Processing an ICA will lose EEG channel information
Space Feature	*CSP*	- Spatial filter in order to separate two classes from multi-channel data - Maximum difference in variance between two classes - It is affected by the spatial resolution - It is not as effective for asynchronous BCIs as for synchronous BCIs due to non-stationary properties of EEG signals
Time-Frequency Feature	*AR*	- AutoRegressive (AR) spectral estimation is a method for spectrum modelling signals - It is used to obtain the filter coefficients and the feature of the signal - Order selection is important in the AR model
	MF	- Match Filtering (MF) is a method that detects a specific pattern of an unknown signal by obtaining the matched filter from predetermined known signals - A higher correlation between the template and the user's command would show better matching - It is effective when the signals have consistent temporal characteristics
	WT	- The Wavelet Transform (WT) provides not only frequency information but also temporal - The Discrete Wavelet Transform (DWT) has a computational advantage compared to the Continuous

		Wavelet Transform (CWT) as wavelets are discretely sampled - WT is suitable for non-stationary signals - However, selecting the appropriate wavelet family is important
Feature Selection	*GA*	- A Genetic Algorithm (GA) is an optimisation procedure to find the most efficient set of features - Requires high computational resources
	SS	- Sequential Selection (SS) is finding the optimal subset of features by adding or removing features sequentially

AutoRegressive (AR) Model

The AutoRegressive (AR) model is a stochastic process which estimates future values based on a weighted sum of past values. In BCIs, EEG signals are assumed as the output of the AR models, where the input is white noise with a mean of zero and a certain variance. Then, the filter coefficients (AR coefficients) are used as the features of the EEG signal as it is assumed that the AR coefficient will be different depending on the mental tasks [6].

$$X_t = \sum_{i=1}^{N} a_i X_{t-i} + \varepsilon_t \qquad \text{Equation 3-1}$$

Equation 3-1 is the definition of AR modelling, where a_i is the AR coefficient, X_t is the assumed EEG signal, N is the order number and ε_t is noise. These N numbers of a_i are used as the feature vectors for classification. In the AR model, choosing an appropriate order is important. If this value is too low or too high, the spectrum is going to be very smooth or it will have excessive peaks [6]. In [110], an order 6 was found to be optimal in imagined speech EEG signals.

However, AR modelling assumes the series X_t is linear and stationary so the performance is not great when the signal is not stationary, especially in an on-line system. For this reason, various adaptive methods have been designed such as the MultiVariate Adaptive AR (MVAAR) model for on-line BCI systems [6, 111].

Common Spatial Pattern (CSP)

The Common Spatial Pattern (CSP) method has been used in various BCIs. One of the first CSP implementations in Motor Imagery BCIs was introduced in [112, 113]. In the following years, much research has been carried out that shows that CSP can be successfully implemented in Motor Imagery BCIs.

CSP is a mathematical technique which separates two classes from multi-channel data [114]. In BCIs, it projects multi-channel EEG data into low dimension space with a projection matrix by using a linear transform [115]. This CSP filter finds optimal variances between classes by maximising one class's variances and minimising the other condition [114].

In this project, the two classes were chosen as matrices X_I and X_H, which represent the idle state and the high tone speech EEG signal after the signal processing procedure.

Matrix *X* has an *S*N*T* dimension, where *S* is the number of sampling points, *N* is the number of electrodes, which is 64 channels, and *T* is the number of trials. The normalised covariance matrix from multiple trials can be described as:

$$R_I = \frac{\sum^T X_I {X_I}^T}{trace\left(\sum^T X_I {X_I}^T\right)}$$

Equation 3-2

$$R_H = \frac{\sum^T X_H {X_H}^T}{trace(\sum^T X_H {X_H}^T)}$$

where X^T denotes the transpose matrix of *X* and trace(*X*) refers to the sum of the diagonal elements of *X*. Let $R = R_I + R_H$ and the whitening transformation, *W* can be calculated by the Ramoser Equation [113] as follows:

$$W = \sqrt{diag(Eval)^{-1}} * Evec^t$$

Equation 3-3

where diag(*X*) is the diagonal matrix of *X*. *Eval* (Eigenvalues) and *Evec* (Eigenvectors) are sorted in descending order. As a consequence, S_I and S_H share common eigenvectors when they are transformed as:

$$S_I = W * R_I * W^T$$
$$S_H = W * R_H * W^T$$

Equation 3-4

therefore, the eigenvector, which has the maximum eigenvalue for S_I, has the minimum eigenvalue for S_H and vice versa as the sum of two corresponding eigenvalues is always 1 [113]. The projection matrix P, which contains spatial filter coefficients, can be described as:

$$P = EV^T * W \qquad \text{Equation 3-5}$$

where EV are the sorted eigenvectors of S_I and S_H in ascending order according to the eigenvalues. The columns of the inverse matrix of P (P^{-1}) are the spatial patterns, which can be used for the neurophysiological mapping projection of the scalp [114]. The first column of P^{-1} represents the largest variance of one class and the smallest of the other and the last column represents the opposite [115]. The original EEG signal can be transformed into E:

$$E = P * X \qquad \text{Equation 3-6}$$

After the spatial filter process, uncorrelated components E (64 according to the samples) were sorted in order to maximise or minimise the variance of one class. Thus, the first and last rows of E represent the maximum difference in the variance between the classes.

Frequency Band Power

Band Power (BP) is one of the simplest spectral domain feature extraction methods in BCIs. In our study, the Fast Fourier Transform (FFT) was applied to the pre-processed EEG signal and separated into multiple pre-defined frequency ranges (the details will be discussed in each experiment-based Chapter). At the next stage, these signals were square powered and averaged within the specific band.

Wavelet Transform (WT)

EEG signal can generally be analysed in the time-domain or frequency-domain. However, the Wavelet Transform (WT) offers time-frequency features, which are very efficient for non-stationary signals such as EEG.

A wavelet is a certain formed waveform, which has a limited duration and an average value of zero [116]. The WT decomposes the signal in the time and frequency domain at multiple resolutions, by shifting wavelets along the signal at various scales [6]. The output of WT is the set of coefficients, which represent the input at the wavelet basis [117].

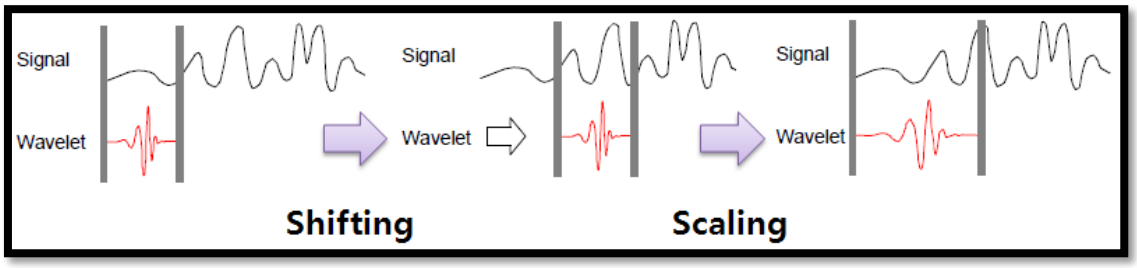


Figure 3.7. Steps of CWT [116, 118].

WT can generally be categorised as Continuous WT (CWT) and Discrete WT (DWT). CWT is the sum of all the times of the signal multiplied by scaled and shifted versions of the wavelet. As can be seen from the Figure 3.7, shifting and scaling simply means moving the wavelet window and stretching or compressing the wavelet. There is a correspondence between the wavelet scale and frequency. The low scale, which is the compressed wavelet, correlates more with high frequencies than with low frequencies. In contrast, the high scale correlates more with the low frequencies of the signal [116]. The mathematical definition of CWT is:

$$w(scale, position) = \int_{-\infty}^{\infty} f(t)\psi(scale, position, t)dt \quad \text{Equation 3-7}$$

where $\psi(scale, position, t)$ is the scaled and shifted version of the mother wavelet $\psi(t)$. CWT generates many wavelet coefficients $w(scale, position)$, which are a function of scale and position [116].

CWT generates enormous numbers of wavelet coefficients as it works at every scale and position. Due to its redundancy and complexity, DWT was introduced.

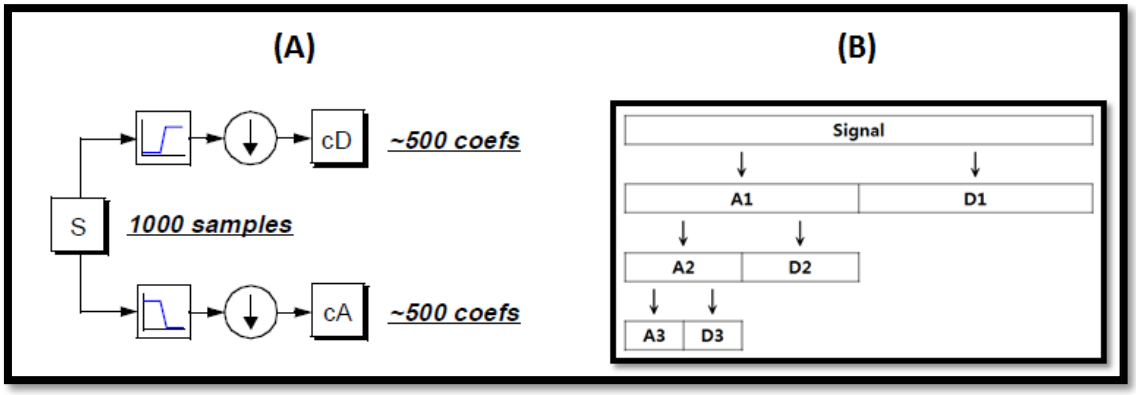


Figure 3.8. (A) DWT filtering process, (B) DWT decomposition tree [116, 118].

The DWT process divides the original signal into approximation (low frequency components) and detail (high frequency components). As can be observed from Figure 3.8, DWT has a filtering process as well as down-sample processes in order to prevent the length of data from increasing twice from the previous level.

There are many different mother wavelets and families. However, finding the suitable wavelet for the specific BCI application is of most importance as the feature coefficients will be significantly varying depending on the choice of wavelets.

Feature Selection

The above feature extraction methods usually produce hundreds of features depending on the number of channels and sample size. Therefore, a feature selection had to be applied in order to reduce the feature set size and class overlap, and to improve computational efficiency. To this end, the Davies-Bouldin Index (DBI) [119, 120] was applied in this thesis. The DBI is a cluster overlap measure. Smaller DBI values indicate a better class separation, with smaller class overlap and a larger distance between classes. DBI is defined as follows:

$$DBI = \frac{S_i + S_j}{M_{ij}} \qquad \text{Equation 3-8}$$

where S_i and S_j is the dispersion of cluster i and j, respectively and M_{ij} is the Euclidean distance between cluster centroids i and j [120, 121]. In our work (two classes), we only require Equation 3-8 for the DBI value calculation. Thus, DBI values for each feature were sorted in an ascending order and the features which had a smaller DBI value than the threshold were selected as a feature set for further classification analysis.

3.9 Classification

BCI systems should be able to classify different mental tasks in order to control machines. However, as it was mentioned in the feature extraction part, EEG signals are noisy, non-stationary and normally have a small training set [109]. For this reason, the selection of an appropriate classifier is important in order to increase the reliability of the system. In [109], various classification techniques for EEG-based BCI systems were reviewed so according to this study, the most common BCI classifiers will be shown and discussed in this section.

The most common BCI classifiers can be categorised as ***Linear classifiers***, ***Neural networks*** and ***Non-linear Bayesian classifiers*** [122]. The linear classifier uses linear functions in order to distinguish classes and these are the most popular algorithms for BCI applications which were also used in this project. Therefore, the LDA and SVM

classifier will briefly be covered in the following section. The neural networks classifier is imitating the neurons nerve system, which has several layers in order to produce nonlinear decision. There are many examples of neural networks such as the Multi-layer perception, Radial basis, Learning vector quantization and Dynamic neural network. The non-linear Bayesian classifier on the other hand, is not common and not fast enough for real-time BCIs but it sometimes shows good classification performance with their non-linear decision boundaries (e.g., Bayes quadratic, Hidden Markov Model) [122].

Linear Discriminant Analysis (LDA)

Linear classifiers use linear functions in order to distinguish between multiple classes and the Linear Discriminant Analysis (LDA) is one of the most popular linear classifiers in BCI systems.

As the dimensionality reduction is one of the objectives of the LDA process, it provides a simpler classification in multiple dimensions of data. LDA uses a linear hyperplane, which is a projection line which maximises the separation between the classes (by increasing the distance between the mean of the two classes) and minimises their overlap (by decreasing the variance within the classes) [109]. The linear discrimination is based on the function:

$$g(x) = w^T x + w_0 \qquad \text{Equation 3-9}$$

where x is the sample to discriminate, w is the weight matrix and w_0 is the bias or threshold [93].

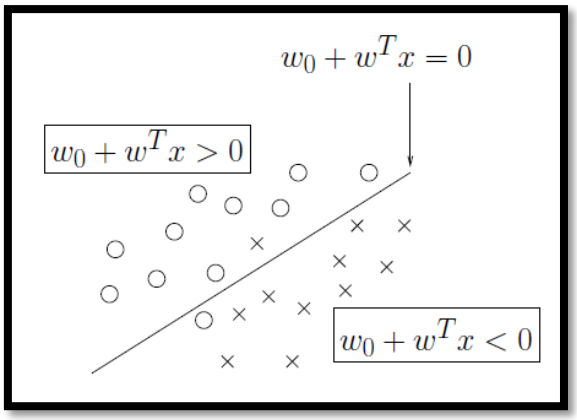


Figure 3.9. LDA hyperplane separating two classes [109].

The advantage of LDA is its simplicity and low computational power, which allows the LDA to be successfully used in online BCI systems [109].

Support Vector Machine (SVM)

The Support Vector Machine (SVM) is another popular linear classifier in BCIs. It also uses a discriminant hyperplane in order to separate classes. However, the SVM chooses the linear line based on the maximisation of the margin between the classes [109].

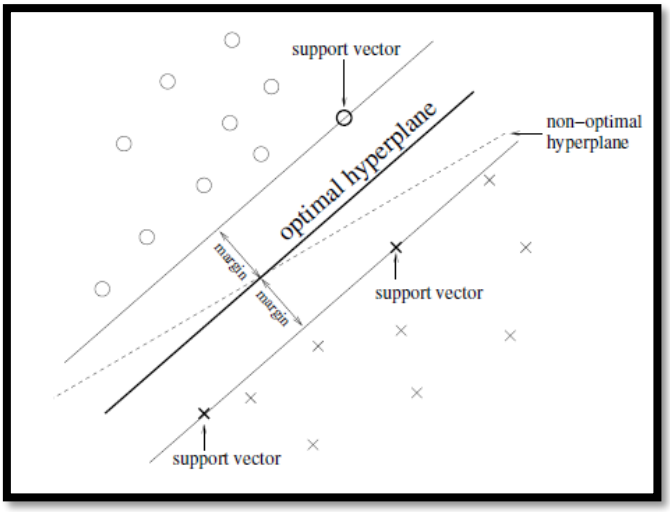


Figure 3.10. SVM hyperplane maximising the margin between classes [109].

As can be seen from Figure 3.10, the set of vectors is optimally separated by the SVM hyperplane if the distance between the closest vectors and the hyperplane is maximal. However, this case only applies when the training data are linearly separable but in most cases they are not. As a solution to this problem, the SVM estimates the noise in the data depending on prior knowledge [123].

3.10 Performance Assessment

There are many common performance evaluation measures for BCIs such as the Area Under the Curve (AUC), Receiver Operating characteristic Curves (ROC) and Mean Squared prediction Error (MSE). However, these measures are not taking into account the system operation time, especially the idle period length, which is a very

important and critical evaluation criterion for a self-paced onset detection system. Therefore, they will not be covered in this thesis.

In this section, the confusion matrix performance evaluation measures will be briefly described as background knowledge for the better understanding of this thesis. In addition, the mutual information and information transfer rate, which are very common evaluation systems will be discussed, along with how their issues can be applied to online self-paced systems. Based on the issues of the current evaluation systems, our novel performance evaluation metric will be proposed and discussed in Chapter 6 - **Performance Assessment Score**.

Confusion Matrix

A confusion matrix is a table layout that is used to visualise the performance of a classification model. In the common BCI two-class problem, classes are often labelled as 'target' and 'non-target'. In this case, the True-Positive (TP), True-Negative (TN), False-Positive (FP) and False-Negative (FN) cases can be described as follows [124]:

TP: targets that were correctly classified.

TN: non-targets that were correctly classified.

FP: non-targets that were incorrectly classified as targets.

FN: targets that were incorrectly classified as non-targets.

The TP rate (***TP / TP + FN***) and the FP rate (***FP / FP + TN***) are usually used in BCI systems. They show the proportion of positive items that were correctly classified and the error rate of false alarms, respectively. In addition, the classification accuracy (***TP + TN / TP + FN + FP + TN***) is a common performance measure in BCIs that shows the percentage of the total correctly classified patterns.

		Predicted condition	
	Total population	Predicted Condition positive	Predicted Condition negative
True condition	condition positive	**True positive**	**False Negative** (Type II error)
	condition negative	**False Positive** (Type I error)	**True negative**

Figure 3.11. Confusion matrix [125].

Mutual Information & Information Transfer Rate (ITR)

Mutual information finds out how many bits of information are effectively sent on average when the classifier makes decisions about all possible situations of the inputs. In BCIs, the formula is defined as [124]:

$$B(bits) = \log_2 N + P \log_2 P + (1-P) \log_2 \left(\frac{1-P}{N-1}\right) \quad \text{Equation 3-10}$$

where *N* is the number of classes and *P* is the accuracy. However, this formula holds under the following conditions [126]:

- *N* selection of classes is possible and the classes have the same probability.
- The accuracy is the same for each class.
- Errors of the classifier must have the same probability for each class.

This mutual information does not take into account the speed of the system. Therefore, the Information Transfer Rate (ITR) is used more in BCI systems. It simply divides the number of bits *B* by time *T* (either in minutes or seconds), which is the average time of a single trial ($ITR = B/T$).

However, there is one major issue which concerns the mutual information evaluation of online self-paced BCI systems. As can be observed from Figure 3.12, the performance *B* (in bits) presented in the y axis shows the highest performance results not only when it has the highest number of true-positives with 0 false-positives but also with 0 true-positives when it has the highest number of false-positives, which should give the lowest performance results. This happens because if the 100% false output is inversed manually by the user, it gives a 100% correct output. However, in online self-paced systems, users would not be able to know whether the system gives a false output or a correct output. This will cause a problem in the case where the performance output shows results of 80% accuracy, when they actually have a 20% accuracy. In addition to this problem, mutual information and ITR measures are not taking into account the idle period length, which is an important assessment criterion for self-paced onset detection systems. For these reasons, we proposed a novel assessment score metric for self-paced BCIs and it will be discussed in Chapter 6 **- Performance Assessment Score**.

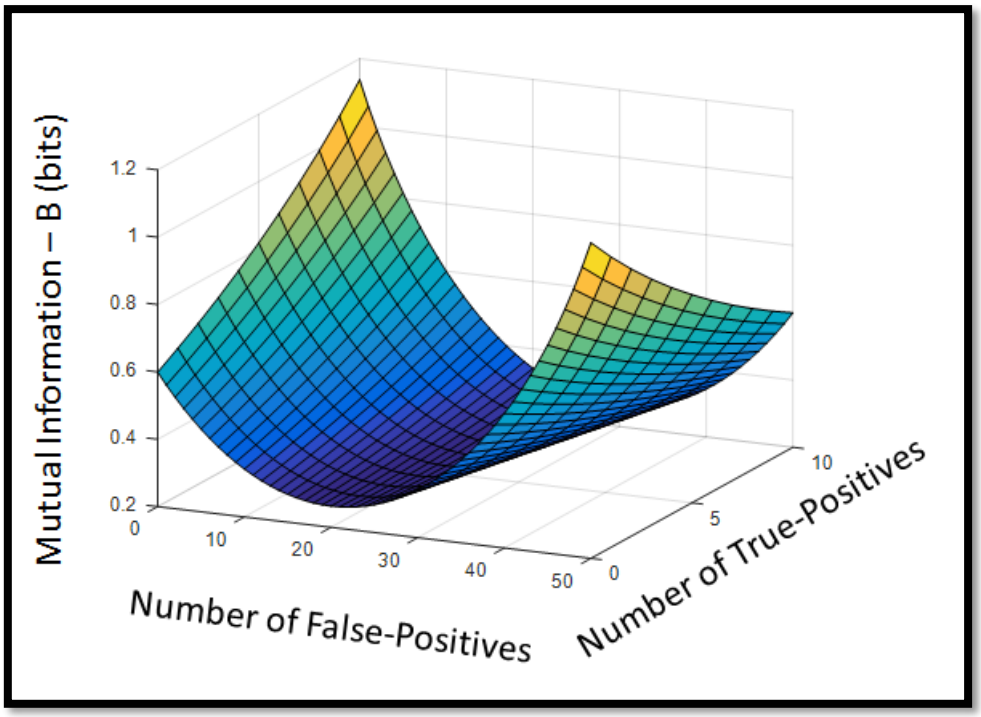


Figure 3.12. Mutual information performance graph based on various number of false-positives and true-positives. $B(bits) = \log_2 N + P\log_2(P+\alpha) + (1-P)\log_2\left(\frac{1-P+\alpha}{N-1}\right)$, α = 0.1 was added to P (accuracy) in order to prevent $\log_2(0)$.

3.11 Summary of Existing Challenges of Onset Detection

In this section, the existing challenges of onset detection will briefly be summarised as a conclusion of the introductory sections and before we move to the next chapters (experiment-based).

BCIs have rapidly been developed in the last few years but they are mostly restricted only in laboratory settings due to the lack of practicality and usability which is derived by the cue-based approach. Therefore, self-paced modality is more suitable

for applying BCIs to real-world uses. However, there exists a great challenge for SP-BCIs. They are more complicated to analyse than the CB-BCIs as they have lack of knowledge about the precise time location of the user's command. The user's control intention and timing are usually unknown to the machine [5, 7]. SP-BCIs should therefore continuously analyse the ongoing brain activity and they should be able to distinguish between the Intentional-Control (IC) and Non-Control (NC) states (called the onset detection problem).

Based on our literature review, the motor imagery task was mostly used (e.g. [7, 8, 10, 11]) for onset detection. However, one crucial issue emerges if this is used outside the indoor laboratory settings. The motor imagery mental procedure largely overlaps with other common spontaneous cognitive states. For example, the classifier would not be possible to identify whether the onset detection was from the actual command or from other daily-life gestures such as shaking hands, grabbing an object or cycling. This is the main challenge in the current onset detection system.

For this reason, we suggested sound-production related cognitive tasks (sound imagery) for the onset detection method. The chances of IC false positives also need to be reduced, but this can be addressed by choosing cognitive tasks that do not significantly overlap with other common, spontaneous and frequent cognitive states [4]. The use of specific words/syllables/letters for onset detection would likely increase both the onset false positives as well as the task-related false negatives due to the large overlap with the continuous internal speech in normal thought processes. As a result, we have chosen imagining a high tone or siren-like sound production tasks to be onset switches, both of which are unlikely to overlap with normal thought processes.

Another big issue in the current onset detection system is its low performance. This is mainly because of the variety of the idle state mental pattern. Task vs. task is usually easier as there are two different clear patterns which can simply be classified. However, the idle state has no clear spectral/spatial characteristic, therefore it makes it difficult to be classified. In order to address this issue, some papers (e.g. [90]) defined the idle state as a relaxed state with the eyes closed (alpha wave generation). However, it is not an appropriate approach if it is applied to a real-world system. In this thesis, the effect of the variety of idle states has not been investigated, however the sound imagery onset detection achieved significantly better results in terms of performance than the motor imagery task.

4 A Novel Technique of EMG Artefacts Contaminated EEG Channel Selection and Processing (Based on a paper [2])

4.1 Introduction

This chapter is based on a journal paper (submitted / under review) [2].

As it was discussed in Chapter 1, artefact handling is an important challenge for the onset detection system as EEG signals usually have low signal-to-noise ratios and may contain electromyography (EMG) and/or electrooculography (EOG) artefacts [6]. In the onset detection study in particular, users stay calm and relaxed during the idle period and then they execute the task for the onset. During this procedure, it is likely that the participants may unintentionally generate some artefacts. Therefore, it is important to handle these artefacts in advance in order to ensure that the system performs purely based on the cognitive tasks and not the artefacts.

However, previous BCI and EEG studies applied artefact handling techniques to all EEG channels. As such, e.g., in the case of artefact removal with the use of a blind source separation, a common approach in BCIs, there may be significant loss of useful EMG-free EEG information [19-22]. In order to minimise this information loss, we

propose an algorithm for the rejection of any EEG channels that are contaminated by class-related EMG artefacts, and which have called EMG-CChs. Typical EMG handling techniques such as ICA, PCA and BSS-CCA were only applied to the EMG-CCh and not to all channels in this study. This combined approach was then compared to other existing methods without doing any channel elimination, i.e., methods that apply EMG artefact removal to all EEG channels. The comparisons were made against our own onset detection data as well as against a BCI competition data set.

In our data set, the comparison results against the BSS artefact removal approach, which was applied in two ways, one to all channels and one only to EMG-CCh, showed that ICA, PCA and BSS-CCA can yield a significantly better ($p<0.05$) class separation with the proposed method (79% of the cases for ICA, 53% for PCA and 11% for BSS-CCA). In the BCI competition data, we saw an improvement in 60% of the cases for ICA and BSS-CCA.

The aim of this chapter is to describe a novel EMG artefact contaminated EEG channel selection and handling technique which shows advantages over the common blind source separation EMG handling methods. There are no existing methods for the removal of EMG artefacts based on the correlation between EEG and EMG channels with statistical testing. In addition, the EMG-CCh selection can be used on its own or it can be combined with pre-existing artefact handling methods. For these reasons, we believe that this method can be useful for other EEG studies (e.g. BCIs, brain mapping, and clinical areas).

4.2 Methodology

Experimental Paradigm and Data Set Description

In this paper two different experiment data sets were tested. One set was from our covert and inhibited overt sound-production onset detection study [1, 5] and the other was from BCI competition IV data set 2a [23]. Both sets are discussed below.

i) Sound-production Related Cognitive Tasks for Onset Detection:

In that study, four different mental tasks were tested for the onset detection: Inhibited-overt high tone; inhibited-overt siren-like sound; covert high tone; and covert siren-like sound. Inhibited-overt sound production involves tensioning of the vocal cords but there is no actual sound production that can clearly be heard. On the other hand, covert production is a pure imagination process which include covertly making the sound as well as imagining hearing the sound (auditory imagery) [1, 5]. These four-different sound production cognitive tasks were classified against the idle (a.k.a. non-control or null) state as an onset switch.

The study had seven subjects and they all completed four different mental tasks. Thus, the total number of runs was 28 (7 participants * 4 tasks). Each task run had 30 trials. In a single trial, there were 3-30 seconds of idle state followed by 3 seconds of an intentional command (IC, i.e., one of the four cognitive tasks) state. The idle state length was freely chosen by the user as the study was focused on self-paced activation of the BCI system. The recorded EEG data were segmented using a 0.5s window without overlapping and were then separated into idle and task state categories. If the 0.5s window happened to include both idle and task states, it was discarded from the analysis. These separated idle and task segments were pre-processed and applied with

various feature extraction method. Then, Davies-Bouldin index (DBI [119, 120]) was calculated to test class separation. There were 64 electrodes placed based on 10-10 layout system with 2 reference channels on both left and right earlobes. In addition, 4 facial electrodes were placed to detect EOG and EMG artefacts (more details of the setup will be discussed later in section B below). A Biosemi ActiveTwo system was used with 512 sample/s.

ii) BCI Competition Data:

The BCI competition IV data set 2a was recorded from 9 participants to classify four different motor imagery tasks in a cue-based BCI study [23]. In order to simulate a self-paced scenario using these data, we only selected specific time segments (Figure 4.1). Figure 4.1 represents their timing scheme of a single trial for the competition data. The first 0-2s segment `was the preparation state (fixation cross) that corresponded to the idle state in our data. However, there was a short acoustic warning tone at 0s, which could produce event-related potentials. In addition, the length of fixation cross state is always 2 seconds, which participants can anticipate. Thus, we defined this state as ***non-specific / expectation state (NE state)***. However, the NE state would be expected to be somewhat similar to that preceding a self-paced task event. For this reason, we treated the NE state as corresponding to the idle state. Caution is needed when analysing results stemming from this, but the choice is still suitable for this study as the data set was chosen merely to extend our analysis to include an existing and well accepted data set. Also, this data set was chosen as it was the only BCI competition set that includes facial channels, which were needed for a suitable comparison with results from our data.

The four-different motor-imagery tasks were regarded as being in the ***Task state*** class. To minimise ERP-related effects, we only used the 4-6s segment for the Task state.

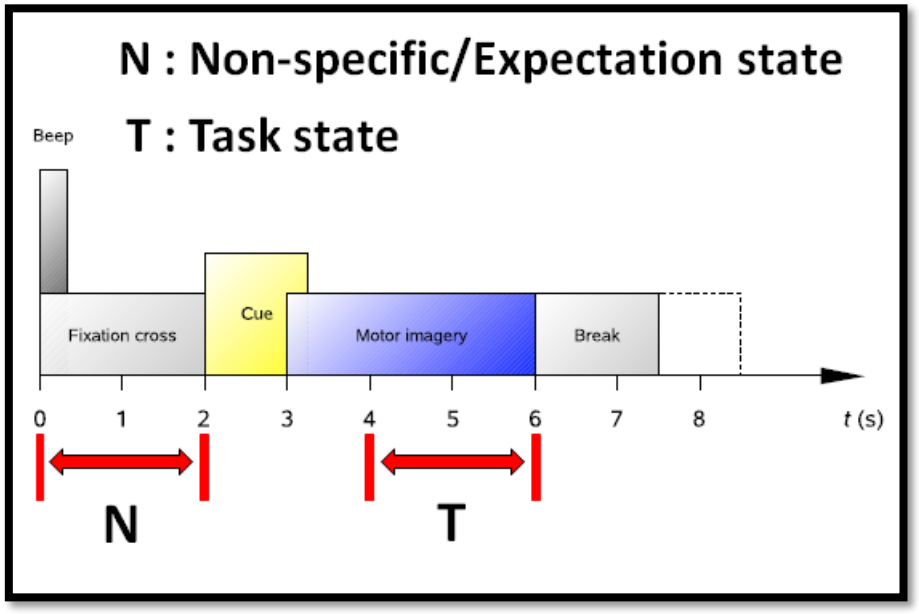


Figure 4.1. Timing scheme of a single trial, BCI competition data (modified from [23]). N (null) and T (task) represent selected areas that were used in the present study.

EMG Channel Placement

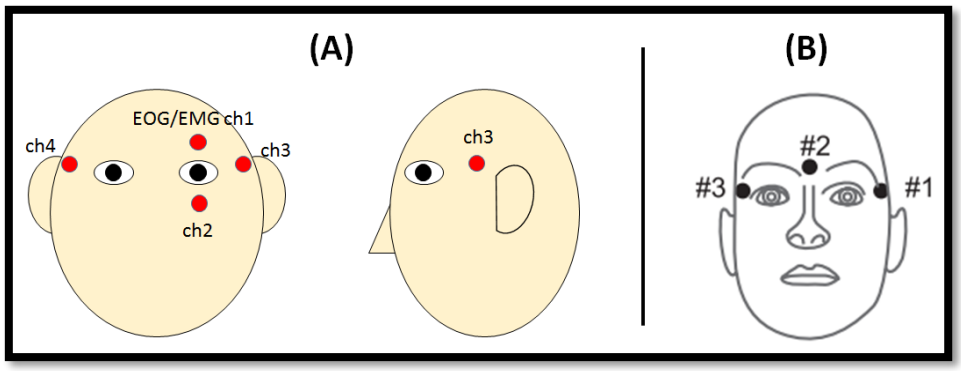


Figure 4.2. (A): Four facial EOG/EMG electrodes placement for our onset detection system. (B): Three facial electrode channels for BCI competition data set **[23]**.

Figure 4.2 displays four EMG channel locations. This placement was based on [127]. **EMG_ch1** was located above the corrugator muscle and it was used mainly to detect forehead EMG and eye blink artefacts. **EMG_ch2** was set above the levator labii and zygomaticus muscles to detect facial twitches and upper limbs artefacts as well as up/down EOG, conjointly with **EMG_ch1**. **EMG_ch3** and **EMG_ch4** were placed around the anterior-most edge of the temporalis muscle to detect left/right EOG and temporal EMG artefacts.

Spectral Domain EMG Artefact Content

Spectral domain EMG artefacts were analysed to find out whether these artefacts affect class separation and, if so, in which frequency range. This analysis was done with our self-paced onset detection data set [1, 5]. In the analysis, the frequency bands were separated into eight different ranges (Table 4.1).

Table 4.1. The eight different frequency bands used in this study.

Name	**Freq1**	**Freq2**	**Freq3**	**Freq4**
Frequency Range	2-4 Hz (Delta)	4-8 Hz (Theta)	8-12 Hz (Alpha)	12-16 Hz (Low Beta)
Name	**Freq5**	**Freq6**	**Freq7**	**Freq8**
Frequency Range	16-20 Hz (Beta)	20-30 Hz (High Beta)	30-42 Hz (Low Gamma)	42-100 Hz (High Gamma)

Figure 4.3 illustrates the absolute Pearson correlation coefficient (CRC) values between 64 EEG channels and 4 EMG channels for each of the eight frequency bands.

The CRC values shown were averaged across all subjects, all four mental tasks, and all trials, for illustration purposes. Panel (A) shows data from the idle state while panel (B) shows data from the IC task state. Also, all data was submitted to the eye blink and EOG artefact removal procedure (explained below). Thus, the data shown contains only EEG and EMG artefacts. Red and orange areas indicate EEG channels that are highly correlated with EMG channels, i.e., they had high contamination by EMG artefacts. As expected, the plots show that frontal areas have EMG contamination as seen by the high correlation with EMG_ch2 throughout all the frequency bands. On the other hand, EMG_ch3 and EMG_ch4 show high correlation with temporal area EEG.

In [128] it was shown that EMG can contaminate all EEG bands and its contamination level differs with scalp location. We observe a similar pattern in our data. In addition, EMG contamination for the idle (panel A) and IC (panel B) states shows quite similar results for both cases. This indicates that EMG artefacts can appear in the idle state as well as in IC states. I.e., **EMG contamination per se may not be idle vs. IC class-dependent**. For this reason, we statistically checked which channel locations were affected by EMG artefacts during IC states more than during idle states. Such channels (i.e., EMG-CCh) were then removed from the analysis and the existing EMG removal techniques were applied only to the remaining channels. Thus, the system's performance result will be based on the only with the mental tasks but artefacts. The details of the procedure for EMG-CCh selection method will be discussed in section '*EMG Artefacts Channel Selection'* below.

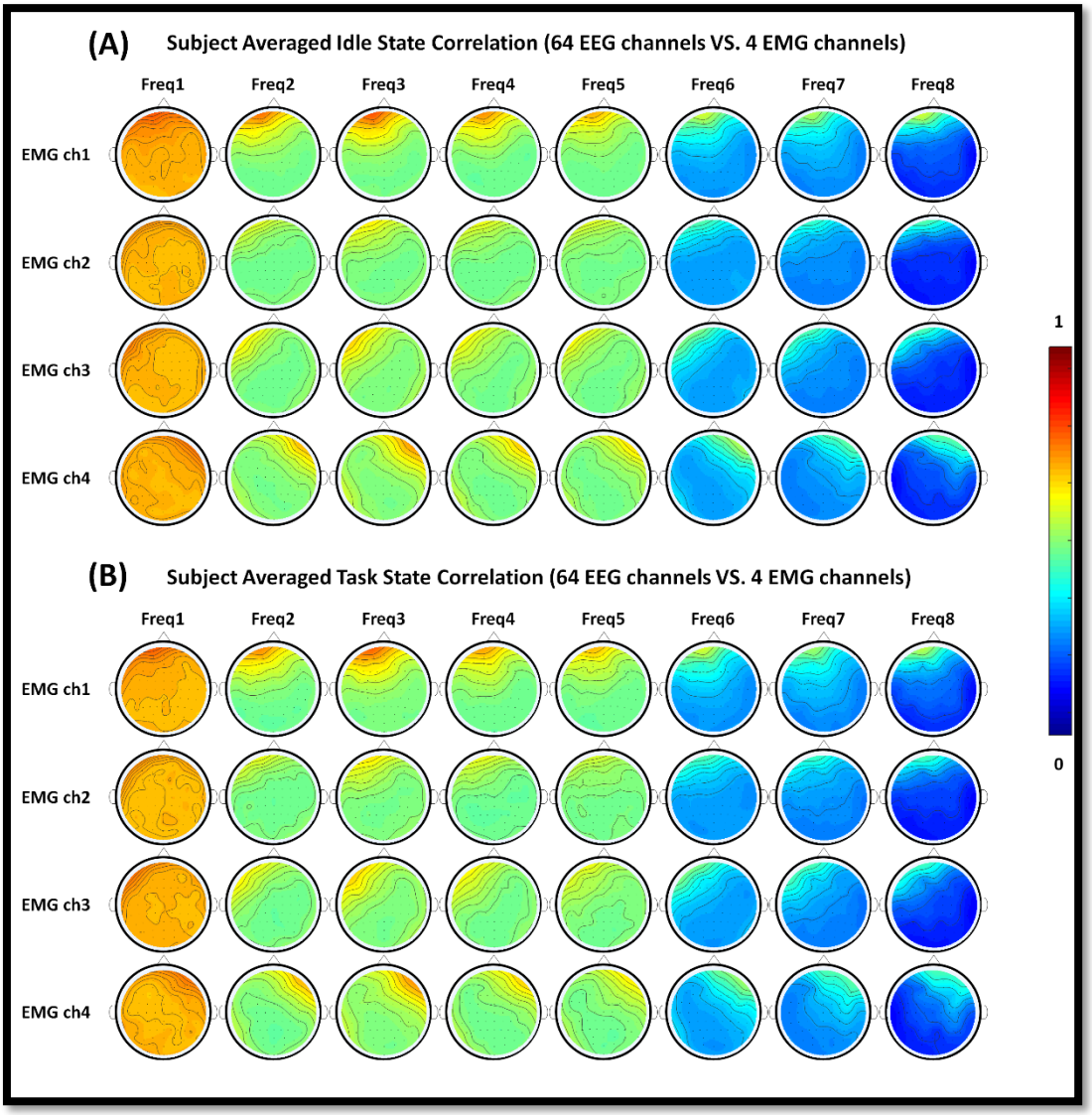


Figure 4.3. (A): Topographic map of absolute Pearson correlation coefficients between 64 scalp EEG electrodes and the 4 EMG channels shown in Figure 4.2. The correlation values in each channel were averaged over all subjects, four onset tasks and all trials. Panel (A) represents during idle state while (B) shows during IC task state. Freq1-Freq8 ranges from Table 4.1 [2].

