UNIVERSITY OF SOUTHAMPTON

FACULTY OF ENGINEERING AND THE ENVIRONMENT

Institute of Sound and Vibration Research

PATTERN IDENTIFICATION OF MOVEMENT RELATED STATES IN BIOSIGNALS

by

Khondaker Abdullah-Al-Mamun

Thesis for the degree of Doctor of Philosophy

June 2012

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND THE ENVIRONMENT
Institute of Sound and Vibration Research

Doctor of Philosophy

PATTERN IDENTIFICATION OF MOVEMENT RELATED STATES IN BIOSIGNALS

by Khondaker Abdullah-Al-Mamun

The advancement in biosignal processing and modelling has led to exploring the human brain and developing assistive Human Machine Interface (HMI) as well as Brain Machine Interface (BMI). HMI and BMI require specialised techniques for signal processing and pattern recognition to reliably translate information from complex non-stationary dynamics of biosignals into controlling commands. The information translation process consists of signal pre-processing, feature identification and classification. Even though there is continuous progress in biosignal processing research, the critical requirement for HMI and BMI has raised significant challenges for current state-of-art translation methods, such as high accuracy, reliability, and robustness in noise, provided that only small amount of data is available in practice. Therefore, analysing biosignals with novel feature enhancement, feature selection and classification methods are important for decoding of movement intention towards development of reliable assistive HMI as well as BMI. It is particularly valuable for neural signal analysis to understand the neural circuit mechanisms. This research project aims to design decoding algorithm with improved classification performance in robustness and accuracy to recognise movement related states from tongue movement ear pressure (TMEP) signals and deep brain local field potentials (LFPs) by integrating features extracted through multiple domains, and applying pattern classification methods. To achieve the above aim, this project addresses a number of research issues by utilising conventional and efficient signal information extraction, selection and pattern classification techniques.

The first part of this research project successfully developed a robust decoding technique for identifying tongue movement commands from TMEP signals in adverse environment for designing an assistive HMI. This decoding strategy utilised wavelet method for optimal feature enhancement and achieved high accuracy in real time with pattern classification methods of Bayesian and support vector machine (SVM). In the second part, the movement commands are decoded from deep brain local field potentials (LFPs) from basal ganglia (Subthalamic Nucleus (STN) or Globus Pallidus interna (GPi)). An efficient translation algorithm is developed to decode deep brain LFPs for identification of movement activities. Neural synchronisation measures including event related desynchronisation and synchronisation, and functional coupling are utilised to extract discriminatory information as features. We further developed a new feature selection strategy named as weighted sequential feature selection (WSFS) to select an optimal feature subset, which is proved robust for high dimensional, small size dataset. Together with WSFS and pattern classification methods (Bayesian or SVM) high decoding performance for identifying movements was achieved. This research work not only assists decoding movement activities for the application of BMI, but also may help to advance understanding of the neural circuit mechanisms related to motor control as well as development of more efficient therapeutic techniques for neuromotor diseases, such as Parkinson disease.

Table of Contents

List of Figures

List of Tables

Declaration of Authorship

I, KHONDAKER ABDULLAH-AL-MAMUN,
declare that the thesis entitled
PATTERN IDENTIFICATION OF MOVEMENT RELATED STATES IN BIOSIGNALS
and the work presented in this thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University;
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- Where I have consulted the published work of others, this is always clearly attributed;
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- Parts of this work have been published as conference and journal papers and listed in ***Appendix A.***

Signed: ……

Date:……

Acknowledgements

First and foremost I would like to express my heartiest gratitude and honour to my supervisor, Dr. Shouyan Wang for his advice, constructive criticisms and guidance. His encouragement helped in every stage of accomplishment of my PhD research work and made it much more enjoyable, and productive. He made me more critical about my ideas and thinking beyond the boundary. He is not only supervising my research work during my PhD study but also advised me to efficiently maintain work and family life balance. I would like to take the opportunity to thank my co-supervisors Prof. Mark Lutman and Dr. Ravi Vaidyanathan for their continuous advice and support during my course of study. I would also like to thank Prof. Steve Gunn's help on my understanding of support vector machine classification and his course on advance machine learning. I am grateful to Prof. Stephen Elliott, Prof. Christopher J. James, Dr. David M. Simpson and Dr Steven Bell for their constructive criticisms and advice on my work, which helped me to shape up my research better. I would like to thank Dr Stefan Bleeck, Dr. Ghulam Muhammad, and Dr. Guoping Li for valuable discussion and suggestion, which gave me a lot of inspiration. My thesis would be impossible without the financial support from ISVR and EPSRC. My special thanks to Michael Alec Vere Mace, colleague in my project for his co-operation. Also thanks to my colleagues Hadeel Alsaleh, Eliza M. Tucker, Hala Alomari, and Tayyaba Azim.

Finally, special thanks to my grandfather, my parents and my uncle for their unconditional love, understanding, great support and encouragement to pursue higher study. I also like to extend my deepest thanks to my wife for her love and support over the past three years and my two years old son Maahir for making the job of working towards PhD thesis that bit more difficult but extended the joy in my life. My ultimate tribute goes to the Almighty God for bringing this work on light.

List of Abbreviations

AI	Artificial Intelligence
AIC	Akaike information criterion
ANC	Activity of neural cell
AR	Auto regressive
BC	Bayesian classifier
BCI	Brain computer Interface
BMI	Brain machine Interface
CA	Classification accuracy
CT	Classification time
CWT	Continuous wavelet transform
DBS	Deep-brain stimulation
DNI	Direct neural interface
DTF	Directed transfer function
DWT	Discrete wavelet transform
EAR	Eye-based activity recognition
ECoG	Electrocorticography
EEG	Electroencephalogram
EMG	Electromyogram
EOG	Electrooculogram
ERD	Event-related desynchronisation
ERS	Event-related synchronisation
FFT	Fast Fourier transform
fMRI	Functional magnetic resonance imaging
GC	Granger causality
GP	Globus pallidus
GPi	Globus pallidus interna
HMI	Human machine interfaces
HMM	Hidden Markov model
ICA	Independent component analysis

ISF	Interclass separability based on the F-score
IT	Information transfer
LDA	Linear discriminant analysis
LFPs	Local field potentials
M1	primary motor cortex
MI	Motor imagery
MLN	Multi-layer neural network
MRI	Magnetic resonance imaging
MRP	Movement-related potential
MVAR	Multivariate auto regressive
NIRS	Near infrared spectroscopy
NN	Neural networks
PCA	Principal component analysis
PD	Parkinson disease
PSD	Power spectral density
RBF	Radial basis function
SBFS	Sequential backward feature selection
SCP	Slow cortical potential
SD	Standard deviation
SFFS	Sequential forward feature selection
SFS	Sequential feature selection
SNR	Signal-to-noise ratio
STFT	Short time Fourier transform
STN	Subthalamic nucleus
SVM	Support vector machine
SWT	Stationary wavelet transform
SWT-DS3	SWT coefficients at scale 3
SWT-DS4	SWT coefficients at scale 4
TIME	Time information
TMEP	Tongue-movement-ear-pressure
WPT	Wavelet packet transform
WPT-DS3	WPT coefficients at scale 3
WPT-DS4	WPT coefficients at scale 4
WSFS	Weighted sequential feature selection
WT	Wavelet transform

Chapter 1 : ***Introduction***

1.1 Background

Research on bioengineering has been an exciting and challenging scientific research discipline that advances our knowledge through integrating the field of engineering with the biological sciences and clinical medicine. Different discoveries in bioengineering research have made more or less significant impact in our society that improved human health and enhanced human living. Good examples of such discoveries include battery powered cardiac pacemaker that substitutes the natural cardiac pacemaker, helps to improve heart's electrical activity of the people with heart dysfunction; neural prosthetic device of cochlear implant or bionic ear that helps to recover hearing sensation for people with deafness. Beside the various successful inventions in the last century, biological systems, particularly function of complex brain network are still mysterious research field to uncover.

Understanding the biological system through investigating physiological and neurophysiological activities of human has been stimulating multiple research fields. One of the effective ways of studying these activities is acquiring and processing the respective biomedical and biological signals (biosignals), which represents the underlying characteristics of the physiological processes. Advances of processing biosignals incorporating with artificial intelligence (AI) provides the opportunity to discover new mechanisms, invent new techniques, develop new devices for clinical and physical intervention that improves quality of life for people with disabilities.

Disability creates severe obstacle in human life and disrupts the human day to day activities. A conservative estimate of the overall prevalence is that more than hundred millions of people around the world are suffering different kinds of disabilities (Coyle 2006; Tibble 2004; Brault 2008). Many of them are suffering from different forms of movement or motor impairment, which is caused by dysfunction of the brain, and body

disability due to various brain injury and/or neurodegenerative diseases, such as spinal cord injury (SCI), amyotrophic lateral sclerosis (ALS) or stroke (Vaidyanathan et al. 2007; Stroke 2008; Center 2008; Coyle 2006). These injuries lead to conditions of paralysis, quadriplegia and some cases locked-in syndrome that limits human independence, mobility and communication. To assist people with the above conditions, various bioengineering technologies have been invented so far. The area of such technological systems is broadly categorised as assistive Human Machine interfaces (HMI), which develops an interface that mediates the communication between disable people and assistive devices (Cook & Hussey 2002). With the help of assistive HMI, people with severe disabilities can recover a certain degree of autonomy to lead self-supportive, independent and quality living, as well as it also reduces the social burden and cost of living (LoPresti et al. 2008; Kumar et al. 1997). The idea underlying assistive HMI is to measure biosignals (electric, magnetic, or other physical manifestations) of physiological or neurophysiological activity and to translate these into commands for assistive devices. More specifically, the idea underlying assistive HMI is to detect patterns of biosignals activity and to transfer these patterns to commands executed by assistive devices, such as computer or other electromechanical devices. Prototype systems allow to control power wheelchair or to operate prostheses or artificial limbs, solely based on biosignals related to user intention.

When the development of assistive HMI system specifically considers the pattern of neurophysiological activities of the brain, i.e. neural signals of brain activities, it is further categorised as Brain Machine Interface (BMI) (Wolpaw et al. 2002). In other words, BMI introduces a new augmentative communication and control channel for a user, which is not dependent on the peripheral neuronal or muscular activity. With the technological advancement, BMI creates hope to perform such activity through electromechanical devices using the human brain signal to translate intention into actions (Lebedev & Nicolelis 2006). In addition to that, BMI will not only help disabled people to activate their daily life but also it will open up the possibility to recover patients from ranges of neurological disorders through neuro-rehabilitations. Recent advancement in biosignals processing, particularly processing of neural signal and brain machine interface have been reported in the scientific literature (Pasqualotto et al. 2011; Nicolas-Alonso & Gomez-Gil 2012; Bashashati et al. 2007; Ince et al. 2010; Micera et al. 2010; Huang et al. 2009).

1.2 Motivation and focus of the thesis

A wide range of research has been conducted for developing various assistive HMI based on biosignals for assisting people with physical impairments. In spite of the significant progress made to develop various techniques and devices for assistive HMI, current techniques have not fully addressed users' requirements and better interface between users and assistive devices are still greatly expected. Most of the interfacing techniques need bodily movements to operate the devices, which may be difficult for the disabled people with severe movement disability. Therefore, innovative techniques are in demand to support day to day communication for people with severe movement impairments. One of the promising ways to establish such communication is the brain machine interface that translates neurophysiological activities of the brain through recording of neural signals to operate the assistive devices. On the other hand, such communication can be established by identifying and translating potential biosignal related to physiological activities of the human to operate assistive devices. It is noted that in both cases users do not require bodily movements to perform the communications.

To address the later option, recently a non-intrusive approach was introduced based on tongue-movement-ear-pressure (TMEP) signal to control assistive HMI devices (Vaidyanathan et al. 2007). In this approach, users express their intention by making impulsive actions of the tongue, which create unique acoustic pressure signals within the ear canal. These pressure signals can be recorded easily using a microphone earpiece positioned non-invasively within the ear canal. The advantage of utilising the tongue is that it has an inherent capability for fine motor control, involving multiple degrees of freedom, as it has evolved to perform sophisticated motions during speech and mastication. The system also has the additional benefits of being simple, cheap and non-invasive. For people with limited control of their limbs, paraplegia or even quadriplegia, if they still have the ability to perform their tongue movement in daily life, they can use different prescribed tongue movements to communicate with computers and control assistive devices through the sensing of bio-acoustic pressure signals.

Continuing research on cognitive neuroscience has made substantial progress to understand the neurophysiological process of the human brain and its applications to the clinical interventions as well as the development of novel BMI system (Kringelbach et

al. 2007; Gerven et al. 2009; Engel et al. 2005). BMI provides a direct communication between the brain and machine through measuring neural information from the human brain and translating that information into actions to support people with severe movement disabilities. The studies on BMI will also lead us for greater understanding of the brain and how brain processes or transfers information.

Generally in the assistive HMI system, biosignals are recorded and analysed for identifying the state of the physiological activities. However, when considering the brain machine interface, the system not only records and analyses the neural signals for identifying the states of brain neurophysiological activities but also the system able to feed signals into the brain. In other words, the aim of BMI is to establish bidirectional communication, in which it can *write-in* signals to the brain, typically through electrical stimulation, or *read-out* signals by recording neural activity. A good example of *write-in* systems is cochlear implants, i.e. devices that transform sounds from the environment into electrical impulses, which are in turn used to directly stimulate auditory nerves. Another *write-in* example is deep-brain stimulation (DBS), which is a surgical implantation of a medical device called a brain pacemaker that sends electrical impulses to specific parts of the brain to suppress abnormal neural activity. DBS made remarkable therapeutic benefits for patients with neurological diseases, such as chronic pain, Parkinson's disease, tremor and dystonia (Kringelbach et al. 2007). This technique is not only clinically useful, but also it can provide new insights into fundamental brain functions through direct manipulation of both local and distributed brain networks. It also provides the opportunity to *read-out* the human neural activities directly from the deep brain structure, which can be possible to integrate in the development of BMI for assisting people with disabilities or develop more advanced therapeutic interventions. However, due to the surgical implantation procedure the considerstion to establish BMI for assistive communication based on DBS is very limited, it can only be consideriable within the disabled population who already have DBS implantation.

The usual target of the DBS implantation for movement disorders is in the subcortical structure of the brain, basal ganglia, which is a part of the brain circuit involved in motor control. As mentioned above DBS offers an unique opportunity to sense the basal ganglia by recording of local field potentials (LFPs) activity related to movement control (Brown & Williams 2005). It is also evident that the cortex and the basal ganglia are involved in the decision making or action selection as well as the preparation,

execution and imagining of movements (Bogacz & Gurney 2007). Therefore basal ganglia neural activity, deep brain LFPs related to movement control could be an alternative or supportive source for designing potential neural interface system to establish communication between brain and machine. More importantly, early prediction of the onset of the abnormal neural activity from deep brain LFPs may provide the opportunity to develop advanced therapeutic interventions, for instance demand driven or close loop DBS. Such devices may activate the deep brain stimulation adaptively based on the demand or feedback to suppress abnormal neural activities (Rosin et al. 2011; Santos et al. 2011; Gasson et al. 2005). This advancement will substantially improve the present continuous or open loop DBS procedure (i.e. pre-defined fixed stimulation pattern and intensity) for the treatment of the patients with Parkinson's disease, tremor and dystonia as well as other diseases.

Motivated by the aforementioned source of biosignals, techniques and their benefits, this thesis focuses on identifying movement related states from biomedical signals (TMEP or LFPs) in order to improve the robustness, accuracy, efficiency and reliability in HMI or BMI systems. An important component of any HMI or BMI system is signal translation, i.e. signal processing and pattern recognition methods that decode biosignals (TMEP or LFPs) related to different types of movement activities. Translation of biosignal, particularly neural signal requires specialised signal processing and pattern recognition methods to extract inherent information reliably from its complex non-stationary dynamics. Hence, in this thesis special emphasis is given to the algorithms that learn from a set of training data and how to discriminate biosignal segments containing a movement activity from one another as well as from interferences. In particular, this research aims to design a decoding algorithm with improved classification performance in robustness and accuracy to recognise movement related states from TMEP signals and deep brain LFPs by integrating features extracted through multiple domains, eliminating redundant information from the feature space, and applying pattern classification methods.

1.3 Challenges and contribution to knowledge

The overall research in this project are the extraction of discriminative feature, reduction of large amount of redundant or least discriminative feature from the feature space, and classification of TMEP signals and deep brain LFPs to identify movement

related states. The specific practical issue makes the project challenging in several aspects. The non-stationary properties of the biosignals causes large variations in features, trial to trial variability or inconsistency within the same class of features, as well as small variations in features between the different classes. Practically only limited or small size of dataset can be collected from healthy people or patients. Such small amount of data makes the training of classifiers less generalised. Again unbalanced size of the dataset in different classes of biosignal further makes the poor generalisation of classifier estimation. The artifacts or noise in real environment requires the classification and pattern identification to be more robust. On the other hand, different biosignals have their own unique characteristics; this is due to the fact that the design of decoding algorithms for particular biosignals requires adapting such characteristics to ensure robustness in classification. The recorded neural activity of LFPs reflects the integration of multiple complex brain functions and there are usually high level background neural activities. It may need a large amount of features to decode the neural activities from such single channel signal. The feature selection and dimensionality reduction is also difficult given the small amount data available. In the following sections challenges in movement decoding from biosignals (TMEP and LFPs) and major contribution will be addressed.

1.3.1 Decoding of TMEP signal

In the TMEP signal identification, several challenges need to be overcome to make it robust and reliable for the application of assistive human machine interface to support patients with disability. One challenge to identify TMEP signals in the real environment is the influence of interferences, including external noise from the surrounding environment, motion artifacts, internal noise or artifacts due to natural tongue movements. Such interference problems are generally challenging in any assistive HMI system. The other challenge is that only a limited number of signals are available to train and calibrate the classifier in real environment. The accuracy and robustness of classification algorithms highly depends on optimal feature enhancement, especially in noisy environments. To address the above challenges notable technical contributions have been made in this thesis are listed in the following.

- A wavelet transform based approach developed for feature enhancement of TMEP signals. As the TMEP signals of movement actions exhibit transient behaviour in the

order of tens of milliseconds, the wavelet approaches can reliably extract features in a time-frequency domain for the classification. Experiment shows that wavelet domain feature outperforms time domain feature in terms of classification accuracy in both clean and noisy environments. In addition to that, wavelet based approaches efficiently extracted discriminative features for identifying TMEP actions from various potential oral interferences.

- The robustness of the TMEP movement action identification method was evaluated with various types of interference signals in subject specific and generalised interference setting, and high identification accuracy was achieved. Feasibility of this identification method was also evaluated in a real-time setting while considering a wide range of potentially interfering factors, for example free speech. The proposed method is still able to maintain a good performance level in these varied environments.
- To overcome the limitation of small size training data, very limited, small and large sizes of training dataset, were evaluated and found that with small size training data, it is possible to get satisfactory performance based on the developed decoding method for classification of TMEP movement actions.

1.3.2 Decoding of deep brain LFPs

Decoding of brain activity requires to investigate the neural circuit mechanisms or to develop a brain machine interface system. As neural signals are very complex in nature, consequently the decoding task of neural signals are more challenging than TMEP signal as well as other biosignals. It requires specialised approaches to reliably decode information for identifying the inherent activity. Neural signals are highly non-stationary consequently extracted signal information has high variability and also contain highly redundant information of the activity. It is also noted that due to the lack of user concentration and training, in most cases neurophysiological activity related neural signal contains least discriminative information. On the other hand, neural activity of local field potentials from basal ganglia is only investigated in clinical perspective. To the best of our knowledge, there is no study to decode basal ganglia local field potentials for identifying its feasibility towards development of BMI. Therefore, investigation of deep brain LFPs by identifying discriminative neural features for decoding movement activity could open up potential possibility to use it for neural interface application as well as for identifying better treatment of diseases through DBS.

To address the challenges in deep brain LFPs decoding of voluntary movements and acquisition of new knowledge, the notable technical contributions have been made in this thesis are listed in the following.

- A neural decoding algorithm is developed for identifying movement activity from deep brain LFPs. Decoding results from twelve subjects show that it is possible to identify voluntary movement activity with high accuracy. These will enhance our understanding of human brain circuit for motor operation and also facilitate to develop advanced neural interface for brain machine communication.
- Beside the wavelet based feature extraction approach, an alternative neural feature extraction approach called neural synchronisation is introduced based on the analysing Granger causality, which is strongly contributed in decoding process and achieved superior performance when incorporated with instantaneous power features.
- A new feature selection strategy, weighted sequential feature selection (WSFS) is developed to select optimal feature subsets by significantly reducing the redundant or least discriminative feature in the feature space. Experimental results show that WSFS significantly improves decoding performance in movement classification.

1.3.3 Publications

During the period of my doctoral study, the research work performed in this thesis has been published in a number of peer reviewed journals and conferences. A list of publications of this research work is included in *Appendix A*.

1.4 Outline of the thesis

This thesis is organised into seven chapters. Chapters 2 and 3 contain background materials, chapters 4 to 6 mainly describe the research specifics of this thesis, in which chapters 4 and 5 describe the research work on TMEP signals and chapter 6 describes the research work on deep brain LFPs. Chapter 7 draws the conclusion of the thesis. The contents of each chapter are listed in the following.

- Chapter 1 introduces the background, motivation and focus of this thesis. The challenges and contributions of this research are summarised.
- Chapter 2 gives a review of assistive HMI and biosignals. Related topics are reviewed including different methods for measuring biosignals i.e. physiological and neurophysiological signals, the types of biosignals that are already investigated or can

be used for assistive HMI and BMI systems, structure of assistive HMI system and its performance measures.

- Chapter 3 provides a review of signal processing and pattern recognition methods for biosignals that includes details of feature extraction, feature selection and classification, and also brief description of these methods.
- Chapter 4 describes pattern classification of tongue movement action from TMEP signals. At first, it illustrates background of TMEP signals and its previous research. Then TMEP signal feature extraction methods based on wavelet packet transform and stationary wavelet transform, and classifications are presented. After that details result and discussion of tongue movement action classification are explained.
- Chapter 5 illustrates robust identification of tongue movement action from interferences. At the beginning describes the acquisitions of TMEP action and interferences related signals. Then, presented the feature enhancement method developed for TMEP action and interferences classification. The offline and online results of TMEP action and interferences classification are also presented. Finally, the discussion including limitation of the TMEP action and interferences classification is presented.
- Chapter 6 presents the detailed research work on decoding movements from deep brain local field potentials. It starts with the background of neural signal including deep brain LFPs and its characteristics. In the second part it describes the signal acquisition procedures. The details of the developed feature extraction methods based on wavelet packet transform and Granger causality, new feature selection methods called weighted sequential feature selection, and classification methods are presented in the third part. Then the detailed comparative results of deep brain LFP decoding are demonstrated. Finally, it discusses the limitation of deep brain LFP decoding.
- Chapter 7 presents the conclusion and states the potential future research.

Chapter 2 : *Review of Assistive Human Machine Interfaces and Biosignals*

2.1 Introduction

Assistive human machine interface (HMI) is an assistive (or augmentative) communication technology that enables individuals with disabilities to perform actions that might otherwise be difficult or impossible. According to the World Health Organization (WHO) assistive technologies are primarily employed to contribute to successful functional outcomes for people with disabilities or impairments (Cook & Hussey 2002). There are many kinds of disabilities that affect people in different ways. Statistics show that a large number of people are suffering with disabilities in the world. In the UK, about 9.8 million (statistics of 2003) people are living with disabilities (Tibble 2004). On the other hand, according to statistics of 2005 in the United States, about 54.4 million (18.7 percent of total population) people have certain level of disability and 35 million (12 percent) have a severe disability (Brault 2008). These figures are only in two developed countries; however there will be more figures of disability in other developed countries as well as in the developing countries. There are several causes of disabilities that vary often in human life. Some people are born with a disability, some people get sick or have an accident that results in a disability and some people develop a disability as they get older. Some of the conditions that cause severe motor disabilities are Alzheimer disease (AD), amyotrophic lateral sclerosis (ALS), cerebral palsy (CP), stroke, multiple sclerosis (MS), Parkinson disease (PD), spinal cord injury (SCI), severe traumatic brain injury, progressive supranuclear palsy (PSP) and so on. People who have the above diseases need support from others like carers or some devices. Research in assistive HMI system are continuously inventing different sort of devices that support people with different disabilities to interface their environments. The assistive HMI system and its interface to establish the communication requires user intention or input to operate the system. Human physiological signals are the common

form of input into most assistive HMI systems that establishes the communication. With the technological advancement, assistive HMI system will not only help the disabled people to manage their daily life but will also open up the possibility to recover patients from ranges of neurological disorders through neuro-rehabilitations using Brain Machine Interface (BMI) (Collinger et al. 2011). BMI is another form of assistive HMI system that translates neurophysiological activity of the brain into device command for establishing communications.

In this chapter, a brief review of assistive HMI technology and potential biosignals to operate the assistive system are presented. Section 2.2 illustrates the assistive HMI and BMI systems. The state of art of biosignals including electrophysiological and neurophysiological signals that have potential to use in assistive HMI are described in section 2.3. A brief review of biosignals used in assistive HMI system is discussed in section 2.4. Assistive HMI system structure and its performance measurement are presented in sections 2.5 and 2.6, respectively, and finally, section 2.7 draws conclusions of this chapter.

2.2 Assistive human machine interface (HMI)

HMI is a technology that enables the interaction between human and machine. In HMI system, it takes user intention as input, processes the input and then generates a command as an output to operate a device. The tasks of HMI are to design, evaluate, and implement a system that will automate human life. When automation is implemented for improving human life, several issues need to be considered such as usability, reliability, safety and also human factors. The communication between human and machine in the HMI systems was categorised into two modes, direct and mediated (O'Malley 2007). Direct human-machine communication can be accomplished via human activities, such as speech, vision, gesture, or different human physiological and neurophysiological signals. On the other hand, mediated communication of HMI can be accomplished via virtual environments, graphical user interfaces, and collaborative software agents. When considering direct human-machine communication for gaining independence to people with physical impairment or disability for living a normal life, then it can be generalised as assistive HMI. Particularly, assistive HMI aims to restore the lost functionality of disabled people (which is otherwise not possible to be performed by them) based on the human physiological activity such as biosignals. A

generic model of assistive HMI deign is shown in figure 2.1. Assistive HMI is generally a broader area of research, which includes type of human physiological activities or signals that can be used as input to drive the system, the decoding process for generating output based on the given input, and the operating devices. Most of the assistive HMI devices investigated or under investigation were based on biosignals that also included brain signals. When the investigation is performed based on the signals of brain activity for developing HMI, then this area is generalised as BMI.

Figure 2.1: A general interface model for assistive HMI. In this model, the functional limitation of a person is shown on the left. The person desires to communicate with his or her environment to perform an activity, which may be to move his or her hand or to interact with the devices. The person's functional limitation results a gap between his or her desired abilities and those abilities required to perform the activity. The resulting inability can be overcome through assistive HMI which can provide the functionality that the person requires to bridge the ability gap to perform the desired activity.

There are ranges of assistive HMI technologies developed and under research for assisting disabled people to perform many activities of daily living thus improving their quality of life. For instance, people with limited hand functionality may use an alternative keyboard or a special mouse to operate a computer; people who are blind may use software that reads text on the screen in a computer-generated voice such as text-to-speech; people with speech impairments may use a device that speaks out loud as they enter text via a keyboard (text-to-speech); people who are deaf may use a TTY (text telephone) or speech-to-text software; or people with low vision may use software that enlarges screen content (O'Malley 2007). On the other hand, advances of assistive HMI technologies are trying to restore human mobility or limbs for assisting paralysed

people to interact with their environment. One of the common ways of assisting human mobility is a power wheelchair. Disabled user can control and operate a power wheelchair based on their ability. Until now, several techniques were invented to control a wheelchair by the respective user that includes conventional joysticks, sip and puff switches, pressure pads, laser pointers, speech recognition systems and force reflecting joysticks (Simpson et al. 2008; Kumar et al. 1997; Vaidyanathan et al. 2007). None of these techniques fully addressed all the ranges of disabled user's requirements, and alternative techniques are still in need that can help severely disabled people to control the wheelchair as well as other devices to manage their daily living. On the other hand, the robotic and prosthetic devices already created hope to restore damaged human limbs for the disabled person (Micera et al. 2010). The prosthetic device can joint with the human body and operate based on body movements or biosignals. It can be re-established the function of one or more limbs. From the literature, it was clear that biosignals and its decoding process are most essential part to develop assistive HMI systems (Kumar et al. 1997; Collinger et al. 2011).

2.2.1 Brain machine interface (BMI)

BMI also known as brain computer interface (BCI) or direct neural interface (DNI), or may be known as other names, however throughout this thesis we refer to it as BMI. BMI is an assistive HMI system that establishes the direct communication pathway between the human or animal brain and an external device. The tasks are usually performed in BMI by detecting complex neurophysiological activity generated from human brain. Later neural signal translates into meaningful control commands to perform the tasks through the device or machine. This field of research may have high impact on disabled individuals who cannot otherwise physically communicate (Mason et al. 2007; Bashashati et al. 2007; Nicolas-Alonso & Gomez-Gil 2012; Pasqualotto et al. 2011). BMI has been based on two different prototypes: *non-invasive,* which measure activity from large groups of neurons with electrodes placed on the surface of the scalp, and *invasive,* which measure activity from single or multiple neurons with miniature wires placed inside the brain. Invasive signal recording is less vulnerable to artifacts and has the advantage of an excellent signal-to-noise ratio (Leuthardt et al. 2009; Lebedev & Nicolelis 2006).

Brain neurophysiological phenomena are specific features of the brain activity that appear in the brain signals and can be categorised in two groups based on the origin of the phenomenon in the brain (Mason et al. 2007; Bashashati et al. 2007). Those neurophysiological phenomena generated as the result of cognitive responses of the brain are called *endogenous*. The ones evoked by an external stimulus are called *exogenous*. Exogenous neurological phenomena based BMI systems usually require less user training. The limitation of these systems is that they require a constant commitment of one of the sensory pathways to an external stimulus. On the other hand, endogenous-based BMI systems rely on the generation of a phenomenon that is more natural and thus expected to cause the users less fatigue. This may be the reason why more than 80% of BMI studies use endogenous neurophysiological phenomena to control BMI systems (Bashashati et al. 2007). To generate a suitable neurophysiological phenomenon, endogenous-based BMI systems usually need extensive user training. This training may take a long time, sometimes even up to a few months. Nonetheless, the use of complex signal processing techniques for detecting weak neurophysiological phenomena can greatly reduce the training process to develop BMI for real-time applications.

2.3 State of art of biosignals for assistive HMI

The human body is made up of many functional systems, for example, the nervous system, the cardiovascular system, and the musculoskeletal system, as well as others. Each of these systems is made up of several subsystems that carry on many physiological processes and also communicate through sending and receiving physiological information. The human physiological processes are complex and they can be measured as physical quantities such as pressure, temperature, electrical pulse (voltages or currents), or biochemical contaminations (Rangayyan 2002). These measured quantities of physiological processes are defined as biosignals. Different types of diseases or disorders alter the normal physiological process of the biological system, and consequently interrupt human life cycle. To identify the diseases or disorders as well as predicting the possible abnormality in the physiological process of the biological system, efficient biosignals recording and monitoring is essential. Particularly, effective monitoring of biosignal has an important clinical need for appropriate diagnosis and/or treatment of diseases. Most of the biosignal acquisition and monitoring techniques are based on measuring of electrical potentials of human body, such as measuring the

heartbeat rate as electrocardiogram (ECG), measuring muscle contraction during neuromuscular activities as Electromyogram (EMG). Some other well defined techniques that measure biosignals are based on ultrasound, or magnetic field. Beside the clinical need, biosignals are playing a substantial role for acquiring information from human body for developing assistive HMI system, such as recording of voluntary movements from muscle activations as form of EMG or from neural activity as form of electroencephalogram (EEG) to control an assistive device that generates a new hope for paralysed people.