EOG Artefacts Removal

EOG artefacts can be separated into two different types. One is eye blink and the other is eyeball rolling. Eye blink artefacts can easily be detected as they have relatively high signal amplitude compared to EEG. However, eyeball rolling EOG is somewhat

different. Thus, in this paper, eye blink EOG artefacts were rejected with a discrete wavelet transform (DWT) denoising method while eyeball rolling artefacts were handled with the EMG & EOG contaminated channel selection method.

The window segments that contained the eye blink artefacts were automatically rejected based on the method shown in [108, 129]. DWT was applied with Haar mother wavelet as it resembles eye blink artefacts. The signal was decomposed up to level 6 and was then, reconstructed using only the approximation coefficients, as in [108, 129]. If the reconstructed signal's standard deviation (std) was higher than 3 times the preceding data's std (using the previous 500ms window), then this data segment (all channels) was regarded as eye blink artefact and was rejected from further analysis.

EMG Artefacts Channel Selection

The method described in this section comprises the main novelty in our study. The technique can be applied as an EMG-contaminated channel rejection method on its own, or it can be combined with other EMG handling algorithms such as ICA, PCA, BSS-CCA. By applying the latter EMG handling methods only to the EMG contaminated channels, useful signal information loss can be reduced.

Figure 4.4 shows the EMG-artefact contaminated channel selection procedure:

1) Calculate the absolute Pearson correlation (CRC) values between 64 EEG channels and EMG_ch1 for each idle and task state. This will generate *64*NI window segments* idle state correlation values and *64*NT window segments* task state correlation values, where *NI* and *NT* are the numbers of idle and task trials, respectively.

2) Calculate the Wilcoxon-test *p*-values between the two lists (*NI* and *NT*) for each of the 64 channels.
3) If an EEG channel's EMG correlation value for the task state (average of *M* correlation values) is higher than for the idle state (average of *N* correlation values), AND the difference is statistically significant ($p<0.05$), THEN this channel was selected as an EMG contaminated channel, EMG-CCh.
4) Repeat steps 1 to 3 with EMG ch_2, EMG_ch3 and EMG_ch4.

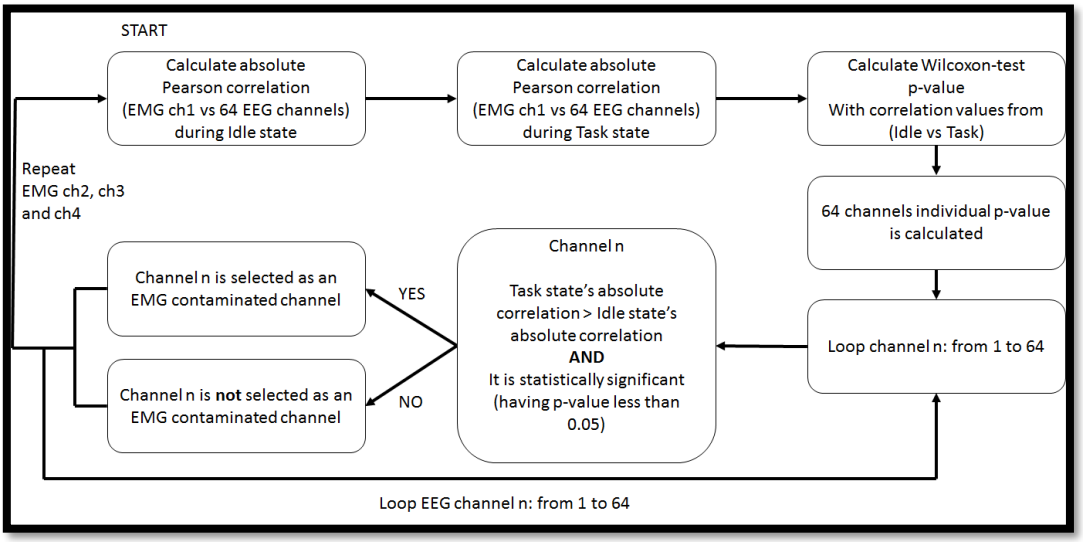


Figure 4.4. EMG artefacts contaminated channel selection procedure [2].

Rationale of the procedure:

The aim of the method is to select EMG-artefact contaminated channels which affect classification (idle vs. IC) results. Participants were instructed to stay calm and relaxed in the idle states. However, when they executed a task state, unexpected EMG artefacts (e.g., facial twitches and eye movements) can contaminate EEG, especially if they are unfamiliar with BCI experiments, even though they were instructed to avoid

any muscle movement. Thus, it is difficult to identify whether the result is purely EEG-based or it is EMG-contaminated. Many other blind source separation EMG removal methods have been suggested to deal with this issue. However, as existing methods are applied all EEG channels, it is possible that important EEG information is being lost in the process. For this reason, our EMG contaminated channel selection method calculates correlations between EMG channels and EEG channels, and contaminated channels are removed before applying other methods, as follows:

- IF the EMG vs. EEG correlation was greater during task states than during idle states, which indicates that there was potentially class-dependent EMG contamination during task execution, AND
- IF the correlation increase was statistically significant,
- THEN the EEG channel can be regarded as having class-dependent EMG contaminated and is therefore removed.

Possible limitation in our method: It is possible that a channel with lower CRC in idle states could still have too much class-dependent EMG. To test whether this may have been the case, a reliability analysis was performed (see section *Reliability of the EMG-CCh Selection Method* below).

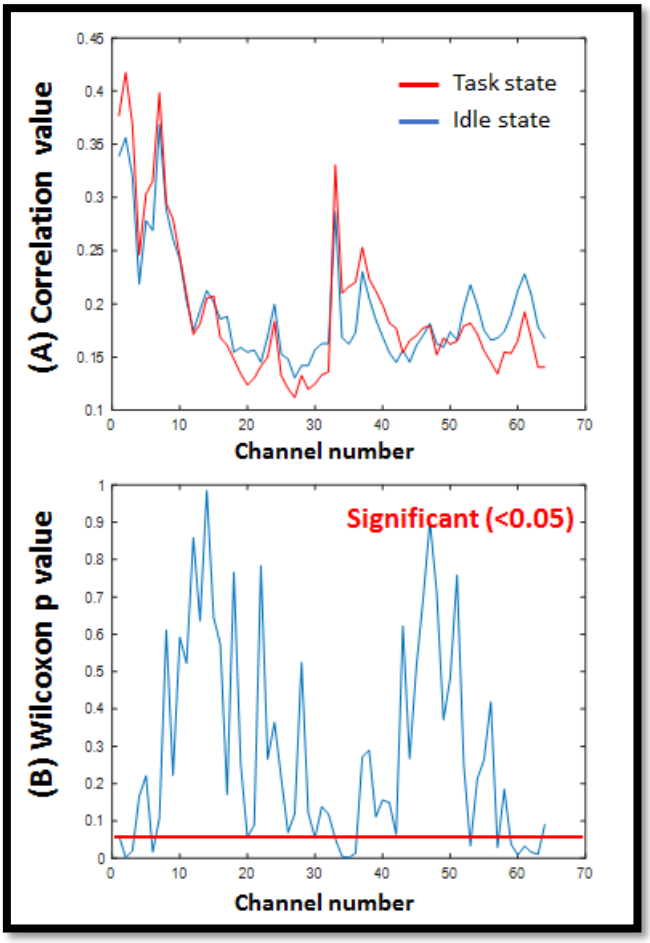


Figure 4.5. An example of EMG-CCh selection procedure for Participant 1's inhibited overt siren task. Figure (A) represents correlation values from 64 scalp EEG channels vs. EMG ch1. Figure (B) shows Wilcoxon test p value between idle and task state's correlation values [2].

This EMG-CCh selection procedure is quite strict, unbiased and reduce valuable data loss when compared with typical thresholding methods presented in [130]. Figure 4.5 shows an example of the proposed EMG-CCh selection process. Figure 4.5A represents correlation values from 64 channels for each idle and task state (averaged from number of window segments). Channel number 35, for example, shows task state's EMG correlation is around 0.23, i.e., it is unlikely to be correlated with EMG artefacts if we cut it based on a typical thresholding method. However, if it is statistically compared with idle state by using the *p*-value from Figure 4.5B, it is certain

that the task state's EMG correlation is significantly higher than in idle state. This indicates that channel number 35 is affected by EMG during the task activation. Thus, this channel has to be processed with further EMG handling methods to remove class-dependent EMG related performance results. Wilcoxon method was used for the statistical test as the correlations are non-Gaussian, so it is suitable for our handling method.

The selected EMG-CChs will be shown separately for our onset detection data set and for the BCI competition data set, respectively.

- **Sound-production Related Cognitive Tasks for Onset Detection:**

Figure 4.6 shows the artefact-contaminated areas from the EMG-CCh selection method for each participant and task. The red area represents the EMG-CCh. Some cases, only a couple of or none of channels were selected. On the other hand, 50 channels were selected as EMG-CCh in the worst case (participant 1 C_Siren task). The EMG affected channels are task-dependent and it are not consistent between subjects. It could be due to the fact that the level of tasks activation depends on the subject and on the mental task. However, this strict EMG handling processing ensures that all the remaining analysis on EEG- based class-dependent information.

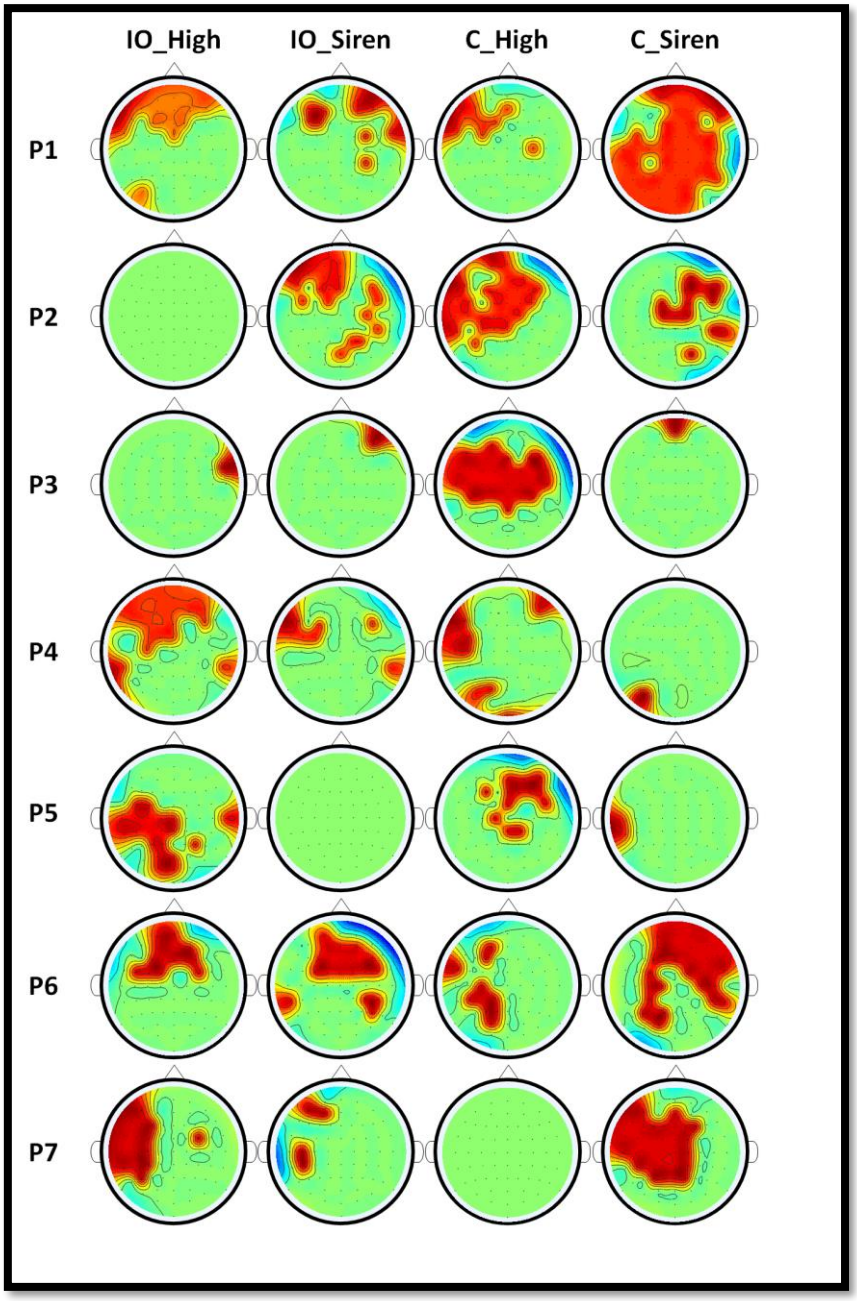


Figure 4.6. EMG contaminated channel area that selected from the EMG-CCh channel selection method. Red area represents the selected channels, which EMG artefacts are contaminated, while green is normal channels. The orange and blue area has no meaning. It is simply caused from drawing algorithm [2].

- **BCI Competition Data**

This sub-section deals with data from the BCI competition set with 6 participants. The default channel placement can be found in [23]. Participant 1 had only channel 13

selected as an EMG-CCh. Participant 4 had fifteen EMG-CChs and the uncontaminated channels were ch14, 15, 16, 19, 20, 21 and 22. Participant 6 had uncontaminated channels ch19 and 22 and the remaining channels were EMG-CCh. Participant 7 had only five EMG-CChs, i.e., ch8, 12, 13, 17 and 18. Participant 9 had seventeen EMG-CChs and the uncontaminated channels were ch17, 18, 20, 21 and 22.

Participant 2 and 3 had all the twenty-two channel selected as EMG-CCh. This means that their data were so contaminated by EMG as to preclude their use. Participants 5 and 8 had no EMG-CChs. Thus, these four subjects' data were not used for further analysis as this did not allow a comparison between typical EMG handling techniques and our method.

EMG Artefacts Handling

After the EMG-CCh selection process, various common EMG artefact handling methods were tested.

- ***Simple Channel Rejection*:** EMG contaminated channels were simply eliminated from further analysis.
- ***Blind Source Separation Canonical Correlation Analysis (BSS-CCA)*:** CCA measures the linear relationship between two multi-dimensional signals [131]. This method can be used to solve BSS problems (proposed in [132]) by taking the EEG multi-channel source vector as the first multidimensional variable and temporally delayed EEG source vectors as the second variable [97]. In this experiment, the threshold for the autocorrelation coefficient ρ was chosen as 0.35 based on the study in [133]. Thus, EEG sources that had $\rho < 0.35$ were removed. If there is no

source that has less than the threshold ρ value, the last source that has the lowest autocorrelation coefficient was removed.

- ***Independent Component Analysis (ICA)*:** ICA is also a widely used technique to solve the BSS problem. It separates EEG sources into multi components, which maximizes their mutual independence [99]. By removing the artefact ICs, it produces EMG free signal. However, it normally requires manual/visual inspection to check artefact ICs. In this paper, automatic artefact ICs were detected using Kurtosis and Entropy, which was suggested in [102, 103]. Both were computed for all the ICs and were normalized to a 0 mean and 1 standard deviation. The threshold was set at ±1.64 (based on [102]). If the IC exceeded the threshold, it was regarded as an artefact component and was removed.
- ***Principal Component Analysis (PCA)*:** PCA is a statistical procedure that identifies patterns in high dimension data by finding their similarities and differences [104]. It uses linear transformation and it generates a set of principal components (Eigenvalues). The higher value of the PC represents strong correlation between multiple signals (e.g., channels or trials) and the smaller value of PCs can be regarded as a noise component. In this study PCs that accounted for 95% of the total power were used (based on [134]) and the remaining PCs were removed.

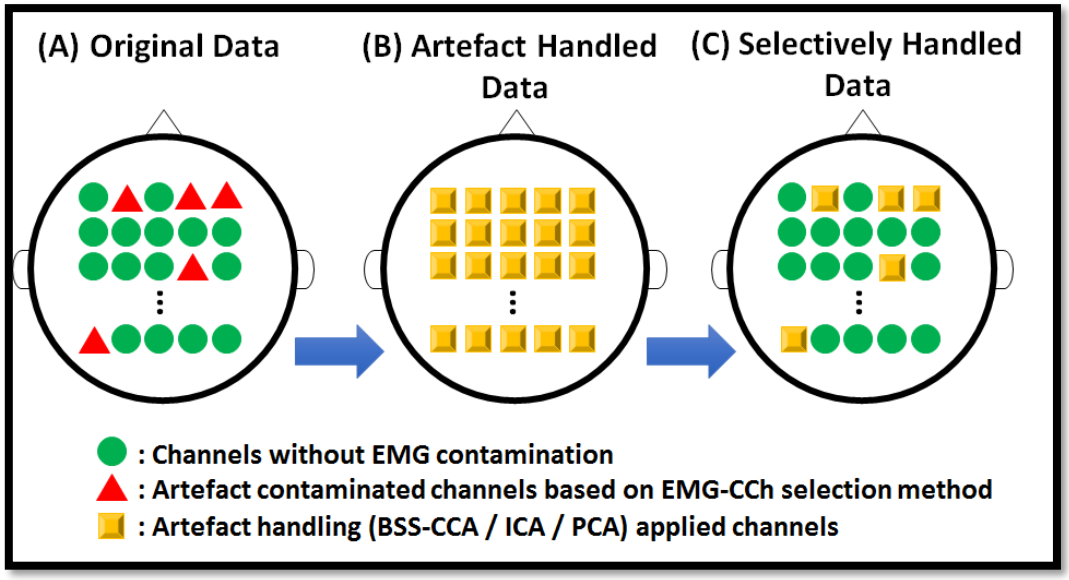


Figure 4.7. Examples of artefact-contaminated channel handling procedure [2].

These typical artefact handling techniques in BCIs were only applied to the EMG-CCh selected in the earlier stages of the analysis. Figure 4.7 shows an example of how to selectively apply EMG handling methods to those EMG-CCh. Panel (A) shows the original data. The green circles represent normal channels that have no class separation effect from artefacts, based on the statistical test. The red triangles depict channels with EMG/EOG artefact contamination. The artefact handling methods (BSS-CCA, ICA or PCA) were applied to all the channels (64 for our onset detection data and 22 for the BCI competition data). This is represented in panel (B) with yellow squares. Panel (C) shows the channel data that were used for the final analysis. Channels without artefact contamination were kept and only the EMG-CChs were selected from (B). BSS-CCA, ICA and PCA separate *N* number of source components if they have *N* number of channels. Thus, applying these techniques only to EMG-CChs, which sometimes could be in very low numbers of channels, would not be suitable to separate artefact components (and it is impossible to do it if the number of EMG-CChs

is 1). For this reason, artefact handling processes were applied to all channels first and then selectively taking channels if they were EMG-CCh. As a result, panel (C) has EMG free data (i.e., statistically no artefact-related class-dependent effects), which were then used for further analysis.

Feature Extraction

The artefact-free data were submitted to four separate feature extraction methods, as follows:

- ***Autoregressive Model (AR)*:** AR model extraction was applied to all channels and the obtained coefficients were used as features. For the coefficient estimation, Burg's method [135] was used. In [12], AR model order number 6 was suggested as optimal for imagined speech EEG signal, so order 6 was chosen.
- ***Band Power (BP)*:** Fast Fourier Transform (FFT) was applied to the pre-processed signals with seven different frequency ranges (Freq2 to Freq8 as described in *'Spectral Domain EMG Artefact Analysis' above*). Freq1 (Delta band) was removed as it highly correlates with EMG channels (as seen in Figure 4.3). Each bands FFT was squared and these were used as features.
- ***Common Spatial Pattern (CSP)*** [115]**:** Using our data idle and sound-production related states were used as the two separate classes. In BCI competition data set, on the other hand, four motor-imagery tasks were regarded as one class and non-specific states for the other class. After the spatial filter process, EEG source components were sorted to maximise

variance for one class and to minimise it for the other class. CSP makes the first and last components represent the maximum variance difference between classes. Thus, the first three and the last three EEG source components were found. Linear regression was applied and its slope was used as a feature.

- ***Discrete Wavelet Transform (DWT)*:** Pre-processed data were decomposed up to level 7 for sound-production related data and up to level 6 for the BCI competition data (due to different sampling rate). Then, the coefficients for the detail components, which represent pseudo frequency bands 4-8Hz, 8-16Hz, 16-32Hz, 32-64Hz and 64-128Hz (up to 100Hz as it was bandpass filtered), were obtained and the variances of the coefficients in each band were used as features. The mother wavelet 'db2', which has just 4 coefficients, was chosen because of its simplicity and because it is commonly used in BCI studies. In our previous study ([4]) we also showed that the choice of wavelet type (db2, coif2 and sym2) did not have a significant effect in onset detection in our (covert) sound-production scenario.

These features were used for DBI calculations for evaluation purposes (details are given below). This DBI value can be different depending on which feature domain is analysed. It is for this reason that various feature extraction methods commonly used in BCIs were tested in order to increase the reliability of the evaluation. In addition, the choice of the feature extraction methods listed above covers the time, frequency, spatial and time-frequency domains.

Evaluation and Feature Selection

In order to evaluate how the EMG-CCh selection and handling method improves class separation compared to the common EMG handling techniques, the Davies-Bouldin Index (DBI) was used. DBI is a cluster separation measure that was suggested in [119, 120]. It measures the average similarity between each data cluster. Lower DBI values indicates better class separation [136]. Note that in our case there is no need for clustering per se as the clusters are simply the classes (idle vs. task states) we which to detect. In this paper, applying artefact handling methods (***BSS-CCA, ICA and PCA***) to all channels will be compared with applying these methods to the selected EMG-CCh only.

The extracted features were sorted in ascending order based on DBI values. The features that had smaller DBI values (i.e., better class separation) were used for further evaluation. In BCIs, as in most human-machine interaction systems, the minimum number of features is recommended to reduce computational loads. Thus, three sets based on a) the smallest DBI, b) smallest 1%, and c) smallest 5% were used for evaluation.

These DBI values between from artefact handling applied to all channels were compared to those from selective EMG-CChs using the Wilcoxon test to compare the means and a *t-test* to estimate statistical power.

Reliability of the EMG-CCh Selection Method

Two questions must be asked concerning the proposed EMG-CCh selection method. First, is it actually selecting the EMG-contaminated channels that have class-

dependent EMG data? And, secondly, is it strict enough to be used? In order to ensure its reliability, we have gone through extra testing by statistically comparing ICA-processed EMG-free data with our suggested method using both our data and the BCI competition data set. The testing procedure is as follows: Figure 4.8 shows how to separate channel data for the test. The testing method will be explained based on the example figure. Firstly, ICA was applied to all channels to remove EMG artefacts. Thus, it can be confirmed that right side of the figure has EMG-free signals. Then, on the left panel, channels were separated based on the EMG-CCh selection method (area ***A1*** for EMG-CCh, and ***A2*** for artefact-free channels). Based on this, the EMG-free data on the right panel is also separated into ***A3*** and ***A4*** using the channel numbers from the left panel. Secondly, feature extraction with an autoregressive model was processed in each area. Then, the data with the smallest 10% of DBI values between idle and task state were selected from sets ***A1*** to ***A4***. Finally, a Wilcoxon test was applied between the best DBIs from ***A1*** and ***A3*** (**Comparison (A)**) as well as between the DBIs from ***A2*** and ***A4*** (**Comparison (B)**). In an ideal EMG-CCh selection scenario, **Comparison (A)** should show significant difference between ***A1*** and ***A3*** as ***A1*** contains EMG artefacts whereas ***A3*** is clean, so the DBI value from ***A1*** should shows significantly lower DBI values, which indicates that EMG artefacts would have played a significant role in class separation. Also, **Comparison (B)** should show no significant DBI difference between ***A2*** and ***A4***. Based on the EMG-CCh selection method, ***A2*** channels were found not to have an EMG-related role in class separation. Thus, it should give similar (i.e, no significant difference) DBI values to those from set ***A4***. Therefore, if **Comparison (A)** shows significant difference and **Comparison (B)** shows no significant difference, it can be concluded that the suggested EMG-CCh selection method correctly chose EMG artefacts contaminated channels that would otherwise have played an EMG-

contaminated role in class separation and the remaining channels did not play any significant class-dependent result.

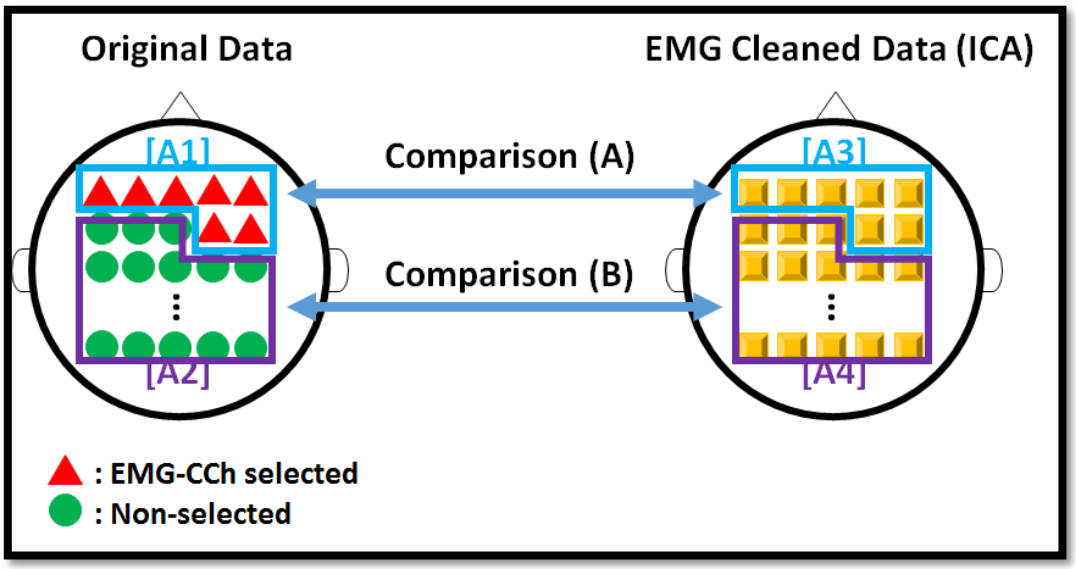


Figure 4.8. Example reliability test of the EMG-CCh selection method [2].

4.3 Results

Sound-production Related Cognitive Tasks Onset Detection Data

Table 4.2 shows an example from participant 1 with the inhibited overt siren (IO_Siren) task onset detection. The smallest 5% of DBI values (384 features x 5% = 19 values) were selected from the AR features. The Wilcoxon test *p*-values and statistical power *t*-test were calculated for ***No EMG handling vs. EMG-CCh removal, ICA vs. ICA (EMG-CCh), PCA vs. PCA (EMG-CCh) and BSS-CCA vs. BSS-CCA (EMG-CCh),*** respectively, with the 19 DBI values. The results show that applying

EMG removal methods only to the EMG-CCh significantly improves class separation between Idle and Task states (p-value < 0.05) when using ICA (EMG-CCh) and BSS-CCA (EMG-CCh), and the power is statistically conclusive (t-test > 0.5). These evaluations were done with all seven subjects and with all four different onset detection tasks. In addition, the single smallest, smallest 1%, and smallest 5% DBI results, respectively were separately evaluated with the four different feature extraction methods.

Table 4.2. Example of Wilcoxon test and Statistical power calculations for features with the lowest 5% DBI values [2].

Average of Smallest 5% DBI values				
AR Model Feature [P1 – Inhibited Overt Siren Task]	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
	4.090 [±0.44]	6.756 [±0.66]	4.179 [±0.59]	4.227 [±0.27]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA (EMG-CCh) [std]**
	4.096 [±0.45]	4.096 [±0.45]	4.086 [±0.44]	3.978 [±0.40]
Wilcoxon-test (p-values)	-0.9069	≈ 0	0.1568	0.0296
Statistical Power (t-test)	0.0502	≈ 1	0.0758	0.7863

- **Smallest 5% and 1% DBI value results**

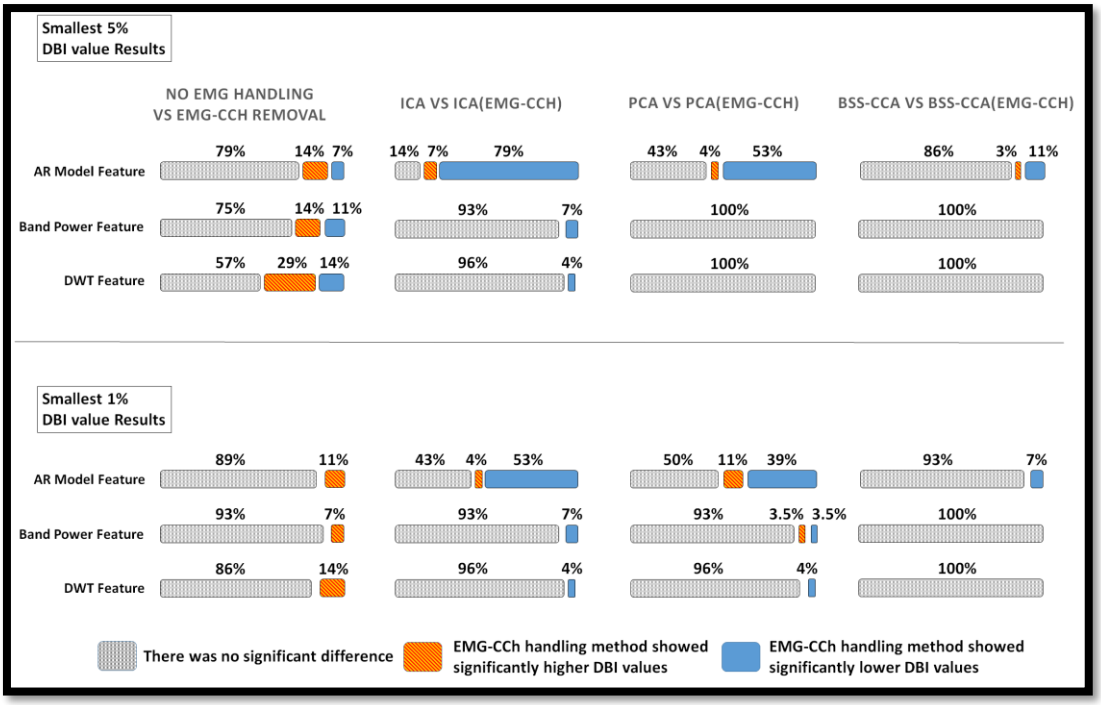


Figure 4.9. Results with features with the smallest 5% and 1% DBI values using the onset detection data set [2].

Figure 4.9 shows DBI comparison results. 5% of the smallest DBI values were taken from each feature domain: 19 (384 AR model features x 0.05), 22 (448 Band power features x 0.05), 16 (320 DWT features x 0.05) DBI values respectively. The blue areas in the figure indicate that applying EMG handling methods only to the EMG-CCh (our proposed method) shows significant improvement (Wilcoxon test p-value < 0.05) in class separation, and power is statistically conclusive (statistical power t-test > 0.5). On the other hand, vertical red areas represent instances in which our proposed method showed significantly higher DBIs and, thus, less class separation between Idle and Task states. The grey horizontal stripe areas showed no significant difference between our method and typical EMG handling techniques when applied to all channels.

As can be seen from the case of ***No EMG handling vs. EMG-CCh removal***, there is no significant difference most of the time. But, average DBI values become higher more times than become lower for all three different feature extraction methods. This an expected result as EMG artefacts could play a role in class separation if they are not handled properly. Thus, removing EMG-CCh reduces class separation, but it shows why artefact handling is required. In the case of ***ICA vs. ICA (EMG-CCh)***, 79%, 7% and 4% out of 28 tests (7 subjects x 4 onset tasks) showed significant class separation improvements with EMG-CCh selection for AR model features, band power and DWT features, respectively. On the other hand, only 7% of the AR features gave higher DBI values with the proposed EMG-CCh method and all remaining tests yielded no significant difference. In the ***PCA vs. PCA (EMG-CCh)*** case, band power and DWT feature made no significant difference but AR features showed significant class separation improvement for 53% of the tests and 43% remained as not significantly different. In the ***BSS-CCA vs. BSS-CCA (EMG-CCh)*** case, 11% showed significant class separation improvement with AR features, while 3% became worse and most tests (86%) showed no significant difference. Band power and DWT features also showed no significant difference in class separation between BSS-CCA applied to all channels vs. only to EMG-CCh.

The smallest 1% DBI results showed similar trend as results based on the best 5% DBIs. While results generally show no significant difference between our method and standard EMG-removal techniques, EMG-CCh selection did yield significantly better class separation more often than it yielded worse separation.

- **The smallest DBI value results**

Table 4.3. Results with the features that gave the smallest DBI value, onset detection data set [2].

AR Model Feature	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
[Average of 7 subjects & 4 onset Detection Tasks]	3.569 [±1.72]	4.351 [±1.86]	4.119 [±2.09]	3.713 [±1.85]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	3.320 [±2.01]	3.769 [±1.87]	3.648 [±1.76]	3.547 [±1.70]
Band Power Feature	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
[Average of 7 subjects & 4 onset Detection Tasks]	2.494 [±1.15]	2.566 [±1.15]	2.592 [±1.20]	2.500 [±1.15]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	2.590 [±1.21]	2.506 [±1.18]	2.514 [±1.19]	2.496 [±1.15]
DWT Feature	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
[Average of 7 subjects & 4 onset Detection Tasks]	2.741 [±1.26]	2.825 [±1.29]	2.740 [±1.20]	2.753 [±1.25]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	2.946 [±1.39]	2.777 [±1.23]	2.762 [±1.27]	2.745 [±1.24]
CSP Feature	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
[Average of 7 subjects & 4 onset Detection Tasks]	3.794 [±7.14]	4.724 [±9.85]	5.977 [±16.10]	3.810 [±7.22]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	3.828 [±6.21]	3.839 [±7.08]	8.320 [±25.52]	3.767 [±7.11]

Table 4.3 lists the (single) smallest DBI value in each case. CSP features were included only in this table because it has just one feature point. Also, as this table shows

only the smallest DBI value, Wilcoxon tests and Statistical power *t*-tests could not be applied. However, even though no inferences can be made concerning statistical significance on a test-by-test basis, the smallest DBI values can give relevant information that might be useful to other BCI studies as minimising the number of data points is always an important goal in BCIs. For this reason, we averaged the smallest DBI values from 28 tests and compared overall results between our EMG-CCh method and standard EMG handling methods (there was no statistical difference $p > 0.05$ from all cases). The averaged smallest DBI values were lower with our EMG-CCh handling method in most cases, except for DWT features and for CSP features with PCA. In terms of CSP featured, some subjects had very large DBI values (i.e., poor class separation) compared to other feature domains. But, in general, our EMG-CCh method gave lower DBI values.

BCI Competition Data Set

The smallest 5% DBIs gave 7, 8 and 6 features from the AR model, band power and DWT domains, respectively. The number of feature was 1 when using the smallest 1% DBIs. Thus, only these two cases were analysed with the BCI competition data set.

In Figure 4.10, 40% of participants showed less class separation with EMG-CCh removal using AR model and band power features in the case of comparisons between ***No EMG handling vs. EMG-CCh removal***. The remaining 60% has no statistical difference in the latter comparison. As before, this is an expected result as EMG would have had a role in class separation.

In the case of ***ICA vs. ICA (EMG-CCh)***, 60%, 20% and 20% of 5 subjects - for each of the three feature domains, respectively - showed significantly lower DBI values

with our EMG-CCh method and none of participants yielded worse class separation with our method. Comparing ***PCA vs. PCA (EMG-CCh)***, only one participant out of five yielded significantly improved class separation with our method. In the ***BSS-CCA vs. BSS-CCA (EMG-CCh)*** similar results were seen for band power and DWT features. For AR features, on the other hand, 60% (3 subjects) showed significant improvement while 1 participant had higher DBI value with the EMG-CCh. Table 4.4 shows a similar trend; it mostly has lower DBI values with the EMG-CCh method except for ICA (EMG-CCh) with CSP feature and PCA (EMG-CCh) with AR model feature.

- **Smallest 5% DBI value results**

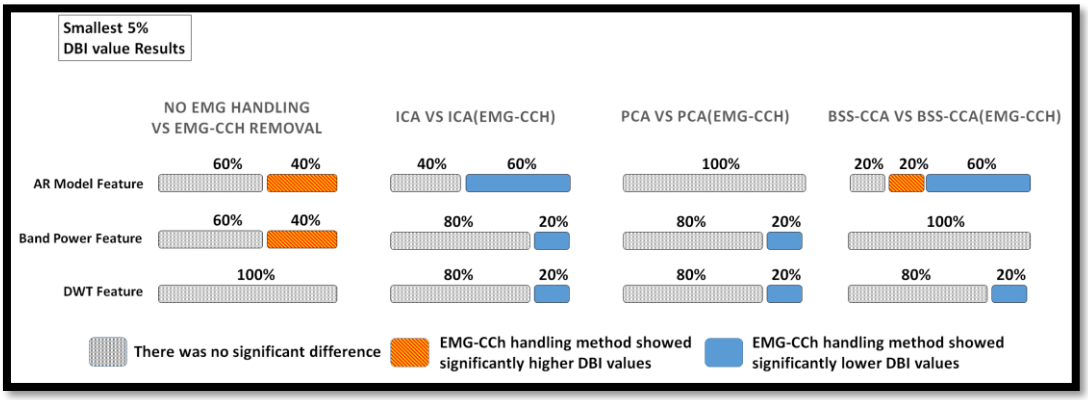


Figure 4.10. Result of smallest 5% DBI values from BCI competition data set [2].

- **The smallest DBI value results**

Table 4.4. Result of the smallest DBI values from BCI competition data set [2].[9]

AR Model Feature	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
	8.623 [±3.23]	11.689 [±4.64]	9.551 [±4.89]	10.753 [±3.13]
[Average of 5 subjects]	**EMG-CCh removal [std]**	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	11.393 [±6.19]	10.362 [±5.34]	9.846 [±5.23]	9.768 [±3.95]

Band Power Feature	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
[Average of 5 subjects]	4.968 [±1.50]	5.492 [±1.89]	5.101 [±1.22]	9.244 [±2.91]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	5.175 [±1.82]	5.175 [±1.82]	4.963 [±1.45]	5.239 [±1.89]
DWT Feature	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
[Average of 5 subjects]	6.756 [±2.52]	7.192 [±2.90]	6.984 [±2.32]	12.408 [±2.95]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	6.957 [±2.84]	6.951 [±2.83]	6.780 [±2.55]	6.940 [±2.70]
CSP Feature	**No EMG handling [std]**	**ICA [std]**	**PCA [std]**	**BSS-CCA [std]**
[Average of 5 subjects]	3.467 [±1.72]	3.722 [±2.03]	5.387 [±4.05]	4.637 [±2.60]
	EMG-CCh removal [std]	**ICA(EMG-CCh) [std]**	**PCA(EMG-CCh) [std]**	**BSS-CCA(EMG-CCh) [std]**
	4.292 [±2.17]	3.806 [±1.88]	3.791 [±1.34]	3.281 [±1.44]

Throughout the thorough literature review, we could not find EMG contaminated EEG channel selection method so far in most EEG studies. Many BCI studies generally applied EMG artefact handling with blind source separation such as BSS-CCA, ICA or PCA [130]. These blind source separation techniques could lose some useful EEG information as well [19-22]. Thus, our EMG-CCh selection method could help this problem. It minimises useful information loss by statistically selecting EMG-CCh, which could cause EMG related class-dependant results, and applying blind source separation techniques only to these channels. This method can keep useful EEG information that could have been eliminated by applying blind source separation to non-EMG-CCh.