The selection of biosignals for assistive HMI system depends on the signal characteristics and its originating source. To enable assistive monitoring and communication, wide varieties of biosignals investigated and have been reported in the literature that include but not limited to ECG, EMG, Electrooculogram, Muscle Force, Blood Pressure, as well as different form of neural signals (Cannan & Hu 2011; Collinger et al. 2011). In the following subsection, brief descriptions of some of the above mentioned biosignals and their acquisition techniques are presented. It starts with the EMG, which measure the muscle activity by detecting surface voltages that occur when a muscle is contacted. Next, techniques for measuring changes of electric field based on eye movement called the electrooculogram (EOG). Then a novel and new type of biosignal for assistive HMI called tongue movement ear pressure (TMEP), which is the measure of pressure signal acquired by monitoring in the ear canal for the different movement of tongue in the oral cavity. Finally different techniques of neurophysiological signal acquisition and their properties are illustrated.

2.3.1 Electromyogram (EMG)

The EMG is an electrophysiological signal that measures electrical current generated in muscles during its contraction, which also represents neuromuscular activities. The nervous system always controls the muscle activity of contraction or relaxation. Hence, the EMG signal is a complex signal, which is dependent on the anatomical and physiological properties of muscles (Oskoei & Hu 2007). In general, the EMG signal is a stochastic (zero-mean random) process, whose standard deviation is proportional to the number of muscle units and the rate of their firing (Wang et al. 2004). The EMG signal amplitude can range from 0 to 10 mV (peak-to-peak) or 0 to 1.5 mV (rms), and useable energy of the signal is limited to the 0 to 500 Hz frequency range, with the

dominant energy being in the 50 - 150 Hz range. It can be recorded non-invasively by using surface electrodes placed on the muscle. When EMG is acquired from electrodes mounted directly on the skin, the signal is a combination of all the muscle fiber action potentials occurring in the muscle(s) underneath of the skin. The electrical activity measured by each muscle electrode and the ground electrode are then amplified and processed to make sense of the underlying activities. The EMG has various applications, which includes biofeedback to train subjects to increase, decrease or stabilise muscle tension, diagnose and treat neuromuscular disorders (Ferreira et al. 2008; Cannan & Hu 2011).

2.3.2 Electrooculogram (EOG)

The human eye can be modelled as a dipole with its positive pole at the cornea and its negative pole at the retina. With such a stable corneo-retinal potential difference, the eye can be considered as the origin of a steady electric potential field. The measured electrical signal from this field is called the electrooculogram (EOG) (Bulling et al. 2011). When the eye moves from the centre position toward the periphery, the cornea approaches one electrode while the retina approaches the opposing one. In the dipole orientation, this variation causes a change in the electric potential field and its measured amplitude of EOG signal. By identifying these variations, eye movements can be tracked and used as control signal. The typical EOG signal frequency content ranges from 0 to 30 Hz. It is a non-invasive recording approach and useful technology to provide support for people with disabilities (Cannan & Hu 2011).

2.3.3 Tongue Movement Ear Pressure (TMEP)

The human ear consists of three parts, i.e. the external ear, the middle ear and the inner ear. The middle ear connects to the back of the oral cavity through the Eustachian tube which is named as auditory tube as well. The Eustachian tube serves to equalise the air pressure on both sides of the eardrum, and allows for drainage of the middle ear by serving as a gateway into the pharynx which is a part of the mouth (Gelfand 2007). The lateral first third of the Eustachian tube begins at the middle ear, whereas the remaining part is enclosed within an incomplete ring of hook-shaped elastic cartilage. The cartilaginous part of the Eustachian tube is normally closed, and the tube opens reflexively by the action of the tensor palatini muscle; which unbends the normally hook-shaped cartilages in response to different activity within the oral cavity such as

swallowing, shouting, eating, drinking, coughing, smoking, and speaking. All of these actions involve active participation of tongue movements, which create a pressure signal that travels through the Eustachian tube to the middle ear. This pressure signal causes a change in the air pressure or airflow within the ear canal, which has been empirically confirmed and characterised as tongue movement ear pressure (TMEP) (Vaidyanathan et al. 2007). This pressure change can be identified by a sensor inserted into the ear canal.

When a person moves the tongue to a certain direction, the generated air pressure changes may be characterised as sound wave or vibration in the ear. As sound or vibration signal spreads, the intensity of the signal falls, which is inversely proportional to the distance from the source. Having taken care of this, placement of the sensor is an important issue to clearly record the desired initiating activity from the pressure signal with a detectable intensity level. With a certain experiment the optimal placement of the microphone to capture the pressure signal can be observed. The sensor consists of an internal microphone and a shielding housing (Vaidyanathan et al. 2007). The placement of the housing shield is similar to the placement of hearing aid housing with proper insulation to protect the internal sound from outside interferences. The insertion of the sensor depends on the individuals' size of ear canal. With a customised sensor for individuals, the internal microphone is capable to record various forms of initiating physical movement of tongue such as clicking against portions of the mouth, touching to certain parts of the mouth etc. The frequency content of the typical TMEP signal ranges from 0 to 200 Hz. It is a non-invasive and unobtrusive recording approach for providing hands-free communication and control for people with disabilities. More details about TMEP signal and its processing are presented in chapter 4.

2.3.4 Neural signals

A normal human brain can generate various responses that include electrical, magnetic, and metabolic responses. These responses can be detected and identified by appropriate sensors for further analysis. There are varieties of ways and sensors exist for monitoring brain activity, these include, non-invasive methods such as electroencephalography (EEG), magneto-encephalography (MEG), positron emission tomography (PET), functional magnetic resonance imaging (fMRI), and optical imaging (i.e., near infrared spectroscopy (NIRS)) as well as invasive electrophysiological methods such as

electrocorticography (ECoG), local field potentials (LFPs), and single unit recording (Lebedev & Nicolelis 2006; Ince et al. 2010; Wolpaw et al. 2002; Mason et al. 2007). However, imaging techniques, MEG, PET and fMRI are still expensive and not handy, which impedes widespread use. Electrophysiological signals, EEG, ECoG, LFPs and single-neuron recordings are the only methods that use relatively simple and inexpensive equipment and have high temporal resolution. Thus, these alternative electrophysiological signals are thought to be potential methods that can offer the possibility for novel assistive communication and control (Collinger et al. 2011; Jerbi et al. 2011). On the other hand neural signals are extensively investigated for identifying different neurological diseases and their therapeutic intervention. In the next subsection, electrophysiological signal measurement techniques of the brain are briefly elaborated.

2.3.4.1 Electroencephalogram (EEG)

Since its discovery by the German psychiatrist Hans Berger (Berger, 1929), the EEG has been extensively analysed to understand the critical function of the human brain and has served as a diagnostic tool in clinical practice. The EEG is an electrical activity induced by the action potentials (firing) of neurons in the brain (Coyle 2006). It is one of the non-invasive and most widely used techniques for recording electrical activity of the brain. The neurons are interconnected in an intricate network and communicate through transmitting and receiving about thousands of spikes (electrical pulses) in each millisecond (Coyle 2006). The voltages carried within the neurons are relatively low, thus the resultant activity acquired through EEG is vague, ranging from 10-20 μV. The brain is the most complex organ in the human body that deals with vast amounts of neuronal activities which involves thinking, action, memory, feeling and perception. The neural activities captured by the EEG signals are extremely complex, providing minimal information about the underlying activity of specific neurons. Signal acquisition device for EEG is simple and cheap, and it takes a small amount of time for preparation to measure the activity. The signals are recorded with small silver/silver chloride electrodes with a radius of about 5 mm, placed on the scalp (figure 2.2) and positioned with 10-20 electrode placement system. To improve the conductivity between scalp and electrodes, conductive gel or saltwater is used. There are a number of techniques used to measure the EEG signals such as measuring the potential difference between pairs of adjacent electrodes (bipolar recording) or measuring the potential differences between electrodes placed at different positions on the scalp with respect to

reference electrode usually attached close to the ear (monopolar recording). The sampling frequency of EEG recording varies although, a standard of about 128 Hz is usually used and the signals are band-pass filtered between 0.5 Hz and 50 Hz. Normally this frequency range contains most of the informative components for utilisation of EEG in BMIs.

2.3.4.2 Electrocorticography (ECoG)

Among the possible ways of recording cortical signals available, electrocorticography (ECoG) offers one of the most clinically feasible options, having better signal quality than EEG and lower technical difficulty compared with other invasive signal recording techniques. The electrical potentials of neural activity recorded in ECoG through surgically implementing array of electrodes, typically an 8×8 grid, on the cortical surface (figure 2.2) (Lal et al. 2005). ECoG has higher spatial resolution than EEG (i.e., tenths of millimetres versus millimetres), broader observable bandwidth (i.e., 0–300 Hz versus 0–50 Hz), and higher characteristic amplitude (i.e., 50–100 µV versus 10–20 µV). Another property is that ECoG signals are barely contaminated with muscle or eye artifacts. Due to the above mentioned advantages, ECoG signals have generated a considerable interest for BMI design. However, due to the necessity for surgery, this paradigm has not been extensively explored. Limited experiments have been performed, mainly with epilepsy patients having ECoG arrays implanted over a period of one or two weeks for localisation of epileptic foci or for pre-surgical monitoring purposes (Leuthardt et al. 2009).

2.3.4.3 Single unit recording

The method of extracting direct electrical information from the neurons of the brain is single unit or neuron recording. This recording performed using microelectrode arrays, has a size of about 5×5 mm and contains around 100 electrodes placed directly into the cortical layers to acquired neural activity from individual neurons (Leuthardt et al. 2009). As like ECoG, brain-surgery is necessary to record neural activity. The difference to ECoG is that electrodes are inserted in the cortex to a depth of several millimetres, i.e. the cortical tissue is penetrated by needle-like electrodes (figure 2.2). Due to the invasive procedure that is needed to record neural signals, single unit recording have been mainly tested in animals (for example monkeys or rat) and very limited experiment performed in human. In single neuron recording, signals are usually

band pass–filtered from 300 to 10,000 Hz and then passed through a spike discriminator to measure the occurrences of spike time (Leuthardt et al. 2009).

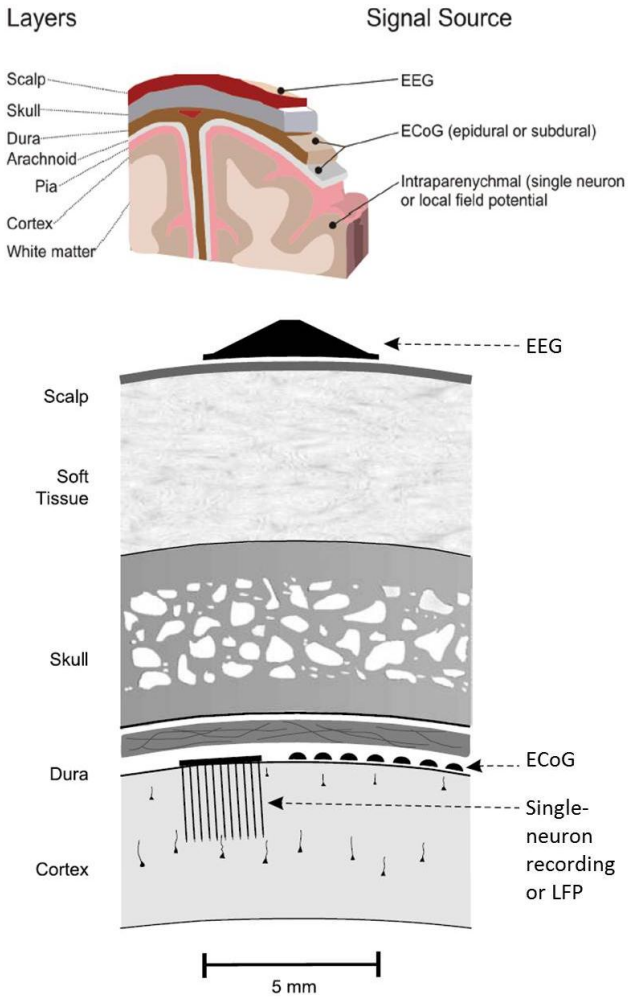


Figure 2.2: Drawing depicting the relationship of the various (EEG, ECoG, LFP, single neuron) signal platforms in terms of anatomy and the population sampled. The bottom image is showing more elaboration of sensor placement for recording of EEG, ECoG, LFP or signal neuron from the brain. From (Leuthardt et al. 2009; Schalk & Mellinger 2010).

Compared with other neural signal recording technologies, the advantages of single unit recording are that signals are acquired at high spatial resolution and the amplitude of the signals are relatively large (for example, 300 μV). There are several disadvantages of this recording that includes risk of infection during surgery, the reaction of brain tissue in the vicinity of the implanted electrodes. Also due to the death of neurons in the region of the microelectrodes, signal quality decays over time and viable data can only be available for a period of several months. Due to the above reasons, there is little hope

of building a long-term BMI system based on single-unit activity (Leuthardt et al. 2009).

2.3.4.4 Local Field Potentials (LFPs)

An alternative technique of acquiring neural activities, the local field potential (LFP) representing the sum of synaptic dendritic potentials within a volume of cortex, is now showing considerable promise for BMI applications (Ince et al. 2010; Mehring et al. 2003). LFPs can be recorded through implanted intracortical electrode arrays or titanium screws (figure 2.2). The bandwidth of LFPs represents neural information content mostly in 0-90 Hz but there may be useful information in high frequency range till 400 Hz. It is easy to record and robust over time, and provides highly precise information on the local brain region. The EEG to single unit recordings, sum activity over different scales varies: millimetres for the EEG, tenths of millimetres for the ECoG and microns for single neuron recording. The LFP lies between the scale of EEG (and/or ECoG) and single unit recordings of sampled activity. This signal comprises the activity of hundreds or thousands of neurons around an electrode tip inserted into the cortex, deep brain or placed over the cortical surface. Thus, like single unit recording, it is invasive; however, it degrades less over time because the listening sphere for LFPs is large, and as a result it is less affected by local scarring (Andersen et al. 2004). It was generally believed that like EEGs, the LFP signal lacks specificity because it is a sum of the activity of many neurons (Andersen et al. 2004). However, recent research has indicated that, using signal-processing methods, a good deal of information can be extracted from LFPs, and thus these signals can be used to augment the BMI with a long listing microelectrode implantations. Also LFP signals are thought to have greater stability in time compared to single neuron recordings and are capable of providing similar amount information as spikes for some parameters, and are even better for others (Andersen et al. 2004; Ince et al. 2010).

Summary

Neural recording techniques have been more or less investigated by the researcher to enable BMI applications as well as some therapeutic treatment of diseases. The properties of the most available techniques including some imaging techniques are summarised in Table 2.1 (Hoffmann 2007). Each technique has its own pros and cons and there is no single technique which entirely represents the neural signature of brain activity (Hoffmann 2007). Therefore, researchers in neuroscience and BMI are

continuing to explore the possibilities to improve the existing techniques or discovering alternative ways to acquire a neural signature for developing real-world BMI applications. However, selection of the technique depends on the experimental conditions, research goal, patient or subject conditions and willingness, options for invasive recordings etc. It is noted that exploration of some existing clinical or therapeutic techniques in respect to BMI application could generate a new hope for decoding neural ensemble for restoring muscular activities through externally controlling devices as well as improving the therapeutic treatment of diseases (Engel et al. 2005; Kringelbach et al. 2007).

Table 2.1: Brain activity measurement techniques and their properties. The relative temporal and spatial resolution of the different techniques indicated with symbols ranging from - - (very low) to ++ (very high). The relative exploration of different techniques indicated with numbers from 1(low) to 3 (high).

Recording Technique	*Type of Measurement*	*Invasive?*	*Temporal Resolution*	*Spatial Resolution*	*Portable?*	*Expensive?*	*Exploration*
EEG	Electrical potentials	No	++	-	Yes	No	3
ECoG	Electrical potentials	Yes	++	+	Yes	No	2
Single unit	Electrical potentials	Yes	++	++	Yes	No	2
LFP	Electrical potentials	Yes	++	++	Yes	No	1
fMRI	Hemodynamic Response (BOLD)	No	-	++	No	Yes	3
NIRS	Hemodynamic Response (optical imaging)	No	-	- -	Yes	No	1
MEG	Magnetic Field Strength	No	++	+	No	Yes	3

2.3.5 Deep brain LFPs from basal ganglia

Despite the success of different invasive and non-invasive neural signal recording and functional imaging techniques that advance our understanding of the human brain, however our knowledge about the neurotransmission process of the deep brain that

related to motor function is largely inferential. Clinical approaches sometimes allow direct recording of brain neurophysiological process at population level (Engel et al. 2005). Most of such recording of deep neural activity in the brain was investigated for identifying diseases related to movement disorders and in some cases a successful treatment procedure was developed to control and alleviate the diseases symptoms. Deep brain stimulation (DBS) is a highly successful therapeutic technique that remarkably alleviates the movement disorders such as Parkinson's disease (PD), tremor and dystonia through implementation of DBS pacemakers in specific brain regions. Fortunately these kinds of therapeutic approach made it possible to carry out invasive recordings of deep neural activity in the human brain and get more insight to advance our understanding. It also created new avenues for improving the therapeutic treatment procedure, as well as incorporating its potential into the development of advanced or alternative BMI (Kringelbach et al. 2007).

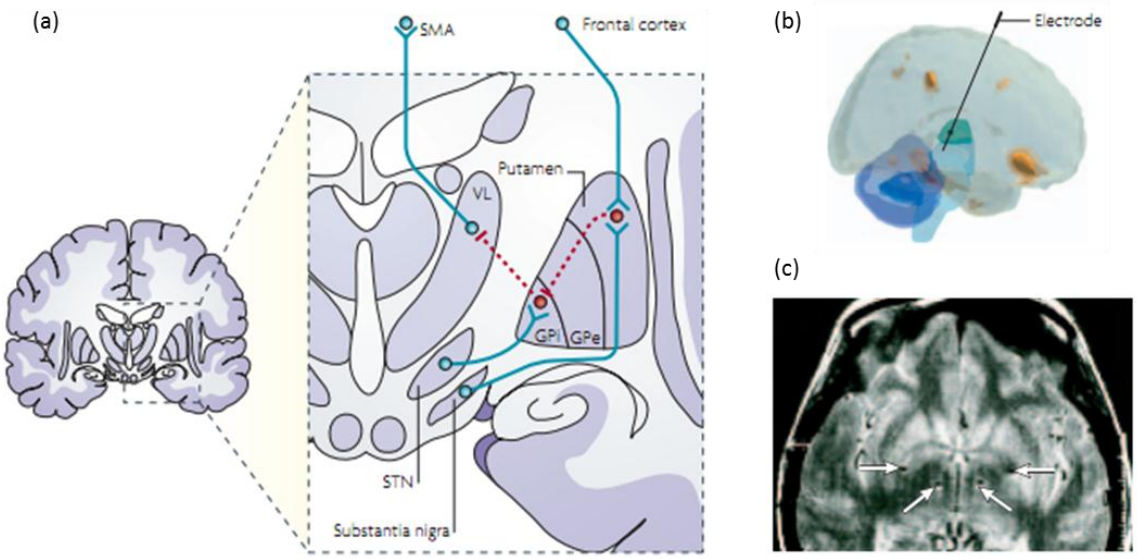


Figure 2.3: (a) Structure of the basal ganglia highlighting the targeted area for DBS, the subthalamic nucleus (STN); the internal segment of the globus pallidus (GPi) as well as other areas and connections. (b) A three-dimensional rendering of DBS stimulation electrode placement using Magnetoencephalography (MEG). (c) A post-operative magnetic resonance image showing the placement of the DBS stimulation electrodes in the bilateral internal globus pallidus (GPi) (upper pair of arrows) and subthalamic nucleus (STN) (lower pair of arrows). From (Kringelbach et al. 2007).

DBS for the treatment of movement disorders has mainly targeted in the human subcortical structure, basal ganglia, through stereotactic operations. Most commonly

DBS electrodes are implanted into the subthalamic nucleus (STN) for PD patients, into the internal segment of the globus pallidus (GPi) for patients with dystonia, and into the ventral intermediate nucleus of the thalamus (Vim) in cases of severe essential tremor (figure 2.3) (Brown & Williams 2005; Kringelbach et al. 2007). These DBS electrodes implantation has allowed us to record neurophysiological activity in the form of local field potentials. Recording of LFPs offers access to the electrical activity of subcortical neuronal populations in human at millisecond resolution. More details about deep brain LFP signal and its processing methods are presented in chapter 6.

2.3.6 Brain activities for brain machine interfaces

The neural signature of brain activity is divided into frequency bands known as rhythms, such as the delta (0.1-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30-90 Hz). These frequency bands are related to different brain states, regions, functions or pathologies. Table 2.2 shows the number of association of mental states with the EEG rhythms (Coyle 2006). All of the brain oscillations are somehow associated with sensory and motor operations and also some of them are associated with abnormal brain function related to neurological diseases. Delta (δ) waves are characteristic of deep sleep and have not been explored for BMI applications. However, recent research shows that it carries considerable information for decoding neural activity (Ince et al. 2010). Theta (θ) waves are enhanced during sleep in adults and often related to various brain disorders. Accordingly, it may be useful for predicting abnormal brain activities such as epileptic seizure. Alpha (α) waves have moderate amplitude and appear spontaneously during wakefulness under relaxed conditions. The oscillatory activity over the sensorimotor cortex with frequency of about 10 Hz and ranges similar to alpha characterises as mu (μ) rhythm. The mu rhythm is also related to the functions of the motor cortex. Beta (β) waves have less amplitude and are strongly related to motor control. It is also associated with mu rhythm and both mu and beta rhythms have been extensively explored for EEG based BMI application. Gamma (γ) waves are associated with movement related activity of the brain and intensely observed in invasive neural recording. It provides discriminative information and is explored in invasive BMI system. Traditionally beta and gamma bands can be subdivided into low and high bands.

Table 2.2: EEG neural frequency rhythms, ranges and their association with number of mental states (Coyle 2006)

Rhythms	*Frequency Range (Hz)*	*Association with mental states*
Delta (δ)	0.1-4	Deep sleep, comatose state and pathologies
Theta (θ)	4-8	Sleeping, Abnormal in awake adults (epilepsy)
Alpha (α)	8-12	Awake but relaxed
Mu (μ)	8-12	Sensorimotor cortex activity
Beta (β)	12-30	Organisation of brain processes, arousal, anxiety
Gamma(γ)	30-90	High mental activity, anxiety, tension, burst of physical activity

Based on these neural rhythms, several form of neurological modalities are employed in brain machine interfaces to generate user control signals. According to the neuromechanisms and recording technology, the brain electrophysiological modalities are used in BMI categorised into five major groups by Wolpaw et al. (Wolpaw et al. 2002), which is sensorimotor activity, slow cortical potentials (SCPs), P300, visual evoked potentials (VEPs) and activity of neural cells (ANC). Later Bashashati et al. (2007) added another two categories 'response to mental tasks' and 'multiple neuromechanisms' (Bashashati et al. 2007).

Sensorimotor activity as a neural source for BMI can be further subcategorised into rhythmic activity (e.g. the mu and beta rhythms) and movement-related potentials (MRPs). Mu and beta rhythms are present in the sensorimotor cortex when a person is not engaged in processing sensorimotor inputs or producing motor outputs (Gerven et al. 2009). A voluntary movement results in a circumscribed desynchronisation (power decrease) in the mu and lower beta bands. This desynchronisation is called event-related desynchronisation (ERD). After a voluntary movement, the power in the brain rhythms at different frequencies increases. This phenomenon is called event-related synchronisation (ERS). Gamma rhythm is high frequency, and the occurrence of a movement (onset) can increase the amplitude of gamma rhythm. Gamma rhythms are usually more prominent in the primary sensory area. Movement-related potentials are low-frequency that start about 1-1.5s before the movement. They have bilateral distribution and present maximum amplitude at the vertex (Bashashati et al. 2007).

Slow cortical potentials are slow voltage changes generated in the cortex. They reflect potential changes of the EEG recording from 300 ms up to several seconds. SCPs are associated with functions involving movement and cortical activation (Wolpaw et al. 2002).

Infrequent or particularly significant auditory, visual, or somatosensory stimuli, when interspersed with frequent or routine stimuli, typically evoke in the EEG over the parietal cortex with a positive peak at about 300 ms after the stimuli. This peak is called P300 (Bashashati et al. 2007).

Visual evoked potentials are small changes in the on-going brain signal. They are generated in response to visual stimuli such as flashing lights and their properties depending on the types of the visual stimulus. These potentials are more prominent in the occipital area. If a visual stimulus is presented repetitively at a rate of 5–6 Hz or greater, a continuous oscillatory electrical response is elicited in the visual pathways. Such a response is termed steady-state visual evoked potentials (ssVEP). The distinction between VEP and ssVEP depends on the repetition rate of the stimulation (Gerven et al. 2009; Bashashati et al. 2007).

The firing rates of neurons in the motor cortex are increased when movements are executed in the preferred direction of a neuron. When the movements are released from the preferred direction, the firing rate is decreased, it is a characteristic of the activity of neural cell (Bashashati et al. 2007).

Assuming that non-movement mental activity of different mental tasks (e.g., solving a multiplication problem, imagining a 3D object, and mental counting) lead to distinct, task-specific distributions of EEG frequency patterns over the scalp, which is considered as response to mental tasks. On the other hand, combination of two or more of the abovementioned neuromechanisms used in BMI design are categorised as multiple neuromechanisms (Bashashati et al. 2007).

The deep brain LFP activity present in the basal ganglia may be broadly subdivided into three frequency bands, <8, 8–30, and >60 Hz, however, these frequency bands are likely to change due to the behavioural and disease correlation of different activities. The best characterised basal ganglia LFP oscillations are at 8–30 Hz frequency band, and well documented in the human striatum, STN and GPi (Brown & Williams 2005). Also these

frequency band oscillations observed to be temporally coupled between the basal ganglia and motor cortex. It is further subdivided into 8-13 Hz (alpha) and 14-30 Hz (beta) bands to justify disease associations. Recent investigations suggest that beta oscillation pattern of basal ganglia LFP activity shows functional connectivity with similar cortical oscillations (Brown & Williams 2005). Also it was reported that movement-related frequency dependent desynchronisation and synchronisation in the STN and/or GPi during externally cued and self-paced movements (clicking or continuous voluntary movements), suggesting that oscillation may be involved in the preparation of the motor response. Particularly beta activity of LFPs in the STN is a good predictor for task performance. It decreases before movement during cued reaction-time task, and the onset latency of this decrease varies with the patients reaction time (Kühn et al. 2004; Engel et al. 2005). This observation suggests that there is an inverse relationship between beta band synchronisation and motor processing.

2.4 Brief review of biosignals based assistive HMI system

Assistive HMI systems create the possibility for people with disabilities to return their life as fully active, productive and be employed in the society. With the rapid growth of electromechanical devices, several assistive HMI systems were developed towards this goal, however, many of them still fail to meet the user's specific requirements. The assistive HMI systems that take mechanical input, in which the user needs to move part of a device physically to generate a control signal. Examples of such systems include hand-operated joysticks, and head or chin-operated (movement) lever for quadriplegic patients (Vaidyanathan et al. 2007). A video camera and tilt sensors based computer interfaces that can track a facial feature and head movement to operate a computer have been implemented (Betke et al. 2002; Yu-Luen et al. 1999). This system requires constant bodily movements, which is uncomfortable for the user and the frequent contact with the device causes skin irritation. It has another limitation that the user's head must be positioned within the range of the device sensors. The 'sip-n-puff' switch, for example, which control a device by sending binary signals, it is simple, easy-to-learn, and relatively low-cost (Wei & Hu 2011; Huo et al. 2008). However it's inflexible and has relatively low communication bandwidth with only 2~4 direct commands. A group of assistive HMI systems were developed based on head-mounted eyeglass and a tongue touch panel to emulate a computer mouse or cursor movements. These devices are controlled by tracking an infrared beam emitted or reflected from a

transmitter or reflector, which is attached to the user's glass, cap, or headband (Takami et al. 1996; Yu-Luen et al. 1999). Some of these devices were constructed using electrooculographic (EOG) potential to track the location of user's eye-gaze on the computer screen, or jaw motion through measuring electromyogragh (EMG) to operate a computer (Barea et al. 2002; Hutchinson et al. 1989; Law et al. 2002). An inherent drawback of EOG based method is that it affects the user's eyesight by requiring extra eye movements and also interferes with the user's normal visual activities, for instance watching, reading and writing. Recently another eye-based activity recognition (EAR) approach using EOG was also reported and showed the possibility to use EOG for assistive HMI (Bulling et al. 2011).

The surface EMGs are electrophysiological signals generated by muscle, which are recorded non-invasively by surface electrode. Some assistive HMI systems were developed based on EMG signals and showed high degree of accuracy for controlling robotic devices like manipulators (robotic arms and hands) and mobile robots (robotic wheelchairs) (Ferreira et al. 2008; Oskoei & Hu 2007a). For example, an EMG based robotic wheelchair is useful for people with motor disabilities in both lower and upper extremities, due to paralysis or amputation. The muscle signals are acquired from the elevator scapulae muscle, and can be generated by voluntary elevation movements of both left and right shoulders. However, despite many advances, EMG control system is now limited in laboratory due to the inadequate controllability, specifically, the lack of intuitive actuation and dexterous control.

On the other hand, some assistive HMI systems were developed based on speech signal and gained popularity for accessing and controlling devices. The vocal joystick is an example of this kind (Bilmes et al. 2006). To perform activity in noisy environment, voice-activated assistive HMI is unreliable. There is a group of assistive HMI developed to control devices, which is operated by tongue. During early 1990s researchers introduced tongue operated keyboard, named Tongue Touch Keypad (TTK) (Huo et al. 2008; Lau & O'Leary 1993). TTK has not been widely accepted due to its size and obtrusiveness. An isometric tongue operated device, Tongue Point was implemented based on the IBM Track Point, which is a small pressure-sensitive joystick fitted to each individual's upper teeth and hard pallet of the mouth (Zhai & Salem 1997). Even though this device offered proportional control, it was very slow to perform the action, and always restricted to a joystick operation, it cannot perform any selection or clicking

operation. As the size of the joystick is 1 cm, it feels uncomfortable in the mouth and interferes with talking and eating. Some other tongue- or mouth-operated systems have been developed, such as Jouse2[1] and Integra Mouse[2], which also operate using a joystick (Huo et al. 2008). In these devices, the user does not need to hold the joystick all times in the mouth, but to grab the stick, the user requires a certain level of head movement, which may not possible by paralysed such as SCI patients.

Another group of tongue operated assistive HMI systems have been devised based on magnetic sensor and wireless technology, such as Tongue Computer Interface (TCI) (Struijk 2006) and Tongue Drive System (TDS) (Huo et al. 2008). Both systems need to place a piece of magnet and two or more receiver sensors inside the mouth. A permanent magnet needs to be placed on top of the tongue in TDS and movement of the tongue is detected through the receiver sensor which is fitted into the teeth or palatal plate. In TDS, it is possible to place a receiver sensor outside the mouth, in this case the user needs to wear a cap such as helmet to carry it. It was mentioned that the information transfer rate in TDS is faster than TCI. Even though the size of magnet and magnetic sensors are small, it is still an intrusive system and has the same comfortably problem as TTK.

Each type of above mentioned assistive HMI system has certain shortcomings. Some of them are limited in their usefulness due to the need for user motion, lack of portability, limited input modalities and lack of flexibility. On the contrary, oral interface mechanisms offer some potential to overcome these challenges, although there are some other issues need to consider due to its intrusive manner. These include irritation in the mouth, interference in verbal communication, hygiene and limited signal generation capacity to some extent. To overcome these deficits, a novel and non-intrusive tongue-movement based assistive HMI was introduced in (Vaidyanathan et al. 2007). Tongue movements within the oral cavity create unique detectable pressure signals within the ear characterised as tongue movement ear pressure (TMEP) signals. This system provided promising output in a control environment with a limited set of operations; However it needs more attention to make it more robust in a more operational setting to work in real life.