In addition to this, based on the results, combining our EMG-CCh selection method with typical EMG artefact handling methods (ICA, PCA and BSS-CCA) significantly improved class separation by having lower DBI values than normal EMG handling approaches in many cases. Only few percentages of tests were become higher DBIs in our onset detection data set and rests, which is most, were become either better class separation or no significant difference with EMG-CCh selection approach. In BCI competition data set, only one test (BSS-CCA with AR model feature) had higher DBI values with our methods. They showed either significantly lower DBI values or no significant difference. These results indicate that our suggested EMG-CCh selection and handling method improves (at least remain same) class separability compared to typical EMG handling approaches in BCIs.

Reliability of the EMG-CCh Selection Method

As explained in '*Methodology*' section '*Reliability of the EMG-CCh Selection Method*' above, we applied two types of comparisons, A and B, respectively, to test for the possibility that we did not eliminate all EEG channels with significant class-dependent EMG contamination.

Using our sound-production onset data and the average of the smallest 10% DBI values from 28 cases (7 subjects x 4 tasks), **Comparison A** showed significant difference (p=0.011, mean value of ***A1=6.26,*** mean value of ***A3=8.11***) while **Comparison B** showed there was no significant difference (p=0.124, mean value of ***A2=4.85, A4=5.00***). The results were similar when we used the BCI competition data set (5 subjects). In this case results also showed a significant difference in **Comparison A** (p=0.016, mean value of ***A1=12.61, A3=25.88***) and no significant difference in

Comparison B (p=0.222, mean of ***A2=18.27, A4=20.94***)**.** From these results, it can be said that our EMG-CCh selection method is indeed correctly identifying channels that could affect class-dependent results.

4.4 Discussion

Electromyography (EMG) artefacts are a well-known problem in Electroencephalography (EEG) related studies such as BCIs, brain mapping, and clinical areas. In order to handle these artefacts, Blind Source Separation (BSS) techniques (e.g. BSS-CCA, ICA and PCA) were commonly used. However, these BSS techniques may remove not only EMG artefacts but also some useful EEG sources [19-22]. To reduce this loss of useful information, a new technique for statistically selecting EMG artefacts Contaminated EEG Channels (EMG-CCh) was proposed in this chapter.

The EMG-CChs are selected based on the correlation between EEG channels and facial EMG channels. The correlation coefficient values were calculated for the idle and task states, respectively, and compared (using a Wilcoxon test) against each other in order to determine whether the artefacts played a significant role in class separation. If a channel's (EMG vs. EEG) correlation value was significantly higher during a task than during the idle states, then this channel was chosen as an EMG-CCh; i.e. it was treated as being significantly contaminated and was therefore removed from the signal classification stages. In order to ensure that our EMG-CCh selection approach does reduce EMG artefacts and the results are not produced simply because of a weak EMG removal, reliability tests on the EMG-CCh selection method were made by comparing

our EMG-CCh selection data with typical ICA-EMG removal data. As it was discussed in section 4.3 '*Reliability of the EMG-CCh Selection Method*, the statistical comparison (Wilcoxon test) with the ICA-EMG removal method showed that the selected EMG-CCh has significantly lower DBIs than its EMG-free (ICA applied) data (our dataset: p-value of 0.0108, BCI competition dataset: p-value of 0.0159), while non-selected channels had no significant difference with the EMG-free data (our dataset: p-value of 0.1241, BCI competition dataset: p-value of 0.2222). This indicates that the proposed EMG-CCh selection method is indeed correctly identifying channels where EMG artefacts played a significant role for class separation, whereas the rest of the unselected channels have no EMG related class separation effect.

The statistical testing results showed that our EMG-CCh selection method correctly found the artefact-contaminated channels. The application of this channel selection method to typical artefact handling techniques, such as ICA or BSS-CCA, would result in the elimination of artefacts in the same manner as typical methods. Furthermore, it reduces information loss, as it only applies to the EMG-CCh. A comparison of signals before and after applying this method would not be necessary, as widely used and validated common blind source separation techniques would guarantee the same result. Thus, it is out of scope for this EMG-CCh selection procedure.

The EMG-CCh selection technique was tested with our own onset detection data set as well as with the BCI competition IV data set 2a. The comparison of the results between the typical BSS artefact removal approach, when applied to all channels and when only to the EMG-CCh, and our method, showed that ICA, PCA and BSS-CCA can yield a better class separation with the proposed method. In particular, a significant improvement ($p<0.05$) in class separation was recorded when autoregressive

coefficients were used, which were extracted from our onset data in 79% of the cases for ICA, 53% for PCA and 11% for BSS-CCA. Only 7% (ICA), 4% (PCA) and 3% (BSS-CCA) of tests became significantly worse with our approach; the rest of the cases yielded no statistically significant differences in terms of class separation performance. In the BCI competition data we reported an improvement in 60% of the cases for ICA and 60% for BSS-CCA when the autoregressive coefficients were used as features. To the best of our knowledge (thorough literature review up until 2017), while there are few EOG artefact removal methods using simple cross-correlation (e.g., Quilter et al. 1977. [137]), there are no existing methods for removing EMG artefacts by applying a statistical approach with correlation values between EEG and EMG channels. The simple method which was proposed in this chapter showed improvement with the use of both our data and the BCI competition data. Furthermore, the EMG-CCh selection method can be used on its own or it can be combined with pre-existing artefact handling methods.

In some neurophysiology studies our method may not be suitable, as it does not remove artefacts from all channels. Even though this represents a limitation in this method, the class dependent artefacts effect and selective removing process would prove beneficial in most EEG-related studies. For these reasons, we believe that this method can be of use for other EEG studies and in the long term, it will have an impact on the field in terms of artefact removal techniques.

4.5 Summary and Conclusions

EMG artefact handling is an essential procedure for EEG based studies where one of its issues is that there is some useful information loss with the common established blind source separation techniques. Therefore, we proposed a new technique for selecting EEG channels contaminated with class-dependent EMG artefacts (named EMG-CCh), in order to minimise the information loss and improve class separation. The method can be used on its own for channel rejection or it can be combined with pre-existing artefact handling techniques and it showed a significant class separation improvement (compared to other existing techniques) with the use of both our data and the BCI competition data set in many cases.

The existence of this improved (reduced information loss) EMG artefacts handling procedure will help our onset detection study to expand to real-world BCI uses in terms of performance as well as reliability (purely task based onset detection). In addition, this approach will provide a significant step forward for the other EEG studies (e.g. BCIs, brain mapping, and clinical areas) in terms of the EMG artefact removal procedure.

5 Classifying Sound Production Related vs. Idle State Towards Onset Detection in Brain-Computer Interfaces in Offline System (Based on papers [4] and [5])

5.1 Introduction

This chapter is based on the published conference proceedings; [4] and [5].

Onset detection allows self-paced BCIs to detect when a user wishes to send a command [9, 138]. Aiming specifically at exploring cognitive states that may be potentially used in onset detection, this chapter proposes the use of novel, speech-related cognitive tasks as a first step of the investigation. Speech (of some sort) is the most natural form of human-to-human interaction, it is therefore suggested that such tasks would be suitable for BCIs from a usability point of view. The goal of this chapter is to classify various easy and intuitive speech-related tasks versus the idle state in offline settings in order to investigate our hypothesis whether the sound-production related cognitive tasks can be distinguished from the idle state and to pick the most suitable one before it is used in an online onset detection experiment.

This chapter describes two different experiments. The first experiment (section 5.2) shows high-tone sound production tasks versus the idle state with various speech modes (covert, inhibited overt and overt). The results showed that the *Covert* state produced the best classification rate vs. the *Idle* state (82.41%, 81.20%, 85.12%, and

74.72%, respectively) compared to the *Overt* vs. *Idle* for all participants ($p \ll 0.05$). The *Covert* state also provided better classification results vs. *Idle* than the *I_Overt* state for participants 1, 2 and 3, and equally good results with the *I_Overt* state for participant 4. The second experiment (section 5.3) investigates covert siren-like sound production tasks and it achieved a 76.88%, 79.58%, 76.67%, 80.2% and 82.71% of true-positive performance for participant 1, 2, 3, 4 and 5, respectively. As a first stage of our sound imagery onset detection investigation in offline cue-based settings, these studies and results provided very useful background information for this project in order to move it forward to the online self-paced onset detection study.

5.2 Experiment 1 (High-tone sound production task)

As a first stage of the experimental process and in order to firstly determine the suitability of the chosen tasks for discrimination against the idle states, we present a cue-based study. The cognitive tasks which were explored here and have provided suitable results, will be used in proper self-paced studies in next chapters.

Methodology

(A) Speech-related State Definition

In this experiment, high tone production and idle state were recorded. The high tone production states were generated under the following three conditions:

- Overt speech (*Overt*): A high tone is actually produced by the user and it can be heard clearly.

- Inhibited overt speech (*I_Overt*): This involves two elements: 1) breathing out, and 2) tensioning the vocal cords but without actually producing the high tone sound.

- Covert speech (*Covert*): Imagination of the processes leading to high tone sound production. To facilitate this task, participants were instructed to also imagine hearing the high tone sound.

For high tone production, participants were instructed to aim at producing an 'um' sound in as high a tone as they can comfortably produce for a few seconds continuously. However, imagining tongue, mouth, or any body movements was not allowed in order to reduce the motor imagery signals not related to high tone sound production.

(B) Experiment Protocol

Each experiment consisted of 6 runs in total (i.e., 2 *Overt*, 2 *I_Overt* and 2 *Covert* modes, all intercalated with idle states). The sequence of cognitive tasks was randomly ordered to prevent sequence-dependent results and predictive guesses by users. In each run, there were 40 trials, of which 30 were high tone states and 10 were idle states.

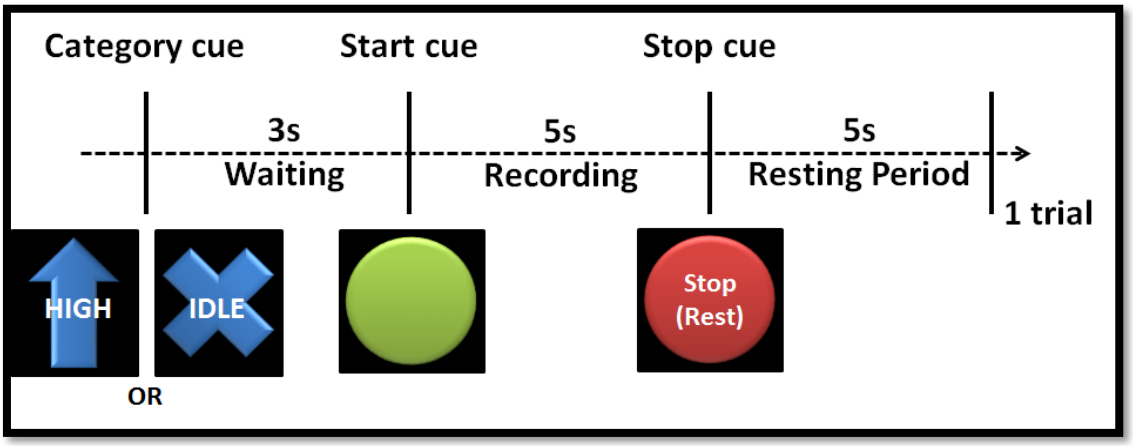


Figure 5.1. EEG data recording timing procedure for each trial [5].

Figure 5.1 shows the recording timing protocol and the visual stimuli. The arrow indicates high tone production and ‘X’ represents the idle state. Either symbol is randomly presented on a screen for 3s to give notice to the participant about the which action to take, but no task is to be executed at this stage yet. When the green circle appears, participants start the designated task (or idle state) and they are to execute it shortly after appearance of the green circle. The latter circle, contrary to the other displayed objects, did not have a 'start' label on it so at to reduce cognitive load. The task is performed for 5s, during which data are recorded. After this a red circle is presented for 5s for the resting period until the next trial begins. This timing protocol (i.e., separate 'category' and 'start' cues) was designed to minimize effects resulting from the different category cues. Visual stimuli generate ERP signals for around 300-400ms after the stimuli. Different cue object shapes/colours could generate different ERP patterns, which must be decoupled from the actual sound production related tasks so as to ensure that we are classifying the tasks, not the stimuli.

The experiments were done in accordance with the University of Essex Ethics Committee guidelines.

(C) Data Acquisition

Four healthy subjects (3 males, 1 female, ages 20-25) with normal or corrected vision participated in this experiment. Each participant sat down on a chair comfortably with their head 1m away from the monitor. 64 electrodes were placed based on 64 channel 10-20 layout and 1 reference channel was recorded from the right side mastoid. The BIOSEMI ActiView software was used for EEG data acquisition and the sample rate was 256 S/s. EEG data was transferred from the BIOSEMI software through TCP/IP communication. On the PC side, a Java application displayed the visual

stimulus and saved the EEG data. The saved EEG signals were used for off-line data analysis with MATLAB.

(D) Signal Processing

To investigate how far after the 'start' stimuli onset useful data can be found, recorded EEG data were segmented with various window lengths, as measured from the 'start' cue onset: 1.0s, 1.5s, 2.0s and 3.0s.

The segmented EEG signal were band-pass filtered (zero-lag Butterworth filter, order 5) at a range of 4-20 Hz. Then, Referencing was applied to reduce common environmental noise by subtracting the right mastoid reference channel from all 64 channels.

(E) Common Spatial Pattern (CSP) for Channel Selection

CSP was used for channel selection by maximizing the difference in variance between classes in multi-channel data [114, 115]. Figure 5.2 shows the common spatial pattern for each subject. For illustration purposes, the plots shown were generated in covert speech mode when the time window is cut at 1.5 second. Channel F5 and nearby areas showed best class separation. As expected, these channels are located around Broca's area, which is related to speech production. The channels with the best CSP-based class separation were used for feature extraction.

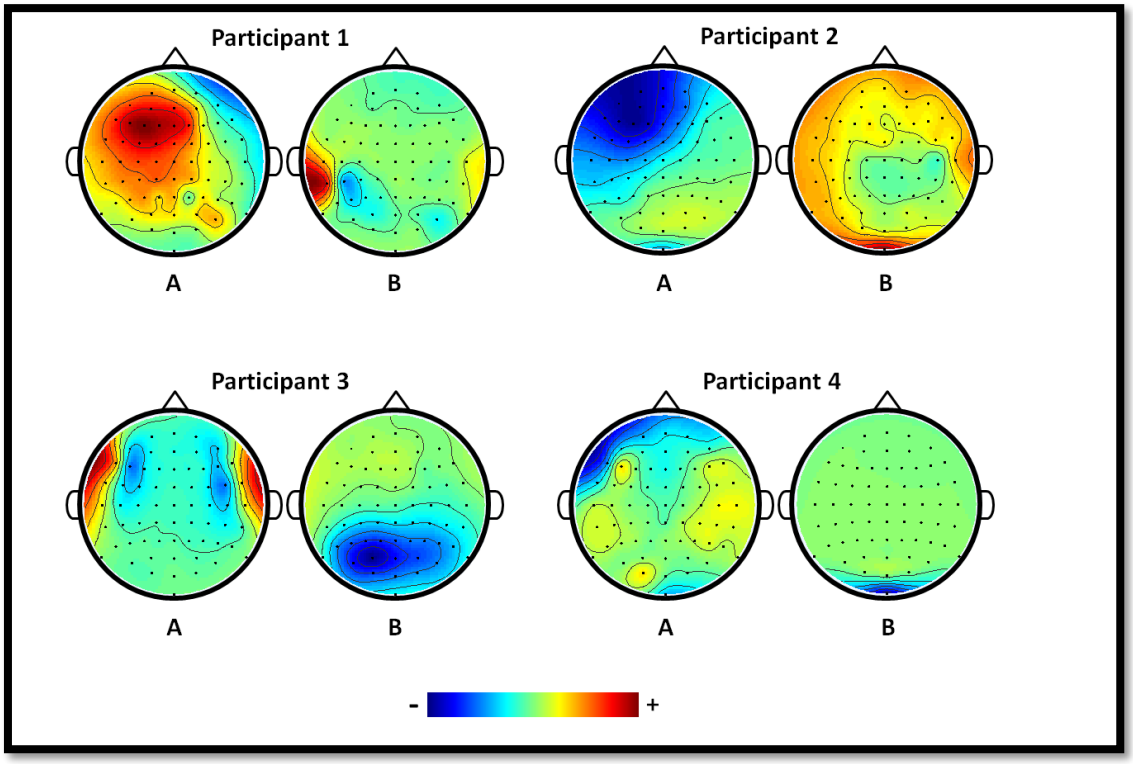


Figure 5.2. Common spatial patterns for each participant (A: Max eigenvalue for covert sound production states; B: Max eigenvalue for idle state) [5].

(F) Autoregressive Model for Feature Extraction

Autoregressive (AR) modelling was applied for feature extraction. The AR coefficients were used as a feature vector. To this end, order number selection is important. In [139] order number 6 was found to be optimal in imagined speech EEG signals. The order number selected was therefore 6. Burg's method ('arburg' function in MATLAB) was used for AR coefficient calculations and these coefficients were used as feature vectors.

(G) Classification

Linear Discriminant Analysis (LDA) was applied for classification. LDA finds a linear projection line that maximizes the separation probability between classes and minimizes the overlap between them [140].

Feature vectors (i.e., AR coefficients) were used as inputs to the LDA. 5-fold cross validation was applied 100 times to improve the classification results reliability.

Results

Table 5.1 shows True Positive classification accuracy results. For each participant, three speech sound production modes (*Overt*, *I_Overt* and *Covert*, in addition to the idle state) and four different time window sizes (1.0s, 1.5s, 2.0s and 3.0s) were tested.

As can be seen from the results, average performances for each of two runs with same speech mode and time window length show that *Covert* sound production mode gave us the best classification accuracy results in most cases (82.41%, 81.20%, 85.12% and 74.72%, respectively; see yellow/highlighted results). The *Covert* vs. *Overt* differences were significant for all participants (Wilcoxon test, $p \ll 0.05$). The *Covert* vs. *I_Overt* differences were significant for participants 1, 2, and 3 ($p \ll 0.05$), but not for participant 4 ($p=0.24$). Likewise, the *I_Overt* vs. *Overt* differences were significant for participants 1, 2, and 3 ($p \ll 0.05$), but not for participant 4 ($p=0.195$).

In terms of bit-transfer rate, the 1.5s window length performed the best for participants 1 and 3. Window_length = 2.0s was best for participant 4, and 3.0s was best for participant 2. From these results, the maximum hypothetical bit-transfer rate (BTR) -- i.e., if the tasks were to be executed only for the duration of the best window length and without the category cue and the rest periods -- can be estimated as:

$$BTR = \frac{60s}{window\_length} \times True\_positive\_rate \qquad \text{Equation 5-1}$$

Thus, the maximum hypothetical BTR was computed as 32.95bit/min, 16.24bit/min, 34.05bit/min and 22.42 bit/min for participants 1, 2, 3, and 4, respectively.

Table 5.1. Performance results for each participant [5].

Participant 1		**Time Windows**							
		1.0 s		***1.5 s***		***2.0 s***		***3.0 s***	
		CSP_CH (Acc %)	***BEST_CH (Acc %)***	***CSP_CH (Acc %)***	***BEST_CH (Acc %)***	***CSP_CH (Acc %)***	***BEST_CH (Acc %)***	***CSP_CH (Acc %)***	***BEST_CH (Acc %)***
SPEECH MODE	Overt run 1	FPz (80%)	FPz (80%)	FP2 (86.25%)	FP2 (86.25%)	PO3 (80.43%)	C4 (82.55%)	PO3 (85.18%)	FPz (85.45%)
	Overt run 2	CP5 (68.43%)	P1 (73.53%)	CP5 (61.28%)	P6 (66.10%)	CP5 (61.83%)	P1 (68%)	CP5 (68.58%)	P6 (74.60%)
	I_Overt run 1	PO8 (68.38%)	Pz (78.59%)	Cz (64.38%)	C6 (73%)	PO8 (67.08%)	O1 (75.15%)	F2 (53.90%)	P9 (76.63%)
	I_Overt run 2	FC4 (32.70%)	TP7 (67.55%)	CP5 (56.25%)	TP7 (72.53%)	TP7 (71.98%)	TP7 (71.98%)	P5 (72.78%)	P5 (72.78%)
	Covert run 1	CP5 (69.65%)	CP4 (80.73%)	F3 (82.88%)	F3 (82.88%)	F3 (78.50%)	AF7 (81.83%)	CP5 (65.98%)	PO3 (78.30%)
	Covert run 2	CP5 (73.43%)	CP5 (73.43%)	P5 (78.48%)	CP5 (81.93%)	P6 (58.95%)	POz (72.30%)	P6 (60.25%)	P9 (76.45%)
Participant 2									
SPEECH MODE	Overt run 1	Iz (67.45%)	P9 (73.48%)	F1 (60.63%)	O1 (77.53%)	TP7 (46.13%)	O1 (80.35%)	AF3 (78.68%)	AF3 (78.68%)
	Overt run 2	FC5 (46.05%)	AF4 (69.15%)	FC5 (45.75%)	Oz (63.75%)	CP4 (66.13%)	CP4 (66.13%)	FC5 (44.63%)	FP2 (75.20%)
	I_Overt run 1	AF7 (58.95%)	P8 (79.70%)	AF8 (82.55%)	AF8 (82.55%)	Fz (63%)	P8 (80.25%)	Fz (69.03%)	T8 (85.05%)
	I_Overt run 2	FPz (58.13%)	PO7 (66.80%)	FP1 (55.25%)	PO7 (72.70%)	FP1 (24.75%)	PO8 (74.48%)	FP1 (45.45%)	P5 (73.35%)
	Covert run 1	F1 (73.98%)	F1 (73.98%)	F3 (53.75%)	PO4 (79.18%)	F3 (60.65%)	CP4 (75.13%)	FP2 (80.75%)	FP2 (80.75%)
	Covert run 2	T8 (75.40%)	T8 (75.40%)	AF3 (45.83%)	P5 (70.20%)	AF3 (59.98%)	AF7 (76.65%)	AF3 (73.93)	T8 (81.83%)

Participant 3									
SPEECH MODE	Overt run 1	C3 (44.55%)	FC4 (67.43%)	C3 (53.85%)	C1 (66.68%)	PO3 (63.20%)	PO8 (78.43%)	PO3 (67.28%)	P10 (73%)
	Overt run 2	F5 (51.55%)	P8 (73.33%)	F5 (60.13%)	FC4 (71.65%)	F5 (64.08%)	F2 (76.93%)	F5 (58.80%)	FC1 (70.28%)
	I_Overt run 1	FC5 (58.85%)	P10 (75.08%)	CP3 (67.30%)	Cz (87.50%)	FC5 (72.48%)	P10 (87.45%)	C4 (72.13%)	C1 (83.73%)
	I_Overt run 2	F5 (59.38)	F8 (71.80%)	F4 (54.48%)	F8 (72.53%)	F4 (60.50%)	P4 (76.35%)	POz (66.88%)	TP8 (73.08%)
	Covert run 1	F5 (83.78%)	AF7 (85.73%)	F5 (91.24%)	F5 (91.24%)	F5 (89.23%)	F5 (89.23%)	F5 (88.15%)	F5 (88.15%)
	Covert run 2	F5 (51.33%)	P8 (72%)	F5 (48.48%)	O2 (79%)	F5 (51.20%)	O1 (78.83%)	F5 (62%)	C5 (81.80%)
Participant 4									
SPEECH MODE	Overt run 1	F6 (59.25%)	T8 (72.15%)	FP1 (74.95%)	T8 (76.05%)	FP1 (58.90%)	P1 (70.23%)	CP5 (55.38%)	TP7 (75.70%)
	Overt run 2	F6 (61.95%)	F8 (71.03%)	F5 (64.05%)	P10 (66.98%)	FP2 (52.75%)	TP8 (65%)	AF7 (71.08%)	AF7 (71.08%)
	I_Overt run 1	P5 (66.55%)	AF8 (77.13%)	Fz (56.33%)	AF8 (71.53%)	Fz (53.65%)	AF8 (72.78%)	Iz (75.48%)	Iz (75.48%)
	I_Overt run 2	F5 (57.80%)	FC6 (69.55%)	PO4 (70.83%)	FC6 (76.95%)	PO4 (70.30%)	Oz (76.10%)	PO4 (67.40%)	F8 (73.40%)
	Covert run 1	F5 (64.38%)	T7 (67.05%)	F5 (67.48%)	F8 (69.33%)	A7 (65.68%)	F8 (75.68%)	FT7 (61.90%)	FT8 (65.20%)
	Covert run 2	FPz (58.68%)	PO4 (63.88%)	FC1 (52.48%)	Oz (75.95%)	AF7 (51.12%)	PO4 (73.75%)	AF7 (56.48%)	PO4 (66.50%)

***ACC**: Accuracy**. CSP_CH**: The channel was chosen from CSP channel selection method. **BEST_CH**: The channel that actually performed the best out of 64 channels. **Yellow (highlighted) results**: The best accuracy results for two runs for each participant.*

For channel selection, the CSP method (see CSP_CH in Table 5.1) yielded rather unsatisfactory results. Classification results with this method were worse than when the best channel was chosen by skipping the CSP stage and using only the best

classification results from all 64 channels (i.e. BEST_CH in Table 5.1). CSP-based channel selection could reduce computational time by not requiring classification using all 64 channels but on average it achieved 16.125% less than the BEST_CH method.

Discussion

In [141], various cognitive tasks were compared with the idle state and the motor imagery tasks achieved the best accuracy with a value of 73% on average for 7 subjects. By using this value as a benchmark, it can be seen that the present study achieved rather satisfactory results, even though there were some differences between their methods and ours, especially with regards to the duration of the active state tasks, i.e. 10 s in [141] vs. 5 s in ours.

The physiological locations which had the maximum differences in EEG signals between the idle and sound production states can be seen in Figure 5.3. The Figure illustrates the distribution of the best channels that are shown in Table 5.1. As can be observed from Figure 5.3, it is difficult to interpret the physiological map, something which can partly be attributed to the well-known poor EEG spatial resolution resulting from volume conduction effects.

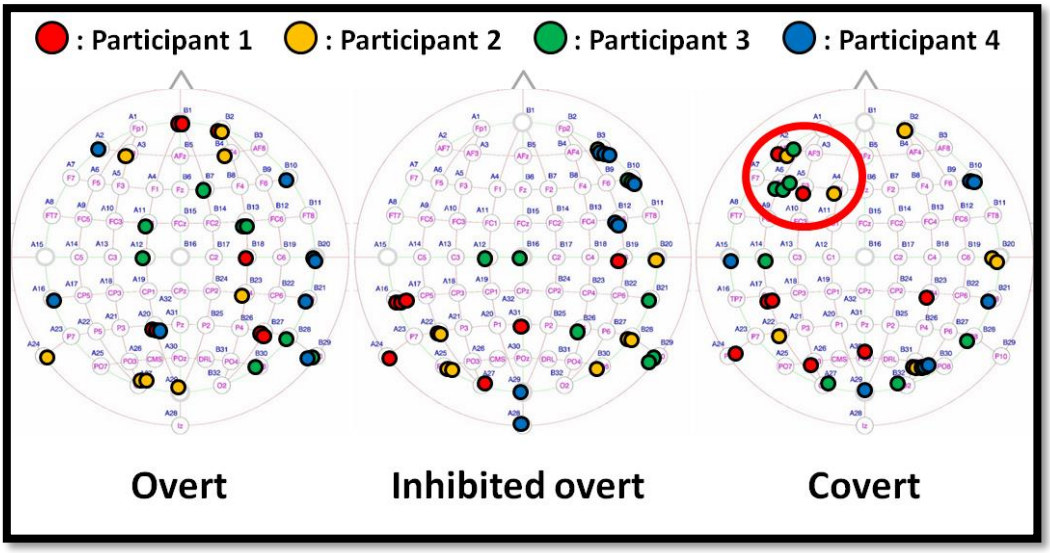


Figure 5.3. Best channels location for each speech mode [5].

However, the map for the *Covert* state does show a higher concentration of channels (AF7, F5, F3 and F1) around Broca's area, which was more in line with our expectations.

In terms of overall classification accuracy, possible biases towards the active or idle states can be investigated. Table 5.2 shows the true positive rates for active and idle states for the best runs (as seen in Table 5.1 above) for each participant. As can be noticed from Table 5.2, active states (*Covert*, *Overt*, and *I_Overt*) have higher true-positive values than the idle state. This result is possibly produced by the biased number of training set trials for each run (30 active state trials vs. 10 idle state trials, the number of idle states is 1/3 of the speech states). This is something that will require to be rectified in future works of this study.

Keeping in mind that the main purpose of this study was to test cognitive tasks for their potential use in an asynchronous onset detection, an obvious limitation of this work was the use of a cue-based protocol. More specifically, the recorded EEG signals

contain stimulus-timing information (via event related potentials) that would not be deliberately present in a true self-paced BCI. However, we are hopeful that the results are still useful in asynchronous BCIs as the same green circle was present in both the active and idle state initiation, and so our promising classification results reflect the actual separation between the active states. In the worst-case scenario, the methods reported here can be easily used in cue-based BCIs as well.

Table 5.2. Idle and Speech states accuracy for each participant [5].

	Run / Channel / Window_length	***Idle state Correct / Input (Accuracy)***	***Active state Correct / Input (Accuracy)***
Participant 1	Covert run 1 / F3 / 1.5 s	161 / 200 (80.5%)	504 / 600 (84%)
Participant 2	Covert run 2 / T8 / 3.0 s	140 / 200 (70%)	506 / 600 (84.3%)
Participant 3	Covert run 1 / F5 / 1.5 s	178 / 200 (89%)	547 / 600 (91.2%)
Participant 4	Covert run 1 / F8 / 2.0 s	116 / 200 (58%)	504 / 600 (84%)

5.3 Experiment 2 (Siren-like sound production task)

In this experiment, a modified cue-based protocol was used in order to focus on distinguishing between the active and idle states. Users were free to spontaneously

control the timing of execution of the active task. Self-reporting by means of a key-press was used in order to determine the time location of the active state data.

Methodology

(A) Covert siren sound production and Idle state definition

In this experiment, the active state and the idle states were defined as follows:

- **Active state** (*Siren*): covert production of siren-like sounds (e.g., 'wee-woo wee-woo') as well as simultaneous auditory recall (imagination of hearing the siren sound). During the imagination of 'wee' syllable, participants were instructed to think of a high pitch sound and they were to think of a low pitch sound during the word 'woo'.
- **Idle state** (*Idle*): This would be equivalent to a null state in a self-paced BCI. Participants were instructed to think of nothing in particular and to keep their eyes open to prevent generating strong alpha waves.

In terms of covert production, participants were not allowed to imagine tongue, mouth, or any body movements in order to reduce the motor imagery related signals.

(B) Experiment Protocol

Figure 5.4 shows the recording procedure for one trial. There were two different protocols:

- **Protocol A** (*marker-task design*): the participant pressed a key on the keyboard at a time determined by him/herself (within a 100s recording period). Then, immediately after pressing the key, the subject performed either the active or idle state.

- **Protocol B** (*task-marker design*): the participant first executed the active or idle task for a few seconds and then pressed the same key as in protocol A.

The same key was pressed in both protocols in order to eliminate class-dependent information related to the motor task of pushing the key.

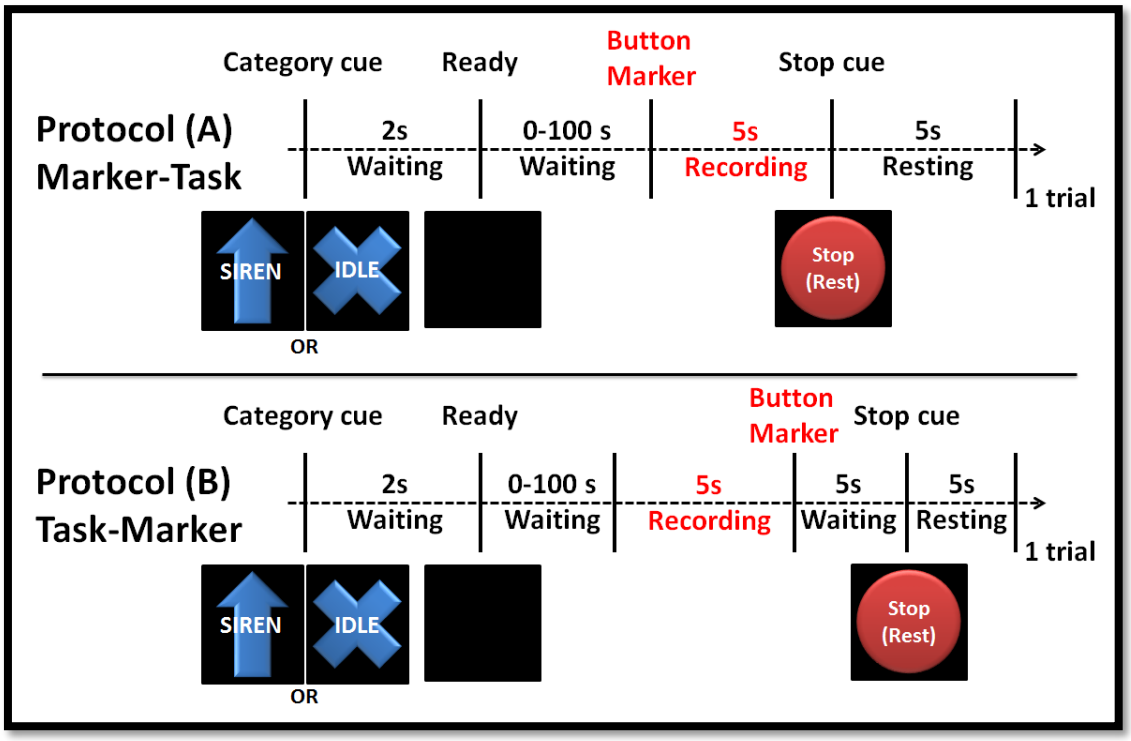


Figure 5.4. EEG data recording protocol for each trial (A: marker-task design, B: task-marker design) [4].

Each experiment had 6 runs, i.e., three for protocol A and three for protocol B. Protocols A and B were mix-ordered (i.e., 'ABABAB' for three subjects and 'BABABA' for two subjects) to minimize sequence-dependent results. In each run, there were 40 trials, which contained 20 siren states and 20 idle states, in random order.

The category cue at the beginning of each trial randomly shows either an up-arrow or an 'X', which indicates active and idle states, respectively. The category cue was presented on a monitor for 2s and disappeared to reduce class-dependent visual event related potentials during the self-paced task. 5s of data post (i.e., protocol A) or prior to (protocol B) the key press were used for epoching.

Compared to our previous study in [5], which used a standard cue-based protocol, the present protocols has the advantage that subjects have freedom (within 100s) as to when to execute the active or idle states. This makes our results more relevant towards onset detection in truly self-paced BCIs.

(C) Data Acquisition & Signal Processing

The experiments were done in accordance with the University of Essex Ethics Committee guidelines.

Five healthy subjects (normal or corrected-to-normal vision, ages 20-25) participated in this experiment. One of the participants had previous experience with BCIs; four participants were naïve subjects. Each subject was sat down on a medical chair and the computer monitor was set 1m away from the participant's face. Based on the 10-20 layout system, 64 electrodes were placed on the head; one reference channel was recorded from the right earlobe. EEG data was acquired via the BIOSEMI Actiview software at 256 samples/s and they were transferred to the PC through TCP/IP communication. Off-line data analysis was performed with MATLAB.

EEG data for classification consisted of one single window pre or post key-press (depending on whether protocol A or B was used). The three different data segment lengths investigated were 1.0s, 2.0s and 3.0s (one single window), as measured from the key-press event. The segmented EEG signals were band-pass filtered (zero-lag

Butterworth filter, order 5) with cut-off frequencies at 4Hz and 120 Hz, and notch filtered (BW filter, order 5) at 49-50 Hz to remove mains interference.

(D) Wavelet Transform

Wavelet transforms offer time-frequency features and they have performed well with non-stationary brain signals [142]. In this experiment, Wavelet Packet Decomposition (WPD) was applied to the pre-processed EEG data. The wavelet decomposition process divides the original signal into low (approximation) and high (detail) pseudo frequency components and down-samples to half of its original sampling rate at each decomposition level [116].

Figure 5.5 shows the WPD tree used in this experiment. The signals were decomposed up to level 5. Initially, each of the 62 nodes (i.e., level_1: 2, level_2: 4, level_3: 8, level_4: 16 and level_5: 32) and from all 64 channels were used for feature extraction and were used individually as inputs for the classifier. The best node and best channels were then determined from the classification results.

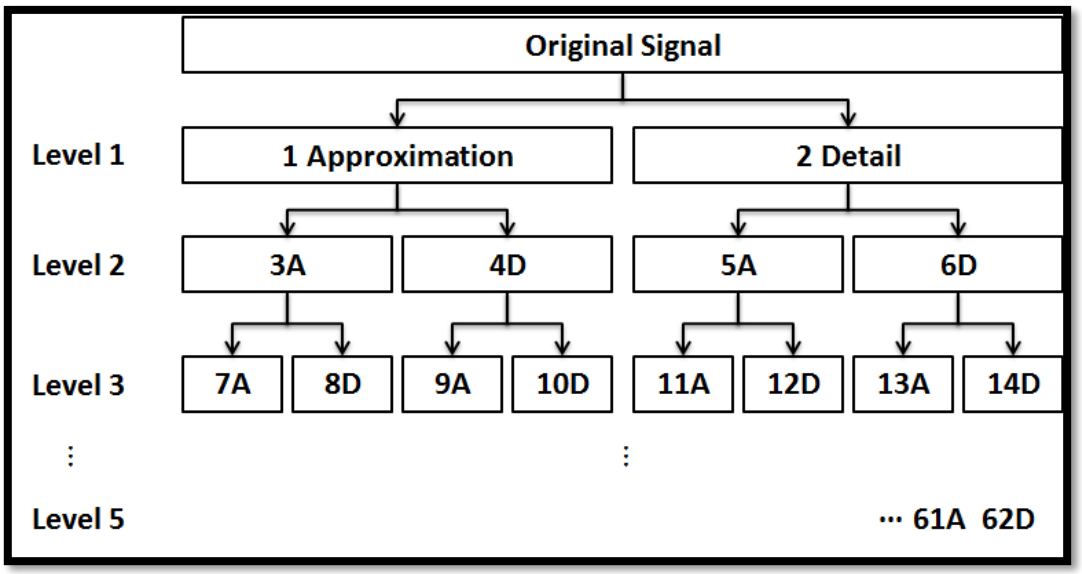


Figure 5.5. Wavelet Packet Decomposition tree [4].