[1] http://www.jouse.com/jouse2/home
[2] http://www.smartboxat.com/accessories/integramouse/

Over the last decade researchers are investigating alternative ways to develop assistive HMI which is named as BMI. BMI are trying to translate human neurophysiological activity into device command, and it showed high impact that does not require any bodily movement. Successful development of BMI may overcome some limitations of other assistive HMI and will provide a promising solution for paralysed people to communicate with the outside world. Extensive research has been undertaken by utilising brain waves, such as EEG to develop BMI for assistive communications (Wolpawet al. 2000; Mason et al. 2007; Wolpaw et al. 2002; Nicolas-Alonso & Gomez-Gil 2012). Due to the low magnitude of the EEG signals, it is also prone to external interference and motion artifacts. As EEG signal based BMI is more susceptible to noise and its bandwidth is limited, recently researchers start to make effort to develop invasive BMI through the recording of subdural electrocorticograms (ECoG) or intracotical recordings (Lal et al. 2005; Leuthardt et al. 2009). This methodology provides neural signals with better quality (less noise and higher bit rate) and has potential for future improvement. At the same time, it carries risks associated with an invasive surgical procedure. Several groups have built BMIs based on neural recording from a single cortical area (Lal et al. 2005; Leuthardt et al. 2009). A single-area BMI decodes neural activity specific for that area, for example motor actions in the primary motor cortex or cognitive tasks in posterior parietal cortex. On the other hand, simultaneous recording from multiple areas may provide an advantage on distributed processing of information in the brain for developing complex BMI (Lebedev & Nicolelis 2006). In certain cases, small groups of neuron are sufficient to provide control information for a BMI. This design suffers from instability related to variability of neuronal activity and changes in the sampled populations of neurons. LFPs activity recording can be possible from single or multiple areas to generate neural signal for BMI. Their advantage is that they reflect population effects such as neural oscillations (Lebedev & Nicolelis 2006).

A recent review of motor rehabilitation BMI based on hand movement kinematics from EEG, MEG and intracranial recording (ECoG and LFP) reported that invasive systems can successfully decode movement trajectories from the spiking activity of neurons in primary motor cortex and posterior parietal cortex (Jerbi et al. 2011). However they also noted that non-invasive (MEG or EEG) brain signals may contain sufficient information for decoding of movement direction and hand kinematics. A single trial EEG based

decoding method for controlling two-dimensional cursor was evaluated both offline and online from event-related desynchronisation (ERD) and post-movement event-related synchronisation (ERS) and reported notable success (Huang et al. 2009).

Moreover, researchers are also trying to develop neuro-motor prosthesis. A neuro-motor prosthesis is a device to provide movement commands from brain signals so that neurologically impaired patients are able to perform daily tasks to interact with their environment. Studies in monkeys (Musallam et al. 2004) and humans (Ojakangas et al. 2006; Patil & Turner 2008) have shown that the primary motor cortex (M1) could potentially provide movement-related signals to control assistive devices for paralysed patients. The majority of human studies for BMI incorporated the contralateral neural representation of movements, however there is increasing interest in designing BMIs with ipsilateral control that can support patients with unilateral hemispheric injury. Ganguly et al. (2009) investigated the possibility to design ipsilateral control based BMI that can help a large group of patients who have motor cortex damage and contralateral weakness (Ganguly et al. 2009). They recorded ipsilateral kinematics from cortical field potentials (i.e., local field potentials (LFPs) in monkeys and subdural electrocorticogram (ECoG) in human subjects) and reported high decoding accuracy for ipsilateral limb kinematics.

In general, not much research has been done on invasive based BMI, and little on LFPs based system. Recent reviews reported that among the invasive recording methods, research on LFPs based decoding is growing, and also showed that LFPs can encode movement parameters at a level comparable to unit recordings and ECoG (Marzullo et al. 2005; Jerbi et al. 2011; Ince et al. 2010). Using conductive skull screws, cortical LFPs were recorded from a locked-in patient. Both EMG and LFP activity increases prior to and during switch activation and showed that movement intentions decoded from cortical LFPs were capable of substituting EMG to control a device (Kennedy et al. 2004). Moreover, BMI based on decoding LFPs may provide a better hope for performing different motor activities for paralysed people.

Based on the biosignals and its decoded states, researchers are trying to reactivate disabled people's lives through automating a range of daily living activities. Interestingly one concept system called intelligent home or smart home has been introduced (Ju-Jang et al. 2007). The goal of smart home is to assist the elderly or

handicapped people in solving their daily living problems. It integrates systems for movement assistance, devices for continuous monitoring of health status, and interfaces for controlling home-installed devices in a human-friendly manner. It is still in its early stage of development in the laboratory.

2.5 System structure of assistive HMI

The basic structure of any assistive HMI including BMI system is similar. The system consists of input (i.e. biosignals from human physiological or neurophysiological activity), output (i.e. control command signal for the device), the machine learning system that translating input into output and the operating protocol (Coyle 2006; Leuthardt et al. 2009). Input is the acquisition of respective physiological or neurophysiological signals such as EMG, TMEP, or LFP while the output generates a mechanical or electronic control signal to operate the featured devices. The translation process between the input and output generally has two stages which is signal feature extraction (also includes pre-processing and feature selection), and classification of the pattern to decode user intended activity. The operating protocol defines overall system operation. Figure 2.4 shows the basic design and operational structure of assistive HMI system.

2.5.1 Signal acquisition

The signal acquisition depends on the type of the biosignal and how it can be acquired to use in an assistive HMI system. Different types of biosignals and their acquisition procedure were discussed in section 2.3. During the signal acquisition through respective sensors or electrodes, signal amplification and filtering may need to be incorporated before being digitised. After digitisation, it is necessary to store the data into a computer for further processing. Based on the research hypothesis, experimental paradigm required to setup for generating cue stimulus and gaining respective responses from the subject to utilise the system. Various experimental paradigms have been implemented depending on the terminology of the assistive HMI or BMI system, such as voluntary tongue movement (left, right, up and down) for TMEP based system. The experimental signal acquisition procedure for TMEP and LFP signal are described in detail in the later chapters.

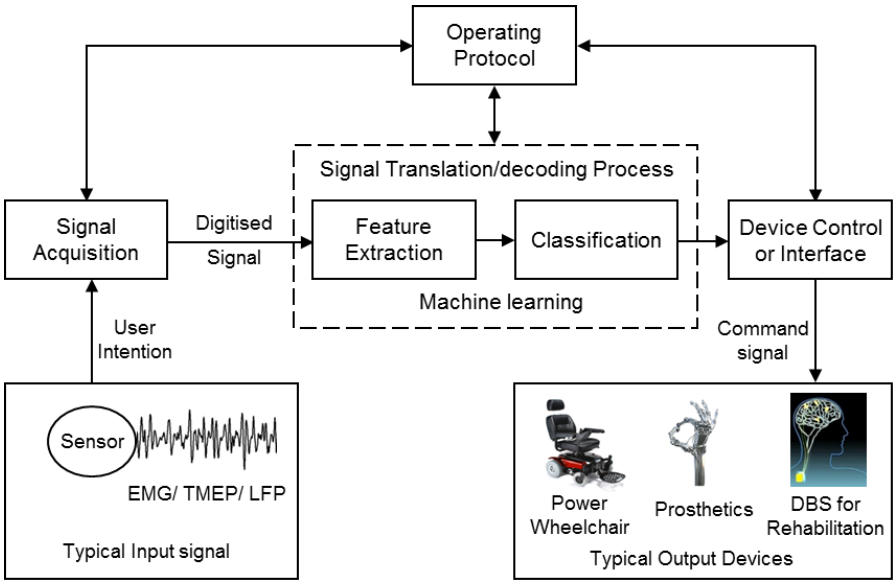


Figure 2.4: Basic design and operational structure of assistive HMI system. Signals for example from the body (EMG), oral cavity (TMEP) or brain (LFP) are acquired by a sensor and processed to extract specific signal features that reflect the user's intent. These features are then translated (classification) into commands that operate a device, for example, a power wheelchair, a neuro-prosthesis, or DBS based rehabilitation. The system works under the operating protocol, which guides its operation.

2.5.2 Signal feature extraction

Feature extraction is the process of enhancing potential information to optimally represent the acquired signal. Each activity signal carries a variety of features in which some features have the potential to discriminate between the activities and some features are similar in all activities. So, a method needs to be selected for extracting signal features which optimally encode the acquired signal. Signal feature extraction process also involves signal pre-processing and feature selection. The details of biosignal feature extraction approaches are presented in chapter 3.

2.5.3 Signal classification

The first part of translation process of biosignal simply extracts specific signal features. The next stage is classification, which translates these signal features into device commands that carry out the user's desired activity. A general classification algorithm consists of two steps. One is training and the other is testing. Training step builds the classifier, i.e., the classifier learns its mapping procedure using the training set. On the other hand, the test step evaluates the performance of any classification algorithm and

computes classification accuracy, i.e. number of correctly classified user's intentions. The percentage of error of the translation system depends on the number of wrong recognition of user's message or command. The error of the system can be minimised by further adaptation or tuning the parameters after the formation of the classifier. The computation cost of the algorithm is also an important issue to design a real time system. The details of biosignal classification methods are presented in chapter 3.

2.5.4 The output device

Until today most of the current assistive HMI researches are tested under the laboratory environment. There is a range of output devices in assistive technology that already have been developed as discussed in section 2.4. The output devices are controlled and operated through the decoded command from biosignals. The operating instruction depends on respective devices, such as for a power wheel chair; it needs to control different movements like left, right, front and back directions. Some of the devices specially in BMI, also use biofeedback in their operation to maintain and improve the accuracy and the speed of communication (Coyle 2006; Leuthardt et al. 2009).

2.5.5 Assistive HMI operating protocol

Any HMI system has a standard protocol that guides its operation which defines how the system is turned on and off, whether communication is continuous or discontinuous, whether message or control transmission is initiated by the stimulus or by the user, what the sequence and speed of interactions between user and system is, whether the system needs user feedback or not, and what types of feedback the system needs, etc. (Coyle 2006; Leuthardt et al. 2009; Wolpaw et al. 2002). The protocols currently used in assistive HMI including BMI research are not completely appropriate for its applications that serve to assist people with disabilities.

The majority of laboratory assistive HMI systems do not give the user on or off control, the investigator turns the system on and off to measure different performance such as communication speed and accuracy. They also define the messages or commands to send to the laboratory system with limited operations. However in real life the user wants to operate the system more flexibly which makes the protocol more complicated in the evolution from research to application.

2.6 Measuring performance of assistive HMI

Performance measurement is essential to improve and evaluate the system. To measure the performance of a system effectively, there are a number of techniques, which include measurement of classification accuracy (i.e. the percentage of correct classifications), classification time (i.e. time required for translation from input to output), and information transfer rate (i.e. bandwidth of transmission channel or amount of information communicated per unit of time) (Coyle 2006; Huo et al. 2008; Wolpaw et al. 2002).

Classification accuracy (CA) determines the percentage of correct identification of the user intentions at the output in particular test trials. Test trials are unknown to the classifier which recognises the behaviour of the classifier. If the system provides 100% accuracy, this implies that all the input signals are correctly classified with their respective classes and the percentage of error rate is zero in this case. The percentage of error rate is simply 100% minus the percentage of CA. The percentage of correct classification and error rate may be calculated using two dimension matrixes called confusion matrices. The number of rows and columns in a confusion matrix is equal to the number of classes in the classifier, i.e. the number of mental or physiological tasks in the system. Classification performance can also be evaluated as sensitivity and specificity. Sensitivity is defined as the ratio of the number of true positives classified to the number of actual total positive cases. Specificity is defined as the ratio of the number of true negative classified to the number of actual total negative cases.

Classification time (CT) is an important issue for designing a real time system. To realise a system in real life, CT should be reasonable. CT depends on the types of operation involved in the translation system from signal acquisition to command generation. It is the sum of the execution time of signal acquisition, signal feature extraction, signal classification, and output generation. In each stage there are several computational operations involved. The total time required for a classification task can be assessed by measuring estimated time required for each operation through the computer.

Information transfer (IT) rate is defined as amount of information transfer per unit of time through communication channel. The channel capacity measured in bits/sec for general data communication, but in HMI, the data transfer rate is very low and it is

measured in bits/min. IT rate also depends on the reliability of the transmission channel. The objective of current assistive HMI including BMI system is to increase the IT rate with accuracy. Various researches have defined the IT rate differently, according to Wolpaw et al.'s definition bit rate can be calculated as follows (Wolpaw et al. 2002).

$$B = log_2 N + P log_2 P + (1 - P) log_2 \left[\frac{(1 - P)}{(N - 1)}\right] \qquad (2.1)$$

where *N* is the number of individual commands or classes that the system can issue, *P* is the classification accuracy (the probability that the desired command will actually be selected), and *B* is the bit rate or bits/trial. The response time of the assistive HMI is *T* and their corresponding IT rate is B/T.

2.7 Conclusion

An assistive HMI including BMI permits a person to communicate with or control the external world without using the normal output pathways of peripheral nerves and muscles. It uses biosignals from physiological activity or neuronal activity as messages or commands to express their wishes. The assistive HMI operation depends on the interaction of two adaptive controllers, the user, who must maintain close correlation between his or her intention or activity, and the system, which must translate the user activity into device commands. The performance of assistive HMI system depends on proper signal acquisition, design of translation or decoding algorithm and generation of commands. The translation algorithm consists of signal feature extraction and classification method, which is crucial for the overall system. Therefore an efficient design of signal translation algorithm can advance the development of assistive HMI as well as BMI system for real life applications.

Chapter 3 : *Review of Signal processing and Pattern Recognition of Biosignals*

3.1 Introduction

In the previous chapter we have illustrated the state of art of biosignals and assistive human machine interfaces (HMI) that let users control a device based on their physiological or neurophysiologic signals corresponding to the respective paradigms. To allow actual control of an assistive HMI system, the respective patterns of the selected signals have to be translated into values that allow discriminating different classes of signals, i.e. biosignals patterns have to be identified through decoding. The prerequisite of the biosignals decoding process is to acquire labelled training data which means recording of biosignals for the prescribed actions. After the biosignals acquisition, machine learning algorithms are employed to decode the given physiological or neurophysiological signals into action commands. Usually machine leaning algorithms are divided into two modules: signal processing and pattern classification.

The signal processing module transforms the raw biosignals into a representative pattern for classification. The signal processing module consists of pre-processing and feature extraction of the biosignals as well as optimal feature selection that makes classification easy and efficient. The main, but not unique, problem is robustness to noise in biosignals decoding, as human physiological or neurophysiological signals data is noisy and/or non-stationary, with data stemming from many sources. The goal of pre-processing and feature extraction is to remove noise and other unnecessary information from the input biosignals, while at the same time retaining important information through extraction that helps to discriminate different classes of biosignals. On the other hand, the goal of feature selection is to reveal the most significant information as well as reduce the dimensionality of the feature space, which optimally represents the

biosignals for classification. Depending on the characteristics of the biosignals appropriate selection of signal processing methods is essential to develop robust decoding algorithm.

After extraction of the optimal features, an efficient pattern classification method is necessary to translate the action commands from the features. The focus of this chapter is to review different signal pre-processing, feature extraction, feature selection and classification methods used in biosignal analysis or decoding. In addition, we also explore some other signal analysis domain such as causality, and incorporate their potential to introduce a new way to generate feature from neural signals. The goal of this review is to develop background knowledge of biosignal analysis, which will eventually lead to design, develop and evaluate efficient decoding algorithms for the biosignals particularly for TMEP and deep brain LFP signals.

Section 3.2, describes different signal pre-processing and feature extraction methods while causality to analyse the functional coupling for feature extraction are presented in section 3.3. Methods for selecting optimal features are described in section 3.4. Finally classification methods for signal decoding are elaborated in section 3.5.

3.2 Pre-processing and feature extraction

To classify the acquired biosignals, the main task is to enhance the discriminative features from raw signals. This can be achievable through applying an appropriate pre-processing and/or feature extraction and/or feature selection methods (Duda et al. 2001). The pre-processing is also called signal enhancement. Biosignals are naturally contaminated with persistent noise or disturbances that make it difficult to understand or reduce the clarity of the signal information. The goal of pre-processing approaches is to reduce the effects of noise and artifacts contamination in acquired signals. Although not all signal translation systems perform pre-processing of the data, however it is an important process, which significantly improves the signal-to-noise ratio (SNR). Pre-processing normally consists of signal filtering, detection (thresholding) and segmentation.

Feature extraction is the process of transforming significant information in the signal to optimally represent the signal. In other words, the feature extraction procedure finds a feature (or set of features) which are common to all signals within each class that

minimises the intra-class variability, while at the same time it (feature or set of features) is not common to any signal contained in the other class that enhances the inter-class variability. As biosignals are non-stationary, classification of the raw time-series data cannot be easy due to the complexity and the similarity among unprocessed signals. To overcome this difficulty several pre-processing and feature extraction techniques are available in the literature which transform the signal from one domain to another domain (Bashashati et al. 2007; Mason et al. 2007). It includes the time domain, frequency domain, time-frequency domain and statistical domain. Time domain features are related to changes in the amplitude of biosignals that are time-locked and occurs based on the stimulus presentation or user actions. Frequency domain features are related to changes in oscillatory activity induced by presentation of stimuli or by intended user actions. The accumulation of information changes in both time and frequency domain together are characterised as time-frequency features. Again by applying statistical linear or nonlinear transformation of the signal, it is possible to generate features in statistical domain.