Further, three different wavelets (db2, coif2 and sym2) were employed due to their simplicity and common use in EEG signal analysis.

(E) Classification

A linear Support Vector Machine (SVM) was applied for classification. SVM finds optimal class separating hyper planes by maximizing the margin between classes [123]. The LIBSVM software [143] was used for training and testing the SVMs. The C value was set to default 1 as different C values showed no significant hyperplane changes in our data set. Feature vectors from the wavelet transform were used as inputs to the SVM. Each training/testing run was performed with 20-fold cross validation to improve the reliability of the classification results.

Results

Table 5.3 shows true-positive classification accuracy results (with standard deviations). There are five subjects and each participant recorded 6 runs in total (i.e., three protocol A runs, and three protocol B runs). The (yellow) highlighted results represent the best averaged performance result out of three different wavelets using a 1second time window length.

Figure 5.6 shows the averaged (including all runs, wavelets and window sizes) accuracy results for protocols A and B, respectively. As can be seen from the figure, protocol B performed slightly better than A, except for participant 1. In detail, the Wilcoxon test p values were 2.6e-04, 0.001, 0.03 and 1.8e-04 for participants 1, 2, 4 and 5, respectively, which means that there are significant differences between the protocols for these participants. Differences in protocol-dependent results were not found to be significant for participant 3 ($p = 0.137$). So, in short, these results show that protocol B gives better classification results for most participants.

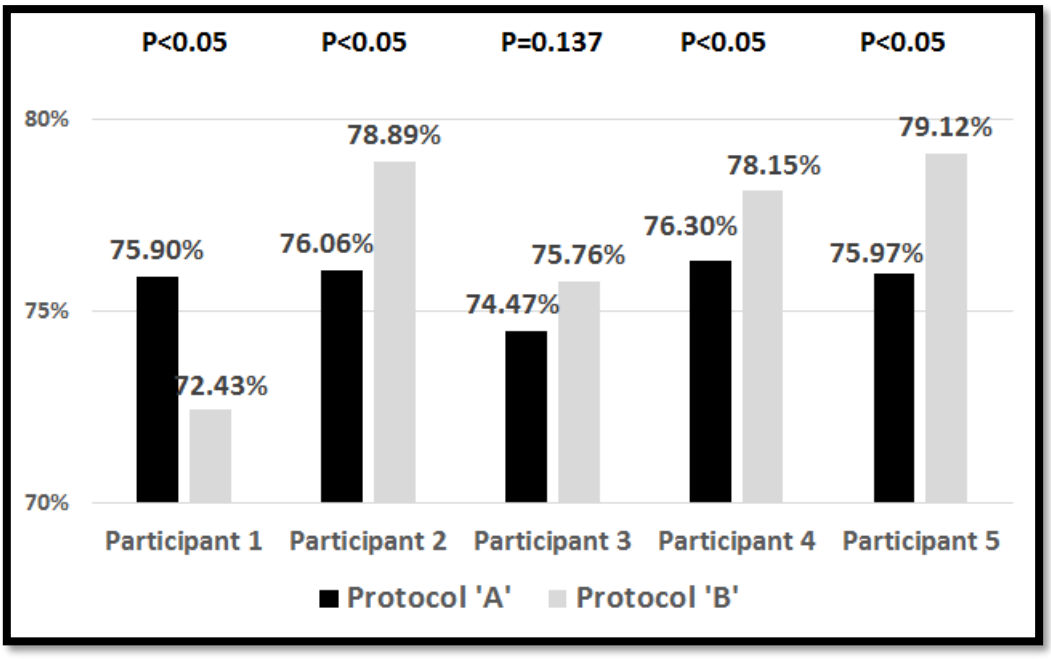


Figure 5.6. The averaged performance results from each participant with different protocol [4].

Table 5.3. Performance results for each participant [4].

		Time Windows								
		1.0 s			*2.0 s*			*3.0 s*		
		Wavelet			Wavelet			Wavelet		
		db2	coif2	sym2	db2	coif2	sym2	db2	coif2	sym2
		Accuracy % (± Standard Deviation *σ)*								
Participant 1	**A Run1**	71.25 (±10.8)	74.37 (±12.5)	73.75 (±10.7)	71.88 (±15.1)	73.13 (±18.7)	70.63 (±14.2)	71.88 (±11.4)	71.87 (±14.5)	72.50 (±13.2)
	A Run2	76.88 (±13.0)	77.50 (±14.4)	76.25 (±11.4)	77.50 (±8.67)	73.75 (±8.98)	80.00 (±11.0)	80.63 (±10.3)	79.38 (±10.9)	78.75 (±12.9)
	A Run3	74.38 (±15.4)	77.50 (±14.4)	77.50 (±12.6)	75.63 (±13.1)	76.88 (±13.6)	78.75 (±15.2)	80.63 (±13.1)	76.88 (±10.9)	79.38 (±8.38)
	B Run4	72.50 (±11.9)	70.00 (±12.4)	71.88 (±14.5)	63.75 (±14.6)	68.75 (±15.4)	67.50 (±14.3)	68.75 (±13.7)	66.88 (±13.0)	75.00 (±11.4)
	B Run5	81.88 (±11.1)	81.25 (±11.8)	78.75 (±10.0)	82.50 (±8.50)	70.00 (±14.2)	71.25 (±13.5)	76.25 (±13.4)	72.50 (±11.2)	78.75 (±16.7)
	B Run6	76.25 (±17.6)	76.25 (±16.2)	70.62 (±13.6)	70.63(± 12.9)	71.88 (±15.1)	73.75 (±12.1)	66.2 5(±16.2)	66.88 (±18.2)	65.00 (±14.9)
Participant 2	**A Run1**	73.75 (±15.1)	66.88 (±12.9)	76.88 (±14.2)	71.88 (±10.6)	76.88 (±15.3)	74.38 (±15.9)	75.63 (±13.1)	67.50 (±15.9)	71.25 (±15.2)
	A Run2	76.25 (±14.5)	80.00 (±11.7)	73.13 (±14.7)	78.13 (±15.1)	72.50 (±7.69)	80.00 (±13.0)	75.63 (±17.4)	80.62 (±13.1)	71.25 (±14.6)
	A Run3	76.25 (±16.7)	80.00 (±11.0)	76.25 (±10.6)	79.38 (±11.6)	77.50 (±13.8)	77.50 (±14.9)	82.50 (±17.4)	81.88 (±14.3)	80.00 (±14.2)
	B Run4	84.37 (±11.4)	78.75 (±13.5)	82.50 (±13.0)	79.38 (±10.9)	80.63 (±13.1)	79.38 (±13.6)	75.00 (±13.4)	81.88 (±11.8)	78.75 (±11.5)
	B Run5	75.63 (±13.1)	82.50 (±14.8)	80.00 (±15.3)	75.63 (±11.8)	75.63 (±12.5)	70.63 (±11.6)	73.75 (±15.1)	73.75 (±13.4)	71.87 (±16.7)
	B Run6	78.13 (±11.3)	77.50 (±11.9)	75.63 (±12.4)	80.00 (±14.8)	83.13 (±10.1)	82.50 (±14.8)	86.25 (±12.0)	87.50 (±9.06)	79.37 (±12.9)
Participant 3	**A Run1**	73.75 (±12.7)	75.00 (±14.0)	81.88 (±15.4)	76.88 (±11.7)	75.63 (±15.4)	80.63 (±10.3)	75.00 (±13.4)	74.38 (±15.9)	83.13 (±10.9)
	A Run2	71.88 (±18.5)	68.75 (±12.5)	81.88 (±12.5)	72.50 (±13.8)	66.25 (±16.2)	71.25 (±17.2)	69.38 (±13.7)	71.88 (±13.4)	70.00 (±12.4)
	A Run3	69.38 (±19.2)	78.75 (±13.5)	66.25 (±13.5)	78.13 (±13.9)	76.88 (±10.9)	72.50 (±11.1)	78.13 (±11.3)	75.00 (±14.0)	75.63 (±11.1)

	B Run4	76.88 (±12.9)	75.63 (±11.8)	71.88 (±16.2)	76.88 (±12.3)	82.50 (±11.7)	74.38 (±14.3)	73.75 (±10.6)	71.25 (±14.7)	72.50 (±16.0)
	B Run5	71.88 (±13.3)	75.63 (±15.9)	73.75 (±17.6)	71.87 (±13.9)	79.38 (±11.6)	75.00 (±12.8)	86.25 (±11.4)	71.88 (±17.1)	85.63 (±11.6)
	B Run6	71.88 (±13.4)	64.37 (±14.2)	75.63 (±13.7)	83.13 (±10.9)	78.13 (±12.7)	80.63 (±15.4)	76.25 (±9.85)	71.25 (±16.2)	77.50 (±13.8)
Participant 4	**A Run1**	80.63 (±11.8)	80.00 (±10.3)	76.25 (±12.7)	81.25 (±11.1)	75.63 (±16.5)	78.75 (±12.2)	73.75 (±13.4)	75.00 (±12.1)	72.50 (±12.5)
	A Run2	66.25 (±19.9)	77.50 (±12.5)	77.50 (±17.0)	70.63 (±17.3)	78.13 (±9.83)	75.63 (±14.3)	71.88 (±16.7)	73.13 (±14.7)	76.25 (±14.5)
	A Run3	79.38 (±11.7)	83.13 (±14.7)	77.50 (±14.4)	80.00 (±11.7)	75.63 (±15.9)	75.00 (±14.0)	75.00 (±12.8)	76.88 (±14.2)	76.88 (±12.9)
	B Run4	76.25 (±16.2)	73.75 (±15.6)	83.13 (±11.6)	70.63 (±10.1)	75.00 (±12.8)	78.13 (±12.0)	77.50 (±12.5)	75.00 (±17.2)	69.38 (±13.1)
	B Run5	71.88 (±12.1)	83.75 (±10.0)	78.13 (±13.3)	88.13 (±9.49)	80.63 (±11.0)	81.25 (±17.4)	81.25 (±13.8)	79.38 (±10.9)	79.38 (±15.3)
	B Run6	81.88 (±11.1)	79.38 (±15.3)	73.13 (±12.3)	76.25 (±10.6)	76.88 (±16.8)	80.00 (±13.7)	-	-	-
Participant 5	**A Run1**	85.63 (±10.9)	77.50 (±10.4)	89.38 (±11.6)	78.13 (±16.1)	81.88 (±11.8)	82.50 (±13.1)	75.63 (±12.5)	80.00 (±13.7)	76.88 (±10.1)
	A Run2	83.75 (±11.5)	75.63 (±15.4)	77.50 (±12.6)	76.88 (±14.7)	80.63 (±11.0)	76.88 (±14.2)	80.63 (±9.49)	70.00 (±16.9)	73.13 (±13.6)
	A Run3	68.75 (±14.9)	71.25 (±14.7)	67.50 (±14.2)	74.38 (±11.1)	69.38 (±11.8)	70.63 (±14.8)	69.38 (±13.1)	65.63 (±9.83)	71.88 (±13.3)
	B Run4	89.38 (±10.1)	89.38 (±11.6)	84.38 (±12.1)	84.38 (±10.6)	87.50 (±11.5)	81.88 (±13.1)	86.88 (±9.49)	86.88 (±13.1)	86.25 (±8.00)
	B Run5	81.25 (±14.9)	73.75 (±11.4)	78.75 (±15.2)	63.75 (±8.00)	71.25 (±15.7)	75.00 (±19.4)	76.88 (±15.3)	78.13 (±15.6)	78.75 (±14.6)
	B Run6	76.25 (±12.7)	76.25 (±18.1)	85.00 (±11.2)	75.63 (±15.4)	77.50 (±12.5)	71.88 (±15.1)	70.63 (±13.6)	78.13 (±10.7)	70.63 (±12.3)

A, B run*: The recording protocol either 'A' or 'B', which are explained in the methodology section.*

Yellow (highlighted) results*: Accuracy from one channel (gave highest accuracy) in 1s window for each participant.*

** Participant 4: B run6, 3.0s time window size were removed due to miss-pressed keyboard marker input*

In terms of data segment sizes, 1s, 2s and 3s segments were compared. Figure 5.7 shows averaged true positive accuracy values including all wavelet type and both protocol types, for each subject.

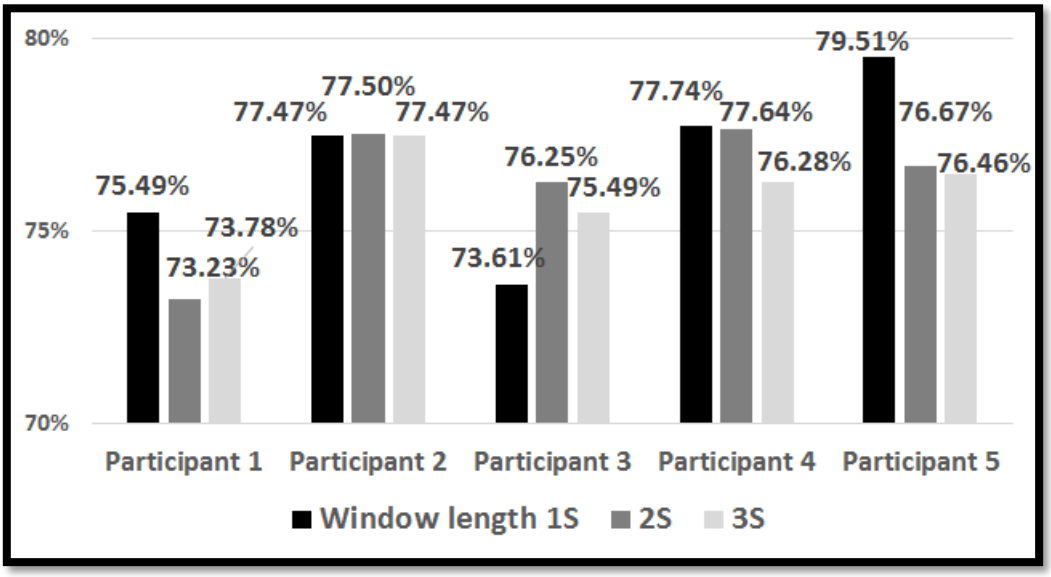


Figure 5.7. The averaged performance results from each participant with different time windows [4].

As can be seen from the Figure 5.7, the 1s second time window segment worked best for 3 of the 5 participants. But most of the time the differences were insignificant. In detail, the *p* values between the 1s and 2s window conditions were 0.04, 0.95, 0.02, 0.85 and 0.006 for participants 1, 2, 3, 4 and 5 respectively. For 1s vs. 3s windows, the *p* values were 0.11, 0.73, 0.08, 0.15 and 0.002. This shows that only subject 5 gave significantly better results with the 1s window. In addition, for the 2s vs. 3s comparison, *p* values were all higher than 0.05. Therefore, there is significance in some cases but not consistently. From this result, the 1s window length can be regarded the most efficient window as it can minimise computational power with small window size and increase system speed compared to longer windows. For this reason, the highest

accuracies were chosen from the results for the 1s window (yellow highlighted). The best true positive rates were 76.88%, 79.58%, 76.67%, 80.2% and 82.71% for participant 1, 2, 3, 4 and 5, respectively.

In terms of electrode location, for participant 1, P5, TP8 and P1 worked the best (yellow highlighted part). For participant 2 channel P8, P6 and P8; for participant 3: FC8, T7 and FT7; for participant 4: CP2, CP6 and TP8; and, for participant 5: AF8, P9 and F2 were found to work the best from the runs with the 1s time window length.

Brain activity related to covert siren sound production appeared all around the head, not only at Broca's and Wernicke's areas. This shows that the physiological location for the covert siren sound production is difficult to identify.

Discussion

This experiment investigated the potential of a new method for onset detection towards asynchronous BCIs. Siren-like sound covert production and recall were classified against the idle (no task) state in an off-line system.

Wavelet packet decomposition was employed for feature extraction and the wavelet type slightly affected the performance results. The best decomposition level and node for each participant was different. For participant 1, the nodes with the highest classification rates were 61, 44 and 40 (the node numbering system can be found in Figure 5.5 above), respectively for the three runs (highlighted in yellow). For participant 2 the best nodes were: 19, 61 and 29; for participant 3: 52, 54 and 62; for participant 4: 55, 54 and 25; and, for participant 5: 6, 11 and 57. The best wavelet varied depending on the participant, window size and protocol type. This suggests that optimised wavelet design should be used. In [144], a simple evolved pseudo-wavelet was introduced. The wavelet sample coefficient was evolved to be optimised to the

specific EMG or EEG signal by a genetic algorithm. In [118], the P300 EEG signal classification rate was improved by around 10% and 12% with the use of evolutionary wavelets. This method could also be used in order to improve performance.

From a performance point of view, five subjects achieved a 76.88%, 79.58%, 76.67%, 80.2% and 82.71% true positive rate, respectively. This showcased competitive results compared to systems based on typical motor-imagery tasks [10, 18]. Therefore, there was the potential of testing in a real online situation for further analysis so we repeated this experiment in real online onset detection settings.

5.4 Summary and Conclusions

The scope of the first experiment was to investigate the use of basic and intuitive speech production related tasks towards their future use in onset detection of asynchronous BCIs. The *Idle* states (i.e. relaxed, non-specific states) were separated against the *Covert*, *I_Overt*, and *Overt* high tone speech sound production states. The *Covert* state produced the best classification rate vs. the *Idle* state (82.41%, 81.20%, 85.12% and 74.72%, respectively) compared to the *Overt* vs. *Idle* state for all participants ($p \ll 0.05$). The *Covert* state also provided better classification results vs. the *Idle* state than the *I_Overt* state for participants 1, 2 and 3, and equally good results with the *I_Overt* state for participant 4.

The best window length values (as measured from the time when the 'start' cue appeared) were 1.5 s for participants 1 and 3, 2 s for participant 4 and 3 s for participant 2, providing a hypothetical maximum bit transfer rate of 32.95, 16.24, 34.05 and 22.42 bits/minute, respectively, for each participant.

Including CSP in the channel selection process produced worse results than skipping the CSP stage and choosing instead the best channel based solely on the classification performance by using individual channels.

The goal of the second experiment was to investigate the use of covert siren sound production in order to distinguish active states from the idle state for onset detection in asynchronous BCIs. Time window segments of 1 s, 2 s and 3 s were tested with two different protocols (A and B) and the results showed that protocol B (*task-marker design*) provided better results than protocol A (*marker-task design)* for most subjects. However, different window segment sizes did not show significant differences. For this reason, the highest accuracy results were chosen from the 1 s window configuration which achieved true-positive performance values of 76.88%, 79.58%, 76.67%, 80.2% and 82.71% for participant 1, 2, 3, 4 and 5, respectively.

Based on these two experiments, speech-related cognitive tasks showed that they can be distinguished from the idle state for an onset detection system in BCIs. In addition, the performance results were very promising in offline settings. For this reason, further experiments were carried out with the speech-related cognitive tasks for an onset detection system, which will be discussed in the following chapters.

The findings in this chapter showed promising results compared to other motor imagery onset detection studies (e.g., [10, 18]). These new sound-production related cognitive task experiments (offline cue-based) could provide useful background information for other BCI studies.

6 Onset Detection Technique for Brain-Computer Interfaces using Sound-production Related Cognitive Tasks in Simulated-online System (Based on a paper [1])

6.1 Introduction

This chapter is based on a published journal paper [1].

In the Chapter 5, new novel high-tone and siren-like sound production related cognitive tasks were proposed and investigated in order to classify them against the idle state for onset detection in offline settings which showed some potential.

Before moving this investigation (offline, cue-based) forward to online onset detection BCI systems and choosing the most appropriate speech related cognitive tasks, simulated-online settings with a self-paced experiment paradigm were tested in this chapter as a prototype towards a practical online system. High-tone and siren-like sound production related cognitive tasks were tested and compared to covert and inhibited overt speech modes.

In addition, a new performance evaluation metric for self-paced BCIs, called the true-false-positive score, was proposed in this research (details in section '*Performance*

Assessment Score') as there is no common and standardised performance assessment method of self-paced BCI systems that takes into account the idle period length, which is a very important aspect of self-paced system.

Averaging the results from the best performing IC tasks for all seven participants, a 77.7% True-Positive (TP) rate was achieved in an offline testing. For the simulated-online analysis, the best IC average TFP score was 76.67% (87.61% TP rate, 4.05% false-positive rate). The results were very promising compared to other motor imagery onset detection studies, which reported the values of 72.0% and 79.7% as their best TP rates and which, crucially, did not take timing errors into account [10, 18]. Therefore, this new sound imagery onset system showed good potential to be used in self-paced BCIs for further online real-life setting investigation (as will be discussed in Chapter 7).

6.2 Methodology

Sound-production Related Tasks and Idle State Definition

In this experiment, there were two different mental tasks for the onset switch, and two modes for each task. Firstly, the modes are separated as in inhibited overt (*IO*) and covert (*C*) sound production. Secondly, high tone (*High*) and siren-like (*Siren*) sound production mental tasks were tested. For the non-control state, idle (*Idle*), i.e., non-specific states were also recorded (to avoid confusion, the term '*idle*' alone will be used in the remaining parts of this paper). The start and duration of the tasks was controlled spontaneously by the user (assisted by a specially designed time-keeping interface,

described below). To minimise artefacts generated from muscle signals, participants were instructed to avoid any unnecessary body movement, but they were still allowed to blink or move their eyes when needed (the artefact rejection methods are explained later in this paper).

In more detail, the states were defined as follows:

- **Inhibited overt sound-production Tasks:**

  Inhibited overt sound-production is different from our normal overt sound-production. Aside from the cognitive effort, it will involve tensioning of the vocal cords but there is no actual sound production that can clearly be heard.

  ***Inhibited overt high tone production (IO_High)*:** participants were instructed to produce an 'um' sound effort with a high pitch that they can comfortably produce for a couple of seconds, but high enough that they think it is an unusual tone and not something they would imagine in a normal situation.

  ***Inhibited overt siren-like sound production (IO_Siren)*:** the siren-like sound effort was defined as 'wee-woo wee-woo'. 'Wee' syllable denotes high notes, whereas 'woo' expresses low pitch. Participants were instructed to produce this sound effort for a couple of seconds.

- **Covert sound-production Tasks:**

  Covert sound-production was a pure imagination process. Thus, there should be no tensioning of any organs related to sound-production. Participants were instructed to imagine making the 'sound', which of necessity included imagining hearing the sound (auditory imagery / auditory recall). Auditory imagery refers to mental imagery in sound perception without an actual external auditory stimuli [145]. In terms of functional neuroanatomy the processes involved in covert

speech have not be fully elucidated, but it is known that it involves the auditory cortex (around Brodmann areas 41, 42 and partially 22 [146]) and, for speech-related imagery, Wernicke's area (Brodmann area 22). Also, the auditory system has been shown to play an important role in overt speech production by giving internal feedback [147], it is possible that a similar role is played in covert sound-production.

***Covert high tone production (C_High)*:** Imagining high tone production (as explained above for the *IO_High* task).

***Covert siren-like sound production (C_Siren)*:** Participants were instructed to imagine making siren-like sounds in covert mode.

- **Idle state (*Idle*):** This is a non-control or null state. The participants were instructed to not think of any of the above IC task states and to stay calm and relax.

During all above tasks, participants were not allowed to imagine tongue, mouth, lips, or any other body movements to avoid motor-imagery related signals.

Experiment Interface Design

While the tasks were controlled spontaneously by the users, it was necessary to provide them with a means to estimate the length of time gone by when executing a task in order to ensure that the IC task lasted sufficiently long to yield enough data to achieve a high classification rate, but not so long that it would lead to such high timing errors as to render the self-paced approach useless. Having in mind the typical task duration in cue-based BCIs, we chose an approximate recommended task duration of

3s, but bear in mind that **the user was still free to start and stop the task at any time that suited them, within a 30s window** (the maximum duration of each NC state within a trial so that the experiments did not run for an unnecessarily long time).

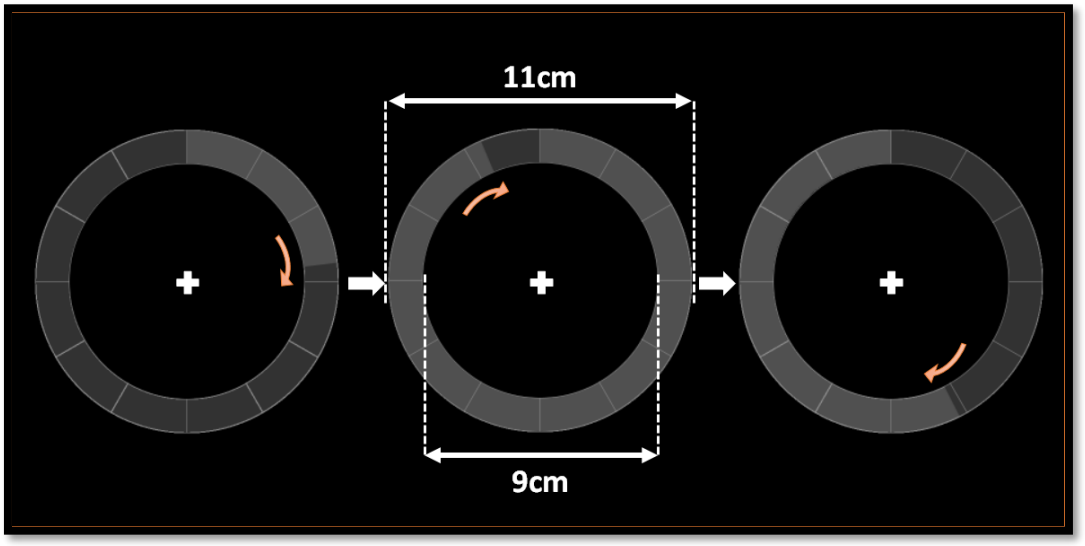


Figure 6.1. The chosen time-keeping interface design. Users fix their eyes on the central cross and estimate their task time as the light grey progress bar grows clockwise [1].

To record onset tasks and idle states for simulated-online scenarios (i.e., treating the data trials sequentially rather than as independent random trials), the time tracking interface needs to be suitable for actual online self-paced onset detection systems even during recording of the training. Thus, there were three main functional requirements: a) The interface should minimise visual event-related potentials (VEP). b) The computer must be able to time-stamp events. And, c) as explained above, the user must be able to estimate task duration. To satisfy these requirements, a few different recording interfaces were considered as candidates and the circular progress bar interface shown here was chosen based on the facts of 1) minimum eye movement, 2)

minimum ERP generation, and 3) usability (defined as ease of use in this study) from three experienced BCI users (i.e., PhD students in our BCI group). To determine the size of the interface, we considered two literature sources. In [148] competing stimuli located less than 5° of visual angle from the central stimulus were shown to affect SSVEP responses. In [149], similar effects were observed in a P300-based BCI. As a result, to avoid these proximity issues, the diameter of the interface's inner circle was set to 9cm and the distance between the monitor and participants was set to 50cm. This leads to about 10° of viewing angle. The viewing angle from the fixation cross to any circular moving object was just above 5°. In addition to this, background and objects colour were chosen to be dark achromatic colours to minimise ERPs. As shown in Figure 6.1, the progress bar in the interface continuously filled with light grey for 12 seconds and then with slightly darker grey (the jump in brightness was small to minimise VEP), followed by light grey again.

Experimental Protocol

Participants performed one run for each task, chosen pseudo-randomly to minimise sequence-dependent effects (randomisation between runs). In each run, participants executed the same task 30 times. They knew which task to perform as they were told about the task, by the experimenter, before each run. Task randomisation within a run was unnecessary and undesired in our case as this is only relevant in a multi-task scenario (e.g., motor imagery for left hand vs. right hand vs. feet vs. tongue, etc.). In our case, on the other hand, the intended task-versus-idle scenario is one in which the end-user would execute the same imagery task every time. I.e., in an onset-detection problem it would make no sense to mix the tasks, as this is not what will be happen in online use.

Between each 30-trial run, participants had short breaks (1-3min, as desired). The total experiment time did not exceed one hour beyond electrode cap set up and explanation of the experiment to the participant.

The experiments were done in accordance with the University of Essex Ethics Committee guidelines.

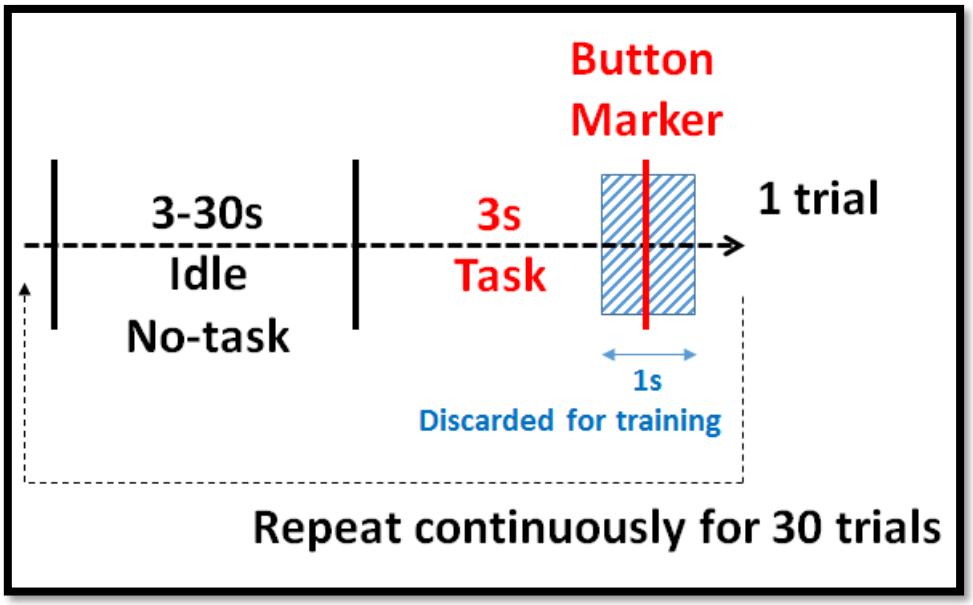


Figure 6.2. EEG data recording procedure for 1 trial [1].

Figure 6.2 represents the recording procedure for one trial. Users were required to stay in the idle state for at least 3s, after which they were free to execute one of the cognitive tasks at any time up to 30s from the beginning of the trial. Immediately after they executed a task for about 3s (aided by the time-keeping interfaces) they were required to press the space key on the keyboard to signal the end of a trial and to provide a time stamp for performance evaluation of the system. The minimum idle state of 3s was chosen to prevent task time-proximity effects in the EEG. On the other hand, the

maximum idle state 30s was chosen based on a previous study [4] that explored different ways of time-stamping active states and in which participants were given a window of up to 100s within which to execute the self-paced task. In that study all participants spontaneously executed a given active task within 15s after a trial began. For this experiment, an extra 15s were included to prevent participants from rushing.

During the whole experiment the same key (the space key) was pressed following a self-paced task to prevent class-dependent information from any motor-related signal. In addition, data 0.5s prior to and 0.5s after the space key (shaded area is Figure 6.2) were discarded from the analysis to avoid motor-imagery related data leading to IC false-positives.

Seven healthy subjects (4 males, 3 females) participated in the experiments. They all had normal or corrected vision and were aged between 22 and 27. Three participants had previous experience with BCIs and two of them had participated in a previous study on covert sound production for onset detection [4]. The other four participants were naïve subjects. Each subject was sat on a medical chair comfortably and a monitor was placed 50cm away from subject's face. A keyboard was placed on their lap to press the space bar for the end-of-trial marker.

Offline and Simulated-online Evaluation Definition

In this experiment, the recorded data were analysed in offline and simulated-online scenarios, as follows.

***Offline evaluation*:** The continuously recorded EEG data was segmented into 0.5s time windows without any overlapping. Then, these segments were separated into task and idle states based on the timing protocol shown in Figure 6.2. If a segmented 0.5s window included both idle and task states, it was discarded, as were the 0.5s before

and after the key-press stamp. After segmentation, half of the epochs for each state were randomly selected for training and the other half were used for testing data. The randomisation-training-testing cycle was repeated 20 times. Offline evaluation gives a preliminary idea about how well the system can distinguish active tasks from idle states for onset detection and the results can more easily be compared to other BCI systems. However, the offline evaluation has drawback towards real onset detection system as it ignores sequence effects (such as possible priming, habituation, etc.) of onset tasks.

***Simulated-online evaluation*:** Data segmentation was done as in the offline study, but with two crucial differences: a) no data windows were discarded unless EOG was automatically detected by the system (using the EOG detection algorithm described below), and b) epoch randomisation was not applied in order to preserve the online-like time structure of the data. Instead, the first 15 trials (half of the recorded trials within a run) were used for training; the subsequent 15 trials were used for testing. Data were not discarded in the manner done in the offline approach because in real online situations there is no end-of-trial marker.

Data Recording and Signal Pre-Processing

A Biosemi (TM) ActiveTwo system was used with the Actiview software for recording data. 64 electrodes were placed based on 10-10 layout system and 2 reference electrodes were placed on the left and right earlobes. In addition, 1 electrode was setup to detect EOG artefacts. Sampling rate 512 samples/s was chosen to ensure recording up to the high gamma band (100 Hz) based on 3dB-point (half power point) of the equipment bandwidth around 104 Hz. In BCI studies, high gamma waves have not been investigated very often due to increased contamination by muscle artefacts, but

previous work by others [65, 67, 150] has shown significant high gamma wave activity associated with some language tasks, hence its inclusion here. On the other hand, recording at a higher rate was not necessary as our interest in EMG was only for artefact removal purposes and, further, sampling at a higher rate could have led to increased EMG-related aliasing in the EEG signals.

Continuously recorded EEG data were segmented with 0.5s window length. The data were band-pass filtered (zero-lag Butterworth filter, order 4) with cut-off frequencies at 2 Hz and 100 Hz. Then a notch filter (zero-lag Butterworth filter, order 4) was applied at 49-51 Hz to reduce mains interference. To remove common environmental noise, the averaged of the two earlobe reference channels was subtracted from all 64 scalp channels.

EOG Artefact Detection

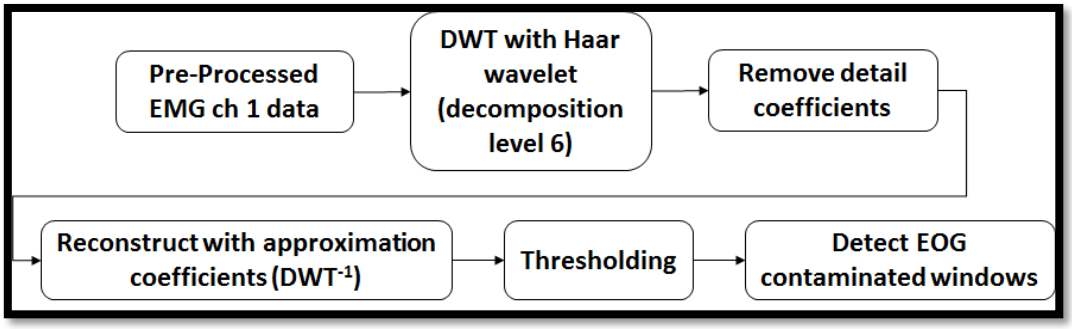


Figure 6.3. Block diagram of EOG artefact detection method [1].

An EOG channel was placed above the corrugator muscle and was used for EOG detection at the forehead region. Figure 6.3 illustrates the procedure for automatic EOG detection. A discrete wavelet transform (DWT) with Haar mother wavelet (because it

resembles eye blink ocular artefacts [108]) was applied to the EOG channel. The decomposition level, 6, was chosen as it showed satisfactory results in [108, 129]. The pseudo-frequency of the level 6 approximation component was 0-8Hz in our case.

In the DWT process, detail parts were removed and then the signal was reconstructed. This is because the papers [108, 129] showed that the EOG artefacts usually appear in lower frequency range (around 0-13 Hz). Thus, one would expect to find the strong wavelet coefficients in the approximation part. For this reason, all of the detail parts were removed, and the signal was reconstructed only with approximation coefficients. The staircase waveform result showed clear EOG artefact detection in their cases as well as in our study (showed in Figure 6.4).

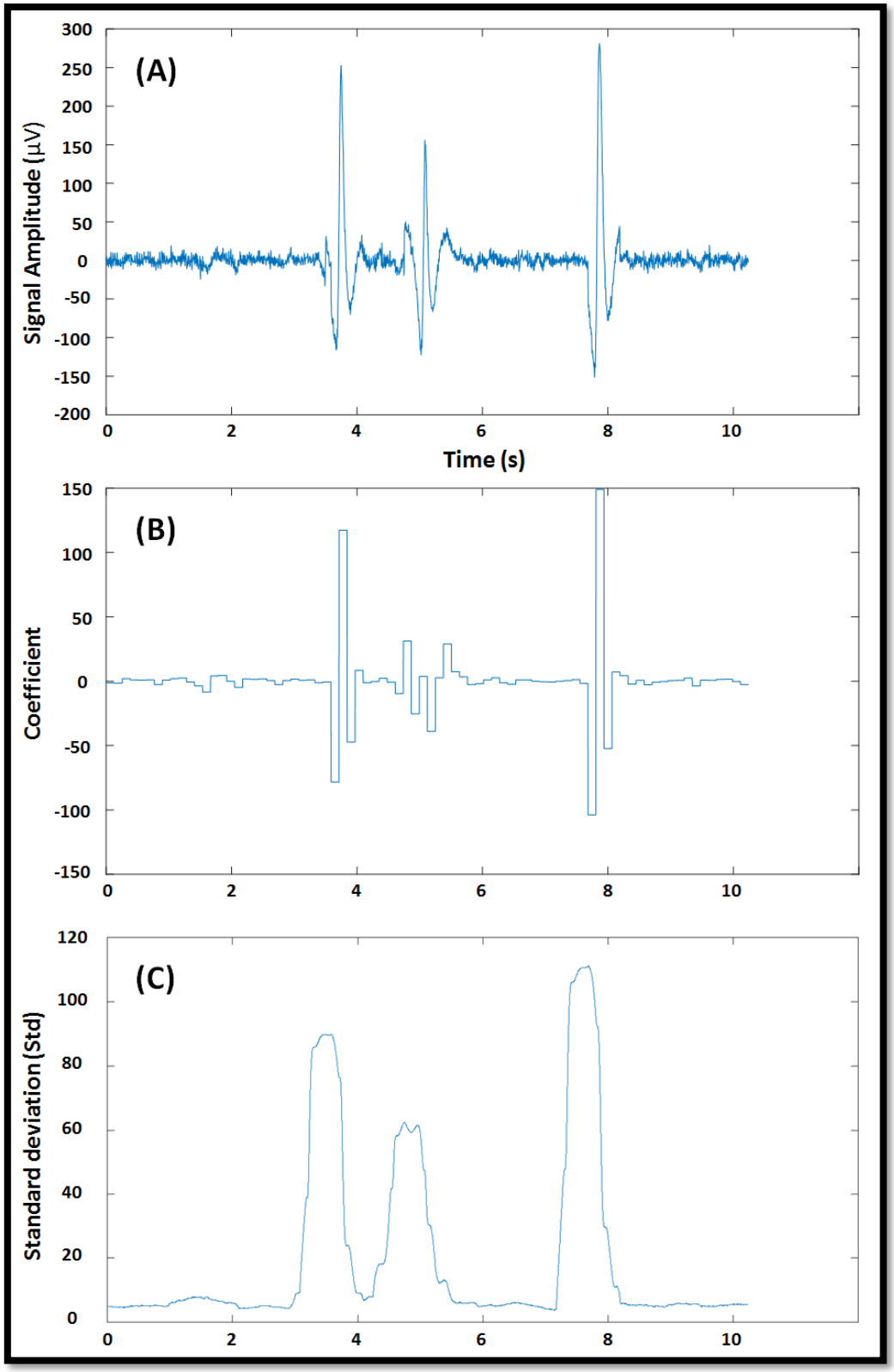


Figure 6.5. Participant 1's first 10s data (A) Pre-processed EOG channel. (B) EOG artefact detection process applied with wavelet transform. (C) Standard deviation of 0.5s data from (A) [1].

To detect EOG artefacts, two conditions needed to be met: i) a standard deviation (**std**, calculated for each from 0.5s non-overlap window segment) jump by a factor of 3, and ii) using a wavelet coefficient threshold, as follows. If we compare Figure 6.5A and B, the EOG detection plot (B, based on the wavelet coefficients at decomposition level 6) can be seen to have large rising/falling edges. When the standard deviation (**std**) was found to jump by a factor of 3, the subsequent data were treated as possible EOG artefact candidates. Within the EOG artefact region, the smallest rising/falling step area was chosen as a threshold in order to avoid discarding false EOG positives that may result from applying only the **3std** condition. E.g., in Figure 6.5B, between 5s and 6s we find a pattern that can be deemed to be border line EOG artefact and, within that region, the smallest step is 20μV. This value was half powered (-3db) and the result was chosen as a threshold. To reduce onset false positives, once the EOG artefact contaminated time-locations are detected, the data for those segments were discarded from further analysis.

EEG Feature Extraction

In order to analyse the EEG signal, two different feature extraction methods were used, band power and wavelets. For the band power a Fast Fourier Transform was applied to the pre-processed EEG signals and its power (i.e., the squared FFT) were selected as features from eight different frequency ranges; Freq1: 2-4Hz (Delta), Freq2: 4-8Hz (Theta), Freq3: 8-12Hz (Alpha), Freq4: 12-16Hz (Low Beta), Freq5: 16-20Hz (Beta), Freq6: 20-30 (High Beta), Freq7: 30-42Hz (Low Gamma) and Freq8: 42-100Hz (High Gamma).

The second feature extraction method was the discrete wavelet transform. It offers time-frequency features and performs well with non-stationary brain signals

[142]. Pre-processed EEG signals were decomposed and their coefficient vectors from levels 6Approximation, 6Detail, 5D, 4D, 3D and 2D (representing the pseudo frequency bands Wave F1: 2-4Hz, Wave F2: 4-8Hz, Wave F3: 8-16Hz, Wave F4: 16-32Hz, Wave F5: 32-64Hz and Wave F6: 64-100Hz respectively) were calculated and their variances (for dimensionality reduction purposes) were used as features. The mother wavelet 'db2' was chosen because of its simplicity and common use in EEG signal analysis [151-153] (also, in our previous study [4] we found that the choice of wavelet type - db2, coif2, or sym2 - did not significantly affect sound-production related onset detection). While it is possible that an extensive study including various other wavelet types and orders (and, for that matter, other JTFA and non-JTFA approaches) could lead to improved results, our study was meant to focus on the use of covert sound-production in onset detection.

Classification

The above feature extraction method produced hundreds of features, i.e., (64ch*7band power + 64ch*5wavelet = 768 features), so feature selection had to be applied to reduce feature set size and class overlap, and to improve computational efficiency. To this end, the Davies-Bouldin index (DBI [119, 120]) was applied. The DBI is a cluster overlap measure. Smaller DBI values indicate better class separation, with lower class overlap and larger distance between classes. Thus, DBI values for each feature were sorted in ascending order and an integer value DBI threshold from 1 to ***N*** was obtained for each participant. The features which had DBI value less than the threshold were selected as a feature set for further analysis. The DBI threshold was chosen as follows.

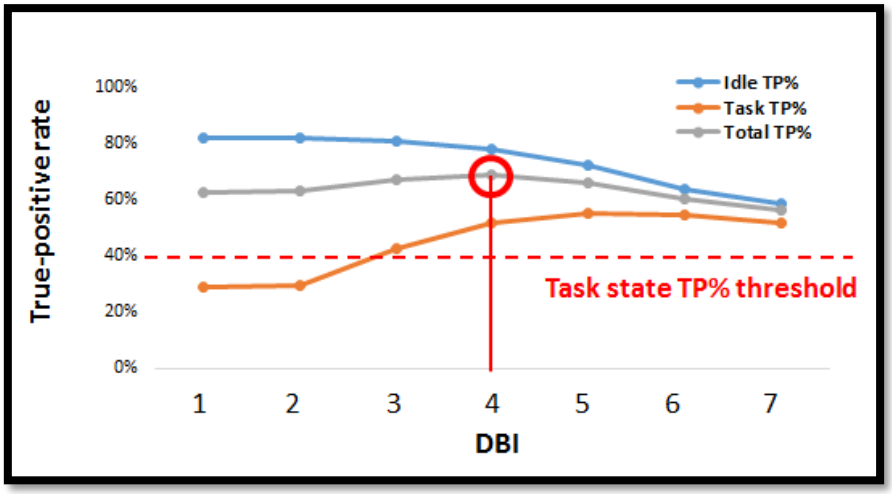


Figure 6.6. Sample training true-positive rates for idle, task periods, and total performance (from participant 1, inhibited-overt siren task). The horizontal axis shows 7 approximate DBI values for illustration purposes [1].

The DBI threshold was chosen based on the training set's classification result (see Figure 6.6). Due to the different sizes of the idle and task states (the idle period is much longer than tasks), classification results could be biased towards the idle state (see points DB=1 and 2 in the figure). By increasing the DBI threshold (e.g., from 2 to 3), the task state's true-positive rate increases while the idle true-positive rate decreases. This behaviour continues until the individual TP rate continuously decreases for both idle and task states. However, in every case there is an optimum DBI value at which the overall TP rate is maximised (e.g., at DBI=4 in Figure 6.6). Thus, the DBI threshold was chosen so that it gave the highest overall true-positive rate for the training data.

After feature selection was performed, Linear Discriminant Analysis (LDA) was applied for classification. LDA was chosen due to its simplicity and low computational power [122] as well due to its widespread use in BCI research. The feature vectors from the feature selection process were used as inputs to the LDA. For the offline analysis,

pseudo-randomisation of the choice of training and testing set epochs was done 20 times and results obtained for each randomisation stage.

EMG Artefact Handling

A challenge in all BCIs, but more so when the gamma band is included in the analysis, is to ensure that classification results are based on brain signals alone, as much as possible, and are not contaminated by potentially class-dependent EMG. In an EMG artefact BCI survey [130], it was shown that 67.5% of the BCI studies included in the survey did not mention whether they handled EMG artefacts or not and 12.1% did not remove EMG artefacts.

EMG artefacts are particularly important for IC state onset detection as switching from an 'idle' state to an IC state may produce involuntary facial twitches that can produce class-dependent EMG artefacts, especially in frontal area EEG, more so than when switching between various IC states. EMG (and other facial artefacts) must thus be minimised.

Independent component analysis (ICA) and blind source separation by canonical correlation analysis (BSS-CCA) are the two mostly used EMG removal techniques in BCIs. Research papers [19, 98] showed BSS-CCA outperformed ICA and it was more suitable for EMG removal thus BSS-CCA was chosen for this experiment.

CCA measures the linear relationship between two multi-dimensional signals [131]. It can be used to solve BSS problem (proposed in [132]) by taking multi-channel EEG as a first variables and temporally delayed version as a second variables [97]. The threshold of autocorrelation coefficient ρ was chosen as 0.35 based on the study in [133]. If there was no source that has less than the threshold ρ, the last source (from descending order sort) that has the lowest autocorrelation coefficient was removed.

Performance Assessment Score

For the event by event performance evaluation, the true-false difference rate was suggested in [154] for self-paced BCIs. However, there are some issues with this approach. Due to the difficulty in measuring true-negatives during idle state, [154] proposed a false-positive rate as ***'FP/(E+FP)'***, where FP is the number of false-positives and E is the number of task state onset events. This false-positive rate was subtracted from the true-positive rate. However, the number of task events and idle events are independent in self-paced system. Yet, the method in [154] would yield the same score even if two different systems have different lengths for the idle states but have the same amount of false-positives. The system with longer idle periods should yield a higher score as this system makes less frequent IC onset false-positives, and is thus more robust, but that is not what the index in [154] would indicate

Thus, to address the limitations in [154], we propose a new performance evaluation score, called *true-false-positive score (TFP_{Score})*, defined as follows:

$$\mathrm{TFP}_{Score}\ (\%) = \frac{(TP + \alpha)}{(tE + \alpha)} * \left(1 - \frac{(FP + \alpha)}{(iE + \alpha)}\right)^2 * 100 \qquad \text{Equation 6-1}$$

where TP and FP are the numbers of true-positive and false-positive 500ms-windows, respectively in this study. tE and iE refer the number of IC task onset events and idle events, respectively. 'α' is set to 0.1, which is a very small number chosen merely to avoid division by zero while still minimising effects on the results. To define iE more clearly, the different online system time periods will be defined as follows:

- ***Recording Time*:** Total recording time for a run without any stops and interruptions.
- ***Task Period*:** Total task activation time, i.e., the sum of all task activation periods (from the beginning of task activation to the stop). This variable includes a timing error tolerance region (described below in Results & Discussion). If the experiment is designed to maintain the task activation state until the user receives feedback, then the tolerance region is not included.
- ***Refractory Period*:** Period in which the signal is ignored after the task activation or false-positive, i.e., the machine ignores incoming data while it executes whatever function is required after onset detection.
- ***Idle Period*:** Total idle state period, ***Idle Period = Recording Time - Task Period - Refractory Period***.

iE will be defined as the number of shifting windows that give classification results as Idle (e.g., assuming a non-overlapping window size of 500ms, a 1s idle period gives iE = 2). The behaviour of Equation 6-1 is shown in Figure 6.7. Ignoring α for simplicity, we obtain the following behaviours, all of which are correct:

- When FP is zero, TFP will vary with iE, so, everything being equal, longer idle periods will yield higher performance scores.
- By multiplying (1-FP rate) to TP rate, the score is reduced if FP is increased.
- If FP is zero, the score will be near TP and will depend on idle period size.
- **FP_rate=top/bottom.** The square power of (1-FP rate) will give more reasonable scores than by removing the power of 2. For example, TP=6 and FP=0 give a TFT score around 60%. The score will be similar when TP=7 and

FP=3 in panel A (i.e., with the power of 2). However, without the power of 2 (panel B), a score near 60% would be obtained with TP=7 and FP=7, which does not make sense as a system with TP=6 and FP=0 is clearly much better than one with TP=7 and FP=7. For this reason, the square power was chosen after investigation with many scenarios.

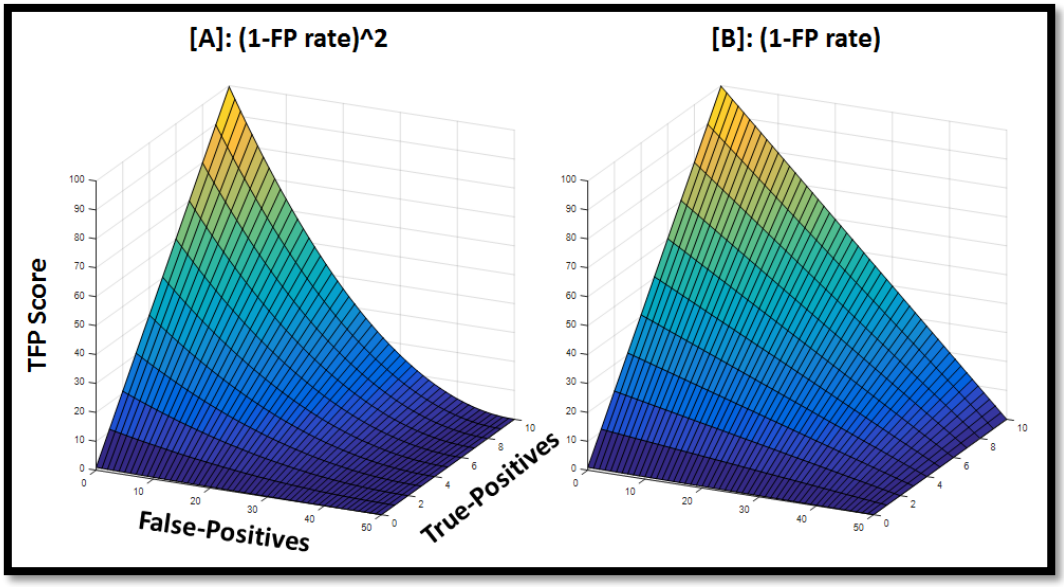


Figure 6.7. TFP Score graph. A) applies (1-FP rate) ^{2}, as in Equation 6-1, while B) is without the square power. Ranges: TP= 0-10, FP= 0-50, tE=10 and iE=50, TFT=0-100. NB: (1-FP rate)^2 refers to (1-(FP+a)/(iE+a))^2 in Equation 6-1 [1].

6.3 Results

Offline Testing Evaluation

Table 6.1. Offline testing accuracy from four different sound-production related onset tasks for all subjects [1].

	Accuracy % (Standard Deviation σ)				
	C_High	C_Siren	IO_High	IO_Siren	Average
P1	***71.98% (±2.66)***	63.56% (±5.32)	63.49% (±3.77)	***70.20% (±3.51)***	67.31%
P2	73.14% (±1.8)	68.44% (±3.17)	***77.44% (±3.51)***	72.41% (±3.83)	72.86%
P3	***87.28% (±1.56)***	82.82% (±2.07)	***87.20% (±2.43)***	78.52% (±1.7)	83.96%
P4	63.42% (±2.66)	64.47% (±2.05)	62.07% (±4.68)	64.89% (±3.25)	63.72%
P5	63.41% (±3.75)	64.01% (±3.69)	60.43% (±4.32)	64.54% (±1.87)	63.10%
P6	73.14% (±1.14)	72.69% (±1.9)	***85.75% (±1.65)***	83.94% (±2.64)	78.88%
P7	***91.84% (±1.2)***	80.08% (±2.79)	87.53% (±1.66)	86.41% (±1.32)	86.49%
Average	74.89%	70.88%	74.84%	74.42%	73.76%

Table 6.1 shows classification accuracy for all seven subjects and four different onset tasks. The ***Bold and Italic*** results represent the highest accuracy out of four different onset tasks for each participant. If there is no significant difference between the highest values (as measured by a Wilcoxon test p-value), both results are marked as ***Bold and Italic.***

For participants 1, 3 and 7 covert high tone sound-production (C_High) achieved significantly higher accuracy (i.e., average true positive rate when discriminating between idle and task) than the other three tasks (p-value $\ll 0.05$). For participants 2, 3 and 6 the inhibited overt high tone sound-production (IO_High) task achieved the highest accuracy. For participants 4 and 5 there was no significant difference between tasks.

Based on the average values shown at the bottom of Table 6.1, the C_High task led to better results, followed by IO_High, IO_Siren and C_Siren. There was no significant difference between C_High, IO_High and IO_Siren but C_Siren showed significant worse result than other tasks. It is thus advisable to determine the best onset task on an individual basis.

In terms of average performance for each subject, participant 3, 6 and 7 achieved relatively high values. Participant 7 had experience in similar experiments from our previous study in [4], so he/she was expected to achieve high performance. However, participant 3 and 6 were naïve subjects. Also, participants 4 (experienced) and 5 (naïve) showed somewhat low performance results compared to other participants, yet they were experienced users. This suggests that previous experience has no significant effect on performance.

The average true positive rate across all tasks and subjects was 73.76%. However, this value rose to 77.7% if only the best task for each subject was considered.

Simulated-online Testing Evaluation

Figure 6.8 shows output testing results for participant 6's IO_High onset task, for illustration purposes. It was chosen because the results contain moderate amounts of true-positive and false-positive events, so it allows us to discuss both cases. The horizontal axis represents the time scale in terms of sample windows, one sample representing a 0.5s window.

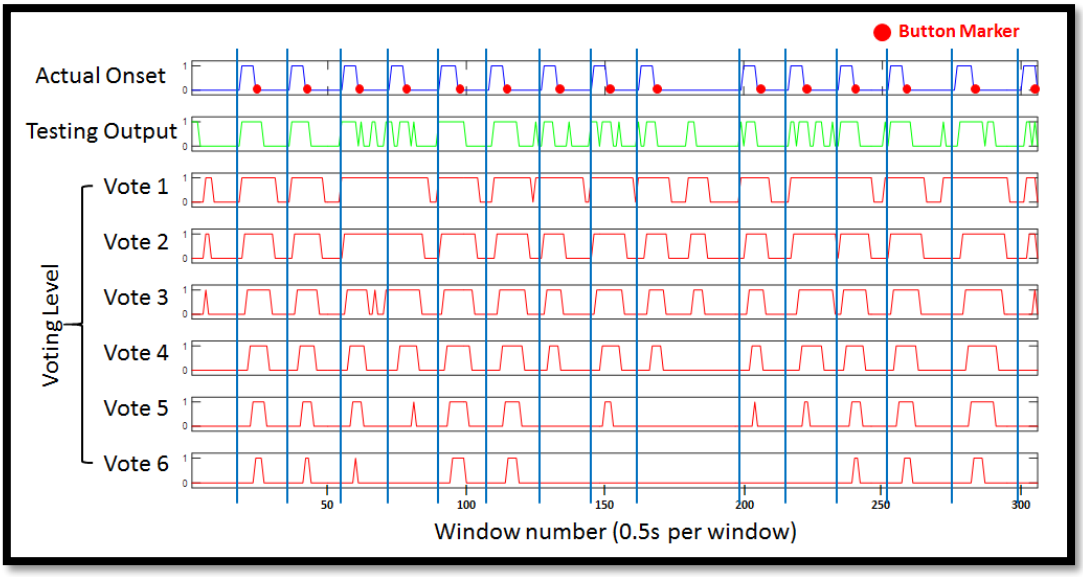


Figure 6.8. Simulated-online output results for participant 6's inhibited overt high tone onset task. The time scale is shown in terms of sample windows, one sample representing a 0.5s window. 'Button marker' denotes a key press after a 3s task was finished [1].

The vertical axis is binary; a value of 1 indicates a non-idle state, while 0 indicates an idle state. The blue, top line depicts actual onset states as determined from the user's input by pressing space bar after executing a non-idle cognitive task. The green plot (testing output) shows the IC task periods as determined by the LDA classifier.

The red plots (Vote 1 to Vote 6) represent results from an applied voting system, designed to assess sensitivity to false and true events, as follows: Six sequential windows (3s data: 0.5s windows*6) from the testing output were selected and a voting process was applied. Within those 6 sequential windows the machine detected ***N*** onset events. 'Vote ***N***' denotes the number of onset windows required for the machine to determine that a real onset has occurred. E.g., 'Vote 2' indicates that the machine required 2 (not necessarily consecutive) of the 6 windows to yield 1 as output in order

to accept an event as being an onset. This process was continuously done by moving a jumping 0.5s windows (i.e., a sliding window with no overlap) from the beginning to the end of the recorded data. As can be seen from Figure 6.8, the incidence of false-positives decreases from Vote 1 to Vote 6. However, true-positives also decreased (and in varying degrees, depending on the participant). For this reason, it was necessary to find an optimum voting level to minimise false events while maximising true ones. This was done based on a true-false-positive score (discussed below).

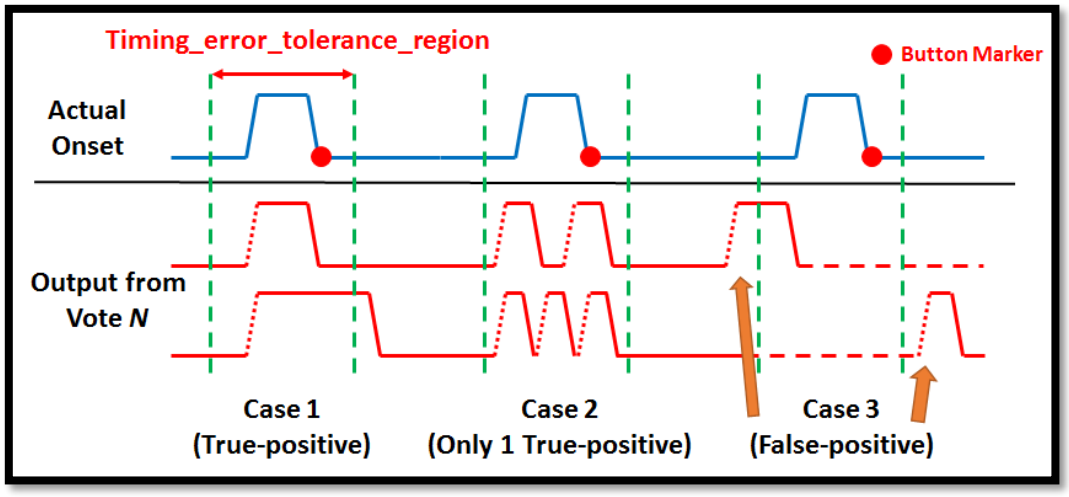


Figure 6.9. True-positive and False-positive definition in the simulated-online situation [1].

For classification performance assessment, it is difficult to achieve sample by sample labelling in self-paced BCIs as well as in this simulated-online recording protocol. Thus, event by event (i.e., one 0.5s window at a time) labelling was adopted. True-positive and false-positive events were defined as shown in Figure 6.9. Although participants were instructed to perform a given task for approximately 3s, we included a timing error tolerance region (TETR) to investigate possible timing errors and their effect on system performance. Two different TETR values were investigated in this

study, i.e.: the original 3s epoch was padded with the following window lengths on each side: 0.5s (i.e., 0.5s+3s+0.5s = 4s TETR) and 1.5s (6s TETR).

In this experiment, only rising edges from the output graph were only considered as onset. There are three different cases depending on the time location of rising edges. Case 1 indicates the machine-detected rising edge appeared within a TETR and this was treated as true-positive. If there were multiple rising edges in a single TETR (as in case 2), only one true-positive was accepted and others were discarded. Case 3 is an example of a false-positive event. If a rising edge appeared outside the TETR, it was regarded as a false-positive even if remaining machine-detected onset window overlapped with an actual event. If multiple rising edges were detected outside the tolerance region, all of them were considered as false-positives.

Table 6.2. Simulated-online performance results. True-false-positive score with optimal voting level. 4s of Timing error tolerance region (TETR) [1].

	TFP Score% [TP, FP]				
4s TETR	C_High	C_Siren	IO_High	IO_Siren	Average
P1	***68.62% [12, 12]***	42.41% [7, 8]	45.65% [8, 12]	68.26% [12, 12]	56.24%
P2	***56.60% [9, 5]***	48.98% [8, 6]	44.95% [8, 12]	53.46% [9, 7]	51.00%
P3	69.95% [11, 4]	46.79% [8, 11]	***78.13% [12, 2]***	72.55% [12, 8]	66.86%
P4	54.06% [9, 8]	64.08% [10, 3]	48.90% [8, 7]	***68.72% [11, 5]***	58.94%
P5	49.51% [8, 6]	49.48% [8, 6]	***61.73% [10, 6]***	54.87% [9, 7]	53.90%
P6	43.46% [10, 5]	68.46% [11, 7]	***74.26% [12, 7]***	54.56% [9, 10]	60.19%
P7	***61.75% [10, 9]***	55.23% [9, 9]	53.54% [9, 6]	43.89% [7, 8]	53.60%
Average	57.71%	53.63%	58.17%	59.47%	57.24%

Table 6.3. Simulated-online performance results. True-false-positive score with optimal voting level. 6s of Timing error tolerance region (TETR) [1].

	TFP Score% [number of TP, number of FP]				
6s **TETR**	C_High	C_Siren	IO_High	IO_Siren	Average
P1	69.57% [12, 7]	58.00% [10, 5]	61.87% [11, 8]	***74.59% [14, 4]***	66.01%
P2	***65.42% [13, 14]***	58.54% [10, 5]	61.13% [12, 7]	58.41% [10, 6]	60.88%
P3	77.08% [12, 2]	58.54% [10, 7]	***88.15% [14, 3]***	81.87% [14, 5]	76.41%
P4	58.85% [10, 4]	***80.89% [14, 6]***	54.66% [9, 6]	79.89% [12, 0]	68.57%
P5	***76.70% [12, 2]***	62.64% [10, 3]	63.64% [10, 3]	73.51% [12, 4]	69.12%
P6	60.05% [10, 5]	80.76% [13, 5]	***86.10% [14, 4]***	65.79% [11, 8]	73.18%
P7	***66.83% [11, 8]***	56.32% [9, 8]	63.87% [11, 13]	56.82% [9, 6]	60.96%
Average	67.79%	65.10%	68.49%	70.13%	67.87%

Table 6.2 and Table 6.3 show simulated-online testing results for each onset task. The values were calculated based on the true-false-positive score (TFP, described above) and the numbers in a square bracket represent the number of true-positives (TP), false-positives (FP). The total number of actual task onset events (tE) was 15 for all runs. The values shown on the tables for voting level are the ones that gave the highest TFP% score out of six votes.

Two different TETR sizes (4s and 6s TETR) were compared. Larger TETRs increase the chances of detecting true-positive events, while at the same time leading to less frequent false-positives. However, the TFP% score takes into account the total idle period length. Thus, if there was no significant difference in the number of true and false-positives for different TETR values, the smaller TETR, which yields a longer idle period, would give a higher TFP% score**.** The average results showed that 6s TETR (from Table 6.3) has higher score than 4s TETR (from Table 6.2). It leaves us further investigation to find out optimal TETR in usability point of view as a system would give quicker response with smaller TETR. It would be our future study to move online

system. In terms of the best voting level sensitivity, results varied widely depending on subject and tasks.

The average TFP score (across participants) for each of the onset tasks were 57.71%, 53.63%, 58.17% and 59.47% (for C_High and C_Siren, IO_High, IO_Siren, respectively) with 4s TETR and 67.79%, 65.10%, 68.49% and 70.13% with 6s TETR. Both results show that IO_Siren task has higher score followed by IO_High, C_High and C_Siren. However, it all vary depends on subjects. When we average the highest TFP scores for each participant, the overall TFP score was 67.12% (TP rate= 72.38%, FP rate=3.78%) with 4s TETR and 76.67% (TP rate= 87.62%, FP rate=4.05%) with 6s TETR.

6.4 Discussion

Self-Paced EEG-based BCIs (SP-BCIs) have traditionally been avoided due to two sources of uncertainty: (1) when an intentional command is precisely sent by the brain, i.e. the command onset detection problem, and (2) how different the intentional command is when compared to non-specific (or idle) states. Performance evaluation is also a problem and there are no suitable standard metrics available. In this Chapter, we attempted to resolve these issues.

Self-paced covert sound-production related cognitive tasks (i.e. high pitch and siren-like sounds) were used in order to distinguish between Intentional Commands (IC) and idle states. The IC states were chosen for their ease of execution and negligible overlap with common cognitive states. Band power and a digital wavelet transform

were used for feature extraction, and the Davies–Bouldin index was used for feature selection. Classification was performed using linear discriminant analysis.

Performance was evaluated under offline and simulated-online conditions. For the latter, a performance score, called true-false-positive (TFP) rate, ranging from 0 (poor) to 100 (perfect), was created in order to take into account both classification performance and onset timing errors. Averaging the results from the best performing IC tasks for all seven participants, a 77.7% True-Positive (TP) rate was achieved in offline testing. For the simulated-online analysis the best IC average TFP score was 76.67% (87.61% TP rate, 4.05% false-positive rate).

Comparison with Other Studies

It is very difficult to directly compare our results with other typical motor-imagery onset detection systems as there is no common evaluation method. In addition, many studies have shown performance results (such as the hit rate) that can only be applied to their own experimental settings (e.g. [8, 10, 24, 25]). Other studies have shown only the classification accuracy. In [82], the average TP rate for three subjects for the idle vs. motor-imagery state was 86.7% and the number of false-positive events was 5.7, however, there was no information regarding the idle period length, and they also calculated the false-positive rate by treating the number of onset events ‘E’ as a true-negative, which is something we do not agree with. . In [10], the motor-imagery versus non-control state achieved a classification accuracy of around 79.67% on average for three subjects. In [18], six different mental tasks versus the idle state achieved between around 55% (Auditory imagery) and 72% (Motor-imagery) values of offline TP rates on average for 5 subjects. In [88], the researchers classified motor-imagery tasks vs. the idle state and they used two two-class classifiers for three different

classes (left hand and right foot imagery vs. idle). If the feature did not belong to motor-imagery tasks, they assumed it belonged to the idle state. They achieved around true-positive rates of 40% in an offline analysis.

Compared to the results from the above studies, our results (i.e. around 76.67% of TFP score: 87.62% TP rate, 4.05% FP rate) look promising. Furthermore, none of the above studies investigated onset timing errors and none attempted to produce a system that would work with a timing error as small as 3 s. In addition, our score system, which was based on TFP, is more complete and more conservative than those of previous approaches, making it suitable for future use in asynchronous BCIs.

It is possible that improved results could be obtained by including other wavelet types and orders as well as other feature domains and classifier types. However, we believe that the fact that such simple features (based on the db2 wavelet) and classifiers (LDA) yielded encouraging results, indicates that the proposed method has a potential for further applications in BCIs.

These results were very promising and competitive to other studies. Moreover, based on our literature review and to the best of our knowledge, there is no reported covert sound-production onset detection system for spBCIs. Results showed that the proposed onset detection technique and TFP performance metric have a strong potential for use in SP-BCIs.

Even though the timing interface we designed was considered to minimise visual event-related potentials by applying a circular progress bar with a high viewing angle based on the published literature, there could still be some VEPs that remained. However, the conditions between the idle and active task periods were the same, therefore we can assume that there were no VEP related class separation effects.

6.5 Summary and Conclusions

This study presented a methodology to address three current issues in self-paced BCIs: a) determining when an Intentional Command (IC) is sent by the brain to the machine, b) reliably discriminating between intentional brain activity and brain states that are non-specific or not relevant to the human-machine interaction and c) the lack of a suitable standard scoring system for performance evaluation in self-paced BCIs.

Averaging all results across all seven participants, the best idle vs. IC offline performance was obtained with the Covert High tone (C_High) sound production imagery (74.89% True Positive (TP) rate). A 77.7% TP rate was achieved when only the best IC task for each individual participant was used for obtaining the average results. These offline results refer to a 3 s timing window, i.e. a 3 s timing uncertainty as to when an actual IC onset occurred. We believe this value is acceptable for most BCI scenarios. For the on-line simulation analysis, the IO_siren yielded the best overall results based on the TFP score (68.49%). The average TFP score considering only the best IC task for each participant was 76.67%. The true positive and false positive rates for the latter TFP score were 87.61% and 4.05%, respectively.

While there are no studies against which our results can be directly compared, previous similar IC onset detection studies using motor imagery have yielded a best classification (true positive rates) of 72.0% [18] and 79.7% [10], without taking into account though timing errors. In this respect, we believe that our results and the proposed methods may be of use to other self-paced BCI researchers.

This chapter showed that our new sound-production related cognitive task onset detection system can be used for online situations based on the results from the

simulated online experiment. Having this as a basis, Chapter 7 will discuss the actual online onset detection system in real-life scenarios.

The new TFP score system is more complete and more conservative than the ones from previous approaches, making it suitable for future use in asynchronous BCIs, which provides a great contribution to the field by giving standardised metrics for evaluation.

7 Comparison Study between Sound-production Related Cognitive Task vs. Motor-imagery for Onset Detection in *Online* Real-life Task Scenarios (Based on a paper [3])

7.1 Introduction

This chapter is based on a published conference proceeding [3].

In Chapter 6, sound-production related cognitive tasks were tested for onset detection of BCIs in a simulated-online setting with a self-paced paradigm before they were used in an online experiment.

In this chapter, the novel intuitive sound imagery self-paced onset detection system is tested in an online situation, which turns a messenger dialogue on when a message arrives by executing an onset task in real-life scenarios (e.g. watching video, reading text).

This chapter consists of two different experiments. Section 7.2 shows a comparison study between our sound imagery and the typical motor imagery approach for the messenger onset task in order to demonstrate the advantages of our proposed method over the motor imagery for a self-paced onset detection BCI system. The results

showed that the sound imagery onset task achieved a 84.04%, 80.84% and 77.17% value of TFP score with the sliding image, watching video and reading text scenario, respectively on average for twelve subjects (significantly better than the motor imagery task except for the reading scenario, which had no significant difference). In addition, the sound imagery task showed a significantly faster response and better usability than the motor imagery onset task.

In Section 7.3 the sound imagery onset detection system was tested at outdoor laboratory settings (an outdoor cafeteria area) in order to demonstrate real-life uses and investigate the limitations of moving BCIs towards real-world settings. The results showed that the outdoor lab task had significantly worse TFP scores than the ones from the indoor lab settings in all daily-life scenarios. It also showed some potential problems from using BCIs in real-world applications, based on users' questionnaires, which can provide useful background information to other self-paced real-world BCI studies.

7.2 Sound-imagery vs. Motor-imagery for Onset Detection

Cognitive Task Description

In this experiment, two different cognitive tasks were tested for the sake of comparison. One was motor imagery, which is a typically used mental task in the BCI field. It is performed by the imagining of limb movements such as from the hands, feet or the tongue [9]. In our experiment, participants were instructed to imagine their primary hand and wrist movement.

The other task was a sound-production related cognitive task (Sound Imagery (SI) proposed in our previous studies [4, 5]). The task had showed encouraging results in an offline semi self-paced onset detection system (Chapter 6). For this reason, we used this SI task in an online experiment in order to test it in real-life task scenarios and to compare it with a typically used motor imagery task. In this experiment, participants were instructed to imagine producing an 'um' sound with a high tone in a covert (i.e. imaginary) manner, which necessarily overlaps with auditory recall (auditory imagery [145]). In addition, there should be no tensioning of any organs related to the sound-production in order to ensure the purely covert task execution. The high pitch tone level was chosen by the participants based on sounds they could comfortably generate for a couple of seconds but high enough to think that they were unusual tones to be used in a normal daily life situation.

As this experiment was about online self-paced onset detection, the idle state (i.e. non-control or null state) had to be defined for training purposes so that the Intentional-Control (IC) task state could be reliably distinguished from the idle state. To this end, participants were instructed not to think of any IC tasks and to stay calm and relaxed for the idle state recordings.

Experimental Paradigm

Twelve healthy subjects (ages between 19 and 27, 10 males and 2 females) with normal or corrected vision participated in the experiments. Four of them (P3, P4, P10 and P12) had previous experience in other BCI experiments and the remaining eight were naïve subjects. Each subject was sitting comfortably on a medical chair and a monitor was placed 50 cm away from the subject. A keyboard was placed on their lap

in order to give feedback to the system. The experiments were conducted in accordance with the University of Essex Ethics Committee guidelines.

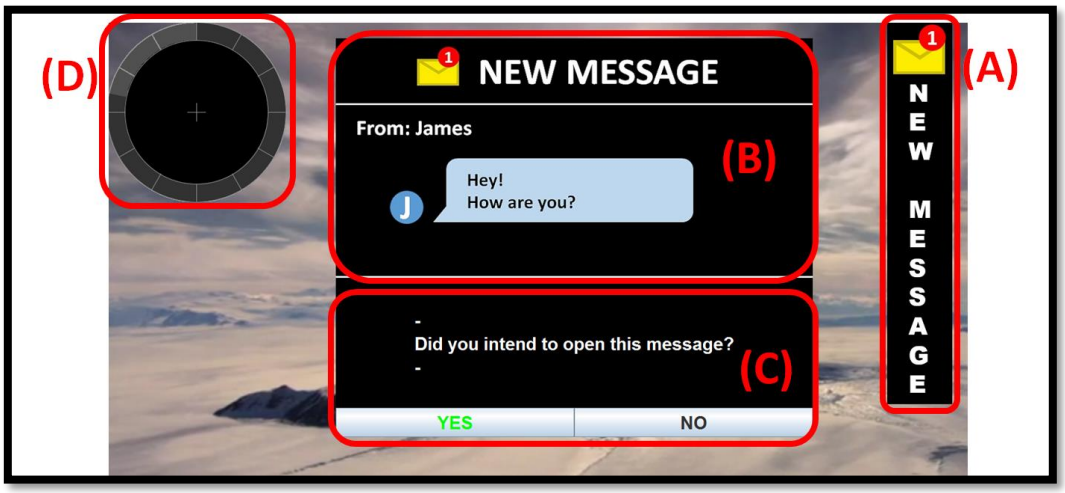


Figure 7.1. Messaging system interface example from the experiment. ((A): new message alert, (B): message dialogue, (C): user feedback panel and (D): time keeping interface) [3].

The experiment was designed to simulate a message opening system when a new message arrived during realistic daily-life task situations (i.e. watching video, reading text from a book). Figure 7.1 shows an example of the system interface. On the background of the screen, a video clip was playing and a subject was watching it while panels (A), (B) and (C) were hidden from the screen. Once the new message arrived (randomly between 5 s and 15 s in order to prevent subject anticipation), panel (A) smoothly and slowly slid into the side of the screen in order to minimise any Visual Event-related Potentials (VEPs). Then, the participants could either keep watching the video without trying to open the message, or they could open the message dialogue (B) by executing the Sound Imagery (SI) task state discussed in previous sections. The participants could perform the SI task action at any point as the system was self-paced.

However, they were asked to execute their onset at least 2 seconds after the new message dialogue had appeared, in order to eliminate any other even-related potentials (e.g. negative potentials or P300) so that the results were entirely based on the SI tasks. While participants were executing their SI task, they could estimate how long it took them to open the message dialogue by referring to the time keeping interface (D). This round shape progress bar continuously filled from a light grey to a dark grey colour for 12 seconds, followed by dark to light grey again. There were small marks at each 1 second interval so participants could estimate their task execution time. As a result, the users could provide feedback to the PC on whether its response was correct (True-Positive, TP) or not (False-Positive, FP) as well as the execution time if it was a TP. After this feedback, panels (A), (B) and (C) disappeared. The process from (A) to (C) comprised a trial and each single run consisted of 15 trials. Each participant had to go through three different runs (different background daily-life tasks). Background daily-life tasks were randomly ordered for testing for each participant in order to prevent any sequence-dependent results.