To enhance the inherent signal characteristics, several methods in the time, frequency and time-frequency domains have been investigated for biosignal feature extraction, and some of them got more attention due to their ability to extract non-stationary dynamics of the biosignals. These include signal averaging, Fourier transform, wavelet transform, and so on. In the following subsection we briefly describe the several signal pre-processing and feature extraction methods, and also a brief review of different biosignal feature extractions using wavelet transforms.

3.2.1 Filtering

In signal processing, one of the most popular pre-processing approaches is filtering which can be applicable in time or frequency domain (Coyle 2006; Wang 2009). The function of a filter is to eliminate unwanted information in the input signal, such as random noise, or to enhance useful information of the input signal. In the literature, there are various methods that can be used to implement the filtering task. The most common types of filtering techniques are finite impulse response (FIR) and infinite impulse response (IIR) (Widrow & Stearns 1985). The overall benefit of using a filter is its simplicity in implementation to eliminate unrelated contents (in particular unwanted frequency components) of the input signal. However, in the case of biosignals, filtering

approaches fail to remove some noise or artifacts. The reason behind of this problem is that some noise and artifacts have the same frequency range that overlaps with biosignals (Coyle 2006). In such cases, statistical signal processing methods, for instance, independent component analysis may be useful.

3.2.2 Signal detection and segmentation

In practice, any biosignals can be monitored continuously to detect the target signal. The detection and segmentation involves identifying the target signal by locating its start and end points based on certain levels of threshold from the continuously recorded signal. The thresholding task depends on the energy of the target signal, which is significantly higher compared to normal or baseline signal activities. Based on this a moving energy averaging approach can detect and segment the event signal (Vaidyanathan et al. 2007; Wang 2009). The energy based signal detection and segmentation approach has been widely investigated and established as an appropriate method for biosignals such as EEG, speech, TMEP (Vaidyanathan et al. 2007; Gupta et al. 1996). When considering a biosignal to drive assistive HMI, accurate detection and segmentation are vital for feature extraction and classification.

3.2.3 Signal averaging

Signal averaging is one of the most commonly used techniques to estimate signals from realisations of a signal-plus-noise random process, which minimises the noise and enhance the signal quality. This process is also called coherent averaging (Wang 2009). Using signal averaging, the SNR of a noisy signal can be improved with the increase of trials, particularly if the signal holds the conditions of (i) noise stationary, (ii) physiological invariability and (iii) no correlation between signal and noise. In practice, direct averaging of repeated event related signals across the several trials will result in a poor blurred estimate because the signals are not exactly aligned in time even after segmentation. This problem can be improved by applying cross-correlation averaging or pairwise cross-correlation-based averaging, which worked well for clean TMEP signal estimation (Gupta et al. 1996; Vaidyanathan et al. 2007). However when the signal was contaminated with noise the performance deteriorated. Nevertheless, in order to achieve good performance, in many cases biosignal needs to transform from time to frequency or time-frequency domain for feature extraction.

3.2.4 Fourier analysis

Biosignal may consist of artifacts or environmental interferences, time domain analysis is not effective to extract discriminative signal features. In that case, frequency domain analysis is useful. One of the most popularly used techniques is Fourier transform and its variation, fast Fourier transform (FFT) (Shin & Hammond 2008). FFT computes frequencies contained in a discrete-time signal without considering their occurrence in time. Later, resultant frequencies of the signal can be used to calculate the power spectral density (PSD) which represents the distribution of power as a function of frequency. Although, Fourier analysis is extremely useful for a stationary signal as it does not vary over time. However, most biosignals are non-stationary; the frequency content of signals can change rapidly over time. Thus, spectral-only methods have the potential to obscure transient or location-specific features within the signal (Addison et al. 2009). In other words, conventional Fourier analysis techniques are insufficient for analysing the time-varying spectral content of biosignals. This limitation can be partly overcome by introducing a fixed length sliding time window to localise the information in time.

The short time Fourier transforms (STFT), is windowed version of Fourier transform, it segments the signal into narrow time intervals, i.e. narrow enough to be considered as stationary, and then take the Fourier transform of each segment (Shin & Hammond 2008). This provides a degree of temporal resolution by highlighting changes in spectral response with respect to time. Although the STFT can resolve the problems to localise the signal energy in time-frequency domain, the window length in STFT is fixed to analyse the complete signal. It may cause a problem of resolution: narrow windows provide good time resolution but poor frequency resolution; wide windows provide good frequency resolution but poor time resolution. To overcome such problem a number of alternative time–frequency and timescale methods are now available for non-stationary signal analysis. Of these, the wavelet transform has emerged as an effective tool for analysing non-stationary signals. Over recent years this method has been explored across a wide variety of areas in science, engineering, and medicine to elucidate local, transient, or intermittent components of the signal in both time and frequency plane (Addison et al. 2009). It has been shown that many biomedical signal processing problems may benefit from time-frequency methods, particularly wavelet

analysis methods (Kiymik et al. 2005; Mahmoud et al. 2006; Samar et al. 1999; Addison et al. 2009).

3.2.5 Wavelet analysis

The wavelet method differs from conventional Fourier techniques by the way in which it localises the information in the time-frequency domain. It has the ability to elucidate simultaneously local spectral and temporal information from a signal in a more flexible way than the STFT by employing a variable window size. Thus, wavelet transforms produce a time–frequency decomposition of a signal over a range of characteristic frequencies that effectively separates individual signal components. This flexible temporal–spectral aspect of the transform permits a local scale-dependent spectral analysis of specific signal features. As a result, it can simultaneously capture both short-term high-frequency and long-term low-frequency information of the signal. Therefore, the method is particularly very useful for the analysis of transients, aperiodicity, and other non-stationary signal features (Addison et al. 2009). Another advantage of wavelet analysis methods is the variety of available wavelet functions or filters, which allows us to select the most appropriate one for the signal of interest. This is in contrast to Fourier analysis that is restricted to only sinusoid functions.

In general, wavelet analysis is the process of transforming a signal from time domain to time-scale domain, which works by breaking up the signal into shifted and scaled versions of the original (or *mother*) wavelet function (Walden 2001). It captures the information in the time-frequency plane, which is especially suitable for the analysis of biosignals. The main difficulty in dealing with biosignals is the extreme variability of the signals and the necessity to operate on a case by case basis (Unser & Aldroubi 1996). Often, one does not have knowledge on what is the pertinent information and/or at which scale it is located. Another important aspect of biosignals is that the information of interest is often a combination of features that are localised in time, space, and scale. Therefore, considering the advantages of wavelet, the analysis of biosignals using wavelet methods can detect and isolate experimental and clinical functional components or events of the investigated signal (Samar et al. 1999).

Other methods were also applied to extract different functionally distinct frequency bands (Addison et al. 2009). In the analysis of biosignals (physiological or neurophysiological signals), frequencies related to the functionally distinct activities can

change over time. In such situations, wavelet methods are able to provide better estimate because the optimal correlation will be obtained when the scaled wavelet filter is similar to the respective activity of interest and is shifted to line up with it in time. By knowing the amount of scaling and shifting we can determine both frequency and time location. The wavelet transform can provide flexibal and optimal filter design. Besides behaving as a bank of band-pass filters, wavelet transform is able to optimally separate components with different properties, for instance, random and patterned components in a signal. The wavelet transform allows us to remove frequency components at specific times in the signals. This provides a powerful capability to remove the unwanted components. Wavelet transform not only provides non-redundant decomposition but also perfect reconstruction. Traditional bank of bandpass filters do not have all of these properties. It was also evident from the literature, wavelet methods can provide better discriminative features than a bank of band-pass filters (Samar et al. 1999; Addison et al. 2009). However, if the analysis of interest is only to extract certain frequency bands from the signals, bandpass filters may be preferable to wavelets, as they are simpler, can be computationally more efficient, and allow for greater flexibility in selecting the desired frequency band.

There are a number of ways to compute a wavelet transform of a signal, which includes, numerically computing the continuous wavelet transform, or using specially designed wavelet filters that generate a highly efficient discrete wavelet transform, also known as a multiresolution decomposition. The idea of continuous wavelet transform (CWT) is to scale and translate the basic wavelet function by very small steps (in fact, infinitely small) in relation to a continuous signal and to compute the wavelet coefficient at each step (Samar et al. 1999). Thus it generates unnecessarily redundant information for analytic signals. Also the computational process of CWT is not very efficient. Therefore, more efficient and computationally simpler wavelet analysis is desirable. Mallat (1989) introduced a multiresolution wavelet decomposition known as the discrete wavelet transform (DWT) (Samar et al. 1999; Addison et al. 2009). Unlike the CWT, the DWT provides a nonredundant, highly efficient wavelet representation of a signal with small number of coefficients, without losing any information. Consequently, it permits perfect reconstruction of the original signal by an inverse discrete wavelet transform.

A DWT algorithm usually decomposes the original signal $x(k)$, k=0,1,…,N-1 into its low- and high-frequency (approximations and details) components with downsampling by two. In practice, the input signal is treated as an initial wavelet approximation. Through low-pass filtering of $x(k)$ by the scaling filter $g(l)$, l=0,1,…,L-1 and high-pass filtering of $x(k)$ by the wavelet filter $h(l)$, l=0,1,…,L-1, generates the coefficient of approximation, $V_1(k)$ and coefficient of detail, $W_1(k)$ respectively in the first level (Wang et al. 2004). In the next level the coefficient of approximation or approximation component obtained from previous level are decomposed into its corresponding approximation and detail component and down-sampled by two. The DWT decomposition coefficient of approximation $V_j(k)$ and detail $W_j(k)$ at level j can be computed as,

$$V_j(k) = \sum_{l=0}^{L-1} g(l) V_{j-1}(2k+1-l \bmod N_{j-1}), \quad N_j \equiv \frac{N}{2^j} \tag{3.1}$$

$$W_j(k) = \sum_{l=0}^{L-1} h(l) V_{j-1}(2k+1-l \bmod N_{j-1}), \tag{3.2}$$

where the $g(l)$ and $h(l)$ are respectively orthogonal scaling and wavelet filters that can be derived from the set of predefined wavelet functions. The decomposition of approximation component can be repeated until the length of the resulting coefficient vector equals 1. The recursive process of decomposition and reconstruction of DWT for biosignal is presented in figure 3.1. In practice, a maximum suitable level of decomposition is chosen, either based on the characteristics of the signal or in order satisfy the entropy criteria. On the contrary, at any level of decomposition, j, it is possible to obtain a perfect reconstruction of the original signal by using (up-sampling) the jth level approximation component and all the detail components from level j to 1. Alternatively, it can be said that the original signal $x(k)$ is reconstructed from the approximation component $V_j(k)$ and detail components $W_j(k), \ldots, W_2(k), W_1(k)$.

The DWT has emerged as a particularly powerful tool for the signal analysis such as signal encoding and compression. However, DWT has a number of limitations, (i) it requires the length of the input time series to be a power of 2 for the full transform, i.e. dyadic, (ii) it is sensitive to the downsampling position of the time series, i.e. lack of translation invariance, and (iii) the number of wavelet and scaling coefficients, N_j,

decreases by a factor of 2 for each increase of the decomposition level that limits the ability to carry out statistical analysis on the coefficients to some extent (Walden 2001). These limitations can be overcome by avoiding the use of downsampling in the DWT and can be achieved by using the stationary wavelet transform (SWT). The SWT is an offshoot of the DWT, whereby the time steps are not subsampled at each level. This destroys the orthogonality in the transform but does provide translation invariance and also leads to generate many more coefficients (sufficient resolutions). Although orthogonality is destroyed, the SWT is very useful for some signal feature extraction and statistical applications (Addison et al. 2009). The SWT is also known by a variety of names in the literature, including the maximal overlap wavelet transform, the redundant wavelet transform or undecimated wavelet transform.

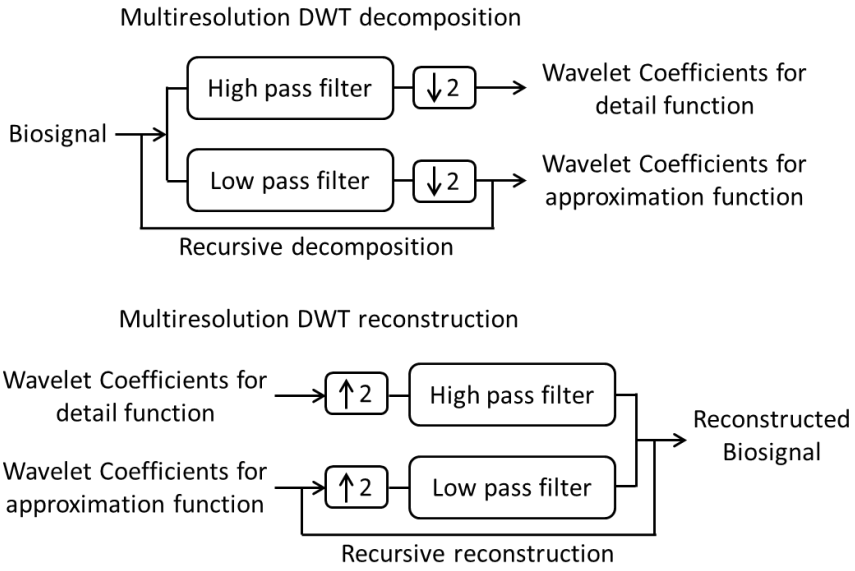


Figure 3.1: Basic recursive decomposition and reconstruction of discrete wavelet transform (DWT) for biosignal. Down-sampling and up-sampling by 2 are represented by the downward and upward arrows with numeral 2, respectively.

On the other hand, to analyse biosignals such as neural signals requires adequate flexibility for partitioning the signal into functionally distinct scales to localise specific information. Wavelet packet transform (WPT) is a generalisation of the DWT that allows the *best* adapted analysis of a signal in a timescale domain. The WPT decomposition of a signal is performed in a manner similar to the DWT, the difference being that, both the approximation and detailed components are further decomposed at each level (Walden 2001; Addison et al. 2009; Samar et al. 1999). At each level, the

WPT algorithm divides the time–frequency plane into rectangles of constant aspect ratio. The selection of WPT coefficients to represent the signal normally depends on an information cost function that aims to retain as much information in a few coefficients as possible. The most commonly used information cost function is the Shannon entropy measure (Addison et al. 2009). Based on the properties of WPT, it is considered an efficient tool to extract time-frequency feature from biosignals. More details of the wavelet transform for biosignal analysis can be found in (Addison et al. 2009; Samar et al. 1999).

3.2.6 Statistical signal decomposition methods

The mostly used statistical signal decomposition methods are principal component analysis (PCA) and independent component analysis (ICA). PCA is a method to extract the principal components, the basis along which the variance of data is maximised. It reduces the dimensionality of multivariate data whilst preserving as much of the relevant information as possible (Bai et al. 2007; Jolliffe 2002). On the other hand, ICA is a member of a class of blind source separation (BSS) methods. In ICA, it extracts statistically independent (source) components from a set of measured (mixture of independent) signals. ICA is also related to PCA, where ICA finds a set of source signals that are mutually independent and PCA finds a set of signals that are mutually decorrelated (Stone 2005). ICA can also be used for noise reduction from biosignals.

3.2.7 Brief review of wavelet based feature extraction in biosignals

The selection of feature extraction methods depends on characteristics of the biosignals. Among various feature enhancement methods, time-frequency based techniques have received more attention to extract discriminative information from biosignals. As described earlier, to extract the discriminative features from different physiological and neurophysiological signals STFT, wavelet transform, stationary wavelet transform and wavelet packet transform were investigated.

Wavelet based methods are effective in both signal noise reduction and feature extraction. A signal de-noising approach based on discrete wavelet transform was investigated for extracting burst and tonic components from surface EMG of dystonia patients (Wang et al. 2004). An adaptive soft thresholding was applied to the wavelet coefficients for separating the superimposed burst and tonic activity, and reported its

effectiveness. Another wavelet based approach for de-noising single-trial visual and auditory event-related potentials was proposed by Quiroga and Garcia (2003) and reported that the approach provided better estimation for event related potentials from EEG (Quiroga & Garcia 2003).

To extract the feature from single trial EEG for classification of left and right self-paced events, discrete wavelet decomposition was investigated (Pires et al. 2007). With DWT based extracted feature, 95% correct classification accuracy has been achieved with a simple classifier. The DWT down sampling approach provides low feature vector and localise good time-scale information. These are advantageous over other methods because they provide several feature vectors with low computation time which is suitable for real-time applications. A new wavelet called sum of neuron action potentials (SNAP) was designed to match neural activity for EEG analysis (Glassman 2005). The coefficients resulting from the DWT with SNAP wavelet were used as features for classification, presenting better or comparable accuracy to other wavelets for voluntary hand movement recognition. A comparative evaluation of spectral feature extraction approaches were performed on EEG-based motor imagery hand movement classification (Herman et al. 2008). According to their results it was observed that wavelet based approach provided comparable accuracy with short-term Fourier transform and power spectral density based techniques. Another two EEG analysis studies were investigated by applying wavelet for feature extraction. One applied CWT to extract feature from the single-trial left and right motor imagery task (Hsu & Sun 2009) and other applied discrete wavelet transform for extracting features from auditory brainstem response (Rui et al. 2006). To analyse the subthalamic local field potentials from the patients with Parkinson's disease, CWT was investigated for feature extraction (Loukas & Brown 2004). This study showed that wavelet based extracted feature are effective for predicting self-paced hand movements from LFPs.

Among the wavelet based methods, wavelet packet transform is more effective to localise specific information from biosignals. To extract the feature from myoelectric signals for continuous classification, time frequency approach, STFT, DWT and WPT were analysed (Englehart et al. 2001; Englehart et al. 1999). These studies showed that WPT based features achieved lowest average classification error compare to STFT and DWT. Similar study was performed by Xie et al. (2009), they also reported that WPT provided better discriminative feature than SWT and STFT in combination with

singular value decomposition for mechanomyogram signal classification (Xie et al. 2009). A comprehensive study on wavelet analysis for EEG signals was presented in (Samar et al. 1999) and reported that WPT is capable of providing greater flexibility and more localised neural information than the DWT. To detect the epileptic seizure from EEG in real-time, a WPT based algorithm was proposed and stated its efficacy in (Zandi et al. 2010). More reviews of the wavelet feature extraction as well as other feature extraction methods for biosignals can be found in (Mason et al. 2007; Bashashatiet al. 2007; Addison et al. 2009; Samar et al. 1999; Quinquis 1998).

3.3 Causality

The multivariate linear regression analysis of time series has been investigated in many areas for signal feature extractions. In the area of biosignals, particularly analysing neural signals, it is an effective means to characterise, with high spatial, temporal, and frequency resolution, functional relations within multichannel recording data. Recent advances of multivariate linear regressive analysis such as the multivariate auto-regressive (MVAR) modelling in neuroscience showed promise for the analysis of large scale network interactions, especially assessing their causal relations (Kaminski & Liang 2005). Parametric methods such as MVAR use the signal to estimate the parameters for a mathematical model to describe the signal. Later signal properties are derived from the parameters of the model. MVAR model have the ability to capture the rapid dynamic changes of a stochastic process within a very short time period, which makes it useful for analysing neural signal with inherent nonstationarity.

The rhythmic occurrence in neural activity can be studied by power spectral analysis while the degree of rhythmic synchronisation between brain regions or structures can be analysed by coherence analysis. The analysis of power spectral density and ordinary coherence for two-channel cases is well established. In multi-channel case the partial and multiple coherences can provide much more specific information about the interactions between channels than ordinary coherences. Partial or multiple coherence is respectively equivalent to partial or multiple correlation analysis in the frequency domain. Partial coherence is capable of measuring the coupling or strength of direct connection between two structures, while multiple coherence provides information about the coupling between the given structure and all the other structures in the investigated system. Ordinary, partial and multiple coherences are very useful method

for characterising various types of synchronisation or coupling; however they are unable to provide information about causal relation (i.e. direction of influence) between the structures or regions such as between two neural activities. It is generally recognised that propagation of neural activity in one structure to other related structure can be understood by identifying directional influence or flow of information.

To analyse the functional coupling between various physiological and neurophysiological processes, particularly in the brain, the MVAR model associated with a powerful set of time- and frequency-domain statistical methods investigated for inferring directional and causal information. The mostly used methods are Granger causality, directed transfer function and directed coherence.

The basic concept of causal interactions between time series was described initially as a probabilistic process. Granger (1969) first formulated this concept and defined that one variable or time series may be caused by the other if it can be better predicted by incorporating knowledge of the second one (Granger 1969). Granger causality was widely used in econometrics, later it was applied to system theory and recently it is widely adapted in neuroscience to investigate the neural mechanism of the brain activity, particularly identifying directional influences or causal synchronisation between different neural activities (Silchenko et al. 2010; Kaminski & Liang 2005). The formation of Granger Causality makes two assumptions about the signal: 1) signals are covariance stationary (i.e., the mean and variance of each time series do not change over time), and 2) time series can be adequately describe by the linear model. Granger causality was first designed to evaluate the causal interaction between two signals, later it was extended for multivariate (more than two signals) cases and referred as conditional Granger causality, which is extremely useful to identify the network connectivity of the brain. Again by applying the Fourier transform, it is possible to analyse Granger causality in the spectral domain. Geweke's work based on the MVAR process first explained Granger causality into the conditional causality and spectral measurement (Wang et al. 2007; Geweke 1982). The original Granger causality has a number of limitations, 1) it can only extract the information about the linear features of signals, however, nonlinear extension is now developed but it is not well evaluated for practical applications; 2) it assumes that analysed signals are covariance stationary, however non-stationary signal can be treated by using a windowing approach assuming that data within sufficiently short windows of non-stationary signal are locally

stationary; and 3) it depends on the appropriate selection of model parameters or variables (Seth 2007).

Saito and Harashima (1981) introduced a method, directed coherence to study relationship between a pair of signals described by a bivariate autoregressive process (Kamiński et al. 2001). It has been applied to the investigation of interdependence between two processes in clinical problems as well as neural mechanisms (Kamiński et al. 2001; Kaminski & Liang 2005). Another method, the directed transfer function (DTF) based on the MVAR was proposed by Kamiński and Blinowska (1991). Depending on signal frequency, DTF can estimate the strength and direction of the influences between various structures of the brain. As the DTF method is based on a MVAR model, it allows to analyse short epochs of signals. The demerit of DTF is that in some cases it may not easily differentiate between a direct and indirect influence or connection. To overcome this shortcoming, a modified DTF (dDTF) method was proposed by combining the DTF and partial coherences (Korzeniewska 2003). Some studies showed that the non-normalised DTF function is equivalent to the Granger causality measure. The details review, and description of different methods for revealing causal influence, can be found in (Gourévitch et al. 2006; Kamiński & Liang 2005; Zou et al. 2009).

3.3.1 Brief review of causal analysis in biosignals

One of the major challenges in neuroscience is the identification of directionality between signals that reflect neural activities. In recent decades it was observed that a remarkable interest has grown in the use of Granger causality to recognise causal interactions in neural processes. To describe causal interactions among different areas in the cat visual cortex, Granger causality was investigated (Bernasconi & König 1999). Brovelli et al. (2004) identified causal influences from primary somatosensory cortex to motor cortex in the beta-frequency band (14-30 Hz) during lever pressing (GO/ NO–GO visual pattern discrimination task) by awake monkeys (Brovelli et al. 2004). In neuroimaging studies, to reveal the directed influences between neuronal populations in functional MRI data, a method called Granger causality mapping (GCM) was proposed in (Roebroeck et al. 2005). Using GCM they investigated the effective connectivity during a complex visuomotor task. Another fMRI study investigated a wavelet variation of Granger causality to identify time-varying causal influences during motor task

(resting vs. finger movement) (Sato et al. 2006). Causal influences between oscillatory subthalamic nucleus (STN) local field potentials (LFPs) and muscular activity (EMG) in Parkinsonian tremor was investigated and showed that predominant directional influence from EMG to LFPs (Wang et al. 2007).

Recently a Granger causality analysis framework was developed for modelling multi-neuron spike train data and showed that it can extract the directed information flow structure among interacting neuronal populations (Krumin & Shoham 2010). They also reported that it can be useful to identify effective connectivity in multi-neuron spike train. To reveal the directional connectivity between the left and right STN LFPs in different frequency bands, partial directed coherence was explored in patients with Parkinson's disease (PD) (Silchenko et al. 2010). From the analysis it was observed that bidirectional coupling exists between the left and right STN in the beta band (10–30 Hz) for an akinetic PD patient and in the tremor band (3–5 Hz) for tremor-dominant PD patients.

3.4 Feature selection

Feature selection is also known as subset selection or dimensionality reduction. It is a process commonly used in machine learning that select an optimal subset of features, which contains the least number of dimensions and have the most contribution to accuracy. During the last decade, the feature selection (or reduction) approaches have become an apparent prerequisite in many applications of biosignals classification for identifying pattern of the event or activity. This is due to the nature of high dimensionality of the biosignals in the feature space as well as limited or small size of data or samples available for designing and evaluating the pattern classifier. Also in the feature space, all the extracted features are not equally important for pattern discrimination, i.e. some features carry significant discriminative information for the particular pattern of the biosignals and some carry little or no discriminative information. Most of the pattern classification methods were originally not formulated to deal with such redundant or irrelevant features. Redundant features significantly affect the pattern classification process and provide poor generalisation. Theoretically to achieve better performance from the designed classifier, all the available features in the feature space quantitatively satisfies that there exist large variance between the classes and small variance within the classes. However, in the perspective of real world

problem such as biosignals, it is almost impossible to satisfy such condition and this leads to poor performance. To this end, it is important to identify a subset of features from the feature space, and provide distinctive information to the pattern classifier for improving performance. Another necessity for feature selection or dimensionality reduction is to minimise the computational complexity of the classification.

The objective of the feature selection techniques is to select a subset of *m* features out of *M* features, $m<M$, that will provide the best classification accuracy. Ideally, feature selection methods search through the subsets of features, and try to find the best one among the competing 2^M candidate subsets based on the given evaluation function. This is a *complete* search procedure to generate the best feature subset. Also it may be very expensive and practically impossible, even for a medium-sized (*M*) feature set. There are two other feature generation procedures that attempt to reduce computational complexity of the search space based on *heuristic* or *random* search methods (Dash & Liu 1997). However, these methods required to compromise the performance and a criterion to prevent an exhaustive search of subsets. In *heuristic* search methods, the generation of subset is the adding or deleting a feature into the existing subset based on the criterion. It is simple and fast in producing one optimal subset. In *random* search methods, the generation of feature subset uses some kind of randomness or probability parameter that tried to convergence based on the given criterion. In this method, the search is not exhaustive; however it does not guarantee to select optimal subset due to lack of global optimisation (Dash & Liu 1997; Zabinsky 2009).

On the other hand, selection of an optimal feature subset from the feature space is always relative to a certain evaluation function. The objective of an evaluation function is to measure the discriminability of a feature subset to distinguish the different classes. In many cases, feature selection techniques use the classifier as an evaluation function with the measure of classification accuracy. There are some other well defined measures that also incorporated as evaluation function into the feature selection techniques, such as distance, information gain.

In the context of pattern classification, feature selection techniques can be considered as three taxonomical categories, (1) *filter* (2) *wrapper* and (3) *embedded* approach depending on the characteristics of the selection techniques and their evaluation function or criterion (Saeys et al. 2007).

The *filter* approaches identify the relevance of features by looking only at the intrinsic properties of the data. In most cases a feature relevance score is calculated based on the distance or information gain measure, and select only the high-scoring features. Subsequently, selected subset of features feed into the classification algorithm. Filter approaches have several advantages, it is computationally simple and fast, easily scalable to very high-dimensional datasets, and also independent to the classification methods. Consequently, once optimal features were selected, it is possible to evaluate different classification methods. However, due to the independence of the classifier interaction, it ignores the feature dependences (i.e. considers each feature separately), which may lead to worse classification performance when compared to other types of feature selection techniques. This is a common disadvantage of filter approach.

The *wrapper* approach overcomes the problem of filter approach through interacting with the classifier during feature selection. i.e., selection algorithm is 'wrapped' around the classification model. Here, all possible subset of features are generated based on search procedure, most cases heuristic, and evaluated through the classifier algorithm with measuring classification accuracy that take into account feature dependences. Finally, the subset with maximum classification accuracy is defined as the optimal feature subset to build the classifier. The common limitation of wrapper approach is that it has a higher risk of overfitting than filter approach and it is computationally very intensive, particularly if computational cost of the classifier algorithm is high. Also the selection of feature is fully dependent on the specific classifier algorithm.

The third category of feature selection techniques termed as *embedded*. The embedded approach tries to overcome the drawback of wrapper approach by addressing the risk of overfitting, however its construction is also like a wrapper through integrating feature dependences with specific learning algorithm. It has the advantage that they include the interaction with the classification model, while at the same time improving the performance.

Based on the feature subset generation (*complete, heuristic* or *random* search methods) and evaluation (*filter*, *wrapper* or *embedded* approach) procedure, a variety of feature subset selection techniques have been developed and investigated (Saeys et al. 2007; Dash & Liu 1997). However, to reduce the complexity of search space for generating feature subset to solve real world problems, heuristic and random search methods are

widely considered. Among them heuristic search has received more attention than random methods due to its simplicity and speed in producing optimal subset. In the heuristic and random search domain all three evaluation approaches were investigated, nonetheless, wrapper approach got more consideration due to its simplicity. The commonly used feature selection techniques in filter approach are *t*-Test, ANOVA, Wilcoxon rank sum test, F-score and correlation-based feature selection (CFS), and in wrapper approach are sequential feature selection (SFS) (which includes sequential forward feature selection (SFFS), sequential backward feature selection (SBFS)), bi-directional search (BDS) and genetic algorithm (GA). Through evaluating different datasets, Wang et al. (2008) have shown that the wrapper approach usually ensures higher classification accuracy than filters (Wang et al. 2008).

Most popular feature selection techniques for biosignals is SFS (sequential forward or backward feature selection), SFFS starts from the empty set, and in each iteration generates new subsets by adding a feature selected by some evaluation function (such as classification accuracy, distance, information gain), while SBFS starts from the complete feature set, and in each iteration generates new subsets by discarding a feature selected by some evaluation function. SFS techniques work under the heuristic search method while genetic algorithm works under the random search method. It is noted that the embedded approach provided higher accuracy in maximum cases. The commonly used embedded feature selection techniques are decision tree and random forest. The more details of different feature selection techniques and their properties can be found in (Liu & Yu 2005; Saeys et al. 2007; Liu et al. 2005; Guyon 2003; Dash & Liu 1997).

3.4.1 Brief review of feature selection techniques in biosignals

In biosignals application, to reduce the dimensionality of the principal component transform and discrete cosine transform (DCT) coefficient features for TMEP and ERP of EEG signal classification, Gupta et al. introduced a feature ranking strategy that initially ranked each feature and then select subset of features using SFS (Gupta et al. 2010). They used four ranking criteria, magnitude, variance, interclass separation and classification accuracy, and showed that performance was significantly improved compared to a previous classification approaches which did not use dimensionality reduction. Another study also used SFS techniques to select optimal TMEP signal feature from DCT coefficients for real-time classification and showed improved

accuracy (Mace et al. 2010). Similarly, a sequential forward-backward search procedure with mutual information investigated for EEG classification (Guerrero-Mosquera et al. 2010). The cross-entropy as filter and genetic algorithm as wrapper approach were investigated for improving the classification of movement-related potentials (MRP) of EEG. It demonstrated that a small number of features selected by the genetic algorithm obtained better performance for classifying limb MRP (Yom-Tov & Inbar 2002). Another EEG classification studies investigated genetic algorithm with mutual information (Fatourechi et al. 2007), and classification accuracy using support vector machine (SVM) (Peterson et al. 2005) as relevance criterion to select subset of features.