Daily-life Task Scenarios

There were three different experimental scenarios. In the first two scenarios, the participants were instructed to open a message dialogue (as explained above) while they were on two different daily life scenarios (one was watching video and the other was reading text). The above message opening onset detection system was tested separately on each of the two daily tasks. The last experimental scenario was the sliding image task. The participants were presented an image and if they wanted to slide the image to see the next one, they executed the mental task. In this scenario, there was no

external stimuli such as a message alert. Consequently, the participants controlled the system in a 100% self-paced approach. These three different experimental scenarios were chosen because they are very common scenarios for most people in real-life situations.

In terms of material, a documentary titled “BBC - The Blue Planet” [155] was used for the video watching task as it requires low cognitive load and emotional neutrality [156, 157]. For the reading task, a book titled “English Fairy Tales” [158] was used as it does not have any complex text and is emotionally neutral as well. Hence, the material had reduced cognitive loads for both native and non-native English speakers. In the sliding image task, natural images from [159] were used. The images, shown in Figure 7.2, were selected in order to keep emotional neutrality (wild background scenery without animals).

Figure 7.2. Example images for the sliding image task [159].

Signal Pre-Processing & Artefacts Handling

An Enobio (dry electrode equipment [160]) system was used for data acquisition. 17 electrodes were placed on the head based on a 10-20 layout and 1 reference channel was recorded on the right-side earlobe. Three extra external channels were placed on the forehead and both the right and left temples (anterior-most edge of the temporalis muscle) based on [23] in order to detect an Electrooculogram (EOG) and Electromyogram (EMG) for artefact removal purposes. The sample rate was 500 S/s (equipment bandwidth: 0-125 Hz) in order to ensure that all the EEG rhythms, up to some high gamma band, can be analysed. High gamma waves have not been widely used in BCIs due to concerns over contamination with EMG artefacts. However, studies have shown high gamma activity associated with language tasks [65, 67, 150]. It was therefore included in the experiments and EMG artefacts handling methods were applied to avoid EMG-related classification results.

EEG data were wirelessly transferred from the Enobio to a PC via Bluetooth. These EEG data were bandpass filtered (Butterworth filter, order 5) with cut-off frequencies at 4 Hz and 100 Hz followed by a notch filter (Butterworth filter, order 5) at 49-51 Hz in order to remove mains interference. Then, the data were segmented with a 0.5 s window length.

The segmented data underwent automatic EOG detection based on [129]. A Discrete Wavelet Transform (DWT) with a Haar mother wavelet (decomposition level 6) was applied to the external channel that was placed on the forehead. If the external channel's data were detected as EOG artefacts, the data segment was rejected from further analysis. If there was no EOG artefact, the EEG data were passed on to the EMG artefact removal process.

For automatic EMG removal, the Blind Source Separation by Canonical Correlation Analysis (BSS-CCA) was used, which is a very common and widely used EMG removal technique in BCIs. The threshold of the autocorrelation coefficient ρ was chosen at 0.35 based on [133]. Then, these pre-processed and EOG/EMG artefacts handled EEG data was used for feature extraction and classification.

Feature Extraction & Classification

In this experiment, four different feature extraction techniques were applied to the artefact-handled data. An ***AutoRegressive Model (AR model, Burg's method)*** with order number 6 was applied based on [12]. The model coefficients were used as features. The second method was ***Band Power (BP)*** extraction with a Fast Fourier Transform (FFT). There were seven different bands: 4-8, 8-12, 12-16, 16-20, 20-30, 30-42 and 42-100 Hz. Each band's FFT value was square powered and it was used as a feature. The third method was the ***Common Spatial Pattern (CSP)***. EEG source components were sorted in order to maximise the variance in one class and minimise it in the other class. Then, the first three and last three EEG source component variances were taken and linear regression was applied. The slope of the fitted line was used as a feature. The last feature extraction method was the ***Discrete Wavelet Transform (DWT)***. The data were decomposed up to level 7 and detailed parts, which represent the pseudo frequency bands of around 4-8, 8-16, 16-31, 31-62 and 62-100 Hz, were taken. From each detail part, the variance was calculated from the coefficients for dimensionality reduction. The mother wavelet 'db2' was chosen because of its common use in BCIs. These four different feature extraction techniques were chosen as together they cover the time, frequency, spatial and time-frequency domains.

These feature extraction processes generated hundreds of feature points for each channel. Therefore, a feature selection method was required. In this experiment, the Davis-Bouldin Index (DBI [120]) was used. The optimal DBI threshold was calculated from the training data for each subject and task. Then, features that had below threshold DBIs were used for classifications.

For the classification, the Linear Discriminant Analysis (LDA) was used. It was chosen because of its simplicity and low computational power [122]. Therefore, it suits the online classification for real-time processing as well as being a widely-used technique in BCIs.

Spatial and Spectral Analysis for Sound-imagery Task

In this section, spatial and spectral characteristics will be analysed for the sound imagery task. As the experiment was online, this analysis was carried out with the training dataset, which was recorded as an offline setting.

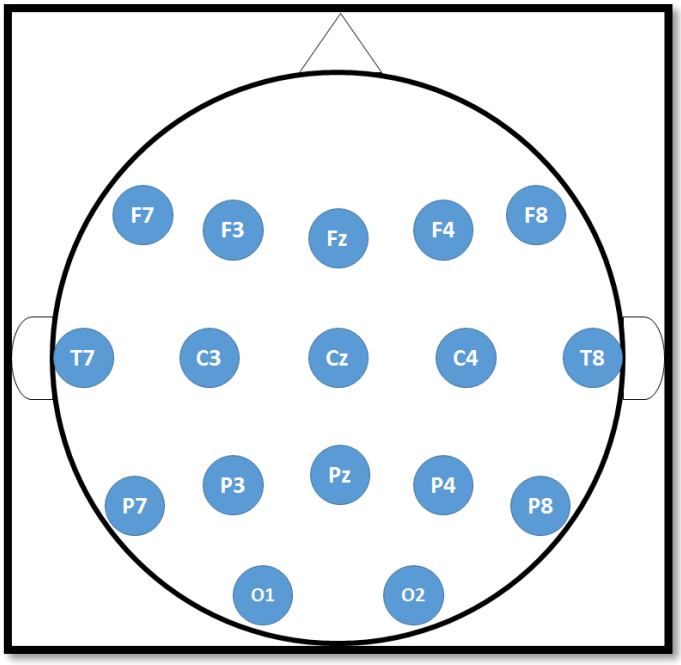


Figure 7.3. 17 electrodes placement used in this experiment.

For the spatial analysis, the Common Spatial Pattern (CSP) was found based on the electrode placement in Figure 7.3. Figure 7.4 shows the visual pattern for each subject and the average result. The pattern varies depending on the subject because of the characteristic of EEG. However, the average result shows that channels around F3, P3 and T7, which are located near Broca's and Wernicke's area that are related to speech, had a big pattern difference between the idle and sound imagery task period. In similar fashion, channel F8 also showed some pattern difference.

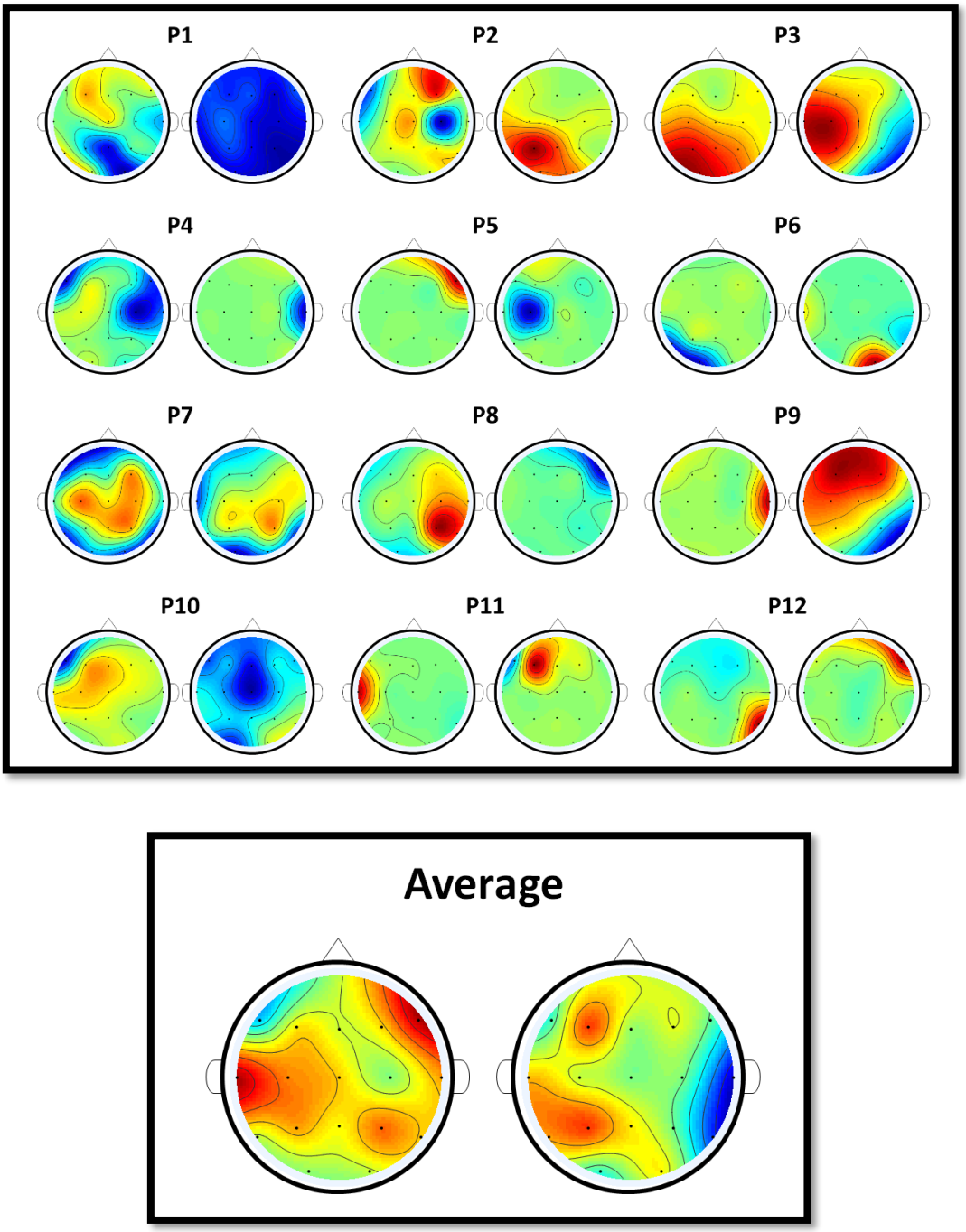


Figure 7.4. Common spatial pattern for each subject and average result. Left: minimum variance for the idle period state. Right: minimum variance for the sound imagery task.

In addition to the CSP analysis, the channels that had the best class separability based on the feature selection method are shown in Figure 7.5. From each participant, the 10 best feature points were selected based on the DBI feature selection method and

their channel numbers were counted and summed up from twelve subjects. As can be seen from the Figure, channel F3 was selected the most amount of times as the best class separable channel, followed by channel T7. It shares some common results with the CSP spatial analysis by having the same F3 and T7 channels that are located near Broca's and Wernicke's area.

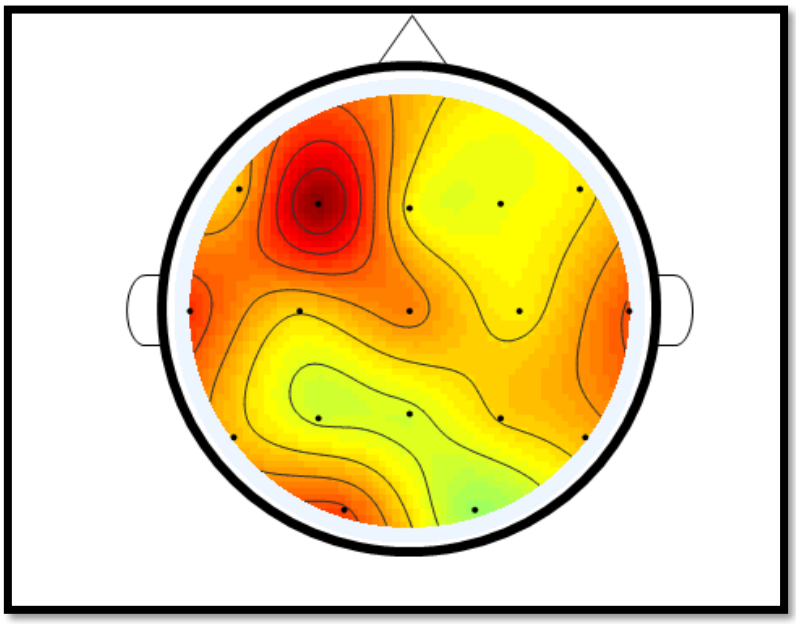


Figure 7.5. Spatial analysis with the DBI feature selection method.

In terms of spectral domain analysis, the frequency band that had the most class separability was found in similar fashion. From the 120 feature points (the best 10 features from each of the twelve subjects), the wavelet transform feature was selected the most times (56 times), followed by the band power feature (44 times) and autoregressive model feature (20 times). The common spatial pattern feature was not selected at all from all subjects. From those 56 DWT and 44 band power features, frequency bands were counted to find out which range was selected the most as the best

class separable frequency band. As can be observed from Figure 7.6 (top part), the pseudo-frequency band of 16-31 Hz was selected the most times followed by the range 31-62 Hz for the DWT feature. On the other hand, in the band power feature in Figure 7.6 (bottom part), the 20-30 Hz band was selected the most times. A review paper [37] reported that some studies suggested that the 30 Hz range should be elicited by linguistic processing of meaningful words but not of meaningless non-words. However, our high pitch sound imagery task showed the best class separability versus the idle state with the range of around 20-30 Hz.

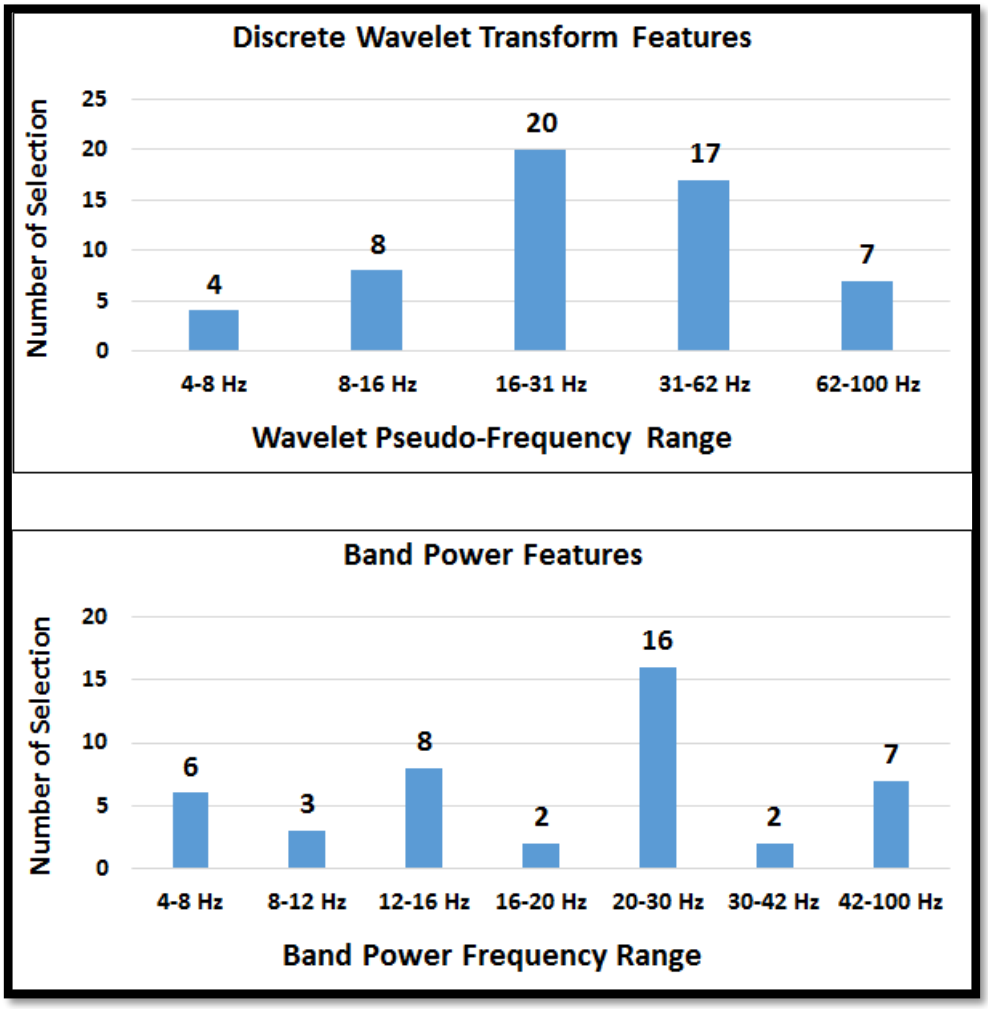


Figure 7.6. Spectral analysis with DBI feature selection method.

Performance Evaluation Method

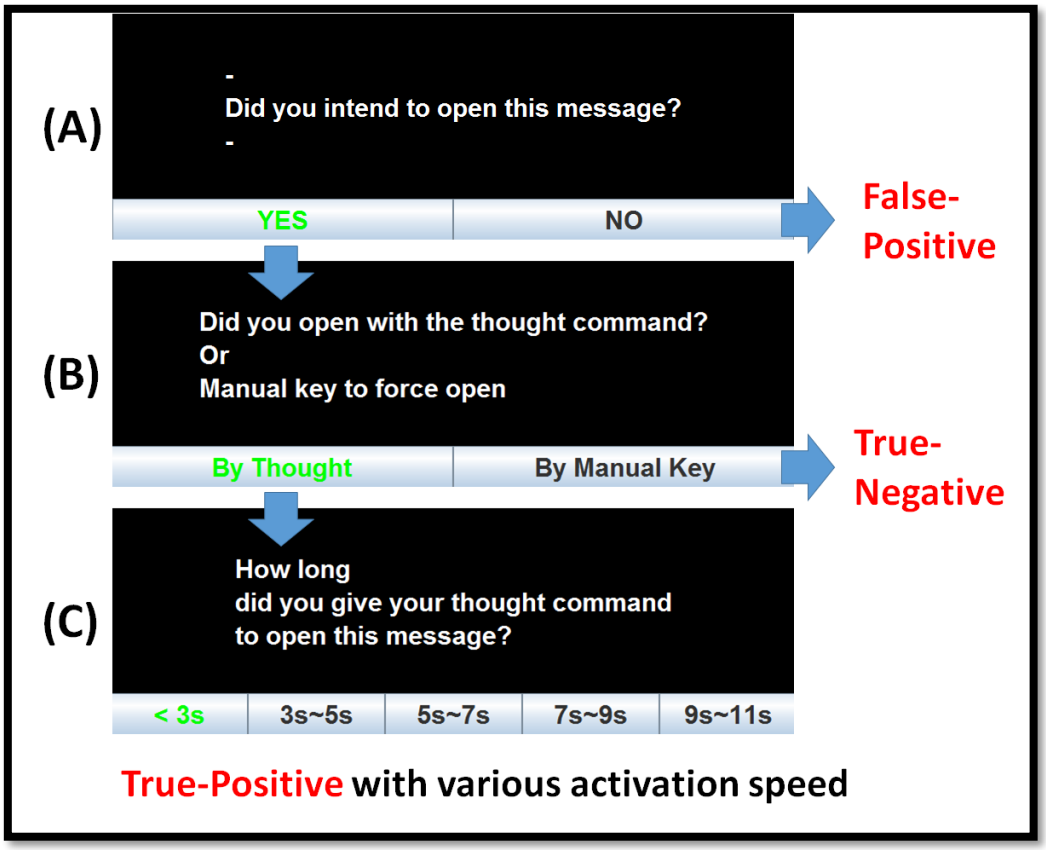


Figure 7.7. User feedback process during the online experiment for performance evaluation [3].

As the study was about an online onset detection system, a performance evaluation took place with the subjects' feedback. Figure 7.7 shows the feedback process. If the machine classified the incoming EEG data as an onset (intentional control) event from the user, a message window would appear on the screen with a feedback panel (A). The panel (A) could also be opened manually by pressing the 'Esc' button on the keyboard to indicate a True-Negative (TN) action. If the event was indeed an intended action, the user would choose 'YES'. Otherwise the user would choose 'NO' for a False-Positive (FP). If the user chose 'YES', the feedback panel would change to (B) in order to clarify whether it was an actual thought command (i.e.

sound-production) or a manually opening of the message (TN action). Subjects were asked to press the 'Esc' button when the continuous onset command (up to 11 seconds) did not work. If it was an intentional thought command, users were directed to panel (C) and were asked how much time had lapsed from the start of the onset until the message dialogue was opened. This will be regarded as a True-Positive (TP) with additional system response speed information (less than 3 s, 3-5 s, 5-7 s, 7-9 s or 9-11 s).

Based on the number of TP and FP, the ***True-Positive rate (TP rate = number of TP / number of TP + TN)*** and ***False-Positive rate (FP rate = number of FP / Idle event)*** was calculated. The definition of the number of idle events is important for the calculation of the FP rate. Firstly, the idle period was defined as: ***total recording time – task activation period – refractory period.*** The refractory period is the period during which a signal is ignored for classification after the TP or FP action (i.e. while the message is opened for user feedback). Therefore, the total number of idle events which can yield output from the classifier was ***idle period (sec) / windows length (sec)***. In our case this was ***idle period / 0.5s***. In addition, the ***True-False-Positive score (TFP_{Score})*** [1] was also calculated in order to take the idle period length into account for the final score in the self-paced system.

Results for section 7.2

Table 7.1 shows the classification performance with the True-Positive (TP) rate and False-Positive (FP) rate on both the Sound Imagery (SI) and Motor Imagery (MI) tasks in the sliding image task scenario. The twelve subjects' average TP rate for the sound imagery task was 88.3% while the motor imagery task had a 73.3% rate. Only

one out of twelve participants (P3) showed that the motor imagery task's TP rate was higher than the sound imagery task's and P5 showed the same TP rate with a lower FP rate in the sound imagery task. In terms of the Wilcoxon test *p* value, the sound imagery onset detection task had a significantly higher (*p* value at 0.033) TP rate than the motor imagery task. Even though the average FP rate in sound imagery had a lower value of 2.6% than the motor imagery at 4.8%, there was no statistically significant difference with a *p* value of 0.451.

Table 7.1. Online onset detection performance results in the sliding image scenario.

	Sliding Image Scenario			
	Sound Imagery		***Motor Imagery***	
	TP rate (%)	***FP rate (%)***	***TP rate (%)***	***FP rate (%)***
P1	66.7	5.9	60	23.0
P2	100	4	93.3	3.2
P3	73.3	1	80.0	2.9
P4	93.3	9.4	86.7	3.9
P5	86.7	0.5	86.7	2.7
P6	80	5.5	60.0	5.2
P7	86.7	4	46.7	0.8
P8	93.3	0	33.3	0.0
P9	100	0.9	93.3	9.1
P10	93.3	0	73.3	0.5
P11	100	0	86.7	2.5
P12	86.7	0	80.0	3.5
Avg	**88.3%**	**2.6%**	**73.3%**	**4.8%**

Table 7.2 shows classification performance results in two different daily-life task scenarios. In the watching video scenario, the 12 subjects' average showed an 86.1% TP rate for the SI task and 63.9% for the MI task. All the subjects had a higher TP rate with the SI than the MI task except for participant 3. The Wilcoxon test *p* value was 0.031, which depicts that SI had a significantly higher TP rate than the MI task. On the other hand, the average result of the FP rate shows that the SI task's FP rate is slightly higher than the one of the MI task but there is no significant difference with a *p* value of 0.259. In the reading text scenario, the average TP rate of the SI task was 81.1% and 77.2% for the MI task. Even though the SI task showed a slightly better TP rate result, there was no statistically significant difference between them with a *p* value of 0.243. The FP rate also showed that the SI task provided a slightly better result (smaller FP rate) but the difference was minor.

Table 7.2. Online onset detection performance results for watching video and reading text in daily-life scenarios.

	Watching Video Scenario				**Reading Text Scenario**			
	Sound Imagery		***Motor Imagery***		***Sound Imagery***		***Motor Imagery***	
	TP rate (%)	***FP rate (%)***	***TP rate (%)***	***FP rate (%)***	***TP rate (%)***	***FP rate (%)***	***TP rate (%)***	***FP rate (%)***
P1	80	7.6	46.7	1.4	53.3	8.2	80	11
P2	93.3	1.5	73.3	2.6	100	1.7	73.3	2.5
P3	46.7	1.7	60	1.2	33.3	2.1	80	2.6
P4	100	6.9	86.7	4.3	93.3	4.2	86.7	6.7

P5	86.7	1.9	46.7	1.5	66.7	3.7	80	3.7
P6	86.7	6.8	60	1.5	73.3	5.5	73.3	2.6
P7	86.7	6.4	20	0.5	86.7	3.9	46.7	0
P8	73.3	0.7	20	0	100	2.6	66.7	0.1
P9	100	0.9	93.3	4.3	93.3	0.6	80	2.7
P10	86.7	0.6	73.3	0.2	86.7	0	86.7	0.5
P11	100	0	100	4.1	100	0.4	100	2.1
P12	93.3	6.5	86.7	1.5	86.7	2.5	73.3	1
Avg	**86.10**	**3.40**	**63.90**	**1.90**	**81.10**	**2.90**	**77.20**	**3.00**

If the two-different daily-task scenarios are averaged, the TP rate of the SI task is significantly higher (83.6%) than the one of MI (70.6%) with a *p* value of 0.0106. There is, however, no significant difference with the FP rate.

Although it is difficult to directly compare our results with other onset detection systems as the experiment environment and tasks are different, our SI task showed a relatively high TP rate. In [10], three subjects produced on average a classification TP accuracy of 79.67% between the motor-imagery task and the non-control state. In [18], six different mental tasks versus the idle state showed TP rates of between 55% (auditory imagery) and 72% (motor-imagery) on an average over five subjects in an offline setting. Compared to these results, our 88.9% (in the video-watching case) and 78.9% (in the text-reading case) TP rates look very promising even though our study was carried out for more realistic scenarios than the ones previously reported by others.

In order to take into account all the true-positives, false-positives and the idle period length at the same time, as they are very important aspects of performance evaluation in self-paced BCI systems, the ***True-False-Positive score (TFP_{Score})*** suggested in [1], was calculated and discussed here. Table 7.3 shows TFP scores for each participant with three different daily-life task scenarios. 83.3% (10 out of 12) of the participants showed that the sound imagery onset detection task performed better in TFP score than the motor imagery task in both the sliding image and watching video scenario. 66.7% (8 out of 12) of the participants showed a higher TFP score with the sound imagery task for the breading text scenario. Participant 3, who had previous experience in BCIs, showed that the motor imagery task performed constantly better than the sound imagery task in all the daily-life scenarios but other participants, such as P4, P10 and P12, who also had BCI experience, did not follow the same pattern. Only two out of nine subjects showed that the MI task had a higher TFP score. From the naïve subjects, 87.5% (21 out of 24 cases) of them showed a higher TFP score with the SI task.

Table 7.3. True-False-Positive score result comparison between the sound-imagery and motor-imagery task in three different daily-life task scenarios.

	True-False-Positive Score (TFP score %)					
	Sliding Image Scenario		***Watching Video Scenario***		***Reading Text Scenario***	
	SI	***MI***	***SI***	***MI***	***SI***	***MI***
P1	59.16	35.66	74.05	45.57	49.03	63.02
P2	92.15	87.40	90.56	69.69	96.60	69.75
P3	72.01	75.45	45.38	58.75	32.22	75.88
P4	76.53	80.00	86.69	79.18	85.50	75.07

P5	85.84	82.12	83.42	45.53	61.87	74.09
P6	71.56	54.05	75.24	58.34	65.50	69.62
P7	79.84	46.25	75.64	20.30	79.98	47.00
P8	93.29	33.75	72.45	20.52	94.88	66.72
P9	98.13	77.17	98.24	85.36	92.14	75.66
P10	93.38	72.80	86.75	73.51	86.76	86.75
P11	99.89	82.32	99.96	91.95	99.13	95.76
P12	86.67	74.56	81.65	84.17	82.43	71.96
Avg	**84.04 %**	**66.79 %**	**80.84 %**	**61.07 %**	**77.17 %**	**72.61 %**

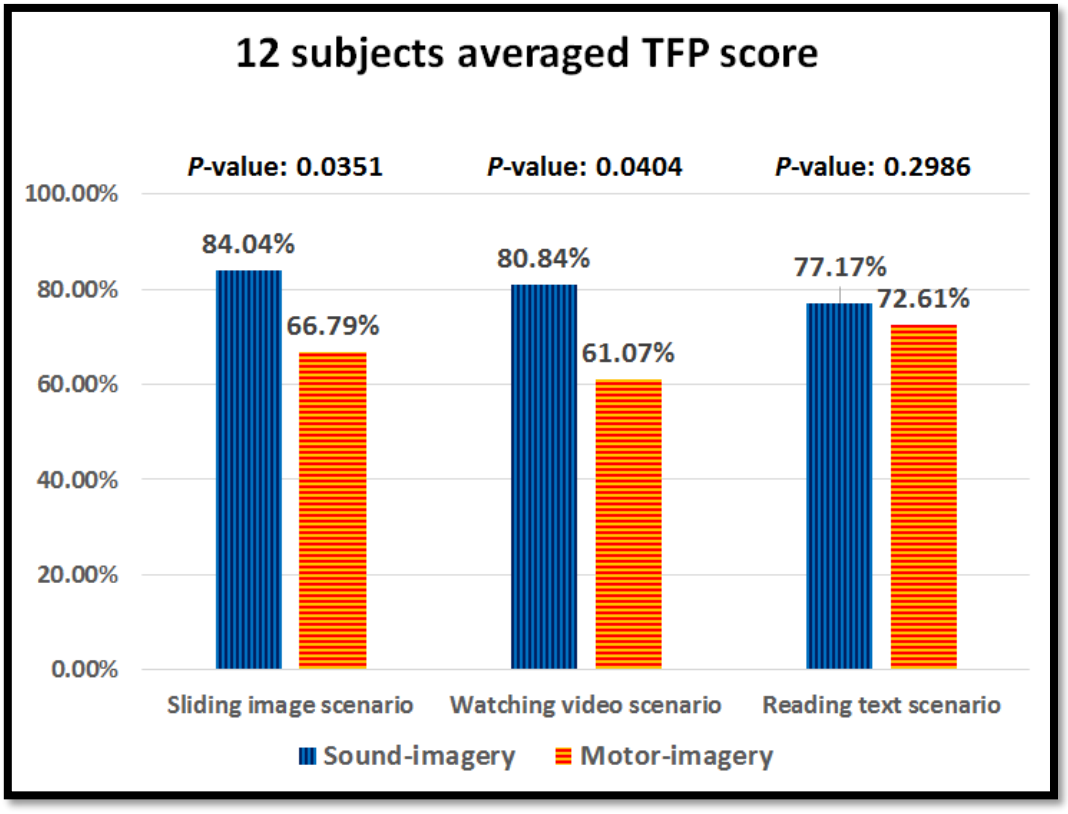


Figure 7.8. Averaged True-False-Positive score result comparison between the sound-imagery and motor-imagery task in three different daily-life task scenarios.

Figure 7.8 shows the twelve subjects' averaged TFP score for each daily-life task scenario. The sound imagery onset detection task produced a significantly higher TFP score than the motor imagery task with a *p* value of 0.035 and 0.04 for the sliding image and watching video scenario, respectively. However, there was no statistically significant difference for the reading text scenario even though the TFP score was higher for the SI task.

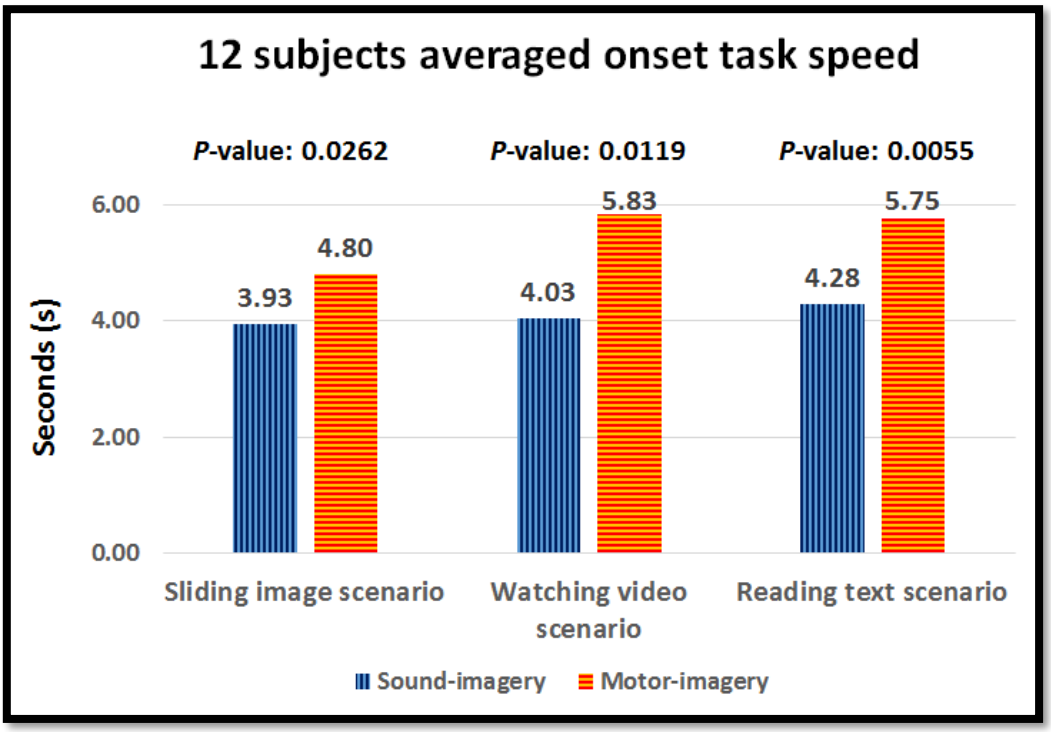


Figure 7.9. Averaged onset system response speed comparison between the sound imagery and motor imagery in three different daily-life task scenarios.

In terms of system response speed, the users' feedback from Figure 7.7 (C) was used in order to calculate the onset response time. Figure 7.9 shows the twelve subjects' averaged onset speed for the SI and MI tasks. The SI task required 3.93 s, 4.03 s and 4.28 s on average for the sliding image, watching video and reading text scenario

respectively, while the MI one required 4.8 s, 5.83 s and 5.75 s. In all of the three-different daily-life task scenarios, the SI task had a significantly faster onset response than the MI task by having a *p* value of 0.0262, 0.0119 and 0.0055, respectively.

Discussion for section 7.2

This experiment investigated an online onset detection method for BCIs by opening a message when it arrived in two different daily-life task scenarios (watching video and reading text) and in the sliding image task. Our new sound imagery task and typical motor imagery task were tested and compared.

In terms of system performance, the sound imagery task achieved a 84.04%, 80.84% and 77.17% value of TFP score for the sliding image, watching video and reading text scenario, respectively on average for twelve subjects. In contrast, the motor imagery task achieved values of 66.79% (significantly worse), 61.07% (significantly worse) and 72.61% (no significant difference), respectively. In addition, the system speed showed a significantly faster response with the sound imagery than the motor imagery task.

From a usability point of view, participants completed a short survey at the end of the experiment regarding the level of difficulty of use of the two different SI and MI tasks. 0 depicts very easy to use and 10 represents very difficult to use. On average, SI received a value of 4.42 while MI received 6.42. Nine out of twelve (75%) subjects marked a lower value (easier to use) for the SI than the MI. Only participants P3, P4 and P12 said that the MI was easier. These three participants were BCI research students, who had experience in MI but not SI. On the other hand, one BCI research student and all the other naïve subjects marked the SI as easier to use. The *p* value of

twelve subjects was 0.0108. Therefore, the SI task was significantly easier to use than the MI one for the onset detection of BCIs.

Based on these results, our new sound imagery task outperformed over the motor imagery task for the self-paced onset detection BCI system not only in performance but also in usability and system speed. Therefore, this prototype of onset detection system showed some strong potential to use BCIs in real-life applications (compared to the typical motor imagery task) and it will move the BCI field a significant step forward once it is developed further by improving current EEG recording issues such as practicality and usability.

7.3 Online Sound-imagery Onset Detection at an Outdoor Laboratory Environment

Experimental Paradigm

Seven out of the previously mentioned twelve subjects participated further in an outdoor laboratory experiment (P2, P3, P5, P6, P7, P8 and P9). They were chosen as they agreed to run an extra experiment in an open space. All the experimental settings were the same as the ones in the indoor laboratory. The sound imagery onset detection task with three daily-life scenarios was only tested in this experiment in order to avoid the total experiment time exceeding 1 and half hours, which meant that the participants could maintain a concentration level similar to the in-lab SI task. The only difference from the indoor laboratory experiment was the recording place. The outdoor laboratory experiment took place at a cafeteria called ‘Zest’ at the University of Essex. Figure

7.10 shows the environment with a 360° photo. During the experiment, there were no restrictions such as people passing by, coffee machine noise or background music. It was a typical cafeteria environment without any external control.

Figure 7.10. Outdoor laboratory experiment environment (Cafeteria) 360 ° photo. More detailed photos can be found in Appendix A.

A new training session was conducted in the cafeteria and all the algorithms and testing sessions were exactly the same as the indoor laboratory experiment. During the experiment, the noise level was recorded with the Cirrus CRL 2.22 sound level meter device [161] and the 'Sound Meter' Android application [162] on a Oneplus2 smartphone [163]. Table 7.4 shows the noise level during the outdoor cafeteria experiment. On average the environmental noise was around 38.6 dB with background music playing, people talking, etc. The maximum noise during the experiment was around 73 dB on average (e.g. coffee machine noise, loud music playing, etc.). On the

contrary, the noise level inside the laboratory experiment was near to 0 as it was a specially designed room for BCI experiments.

Table 7.4. Average and max noise level for each participant during the outdoor cafeteria experiment.

Noise level	P2	P3	P5	P6	P7	P8	P9
Average	38 (dB)	36 (dB)	47 (dB)	42 (dB)	31 (dB)	42 (dB)	34 (dB)
Max	73 (dB)	75 (dB)	72 (dB)	71 (dB)	68 (dB)	75 (dB)	77 (dB)

Results for section 7.3

Table 7.5 shows the true-false-positive score result for the outdoor laboratory experiment. On average the seven subjects achieved scores of 56.54%, 53.46% and 58.01% for the sliding image, watching video and reading text scenario, respectively. The TFP score was dramatically different depending on the subject. For example, P5 achieved a score of 75.78% and 85.77% in the sliding image and watching video scenario, respectively while P6 had a TFP score of 20.24% and 7.19%, respectively.

Table 7.5. True-False-Positive score result for sound-imagery onset detection in three different daily-life scenarios at outdoor laboratory (cafeteria) settings.

	Sound Imagery True-False-Positive Score (TFP score %)		
	Sliding Image Scenario	***Watching Video Scenario***	***Reading Text Scenario***
P2	66.40	53.39	59.99
P3	38.25	58.24	45.36
P5	75.78	85.77	62.37

P6	20.24	7.19	33.52
P7	61.29	57.83	58.04
P8	59.56	49.92	80.57
P9	74.27	61.89	66.20
Avg	**56.54 %**	**53.46 %**	**58.01 %**

Figure 7.11 shows the comparison results between the indoor and outdoor laboratory experiments for the SI onset detection task in three different daily-life task scenarios. From all the scenarios, the outdoor lab task showed significantly lower TFP scores with a *p* value of 0.004, 0.045 and 0.010 for the sliding image, watching video and reading text scenario, respectively (12 subjects for the indoor lab results vs. 7 subjects for the outdoor lab results). Participant 6 in particular, showed a huge decrease in the outdoor settings (-51.32%, -68.05% and -31.98% for the three different daily-life tasks). On the other contrary, participant 3 and 5 showed a slight increase with the watching video and reading text scenarios.

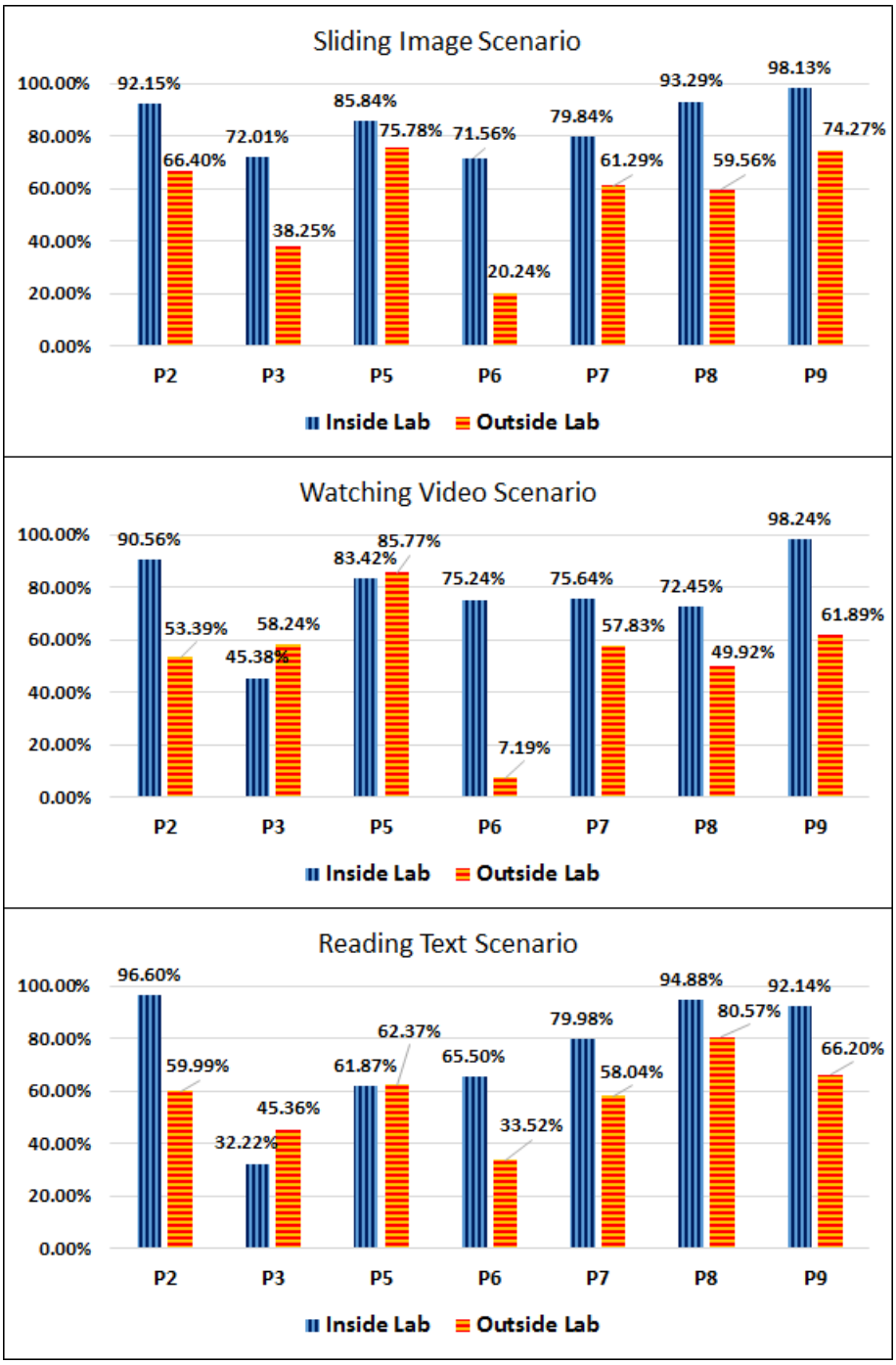


Figure 7.11. True-false-positive score comparison between the indoor lab and outdoor lab experiment in three different daily-life task scenarios.

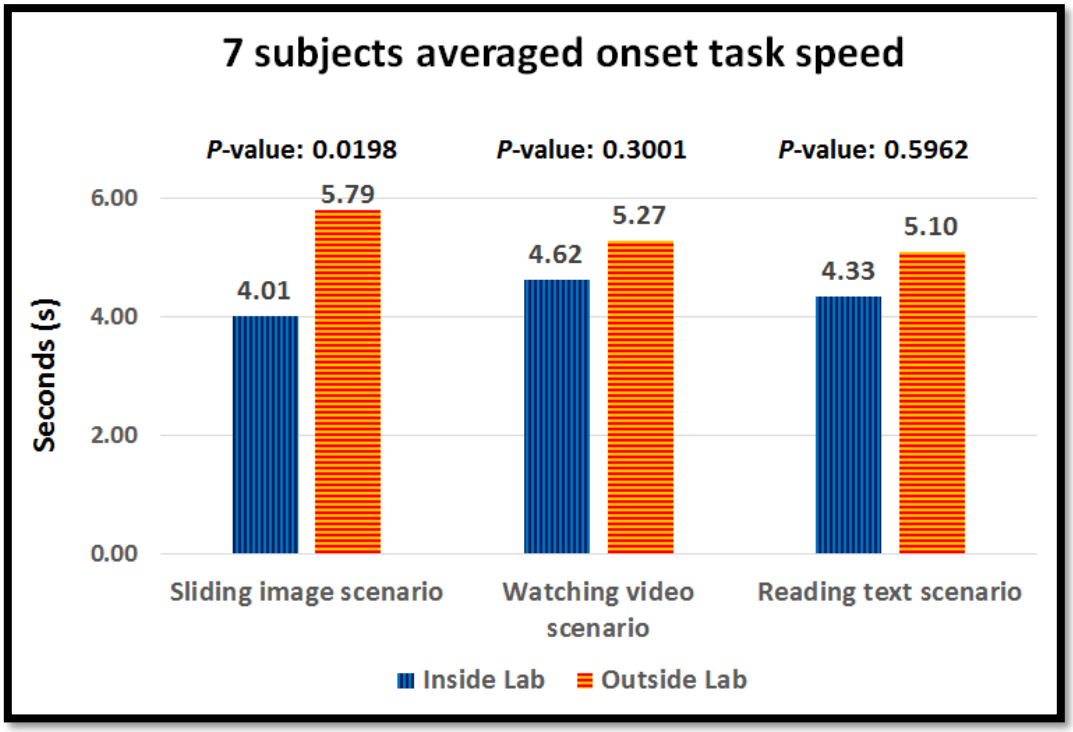


Figure 7.12. Averaged onset system response speed comparison for the sound imagery task between the lab and the outdoor lab experiment.

In terms of system response speed, Figure 7.12 shows the seven subjects' averaged onset task speed comparison between the indoor lab and outdoor lab experiment. The indoor lab setting achieved a response time of 4.01 s, 4.62 s and 4.33 s for the sliding image, watching video and reading text scenario, respectively. In contrast, the outdoor setting reported time values of 5.79 s (significantly slower), 5.27 s (no significant difference) and 5.10 s (no significant difference), respectively. Even though only the sliding image scenario had a significantly slower system response and the others had no significant difference, it was clearly shown that the outdoor setting had a slower onset response time on average, which could make it harder for users to control the system.

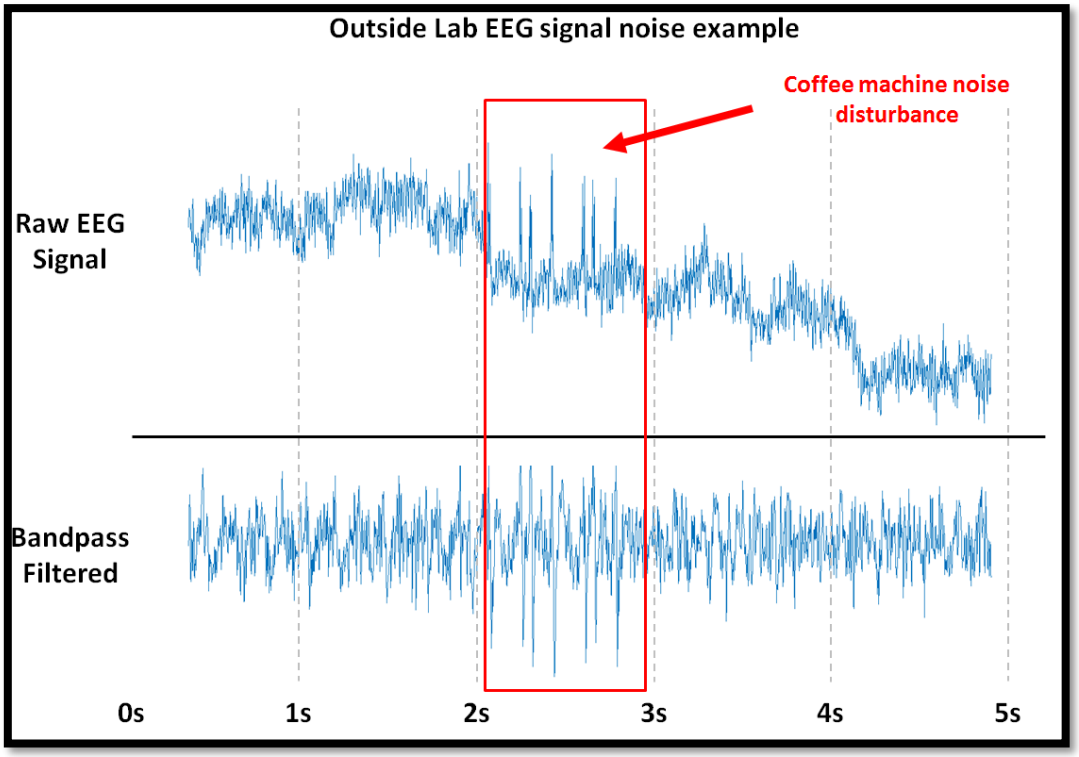


Figure 7.13. EEG signal noise example (Channel Cz) from Participant 6 at the outdoor laboratory. Top figure: raw EEG data. Bottom figure: bandpass filtered data (4-100 Hz). The red rectangle indicates the moment when the participant experiences a coffee machine noise disturbance (noise decibels rose to 71 dB from 32 dB).

With regards to the signal noise at the outdoor lab setting, Figure 7.13 shows its sample. The sample signal which was taken from Participant 6 (channel Cz) represents the moment when the coffee machine noise suddenly disturbed the environment with a noise level of 71 dB (sudden increase from 32 dB). The red rectangle represents this moment. As can be observed, the raw EEG signal (top figure) shows sudden noise interference. Even though the EEG signal was bandpass filtered (bottom figure, 4-100 Hz), the signal noise is still visible. These noisy signals were reported all around the channels quite a few times during the outdoor lab experiment by not only Participant 6 but also by other subjects as well (never reported in the indoor experiment). The signal

noises from environmental interference would be one of the main reasons that makes it more difficult for the users to control the onset task at an outdoor setting than the indoor laboratory experiment.

In addition, at the end of the outdoor experiment, the participants completed a short survey regarding their concentration level. The '0' value indicated that the concentration level at the outdoor laboratory setting was the same as the indoor experiment and the '10' value represented that it was extremely difficult to concentrate compared to the indoor settings. On average, the score was around 5.86 (P2: 8, P3: 7, P5: 0, P6: 9, P7: 2, P8: 6 and P9: 9). Participant 5 and 7 gave a score of '0' and '2', respectively, while other subjects found it was more difficult to concentrate and use the onset detection messaging system. In addition to the score, they also listed their disturbances during the outdoor experiment. They commonly answered the following: 1. Other people's noise (talking, ordering food, etc.). 2. Kitchen noise and food smell (coffee machine, grinder, etc.). 3. Visual distractions (outdoor view, people passing by, etc.). 4. Background music.

Discussion for section 7.3

This experiment extended the previous online message opening onset detection method to outdoor laboratory settings in order to investigate its potential issues and to compare the performance difference between the indoor and outdoor laboratory settings.

From a performance point of view, for all the scenarios the outdoor lab task showed significantly lower TFP scores with a p value of 0.004, 0.045 and 0.010 for the sliding image, watching video and reading text scenarios, respectively. In general, the

outdoor lab setting deteriorates in performance. Some cases (e.g. P6) showed a dramatic decline from around 70.8% to around 20.3%. Furthermore, subjects commonly answered that it was more difficult to concentrate at the outdoor cafeteria than in the indoor lab and they also listed disturbances (e.g. noise, visual distraction). However, P3 and P5 showed a slight increase even if it was not significant. As a result, it can be said that the TFP scores vary according to each subject but they generally produce worse results in the outdoor settings than the indoor laboratory.

In terms of system response speed, the outdoor lab setting experiment had a slower onset speed than the indoor one. Even though only the sliding image task scenario had a significantly slower response speed and the other two, the reading text and watching video scenario, had no significant difference, this may still have happened due to the small sample size (7 participants). Moreover, the slower system speed and lower performance results may have been caused because of the noise of the signal. As it was discussed in the previous sections, the outdoor experimental EEG signals contain some noise from all around the channels and for all the subjects, whereas they had no such noise in the indoor lab experiment. They mainly appeared when the environmental noise suddenly increased (e.g. sudden coffee machine operation). This could be one of the main reasons why participants answered it was more difficult for them to control the onset system in the outdoor experiment.

However, there is one big limitation in this experiment. As was already mentioned, the main goal of this chapter was to investigate the performance difference between our new sound imagery approach and the typical motor imagery. Therefore, the first indoor laboratory experiment was firstly conducted continuously to all the subjects and then the outdoor laboratory experiment was tested. We are well-aware of the fact that the biased recording order could impact the results, therefore it had to be

randomised. However, installing the experiment equipment in an outdoor cafeteria area, returning it to the indoor laboratory and setting it up again requires more time than the opposite. Therefore, in order to minimise the biased recording order effects (e.g. tiredness and boredom), there was a small tea/coffee break in-between the two experiments and the total experiment time did not exceeded 1 and half hours, which could maintain the participants' concentration level similar to one in the indoor lab experiment.

The experiment indicated the problems of moving BCIs towards real-life settings with daily-life task scenarios at an outdoor cafeteria area and the performance results were compared with the ones from the indoor lab experiment. This is a quite important and essential investigation at this stage in the BCI field, in order to use them in real-world applications apart from research-oriented indoor laboratory experiments. For this reason, the investigated experiment can be a useful background study to other self-paced real-world BCIs.

7.4 Summary and Conclusions

The scope of this chapter was to investigate how well our new sound imagery task works for a self-paced onset detection system in real-life scenarios by comparing it to a typical motor imagery task. In addition, it was tested at outdoor laboratory settings in order to explore potential real-world BCI uses. From a performance point of view, our novel sound imagery task showed a significantly better TFP score in the sliding image (84.04%) and watching video (80.04%) scenario (opening message onset

task) than in the motor imagery task (66.79% and 61.07%, respectively). Furthermore, the reading text scenario also reported a higher performance result with our approach (77.17% SI vs 72.61% MI). Moreover, the sound imagery task showed a significantly faster system response (4.08 s SI vs 5.46 s MI on average for the three scenarios) and had a significantly better usability (easier to use) score than the motor imagery. In the outdoor laboratory experiment, the participants listed noises (e.g. kitchen sounds, background music, etc.) and visual distractions (outside view, people passing by, etc.) as the main problem of using BCIs in outdoor real-world settings. From a performance point of view, on average the outdoor laboratory experiment showed a significantly worse TFP score (56%) than the indoor one (80.68%).

Based on these results, our novel sound imagery onset detection system outperformed the motor imagery one and it showed a great potential. This could be a significant step forward for the BCI field which is mainly restricted in research-oriented indoor laboratory settings with the use of motor imagery and cue-based studies.

8 Overall In-depth Discussion

The motivation of this thesis was to develop a more practical and usable BCI system as current BCIs are generally restricted in research at indoor laboratory settings. For this reason, a self-paced (asynchronous) approach needed to be investigated as it provides increased autonomy, flexibility, and interaction with the environment to the users, which is more suitable than cue-based (synchronous) BCIs for the ultimate aim of expanding the use of BCIs not only in laboratory settings but also into the real-world. However, self-paced BCIs have a great difficulty in identifying intentional commands and non-control states, which is called the onset detection problem. Therefore, this thesis has investigated onset detection for self-paced BCIs.

For the onset detection system, a novel sound imagery (sound-production related cognitive tasks) has been proposed as it has a couple of clear advantages over motor imagery (the mostly used method based on the literature review) and other cognitive tasks for onset detection: ***1)*** Intuitiveness; the sound imagery task is easy to produce and control voluntarily as most people constantly 'speak' internally or imagine many words in normal life. ***2)*** It is advantageous for people with motor disabilities for whom motor imagery tasks may not be suitable, which is an important target population for BCIs. ***3)*** Our new sound imagery task does not significantly overlap with other common spontaneous and frequent daily-life cognitive states. On the other hand, motor imagery or word/syllable/letter production tasks would not be suitable for the onset

detection system in practical (real-world) applications, as the machine would not be able to distinguish between whether the onset was the actual command, and other body movements / conversation. ***4)*** The chosen task has no dependence on the user's mother language. In addition to these advantages, Chapter 7 showed that our sound imagery task outperformed the motor imagery in terms of performance, system speed as well as usability score for onset detection in real-life scenarios. This will be discussed in the '*findings and implications of the research*' section 8.1.

In summary, the scope of this thesis was to propose a novel onset detection method for self-paced BCIs and the new sound imagery task was investigated in order to show the potential of moving current laboratory based BCIs to outdoor real-world uses. In order to achieve this goal, various experiments have been carried out. The findings and implications of the research will be shown and discussed in section 8.1, followed by its scientific contributions and limitations in sections 8.2 and 8.3, respectively.

8.1 Findings and Implications of the Research

Distinguishing the Sound Imagery Task vs the Idle State for Onset Detection

In Chapter 5, high pitch tone and siren-like sound imagery tasks were tested against the idle state with various speech modes (Overt, Inhibited overt and Covert) in offline settings (cue-based approach) in order to see the potential of using sound imagery tasks for onset detection. For the high pitch tone sound imagery task, the covert speech mode achieved the highest classification accuracy of around 80.9% on average

for four subjects. In the siren-like sound imagery task experiment, five subjects averaged a true-positive rate of around 79.2%. The results showed a quite promising performance in offline settings so further experiments were conducted.

Chapter 6 compared high pitch tone and siren-like sound imagery tasks in a simulated-online situation in order to choose the most appropriate speech related cognitive task before switching to an online real-life task scenario experiment. In order to simulate the self-paced online recording modality, a new circular progress bar that was recording the interface was suggested, which offers the users the option of self-paced task execution.

In terms of performance results, the high pitch tone sound imagery task showed a slightly better classification accuracy than the siren-like sound imagery for both the covert and inhibited overt speech modes (C_High: 74.89%, C_Siren: 70.88%, IO_High: 74.84 and IO_Siren: 73.76). In the covert speech mode, the high tone sound imagery task showed a significantly higher accuracy than the siren-life sound imagery task. However, there was no significant difference between the two sound imagery tasks in the inhibited overt speech mode. In a simulated-online condition, a new performance assessment method, called the True-False-Positive (TFP) score was suggested for a self-paced system evaluation. The overall TFP score was 76.67% (True-positive rate: 87.62% and False-positive rate: 4.05%) for seven subjects. There was also no significant difference between the sound imagery task modes and the highest performance results would all vary depending on the subject. Therefore, the high pitch tone sound imagery task was chosen for the next online experiment as it showed a higher performance in the offline setting and it was easier to produce than the siren-like sound imagery task. In addition, the covert speech mode was chosen as it restricts

any speech related organ motor imagery so that it can be clearly compared to the motor imagery task.

It was difficult to compare our experiment results with other onset detection studies as there was no common evaluation method. Many of them just showed performance results that can only be applied to their own experimental settings such as the hit rate (e.g. [8, 10, 24, 25]). In [82], the average TP rate for the idle vs. motor imagery task was 86.7% (with false positive events being 5.7 times greater) but there was no information regarding the idle period length. In [10], the motor-imagery versus non-control state achieved a 79.67% classification accuracy on average for three subjects. In [18], six different mental tasks versus the idle state had been tested and they achieved between an around 55% (Auditory imagery) and 72% (Motor-imagery) TP rate on average for 5 subjects in an offline setting. In [88], researchers classified motor-imagery tasks vs. the idle state and they used two two-class classifiers for three different classes (left hand and right foot imagery vs. idle). They achieved true-positive rates of around 40% in an offline analysis.

Compared to these studies, our sound imagery onset detection method achieved quite competitive results (76.67% TFP score, 87.62% TP rate, 4.05% FP rate). Furthermore, none of the above studies considered the self-paced approach. However, our study contemplated all the information about the idle period length and self-paced task execution while recording data, which makes it more suitable for future uses in self-paced BCIs.

Comparison of Sound Imagery vs Motor Imagery for Onset Detection

Based on the previous offline and simulated-online experiment results, the online onset detection system was tested in section 7.2. A message opening / sliding image onset task was designed for watching video and reading text daily-life scenarios. In these simulated daily-life situations, motor imagery and our sound imagery tasks have been tested and compared.

In terms of performance comparison, our covert high pitch tone sound imagery task achieved TFP scores of 84.04%, 80.84% and 77.17% with a sliding image, watching video and reading text scenario, respectively on average for twelve subjects. On the other hand, the motor imagery task produced a significantly worse TFP score in the sliding image (66.79%) and watching video (61.07%) scenarios. However, in the reading text scenario motor imagery gave a lower result of 72.61% but there was no significant difference with the sound imagery task. 83.3% (10 out of 12) of the participants showed better performance results with the sound imagery task than the motor imagery, one participant had the same result for both and only one participant, who had previous motor imagery BCI experience produced a better performance result with motor imagery. However, other three participants who also had BCI experience produced a better result with the sound imagery task.

In terms of system speed, the sound imagery onset response time was significantly shorter than the motor imagery task on average for twelve subjects from all daily-life scenarios. The sound imagery vs motor imagery time responses were 3.93 s vs 4.8 s, 4.03 s vs 5.83 s and 4.28 s vs 5.75 s for the sliding image, watching video and reading text scenario, respectively. This greater response speed is a huge advantage in self-paced BCI systems as it is an important aspect of usability. From a usability point of view, nine out of twelve (75%) subjects answered that the sound imagery task

was easier to use than the motor imagery task for the onset detection. The remaining three subjects were BCI research students, who had experience in motor imagery but not in sound imagery tasks. On the other hand, one research student and all naïve subjects answered that the sound imagery task was easier to use. They marked the score on a scale from 0 (very easy to use) to 10 (very difficult to use) and the average score was 4.42 for the sound imagery and 6.42 for the motor imagery task. The difference was statistically significant.

Based on these results, our new sound imagery task outperformed the motor imagery one for the self-paced onset detection BCI system not only in performance but also in usability and system speed. Therefore, this prototype of onset detection system showed some potential for applying BCIs to real-life uses (compared to typical motor imagery tasks) and it will take the BCI field a significant step forward once it is developed further.

In terms of spatial characteristics for the sound imagery task, a common spatial pattern showed that channels around F3, P3 and T7, which are located near Broca's and Wernicke's area (well-known area for speech related tasks), produced a pattern difference between the idle and the sound imagery task period. In addition, from the feature selection procedure, channel F3 was selected the most amount of times as a best class separable channel, followed by channel T7, which shares a common result with CSP spatial analysis. In the spectral domain analysis, the frequency band 20-30 Hz feature was selected the most amount of times from the feature selection procedure. The work in [37] reported that some studies suggested that the 30 Hz range was elicited by the linguistic processing of meaningful words but not of meaningless non-words. However, our high pitch sound imagery task showed the best class separability versus

the idle state with the range of around 20-30 Hz. This requires some research questions to be investigated further.

Towards Outdoor Laboratory Real-world BCI uses

As it was discussed in previous sections, current BCI studies are generally restricted in research only at indoor laboratory settings or they work only in a specific way once all the circumstances are met. It is very difficult to be used like a daily-based system such as smartphone voice activation (e.g. Android - Google now / IOS – hey siri) due to bad usability and practicality that stems from low accuracy, inconvenient EEG device wear, etc. For this reason, we investigated an online message opening onset detection system at outdoor laboratory settings in order to showcase the potential problems of using BCIs in real-world applications and to compare the performance difference between indoor and outdoor laboratory settings.

In order to simulate the practical applications in a real-life situation, the dry EEG device (Enobio system) was used and the experiment took place at an outdoor cafeteria area. The experimental design was the same as in the indoor laboratory setting. The performance of the outdoor experiment showed significantly smaller TFP scores from all daily-life scenarios by producing values of 56.54% (sliding image), 53.46% (watching video) and 58.01% (reading text) on average for seven subjects. Some participants (e.g. P6) showed a dramatic decrease from around 70.8% to 20.3%. On the other hand, participants 3 and 5 reported some performance increase with the watching video and reading text scenario even though it was a very small amount (no significant difference). Although the performance varied depending on the subject, it generally produced worse results for the outdoor experiment and the participants commonly

answered it was more difficult for them to concentrate and control the task (except P5 and P7 who mentioned it was similar). The participants also listed their discomforts during the outdoor experiment. They commonly answered the following: 1. Other people's noise (talking, ordering food, etc.). 2. Kitchen noise and food smell (coffee machine, grinder etc.). 3. Visual distractions (outside view, people passing by, etc.). 4. Background music.

This study compared results between indoor and outdoor settings and showed the problems of moving BCIs towards real-life applications with daily-life task scenarios, which is an essential investigation at this stage in order BCIs to move forward. Therefore, it could be a useful background study for other self-paced real-world BCI areas of research.

8.2 Scientific Contributions

A Novel Onset Detection Method

In this research, a novel onset detection system was suggested with sound imagery (sound production related cognitive task). As it was discussed, there are a couple of clear advantages of our suggested sound imagery over other cognitive tasks (e.g. motor imagery/word/syllable/letter production tasks): ***1)*** Intuitiveness. ***2)*** Useful for people with motor disabilities ***3)*** No significant overlap with other common spontaneous and frequent daily-life cognitive states. ***4)*** No dependence on the user's mother language. In addition to this, our research results showed a significantly better performance with our sound imagery approach than the motor imagery task for onset

detection in daily-life task scenarios. It had a significantly higher true-false-positive score and it was substantially faster and easier to use (better usability based on a subjects' survey). Furthermore, the new onset detection method had been tested at outdoor laboratory settings in order to test its real-world uses and investigate current issues of moving BCIs forward to outdoor laboratory settings.

Based on our thorough literature review, none of the works on onset detection or self-paced BCIs used a speech or sound-production related approach (they mostly used motor imagery). Therefore, this new finding will impact the BCI field especially in self-paced onset detection studies and lead it towards real-life uses. From a short-term point of view, this study will give some background guidelines about the sound imagery onset detection system. In the long-term, once the system is developed further (like Google now or Siri), this onset detection technique can be implemented to any other cue-based BCI study in order to make it a self-paced system by giving onset freedom to users. Therefore, the current novel study will provide the guidelines and suggest the direction of future BCIs in terms of practical applications in the real-world.

New EMG Artefacts Handling Method

Electromyography (EMG) artefacts are a well-known problem not only for BCIs but also for other EEG related studies such as brain mapping and clinical areas. Therefore, EMG artefact handling is an essential procedure but commonly used Blind Source Separation (BSS) methods (e.g. BSS-CCA, ICA, PCA, etc.) could remove not only EMG artefacts but also some useful EEG sources [19-22]. For this reason, a new novel technique for statistically selecting EMG artefact contaminated EEG Channels (EMG-CCh) was proposed in Chapter 4 in order to minimise useful information loss.

The EMG-CChs are selected based on the statistical analysis (determining whether the artefacts played a significant role in class separation) between facial EMG electrodes and scalp EEG channels. Our method can be applied as a simple EMG contaminated channel rejection or it can be combined with any other pre-existing EMG handing procedure by selectively handling artefacts. In this Chapter, we have discussed comparing results between typical BSS techniques to all channels vs. BSS to EMG-CChs with our onset detection data and BCI competition IV data set 2a (online material [23]). The performance results showed that our EMG-CChs selection and handling method had better class separation than typical BSS approaches. More specifically, significant improvements (p<0.05) in class separation were found when using autoregressive coefficients extracted from our onset data in 79% of the cases for ICA, 53% for PCA and 11% for BSS-CCA. Only 7% (ICA), 4% (PCA) and 3% (BSS-CCA) of the tests became significantly worse with our approach; the rest of the cases yielded no statistically significant differences in terms of class separation performance. With the BCI competition data we saw an improvement in 60% of the cases for ICA and 60% for BSS-CCA when using autoregressive coefficients as features.

The method can be used on its own for channel rejection or it can be combined with pre-existing artefact handling techniques and it showed significant class separation improvement (compared to existing techniques) with both our data and the BCI competition data set in many cases. For these reasons, we believe this method can be of use for other EEG studies and in a long-term view, it will impact the field in terms of artefact removal technique.

New Performance Evaluation Method

So far, there is no common and standardised evaluation method for self-paced BCIs. Therefore, the performance results are all varying depending on the case. For example, some papers (e.g. [8, 10, 24, 25]) just report performance results that can only be applied to their own system such as hit rate, which makes it difficult to compare with other studies. For this reason, we proposed a new performance evaluation metric for self-paced BCIs, called the True-False Positive (TFP) score that considers all the conditions of true-positive, false-positive and the idle period length, which is a very important aspect. For example, if two systems have the same number of true-positives and false-positives but one system has a longer idle period, then it certainly is a better system as it has a smaller false-positive ratio during the experiment. Therefore, we proposed a new performance evaluation metric that takes all these matters into account and can be used from all self-paced BCI systems as a standard evaluation metric (this metric was peer reviewed in a journal paper [1]). Once this metric is widely used in self-paced BCI studies, it would provide a lot of benefits to the field as all systems can be clearly compared with this standard.

8.3 Limitations of the Project

There were a couple of limitations during this research which will be discussed in this section.

Biased Sample Population and Statistical Limitation

A limitation of this thesis was the number of sample sizes and characteristics of the participants who took part in the experiments. Some studies (i.e. Chapter 5: Offline testing, section 7.3: Outdoor testing) had not a large enough sample population, therefore the significant test had to be carried out with a multi-run/trials base. In addition, participants were recruited from the University of Essex, therefore they were mostly academic people (aged 18-28), who were very familiar with working environments where computers are involved (some of them were even from the BCI group). Furthermore, there were no participants with motor disabilities who are also an important target population for BCIs (but it would not be a huge problem in our sound imagery task unlike other motor imagery BCI studies). Therefore, there were some biased sample characteristics in this thesis and it should be taken into account for future study.

Limitation Regarding the BCI System Design

In the final online experiment, the self-paced message opening onset interface was designed in a PC-based rather than an actual mobile messenger application. The reason was that we wished to remove any unexpected artefacts related to mobile network signals and to minimise any event-related potentials (visual and auditory), in order to ensure the onset detection task was purely based on the sound imagery in this investigation stage. For this reason, even though we tried our best to simulate the real-life use of the BCI system, there could be some differences if the system is tested with a real mobile based messenger application.

Online Experiment Recording Order

As it was mentioned in the previous sections, there was one limitation in the online outdoor experiment (Chapter 7) with regards to the recording order. The indoor laboratory experiment was continuously conducted at first to all the subjects, followed by the outdoor experiment. We were well-aware that this biased recording order could impact the final results, therefore it had to be variable in a randomised way for the subjects. However, installing BCI equipment to an outdoor cafeteria area first and then return it indoors requires more time and costs more human resources than the opposite. In addition, the main goal of the Chapter was to investigate and compare the motor imagery and our sound imagery task for onset detection at indoor laboratory settings. Therefore, the indoor experiment was conducted first when the participants were at their best condition. However, in order to minimise the biased recording order effects, there was a small tea/coffee break in-between the two experiments and the total experiment time did not exceed 1 and half hours, which can maintain the participants' concentration level in a similar way as in the indoor lab experiment.

Possible issue with the experimental protocol

In order to minimise Visual Event Related Potentials (VEPs) during the onset task activation, we designed a circular progress bar, which considered a visual angle that would avoid VEPs, in an offline and simulated-online experiment. In online real-life scenario experiments, message arriving notification could generate some VEPs even though we applied a very slow sliding notification (from the right side corner) in order to minimise it. In addition, especially in a watching video scenario, named "BBC - The Blue Planet" [155], this technique was used as it requires low cognitive load and

emotional neutrality [156, 157] but there could be some unknown possible VEPs effects.

Furthermore, there could be a possible overlap of mental tasks. In order to test pure sound-production related cognitive tasks for onset detection, the participants were not allowed to measure the time during the task activation. After this stage, they could estimate their task activation time by referring to the circular progress bar. However, there could still exist some counting task overlap as it was difficult to clearly separate them especially when the participants were asked to give their activation time in order to evaluate their onset response speed. The users would only start monitoring the clock if they had intended to do an onset task. Consequently, while it is possible that the onset detection may have been affected by involuntary counting, the chances of false positives were very low as no counting would occur during the idle period. However, compared to another study [18] (mental arithmetic: subtraction task vs the idle state), which was similar to the counting task, this study produced a lower performance value (around 73% accuracy for 5 subjects on average) than our 80.42 TFP score (85.2% TP rate) according to our results. This shows that the sound imagery task had a clearer role as a main task than the counting task.

However, the possibility of these unknown effects was the same on both of our sound imagery tasks and motor imagery onset tasks as all the experimental settings were the same. Therefore, our results (significantly better for the sound imagery onset task than motor imagery) are still very meaningful in terms of the advantages of sound imagery over motor imagery for the onset detection system.

9 Conclusions & Future Work

This project was undertaken in order to develop a novel method of onset detection in Brain-Computer Interfaces (BCIs). It suggested intuitive sound imagery tasks (sound-production related cognitive tasks) in order to increase the usability and practical uses of BCI systems towards real-world self-paced BCI applications apart from the current research-oriented laboratory settings. Although the contributions and limitations of the research were discussed in detail in Chapter 8, the major achievements can briefly be summarised as:

- A novel onset detection system was suggested with a sound imagery task. It is very intuitive and it does not significantly overlap with other common, spontaneous cognitive states, which makes it feasible to be used in daily-life situations. The performance results showed a significantly better TFP score with the use of sound imagery than with the typical motor imagery task for onset detection. In addition, the proposed system showed a significantly faster onset response and that it was easier to use (better usability based on a subjects' survey) than the motor imagery task.

- A novel EMG artefact contaminated EEG channel selection and handling method was proposed. It showed a significant class separation

improvement (which minimises information loss) compared to current blind source separation techniques with the use of both our data and the BCI competition data set.

- A new performance evaluation metric for self-paced BCIs, called the True-False-Positive score (TFP score) was proposed. So far and to the best of our knowledge, there has been no common and standardised performance assessment method that takes into account all the true-positives, false-positives and the idle period length at the same time. Therefore, this new metric would provide a lot of benefits to the BCI field as all the systems can clearly be compared to this standard.

Despite the substantial progress which has been made in the BCI field with this research, there are still areas for further investigation and research questions. A number of possible future studies could investigate the following:

- Moving to additional real application experiments:

  Our onset detection system was tested in a simulated messenger application in various daily-life task scenarios. However, further investigation with real smartphone-based messenger applications would be of interest. Instead of a few hours of simulated experiment, a couple of days of continuous daily usability test would give us a lot of information with regards to the practical use of the sound imagery onset detection system. It could be compared to other mental tasks (e.g. motor imagery) as well.

- Onset system prototype development for other cue-based BCIs:

  Based on our experimental results, the onset detection application can be developed further and it can be a prototype for the smartphone/mobile BCI as is the 'Google now' or 'IOS Hey Siri' command. Alternatively, it can be extended to a smart-home application system. This prototype development can be used as a tool kit in order to design any cue-based BCIs in a self-paced system.

- More thorough investigation in the brain physiology field for the sound imagery task:

  Besides the engineering technological works, a more in-depth brain physiology study can be carried out. Broca's area and Wernicke's area are well-known parts of language and speech related processing and we also found that these areas were elicited (based on feature selection) for the sound imagery. However, we were not able to see the brain physiological pathway for the sound imagery task execution. The Wernicke-Geschwind model is a widely used pathway model but some literature studies (e.g. [27]) insist that it has some errors and is oversimplified. Therefore, more in-depth brain physiological findings would be an interesting topic for future work.

In addition, there are a couple of remaining challenges in onset detection for BCIs and questions on how some of these challenges may be tackled:

- Better performance:

Even though our new sound-production related cognitive onset task showed significantly better performance results than the typical motor imagery approach, a higher performance accuracy would be needed for a real-world onset detection BCI system. If the BCI system is being used as a daily-life device, the onset detection system should have a near to zero false-positive with acceptable true-positive events. Otherwise, frequent false-positive action would waste resources (e.g. battery life) and lead to low usability. A possible solution to this is adaptive learning. Optimal signal processing and a feature extraction process would surely increase performance. However, an adaptive learning system available to the user, which finds the optimal processing for a specific user, would resolve the issue. If the user uses the onset detection system more frequently, it would improve the system performance. Also, finding the optimal feature extraction parameter would be needed such as optimising the AR model order for future works.

- Better usability:

  Even though we used an Enobio (dry electrode equipment [160]) system with 17 electrodes for better usability compared to the gel type EEG device, it is still difficult to say it is a user-friendly machine designed for daily-use. For this reason, a small mobile (compact) device would be necessary for using the onset detection BCI system in real-world applications. However, the use of only the EEG system would provide limited information if the number of channels is reduced. Therefore, a hybrid device such as a combination of EEG and fNIRS (e.g. [164])

could be a possible answer for this. In addition, shortening the onset time response would increase usability of the system as well. Even though our method showed a faster onset speed than that of the motor imagery approach, its time windows can be further decreased by optimising processing algorithms to individual users (e.g., via adaptive learning / genetic algorithms).

References

[1] Y. Song and F. Sepulveda, "A novel onset detection technique for brain–computer interfaces using sound-production related cognitive tasks in simulated-online system," *Journal of Neural Engineering,* vol. 14, no. 1, p. 016019, 2017.

[2] Y. Song and F. Sepulveda, "A Novel Technique for Selecting EMG-Contaminated EEG Channels in Self-Paced Brain-Computer Interface Task Onset," 2017.

[3] Y. Song and F. Sepulveda, "An online self-paced brain-computer interface onset detection based on sound-production imagery applied to real-life scenarios," in *2017 5th International Winter Conference on Brain-Computer Interface (BCI)*, 2017, pp. 46-49.

[4] Y. J. Song and F. Sepulveda, "Classifying siren-sound mental rehearsal and covert production vs. idle state towards onset detection in brain-computer interfaces," in *Brain-Computer Interface (BCI), 2015 3rd International Winter Conference on*, 2015, pp. 1-4: IEEE.

[5] Y. Song and F. Sepulveda, "Classifying speech related vs. idle state towards onset detection in brain-computer interfaces overt, inhibited overt, and covert speech sound production vs. idle state," in *Biomedical Circuits and Systems Conference (BioCAS), 2014 IEEE*, 2014, pp. 568-571: IEEE.

[6] L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors,* vol. 12, no. 2, pp. 1211-1279, 2012.

[7] C. S. L. Tsui, J. Q. Gan, and S. J. Roberts, "A self-paced brain–computer interface for controlling a robot simulator: an online event labelling paradigm and an extended Kalman filter based algorithm for online training," *Medical & biological engineering & computing,* vol. 47, no. 3, pp. 257-265, 2009.

[8] R. Leeb, D. Friedman, G. R. Müller-Putz, R. Scherer, M. Slater, and G. Pfurtscheller, "Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic," *Computational intelligence and neuroscience,* vol. 2007, 2007.

[9] F. Sepulveda, *Brain-actuated Control of Robot Navigation*. INTECH Open Access Publisher, 2011.

[10] R. Scherer, F. Lee, A. Schlogl, R. Leeb, H. Bischof, and G. Pfurtscheller, "Toward self-paced brain–computer communication: navigation through virtual worlds," *Biomedical Engineering, IEEE Transactions on,* vol. 55, no. 2, pp. 675-682, 2008.

[11] M. Fatourechi, R. Ward, and G. Birch, "A self-paced brain–computer interface system with a low false positive rate," *Journal of neural engineering,* vol. 5, no. 1, p. 9, 2008.

[12] K. Brigham and B. Kumar, "Imagined speech classification with EEG signals for silent communication: a preliminary investigation into synthetic telepathy," in *Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on*, 2010, pp. 1-4: IEEE.

[13] C. S. DaSalla, H. Kambara, M. Sato, and Y. Koike, "Single-trial classification of vowel speech imagery using common spatial patterns," *Neural Networks,* vol. 22, no. 9, pp. 1334-1339, 2009.

[14] M. D'Zmura, S. Deng, T. Lappas, S. Thorpe, and R. Srinivasan, "Toward EEG sensing of imagined speech," in *Human-Computer Interaction. New Trends*: Springer, 2009, pp. 40-48.

[15] L. Wang, X. Zhang, X. Zhong, and Y. Zhang, "Analysis and classification of speech imagery eeg for bci," *Biomedical Signal Processing and Control,* vol. 8, no. 6, pp. 901-908, 2013.

[16] C. Herff *et al.*, "Brain-to-text: decoding spoken phrases from phone representations in the brain," *Frontiers in Neuroscience,* vol. 9, Jun 12 2015, Art. no. 217.

[17] J. Roland, P. Brunner, J. Johnston, G. Schalk, and E. C. Leuthardt, "Passive real-time identification of speech and motor cortex during an awake craniotomy," *Epilepsy & Behavior,* vol. 18, no. 1, pp. 123-128, 2010.

[18] M. Dyson, F. Sepulveda, J. Q. Gan, and S. J. Roberts, "Sequential classification of mental tasks vs. idle state for EEG based BCIs," in *Neural Engineering, 2009. NER'09. 4th International IEEE/EMBS Conference on*, 2009, pp. 351-354: IEEE.

[19] A. Vergult *et al.*, "Improving the interpretation of ictal scalp EEG: BSS–CCA algorithm for muscle artifact removal," *Epilepsia,* vol. 48, no. 5, pp. 950-958, 2007.

[20] H. Nam, T. G. Yim, S. K. Han, J. B. Oh, and S. K. Lee, "Independent component analysis of ictal EEG in medial temporal lobe epilepsy," *Epilepsia,* vol. 43, no. 2, pp. 160-164, 2002.

[21] E. Urrestarazu, J. Iriarte, M. Alegre, M. Valencia, C. Viteri, and J. Artieda, "Independent component analysis removing artifacts in ictal recordings," *Epilepsia,* vol. 45, no. 9, pp. 1071-1078, 2004.

[22] S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, "Independent component analysis of electroencephalographic data," *Advances in neural information processing systems,* pp. 145-151, 1996.

[23] C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, and G. Pfurtscheller, "BCI Competition 2008–Graz data set A," *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces), Graz University of Technology,* p. 16, 2008.

[24] B. Xia, D. An, C. Chen, H. Xie, and J. Li, "A mental switch-based asynchronous brain-computer interface for 2D cursor control," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, 2013, pp. 3101-3104: IEEE.

[25] Y. Chae, S. Jo, and J. Jeong, "Brain-actuated humanoid robot navigation control using asynchronous brain-computer interface," in *Neural Engineering (NER), 2011 5th International IEEE/EMBS Conference on*, 2011, pp. 519-524: IEEE.

[26] P. L. Nunez and R. Srinivasan, *Electric fields of the brain: the neurophysics of EEG*. Oxford university press, 2006.

[27] M. F. Bear, B. W. Connors, and M. A. Paradiso, "Neuroscience: Exploring the brain," 2007.

[28] H. Gamboa. (2005, December). *Electroencephalography*. Available: http://en.wikipedia.org/wiki/Electroencephalography

[29] A. Kübler, B. Kotchoubey, J. Kaiser, J. R. Wolpaw, and N. Birbaumer, "Brain–computer communication: Unlocking the locked in," *Psychological bulletin,* vol. 127, no. 3, p. 358, 2001.

[30] T. Fernández *et al.*, "EEG activation patterns during the performance of tasks involving different components of mental calculation," *Electroencephalography and clinical Neurophysiology,* vol. 94, no. 3, pp. 175-182, 1995.

[31] W. Klimesch, "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis," *Brain research reviews,* vol. 29, no. 2, pp. 169-195, 1999.

[32] L. Venables and S. H. Fairclough, "The influence of performance feedback on goal-setting and mental effort regulation," *Motivation and Emotion,* vol. 33, no. 1, pp. 63-74, 2009.

[33] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," *Proceedings of the IEEE,* vol. 89, no. 7, pp. 1123-1134, 2001.

[34] J. Kropotov, *Quantitative EEG, event-related potentials and neurotherapy*. Academic Press, 2010.
[35] L. Zhang, W. He, C. He, and P. Wang, "Improving mental task classification by adding high frequency band information," *Journal of medical systems,* vol. 34, no. 1, pp. 51-60, 2010.
[36] L. M. A. Valerdi, "Human-Computer Interaction in Synchronous Motor Imagery BCIs," PhD, Computer Science & Electronic Engineering, University of Essex, 2014.
[37] E. O. Altenmüller, T. F. Münte, and C. Gerloff, "31. Neurocognitive Functions and the EEG."
[38] G. Pfurtscheller, C. Brunner, A. Schlögl, and F. Lopes da Silva, "Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks," *Neuroimage,* vol. 31, no. 1, pp. 153-159, 2006.
[39] J. Simmlow, "Biosignal and biomedical image processing," *Matlab based applications*.
[40] S. Sanei and J. A. Chambers, *EEG signal processing*. John Wiley & Sons, 2008.
[41] (December). *User Tutorial: EEG Measurement Setup*. Available: http://www.bci2000.org/wiki/index.php/User_Tutorial:EEG_Measurement_Setup
[42] W. Klonowski, "Everything you wanted to ask about EEG but were afraid to get the right answer," *Nonlinear Biomedical Physics,* vol. 3, no. 1, p. 2, 2009.
[43] *Regions of the brain*. Available: http://www.epilepsy.org.au/about-epilepsy/understanding-epilepsy/human-brain-seizures
[44] H. R. Turner, *Science in medieval Islam: an illustrated introduction*. University of Texas Press, 2010.
[45] J. Ruel, N. Trigueiros-Cunha, and J.-L. Puel. *Auditory Pathways*. Available: http://www.cochlea.eu/en/audiometry/objective-methods/voies-et-centres
[46] D. Purves *et al.*, "Neuroscience," *Sunderland, MA: Sinauer Associates,* vol. 3, 2001.
[47] K. S. Saladin, *Anatomy & physiology*. WCB/McGraw-Hill, 1998.
[48] *The Sense of Touch*. Available: http://www.corpshumain.ca/en/Touche_en.php
[49] *Sensory*. Available: http://droualb.faculty.mjc.edu/Course%20Materials/Physiology%20101/Chapter%20Notes/Fall%202011/chapter_10%20Fall%202011.htm
[50] S. Schwerin. *The Anatomy of Movement*. Available: http://brainconnection.brainhq.com/2013/03/05/the-anatomy-of-movement/
[51] A. C. Brown. (2010, 28 Oct). *NEUROSCIENCE: Speech and Language*. Available: http://www.acbrown.com/neuro/Lectures/Lang/NrLangSpch.htm
[52] M. Wester and T. Schultz, "Unspoken speech-speech recognition based on electroencephalography," *Master's thesis, Universität Karlsruhe (TH), Karlsruhe, Germany,* 2006.
[53] A. S. Dick, B. Bernal, and P. Tremblay, "The language connectome new pathways, new concepts," *The Neuroscientist,* vol. 20, no. 5, pp. 453-467, 2014.
[54] M. Catani and D. K. Jones, "Perisylvian language networks of the human brain," *Annals of neurology,* vol. 57, no. 1, pp. 8-16, 2005.
[55] A. D. Friederici, J. Bahlmann, S. Heim, R. I. Schubotz, and A. Anwander, "The brain differentiates human and non-human grammars: functional localization and structural connectivity," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 103, no. 7, pp. 2458-2463, 2006.
[56] B. Bernal and N. Altman, "The connectivity of the superior longitudinal fasciculus: a tractography DTI study," *Magnetic resonance imaging,* vol. 28, no. 2, pp. 217-225, 2010.
[57] J. Brauer, A. Anwander, D. Perani, and A. D. Friederici, "Dorsal and ventral pathways in language development," *Brain and language,* vol. 127, no. 2, pp. 289-295, 2013.

[58] N. F. Dronkers, "A new brain region for coordinating speech articulation," *Nature,* vol. 384, no. 6605, p. 159, 1996.
[59] F. Agosta *et al.*, "Language networks in semantic dementia," *Brain,* p. awp233, 2009.
[60] C. J. Price, "The anatomy of language: contributions from functional neuroimaging," *Journal of anatomy,* vol. 197, no. 3, pp. 335-359, 2000.
[61] J. R. Binder *et al.*, "Human temporal lobe activation by speech and nonspeech sounds," *Cerebral cortex,* vol. 10, no. 5, pp. 512-528, 2000.
[62] J. A. Fiez and S. E. Petersen, "Neuroimaging studies of word reading," *Proceedings of the National Academy of Sciences,* vol. 95, no. 3, pp. 914-921, 1998.
[63] J.-F. Démonet, G. Thierry, and D. Cardebat, "Renewal of the neurophysiology of language: functional neuroimaging," *Physiological reviews,* vol. 85, no. 1, pp. 49-95, 2005.
[64] E. Edwards *et al.*, "Comparison of time–frequency responses and the event-related potential to auditory speech stimuli in human cortex," *Journal of neurophysiology,* vol. 102, no. 1, pp. 377-386, 2009.
[65] V. L. Towle *et al.*, "ECoG gamma activity during a language task: differentiating expressive and receptive speech areas," *Brain,* vol. 131, no. 8, pp. 2013-2027, 2008.
[66] R. T. Canolty *et al.*, "Spatiotemporal dynamics of word processing in the human brain," *Frontiers in neuroscience,* vol. 1, p. 14, 2007.
[67] X. Pei, E. C. Leuthardt, C. M. Gaona, P. Brunner, J. R. Wolpaw, and G. Schalk, "Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition," *Neuroimage,* vol. 54, no. 4, pp. 2960-2972, 2011.
[68] E. C. Leuthardt *et al.*, "Using the electrocorticographic speech network to control a brain–computer interface in humans," *Journal of neural engineering,* vol. 8, no. 3, p. 036004, 2011.
[69] V. G. Kanas, I. Mporas, H. L. Benz, K. N. Sgarbas, A. Bezerianos, and N. E. Crone, "Joint spatial-spectral feature space clustering for speech activity detection from ECoG signals," *IEEE Transactions on Biomedical Engineering,* vol. 61, no. 4, pp. 1241-1250, 2014.
[70] C. E. Reyes, J. L. C. Rugayan, C. Jason, G. Rullan, C. M. Oppus, and G. L. Tangonan, "A study on ocular and facial muscle artifacts in EEG signals for BCI applications," in *Tencon 2012-2012 ieee region 10 conference*, 2012, pp. 1-6: IEEE.
[71] M. B. Khalid, N. I. Rao, I. Rizwan-i-Haque, S. Munir, and F. Tahir, "Towards a brain computer interface using wavelet transform with averaged and time segmented adapted wavelets," in *Computer, Control and Communication, 2009. IC4 2009. 2nd International Conference on*, 2009, pp. 1-4: IEEE.
[72] S. J. Luck, *An introduction to the event-related potential technique*. MIT press, 2014.
[73] *BCI Competition*. Available: http://www.bbci.de/competition/
[74] X. Gao, D. Xu, M. Cheng, and S. Gao, "A BCI-based environmental controller for the motion-disabled," *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 11, no. 2, pp. 137-140, 2003.
[75] J. R. Wolpaw, D. J. McFarland, and T. M. Vaughan, "Brain-computer interface research at the Wadsworth Center," *IEEE Transactions on Rehabilitation Engineering,* vol. 8, no. 2, pp. 222-226, 2000.
[76] B. Blankertz, F. Losch, M. Krauledat, G. Dornhege, G. Curio, and K.-R. Müller, "The Berlin Brain--Computer Interface: accurate performance from first-session in BCI-naive subjects," *IEEE transactions on biomedical engineering,* vol. 55, no. 10, pp. 2452-2462, 2008.
[77] G. Pfurtscheller *et al.*, "Graz-BCI: state of the art and clinical applications," *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society,* vol. 11, no. 2, p. 177, 2003.

[78] N. Birbaumer *et al.*, "A spelling device for the paralysed," *Nature,* vol. 398, no. 6725, pp. 297-298, 1999.
[79] M. Dyson, F. Sepulveda, and J. Q. Gan, "Localisation of cognitive tasks used in EEG-based BCIs," *Clinical Neurophysiology,* vol. 121, no. 9, pp. 1481-1493, 2010.
[80] C. Herff, D. Heger, F. Putze, C. Guan, and T. Schultz, "Self-paced BCI with NIRS based on speech activity."
[81] C. Herff, F. Putze, D. Heger, C. Guan, and T. Schultz, "Speaking mode recognition from functional near infrared spectroscopy," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 1715-1718: IEEE.
[82] C. Tsui, A. Vučković, R. Palaniappan, F. Sepulveda, and J. Gan, *Narrow band spectral analysis for movement onset detection in asynchronous BCI*. na, 2006.
[83] L. Y. Ganushchak, I. K. Christoffels, and N. O. Schiller, "The use of electroencephalography in language production research: a review," *Frontiers in psychology,* vol. 2, 2011.
[84] A. Porbadnigk, M. Wester, and T. S. Jan-p Calliess, "EEG-based speech recognition impact of temporal effects," 2009.
[85] K. Brigham and B. Kumar, "Subject identification from electroencephalogram (EEG) signals during imagined speech," in *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, 2010, pp. 1-8: IEEE.
[86] B. A. S. Hasan and J. Q. Gan, "Unsupervised movement onset detection from EEG recorded during self-paced real hand movement," *Medical & biological engineering & computing,* vol. 48, no. 3, pp. 245-253, 2010.
[87] F. Popescu, B. Blankertz, and K.-R. Müller, "Computational challenges for noninvasive brain computer interfaces," *IEEE Intelligent Systems,* vol. 23, no. 3, pp. 78-79, 2008.
[88] D. Zhang, Y. Wang, X. Gao, B. Hong, and S. Gao, "An algorithm for idle-state detection in motor-imagery-based brain-computer interface," *Computational intelligence and neuroscience,* vol. 2007, pp. 5-5, 2007.
[89] Y. Huang, Q. Wu, X. Lei, P. Yang, P. Xu, and D.-Z. Yao, "An algorithm for idle-state detection and continuous classifier design in motor-imagery-based BCI," *Power,* vol. 2, no. 4, p. 6, 2009.
[90] J. del R Millán and J. Mouriño, "Asynchronous BCI and local neural classifiers: an overview of the adaptive brain interface project," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on,* vol. 11, no. 2, pp. 159-161, 2003.
[91] K. Qian, P. Nikolov, D. Huang, D.-Y. Fei, X. Chen, and O. Bai, "A motor imagery-based online interactive brain-controlled switch: paradigm development and preliminary test," *Clinical neurophysiology,* vol. 121, no. 8, pp. 1304-1313, 2010.
[92] J. Osselton, "Acquisition of EEG data by bipolar unipolar and average reference methods: a theoretical comparison," *Electroencephalography and clinical neurophysiology,* vol. 19, no. 5, pp. 527-528, 1965.
[93] J. Asensio-Cubero, "Multiresolution Analysis Over Graphs for Brain Computer Interfacing," PhD, Computer Science & Electronic Engineering, University of Essex, 2014.
[94] J. Minguillon, M. A. Lopez-Gordo, and F. Pelayo, "Trends in EEG-BCI for daily-life: Requirements for artifact removal," *Biomedical Signal Processing and Control,* vol. 31, pp. 407-418, 2017.
[95] J. A. Urigüen and B. Garcia-Zapirain, "EEG artifact removal—state-of-the-art and guidelines," *Journal of neural engineering,* vol. 12, no. 3, p. 031001, 2015.
[96] M. Borga, "Canonical correlation: a tutorial," *On line tutorial* http://people*. imt. liu. se/magnus/cca,* vol. 4, p. 5, 2001.

[97] D. Safieddine *et al.*, "Removal of muscle artifact from EEG data: comparison between stochastic (ICA and CCA) and deterministic (EMD and wavelet-based) approaches," *EURASIP Journal on Advances in Signal Processing,* vol. 2012, no. 1, pp. 1-15, 2012.

[98] W. De Clercq, A. Vergult, B. Vanrumste, W. Van Paesschen, and S. Van Huffel, "Canonical correlation analysis applied to remove muscle artifacts from the electroencephalogram," *Biomedical Engineering, IEEE Transactions on,* vol. 53, no. 12, pp. 2583-2587, 2006.

[99] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural networks,* vol. 13, no. 4, pp. 411-430, 2000.

[100] T.-P. Jung *et al.*, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology,* vol. 37, no. 02, pp. 163-178, 2000.

[101] Y. Wang and T.-P. Jung, "Improving brain–computer interfaces using independent component analysis," in *Towards Practical Brain-Computer Interfaces*: Springer, 2012, pp. 67-83.

[102] G. Barbati, C. Porcaro, F. Zappasodi, P. M. Rossini, and F. Tecchio, "Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals," *Clinical Neurophysiology,* vol. 115, no. 5, pp. 1220-1232, 2004.

[103] A. Greco, N. Mammone, F. C. Morabito, and M. Versaci, "Kurtosis, Renyi's entropy and independent component scalp maps for the automatic artifact rejection from EEG data," *International Journal of Signal Processing,* vol. 2, no. 4, pp. 240-244, 2006.

[104] L. I. Smith, "A tutorial on principal components analysis," *Cornell University, USA,* vol. 51, p. 52, 2002.

[105] V. Krishnaveni, S. Jayaraman, L. Anitha, and K. Ramadoss, "Removal of ocular artifacts from EEG using adaptive thresholding of wavelet coefficients," *Journal of neural engineering,* vol. 3, no. 4, p. 338, 2006.

[106] P. Senthil Kumar, R. Arumuganathan, K. Sivakumar, and C. Vimal, "Removal of artifacts from EEG signals using adaptive filter through wavelet transform," in *Signal Processing, 2008. ICSP 2008. 9th International Conference on*, 2008, pp. 2138-2141: IEEE.

[107] A. Al Jumah, "Denoising of an image using discrete stationary wavelet transform and various thresholding techniques," 2013.

[108] M. Kirkove, C. François, and J. Verly, "Comparative evaluation of existing and new methods for correcting ocular artifacts in electroencephalographic recordings," *Signal Processing,* vol. 98, pp. 102-120, 2014.

[109] F. Lotte, M. Congedo, A. Lécuyer, F. Lamarche, and B. Arnaldi, "A review of classification algorithms for EEG-based brain–computer interfaces," *Journal of neural engineering,* vol. 4, 2007.

[110] K. Brigham and B. V. Kumar, "Imagined speech classification with EEG signals for silent communication: a preliminary investigation into synthetic telepathy," in *Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on*, 2010, pp. 1-4: IEEE.

[111] J. Wang, G. Xu, L. Wang, and H. Zhang, "Feature extraction of brain-computer interface based on improved multivariate adaptive autoregressive models," in *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on*, 2010, vol. 2, pp. 895-898: IEEE.

[112] J. Müller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, "Designing optimal spatial filters for single-trial EEG classification in a movement task," *Clinical neurophysiology,* vol. 110, no. 5, pp. 787-798, 1999.

[113] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *Rehabilitation Engineering, IEEE Transactions on,* vol. 8, no. 4, pp. 441-446, 2000.

[114] B. Blankertz, R. Tomioka, S. Lemm, M. Kawanabe, and K.-R. Muller, "Optimizing spatial filters for robust EEG single-trial analysis," *Signal Processing Magazine, IEEE,* vol. 25, no. 1, pp. 41-56, 2008.

[115] Y. Wang, S. Gao, and X. Gao, "Common spatial pattern method for channel selelction in motor imagery based brain-computer interface," in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, 2006, pp. 5392-5395: IEEE.

[116] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, "Wavelet toolbox," *The MathWorks Inc., Natick, MA,* 1996.

[117] G. Strang and T. Nguyen, *Wavelets and filter banks*. SIAM, 1996.

[118] Y. Song, "Controlling Android-Base Smartphone Using P300 Brain-Computer Interface (BCI): Feasibility Study and Prototype Development," MSc, Computer Science and Electronic Engineering, University of Essex, 2013.

[119] J. C. Bezdek and N. R. Pal, "Some new indexes of cluster validity," *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on,* vol. 28, no. 3, pp. 301-315, 1998.

[120] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* no. 2, pp. 224-227, 1979.

[121] F. Sepulveda, M. Meckes, and B. Conway, "Cluster separation index suggests usefulness of non-motor EEG channels in detecting wrist movement direction intention," in *Cybernetics and Intelligent Systems, 2004 IEEE Conference on*, 2004, vol. 2, pp. 943-947: IEEE.

[122] F. Lotte, M. Congedo, A. Lécuyer, and F. Lamarche, "A review of classification algorithms for EEG-based brain–computer interfaces," *Journal of neural engineering,* vol. 4, 2007.

[123] S. R. Gunn, "Support vector machines for classification and regression," *ISIS technical report,* vol. 14, 1998.

[124] R. Poli, D. Valeriani, A. Matran-Fernandez, C. Cinel, and L. Citi, "Performance Measures for Continuous Brain-Computer Interfaces."

[125] *Confusion matrix*. Available: https://en.wikipedia.org/wiki/Confusion_matrix

[126] G. Dornhege, *Toward brain-computer interfacing*. MIT press, 2007.

[127] M. S. Reddy, B. Narasimha, E. Suresh, and K. S. Rao, "Analysis of EOG signals using wavelet transform for detecting eye blinks," in *Wireless Communications and Signal Processing (WCSP), 2010 International Conference on*, 2010, pp. 1-4: IEEE.

[128] I. Goncharova, D. J. McFarland, T. M. Vaughan, and J. R. Wolpaw, "EMG contamination of EEG: spectral and topographical characteristics," *Clinical Neurophysiology,* vol. 114, no. 9, pp. 1580-1593, 2003.

[129] V. Krishnaveni, S. Jayaraman, S. Aravind, V. Hariharasudhan, and K. Ramadoss, "Automatic identification and Removal of ocular artifacts from EEG using Wavelet transform," *Measurement science review,* vol. 6, no. 4, pp. 45-57, 2006.

[130] M. Fatourechi, A. Bashashati, R. K. Ward, and G. E. Birch, "EMG and EOG artifacts in brain computer interface systems: A survey," *Clinical neurophysiology,* vol. 118, no. 3, pp. 480-494, 2007.

[131] H. Hotelling, "Relations between two sets of variates," *Biometrika,* vol. 28, no. 3/4, pp. 321-377, 1936.

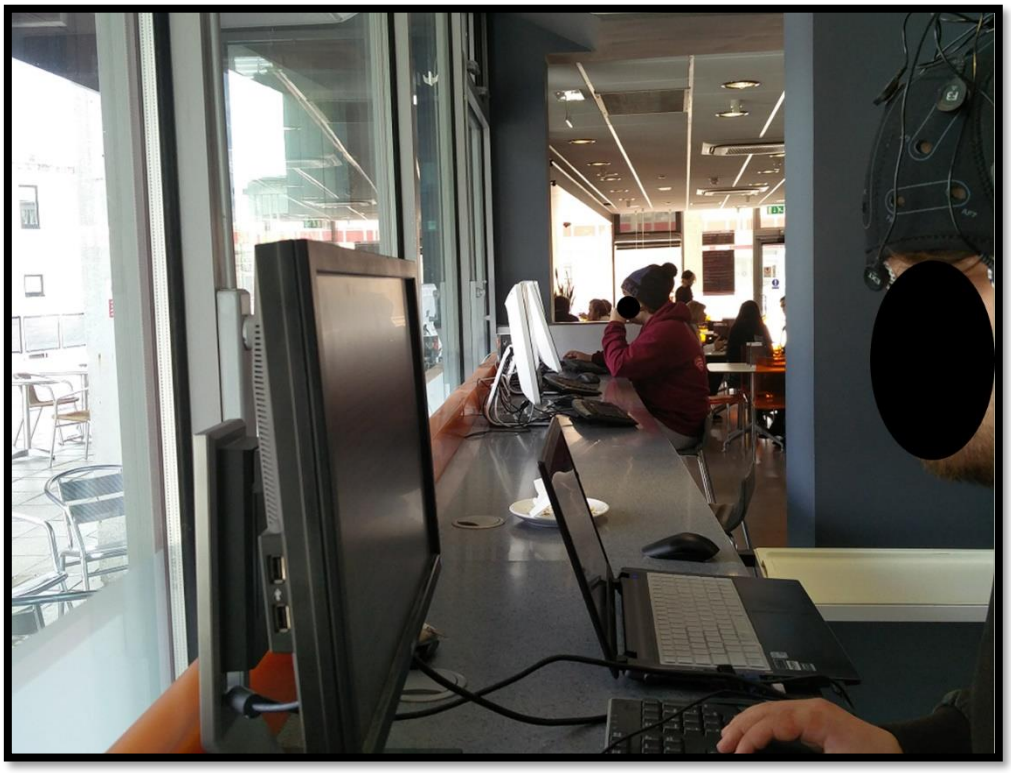
[132] O. Friman, M. Borga, P. Lundberg, and H. Knutsson, "Exploratory fMRI analysis by autocorrelation maximization," *NeuroImage,* vol. 16, no. 2, pp. 454-464, 2002.

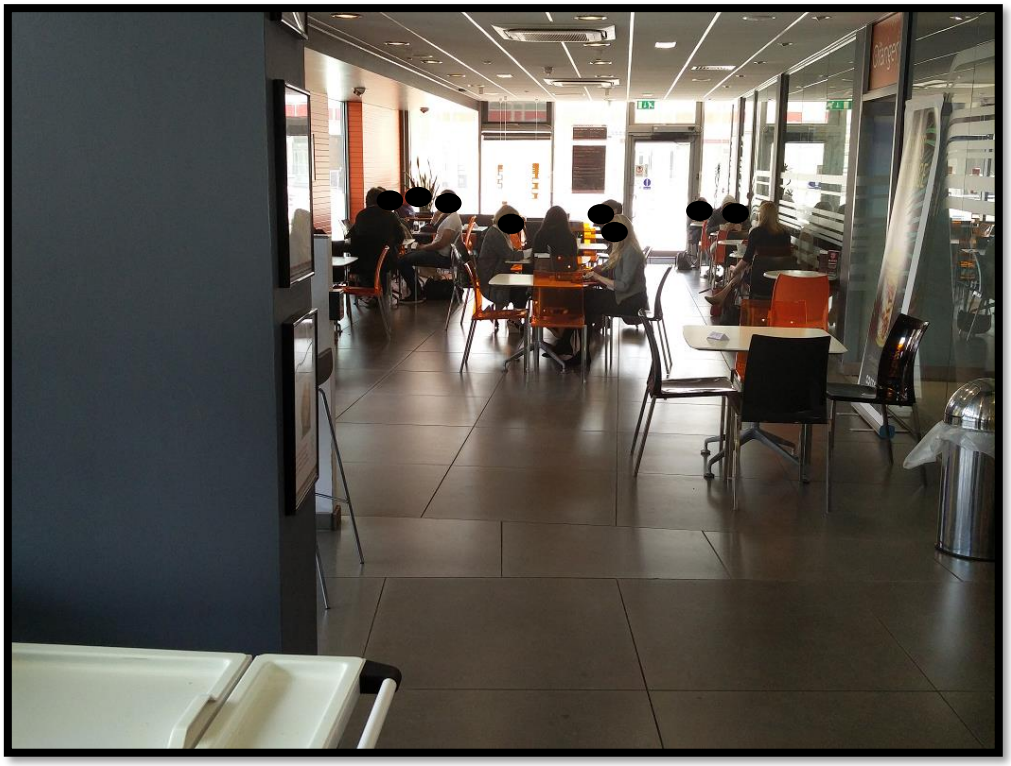
[133] J. Gao, C. Zheng, and P. Wang, "Online removal of muscle artifact from electroencephalogram signals based on canonical correlation analysis," *Clinical EEG and neuroscience,* vol. 41, no. 1, pp. 53-59, 2010.
[134] P. Sharmilakanna and R. Palaniappan, "EEG artifact reduction in VEP using 2-stage PCA and N4 analysis of alcoholics," in *Intelligent Sensing and Information Processing, 2005. ICISIP 2005. Third International Conference on*, 2005, pp. 1-7: IEEE.
[135] J. P. Burg, "A new analysis technique for time series data," 1968.
[136] F. Kovács, C. Legány, and A. Babos, "Cluster validity measurement techniques," in *Proceedings of the 6th International Symposium of Hungarian Researchers on Computational Intelligence, Budapest*, 2005, pp. 18-19: Citeseer.
[137] P. Quilter, B. MacGillivray, and D. Wadbrook, "The removal of eye movement artefact from EEG signals using correlation techniques," in *Random Signal Analysis, IEEE Conference Publication*, 1977, vol. 159, pp. 93-100.
[138] C. Tsui, A. Vuckovic, R. Palaniappan, F. Sepulveda, and J. Gan, "Narrow band spectral analysis for movement onset detection in asynchronous BCI," in *The 3rd international workshop on brain–computer interfaces, Graz, Austria*, 2006, pp. 30-31.
[139] K. Brigham and B. Vijaya Kumar, "Imagined speech classification with EEG signals for silent communication: a preliminary investigation into synthetic telepathy," in *Bioinformatics and Biomedical Engineering (iCBBE), 2010 4th International Conference on*, 2010, pp. 1-4: IEEE.
[140] S. Balakrishnama and A. Ganapathiraju, "Linear discriminant analysis-a brief tutorial," *Institute for Signal and information Processing,* 1998.
[141] M. Dyson, "Selecting optimal cognitive tasks for control of a brain computer interface," Ph.D., School of Computer Science and Electronic Engineering, University of Essex, 2010.
[142] B. Perseh and A. R. Sharafat, "An efficient P300-based bci using wavelet features and IBPSO-based channel selection," *Journal of medical signals and sensors,* vol. 2, no. 3, p. 128, 2012.
[143] C. Chang and C. Lin. (2014). *LIBSVM: A Library for Support Vector Machines*. Available: http://www.csie.ntu.edu.tw/~cjlin/libsvm/
[144] M. R. Al-Mulla, F. Sepulveda, and M. Colley, "Evolved pseudo-wavelet function to optimally decompose sEMG for automated classification of localized muscle fatigue," *Medical engineering & physics,* vol. 33, no. 4, pp. 411-417, 2011.
[145] S. Martin *et al.*, "Decoding spectrotemporal features of overt and covert speech from the human cortex," *Frontiers in neuroengineering,* vol. 7, 2014.
[146] J. O. Pickles, *An introduction to the physiology of hearing*. BRILL, 2012.
[147] K. R. Sitek, D. H. Mathalon, B. J. Roach, J. F. Houde, C. A. Niziolek, and J. M. Ford, "Auditory cortex processes variation in our own speech," 2013.
[148] K. B. Ng, A. P. Bradley, and R. Cunnington, "Effect of competing stimuli on SSVEP-based BCI," in *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 2011, pp. 6307-6310: IEEE.
[149] C. Cinel, R. Poli, and L. Citi, "Possible sources of perceptual errors in P300-based speller paradigm," 2004.
[150] N. Crone *et al.*, "Electrocorticographic gamma activity during word production in spoken and sign language," *Neurology,* vol. 57, no. 11, pp. 2045-2053, 2001.
[151] I. Güler and E. D. Übeyli, "Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients," *Journal of neuroscience methods,* vol. 148, no. 2, pp. 113-121, 2005.
[152] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Systems with Applications,* vol. 32, no. 4, pp. 1084-1093, 2007.

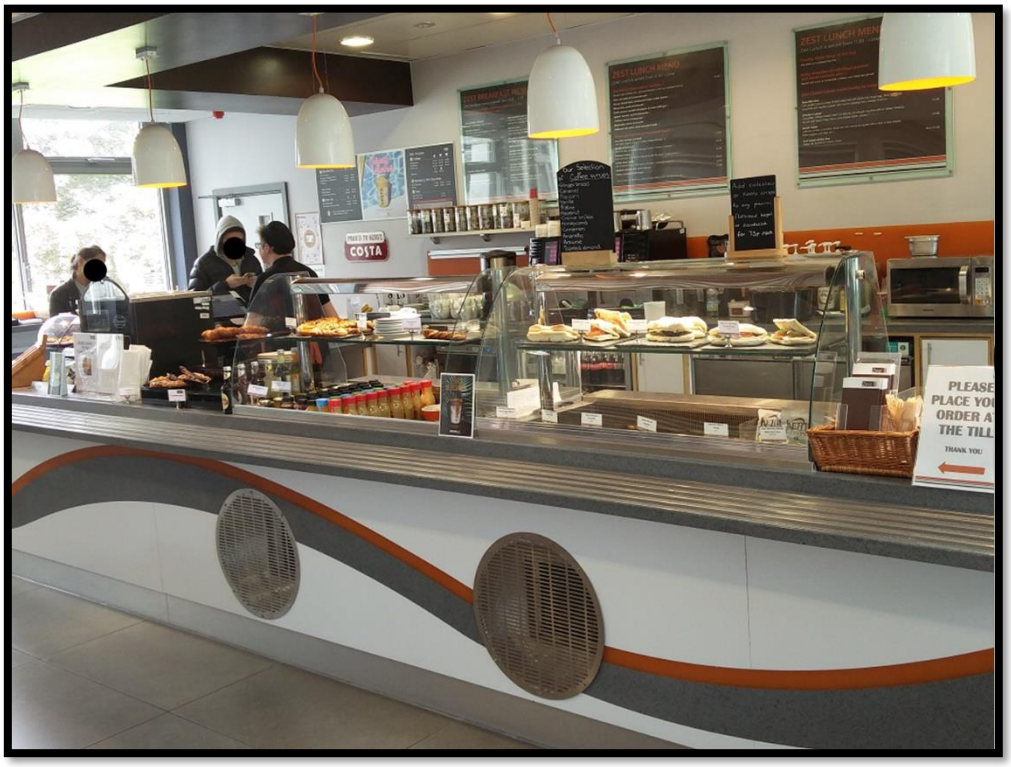
[153] T. Gandhi, B. K. Panigrahi, and S. Anand, "A comparative study of wavelet families for EEG signal classification," *Neurocomputing,* vol. 74, no. 17, pp. 3051-3057, 2011.
[154] G. Townsend, B. Graimann, and G. Pfurtscheller, "Continuous EEG classification during motor imagery-simulation of an asynchronous BCI," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on,* vol. 12, no. 2, pp. 258-265, 2004.
[155] A. Fothergill, "The blue planet," *Á London: BBC Worldwide Publishing (DVD),* 2001.
[156] S. Eltiti *et al.*, "Does short-term exposure to mobile phone base station signals increase symptoms in individuals who report sensitivity to electromagnetic fields? A double-blind randomized provocation study," *Environmental health perspectives,* pp. 1603-1608, 2007.
[157] S. Eltiti *et al.*, "Short-term exposure to mobile phone base station signals does not affect cognitive functioning or physiological measures in individuals who report sensitivity to electromagnetic fields and controls," *Bioelectromagnetics,* vol. 30, no. 7, pp. 556-563, 2009.
[158] J. Jacobs, *English fairy tales*. 1st World Publishing, 2004.
[159] G. Tkačik *et al.*, "Natural images from the birthplace of the human eye," *PLoS one,* vol. 6, no. 6, p. e20409, 2011.
[160] Neuroelectrics. *Enobio Neuroelectrics User Manual*. Available: http://www.neuroelectrics.com/products/enobio/
[161] Cirrus. Available: http://www.cirrusresearch.co.uk/products/sound-level-meters/
[162] A. Apps. *Sound Meter on Google Play Store*. Available: https://play.google.com/store/apps/details?id=com.gamebasic.decibel&hl=en
[163] Oneplus. *Oneplus2*. Available: https://oneplus.net/uk/2
[164] J. Shin, K.-R. Müller, C. H. Schmitz, D.-W. Kim, and H.-J. Hwang, "Evaluation of a Compact Hybrid Brain-Computer Interface System," *BioMed Research International,* vol. 2017, 2017.

Appendix

A. Online outdoor laboratory experiment environment photos






COSTA
PLEASE
THANK YOU