Another feature ranking criterion is F-score that also widely used in feature selection techniques (Zhao et al. 2010). Combining SVM classifier with F-score was evaluated and compared its performance with the SVM integrating F-score plus random forest (RF), and RF plus radius margin approach (Chen & Lin 2006). F-score was also used for selecting feature to classify multi-class obstructive sleep apnea syndrome and reported its effectiveness (Güneş et al. 2010). A classification study based on EEG time series of multilevel mental fatigue, two feature selection techniques, RF combined with the heuristic initial feature ranking scheme and with the recursive feature elimination (RFE) scheme were evaluated. It was reported that both techniques performed better in reducing the error rate and the number of features (Shen et al. 2007).

Recently an evaluation was performed to select appropriate feature selection scheme for high dimensional, small sample size biomedical datasets (Golugula et al. 2011). They argued that an ideal feature selection scheme must be robust enough to produce the same results each time even though there are changes to the training data. Five different filter feature selection schemes (T-test, F-test, Kolmogorov-Smirnov Test, Wilks Lambda Test and Wilcoxon Rand Sum Test) were quantitatively compared on five gene and protein expression datasets corresponding to ovarian cancer, lung cancer, bone lesions, celiac disease, and coronary heart disease, and found that Wilcoxon Rand Sum Test outperformed than other schemes in terms of classification accuracy and robustness. Another study compared the linear (PCA, linear discriminant analysis, classical multidimensional scaling) and nonlinear (isometric mapping, locally linear embedding, Laplacian eigenmaps) dimensionality reduction methods for classifying high dimensional gene and protein expression (Lee et al. 2008). They stated that

nonlinear method outperformed than the linear method, however considering the complexity linear methods are more applicable for the biosignal analysis.

3.5 Classification

Machine learning is considered as a branch of artificial intelligence and it is concerned with the design and development of classification techniques which enable the computer to learn, similar to human learning. A major objective of machine learning research is to develop pattern classification technique, i.e. an automatic model that produces rules and patterns from given input data for classification. Over the period of time many techniques and methodologies were developed for machine learning tasks to classify input data (Muller et al. 2004; Lotte et al. 2007; Jain & Duin 2000). The pattern classification of biosignals, which translates input signal features into device commands that carries out the user's intent. The pattern classification procedure might be linear (e.g. classical statistical analyses) or nonlinear (e.g. neural network); might be supervised (e.g. support vector machine) or unsupervised learning (e.g. clustering, blind source separation) (Jain & Duin 2000; Muller et al. 2003). Whatever its nature, each algorithm does the same basic translation, it transform independent variables (i.e. features of the signal) into dependent variables (i.e. control commands for device) (Bishop 2006).

Generally, pattern classification algorithm consists of training and testing cycles. Training set is responsible for building the classifier while testing set is independent and does not involves in the classifier formation. Some pattern classification algorithms also implement with the validation cycle. Depending on the partitioning criteria of the input data, validation cycle is used to adjust the classifier parameters during its formation, i.e. classifier parameter optimisation that improves the robustness in classification. After validation, classifier is evaluated using test set (Duda et al. 2001; Bishop 2006). The size of training, validating and testing set is also considerable issue for designing optimal classifier; generally large training set provides better generalised classification model. However in reality there is a limitation to collect large dataset of biosignals, i.e. human physiological or neurophysiological signals. According to hold-out estimation one third of the data needs to be reserved for the test set. In this case, the classifier can be made reliable by repeating the training and testing process through randomly partitioning the dataset, and average the error rate of all repetition to produce the overall

error rate. This is also called the repeated hold-out method for classification (Duda et al. 2001). Nonetheless, it is not optimal partitioning approach because different tests overlap each other. The overlapping could be prevented by cross-validation procedure (Duda et al. 2001; Bishop 2006). Here, data is split into *k* subsets of equal size set and then each subset in turn is used for testing and all the remaining sets are used as training set, and finally error of each turn are averaged to estimate the overall error rate. This procedure is called *k*-fold cross-validation. The widely used cross-validation method is 10-fold cross-validation. Another variation of cross-validation procedure is leave-one-out approaches for very limited dataset. Here dataset is partitioned into *n* subset, where *n* is size of the dataset. The leave-one-out cross-validation procedure can be named as *n*-fold cross-validation and depending on the data set size it is computationally very expensive.

3.5.1 Methods for biosignal classification

To identify the biosignals patterns from various physiological or neurophysiological activities of human, various pattern classification methods have been investigated (Muller et al. 2003; Bashashatiet al. 2007; Lotte et al. 2007; Jain & Duin 2000). In order to select the most appropriate classification method to decode biosignal patterns from the respective set of extracted features, the properties of the available classification methods needs to be considered (Lotte et al. 2007). The appropriate selection of classification methods is dependent on the properties of both biosignal features and the classification methods itself.

There are several properties that commonly describe different kinds of classification methods, include generative or discriminative, stable or unstable (Lotte et al. 2007). Generative (also known as informative) classification methods compute the likelihood of each class and choose the maximum likelihood to classify a feature vector (e.g. Bayesian). In contrast, discriminative classification methods only learn the way of discriminating the classes or the class membership in order to classify a feature vector directly (e.g. support vector machine). Stable classification methods have a low complexity, and with small variations in the training set do not considerably affect the performances (e.g. linear discriminant analysis). In contrast, unstable classification methods have a high complexity and with small variations of the training set may lead to important changes in performances (e.g. neural network). Another property is

regularisation that controls the complexity of a classification method in order to prevent overtraining (Jain & Duin 2000). A regularised classification method has good generalisation performances and is more robust with respect to outliers. On the other hand curse-of-dimensionality is another issue for performing pattern classification tasks (Gupta et al. 2010; Lotte et al. 2007). It defines as the amount of data required to properly describe the different classes increases exponentially with the dimensionality of the feature vectors. If the number of training data is small compared to the size of the feature space, the classifier will most probably provide poor performance. It is recommended to use, at least, five to ten times as many training samples per class as the dimensionality (Raudys & Jain 1991; Lotte et al. 2007). Unfortunately this cannot be possible in biosignals pattern classification as biosignals generally have high dimensional feature space with small training set. Therefore this issue is a major concern for biosignals classification, however feature selection or dimensionality reduction can elevate this problem (cf. 3.4).

Linear and non-linear, both types of classifier methods applied in different biosignal decoding for assistive HMI as well as BMI system. However, both have different pros and cons depending on analysed signal, feature extraction method and the problem itself. According to *Second International Meeting on Brain–Computer Interfaces* (held in June 2002), it was agreed that simplicity is generally best and, therefore, the use of linear methods is recommended wherever possible (Vaughan et al. 2003). It was also argued that nonlinear methods in some applications can provide better results, particularly with complex and/or other very large datasets (Wolpaw et al. 2000; Muller et al. 2003), however for real-time application linear methods would be more preferable. Some of the linear and non-linear classification methods that have been extensively explored for biosignals decoding include Bayesian classifier, neural networks (NN) and support vector machine (SVM).

3.5.1.1 Bayesian classifier (BC)

The Bayesian classifier also called Gaussian maximum likelihood classifier and it is one of the most simple and common statistical methods used for pattern classification. It is a linear classifier based on Bayes' theorem, which aims at assigning a feature vector to the class that it belongs to with highest probability (Duda et al. 2001). The Bayes rule is used to compute a posteriori probability that a feature vector is belonging to a given

class. Using the maximum a posteriori (MAP) rule and these probabilities, the class of this feature vector can be estimated for classification (Lotte et al. 2007). The detailed illustration of the Bayesian classification is presented in chapter 4 (cf. section 4.3.4.3).

3.5.1.2 Neural network (NN)

Neural network (NN) also known as Multi-layer neural networks (MLN) are the most popular class of artificial neural networks (ANNs) and have been widely applied to pattern recognition, signal processing, time series prediction, non-linear control and identification problems. The advantage of the neural networks for pattern classification is their automatic training capacity, and ability to implement nonlinear decision functions (Looney 1997; Jain & Duin 2000). NN consists of several simple parallel computational units called neurons assembled in a logical way and constituting several layers. These neurons and layers form a neural network that resembles a biological neural system. A neuron is an information processing unit, consisting of three main elements: synapses (links), a linear combiner, and an activation function. Each synapse (link) contains a weight factor. The inputs of the neurons are connected through the synapse, which is multiplied by the synaptic weight. The linear combiner adds the neuron's weighted inputs together and the activation function limits the neuron's output. Hence, the output of a neuron depends on its inputs and its activation function. There are different types of activation functions, the most commonly used activation functions are hard limit, linear, or sigmoid functions.

The neurons in the network are arranged into three distinct types of layers. The input layer is not itself a processing layer but is simply a set of neurons acting as source nodes which supply input feature vector components to the second layer. Typically, the number of neurons in the input layer is equal to the dimensionality of the input feature vector. Then there are one or more hidden layers, i.e. layers of computing nodes, each of these layers comprising a given number of neurons called hidden neurons. Finally, the output layer provides the response of the neural network to the pattern vector submitted in the input layer. The number of neurons in this layer corresponds to the number of classes the neural network should differentiate. Normally the neurons are completely connected in-between layers, so that each neuron in each layer is connected to every neuron in the next layer.

The widely used approach to train neural networks is the back propagation algorithm (BPA) (Looney 1997). BPA consists of two stages: forward pass and backward pass through the network. In the forward pass, the input is conveyed layer by layer all the way to the output neuron, which produces the true output of the network. In the backward pass, an error signal is produced by deducting the desired output from the actual output. This error signal is conveyed backwards through the network, layer by layer, simultaneously modifying the values of the network weights, thus bringing the actual output closer to the desired output. Although the neural networks are thought to be a good classification method for biosignal decoding, high computational load sometimes limits the use of this method in practical or real-life applications.

3.5.1.3 Support vector machine (SVM)

The support vector machine (SVM) is a new class of a supervised learning technique developed based on statistical learning theory (Vapnik 1999). In recent years, SVM has been considered as a powerful tool for data classification and function estimation, and it is often used for pattern recognition of biosignals (Lotte et al. 2007). The SVM estimates the optimal boundary in the feature space by combining a maximal margin strategy with a kernel method. The SVM is trained according to the structural risk minimisation criterion (Gunn 1998). The decision boundaries are directly derived from the training dataset by learning. The SVM maps the inputs into a high-dimensional feature space through a selected kernel function. It then constructs an optimal separating hyper-plane in the feature space for classification. The major strengths of SVM classification is that the training is relatively easy with few parameters and less possibility to get local optima. Unlike neural networks, SVM scales relatively well to high dimensional data and the trade-off between classifier complexity and error can be controlled explicitly (Gunn 1998). One of the major challenges of SVM is to choose an appropriate kernel function for the given problem. There are standard choices such as a linear, radial basis function (RBF) or polynomial kernel function that are the default options. The optimal kernel function is dependent on the specific data and linear or RBF kernel is generally used in biosignal classification (Lotte et al. 2007; Bashashatiet al. 2007). More details of the SVM classification is illustrated in chapter 5 (cf. section 5.3.1).

3.5.2 Brief review of classification methods for biosignal decoding

Lotte et al. (2007) performed a survey for identifying appropriate classification methods for decoding event related neural signals (Lotte et al. 2007). They divided the classification methods into five different categories: linear, neural networks, nonlinear Bayesian, and nearest neighbour classifiers as well as combinations of classifiers. Linear classifiers are discriminant methods that use linear functions to distinguish classes. Linear discriminant analysis (LDA), and SVM have been considered as linear classifiers. The main drawback of LDA is its linearity that provides poor results on complex nonlinear EEG data which can be overcome by SVM through selecting an appropriate kernel. Bayes quadratic and Hidden Markov Model (HMM) have been considered as nonlinear Bayesian classifiers, which were not very widely used in brain machine interface. HMM is a popular dynamic classifier in the field of automatic speech recognition. HMM is a kind of probabilistic automaton that can provide the probability of observing a given sequence of feature vectors. The probabilities usually considered for HMM is a Gaussian mixture models (GMM) and it is generally suitable for the classification of time series. The nearest neighbour classifiers are relatively simple, which consist of assigning a feature vector to a class according to its nearest neighbour(s). *k*-nearest neighbors (*k*-NN), Mahalanobis distance, Euclidean distance or Manhattan distance are the common nearest neighbour classifiers. Nearest neighbour algorithms are not very popular in the biosignal classification, probably because they are known to be very sensitive to the curse-of-dimensionality, however, with low dimensional feature vectors, it may provide good accuracy.

In general, it was observed that neural network classification method is widely used for classifying EEG signals (Mason et al. 2007; Bashashati et al. 2007). A similar scenario was also observed for the classification of cortical neural recordings, ECoG, single unit and LFPs (Loukas & Brown 2004; Bashashati et al. 2007). In addition to these Bayesian, SVM and LDA classification were investigated for neural signals pattern classification. LDA classification method was extensively investigated with feature extraction based on auto-regressive parametric modelling. On the other hand, with time-frequency based feature extraction methods, neural network and SVM classifications were explored for decoding events from different neural signals. These classification methods were also analysed with combining two or more feature extraction methods (Bashashati et al. 2007).

To explore effective combinations of feature extraction and classification methods for predicting self-paced right and left hand movements of single trial EEG, an investigation was performed in (Bai et al. 2007). 128 channels EEG recording were performed from twelve subjects. They applied several feature extraction and classification methods, and showed that the combinations of feature extraction using ICA, power spectral density estimation and discrete wavelet transform with SVM classification methods provided higher accuracy. They also reported that Bayesian classifier also provided comparable discrimination accuracy. Another comparative study of different classification techniques for recognising mental tasks was investigated in (Rezaei et al. 2006). Five classification techniques Bayesian graphical network, neural network, Bayesian quadratic, Fisher linear discriminate and HMM were compared based on two known EEG datasets (Graz dataset, Purdue dataset). They reported that the Bayesian network appeared to have a significant accuracy and more consistent classification compared to the other four.

Due to the simplicity of the classification process, Bayesian methods can be considered as an alternative compare to more complicated methods such as neural networks. Kohlmorgen et al. (2004) designed a Bayesian classifier to discriminate single-trial event-related potentials (ERP) (Kohlmorgen & Blankertz 2004). The classification worked based on average signal and its variance as a generative model for each event class. The correct recognition rate was achieved 95%, which is comparable to recurrent neural network (96%). Bayesian classification can be used for supervised as well as unsupervised classification (Bai et al. 2007). A study on Bayesian networks with wavelet analysis feature was investigated for classifying auditory brainstem responses (Rui et al. 2006). On the other hand, an adaptive online Bayesian classifier has been developed based on least mean square (LMS) algorithm called Decorrelated LMS to classify the EEG signals (Shiliang et al. 2005). ICA with Bayesian classification was explored to design an automatic removal of artifacts from EEG (Levan et al. 2006). The classifier was trained using numerous statistical, spectral, and spatial features of the EEG. In that system, an ICA component was considered as EEG activity, when it exceeded threshold otherwise it represented an artifacts. Another investigation on Bayesian classification method for classifying normal and abnormal swallowing sounds based on wavelet features was reported in (Spadotto et al. 2009).

An imaginary movement decoding experiment was performed based on intracranial EEG (ECoG) recordings through placing electrode grids on the motor cortex from three epilepsy patients (Lal et al. 2005). The patients were asked to repeatedly imagine movements of tongue, little finger and hands. With the sequential feature selection, SVM classifier was trained and evaluated for movement classification. A P300 based signal classification experiment was investigated in (Thulasidas et al. 2006). They applied PCA to reduce input dimension and SVM with Gaussian kernel for classification of different characters of a word and achieved 95% average accuracy. An approach of multiclass SVM based on DWT feature with the error correcting output codes was presented for classification of four types of electrocardiogram (ECG) beats (normal beat, congestive heart failure beat, ventricular tachyarrhythmia beat, atrial fibrillation beat) and reported high accuracies (Ubeyli 2007).

In most biosignals decoding research, the classification was achieved using a single classifier. However, combination of multiple classification methods using voting strategy, such as ensemble classification may provide better classification performance (Lotte et al. 2007; Mace et al. 2011).

3.6 Conclusion

This chapter illustrated the different pre-processing, feature extraction, feature selection, and classification methods that already employed in biosignal processing for assistive HMI including BMI. The performance of any proposed biosignal decoding system greatly depends on proper selection of feature extraction, feature selection and classification methods, and these methods need to be robust and efficient. For this reason, exploration of existing methods and their processing characteristics for the given input biosignals are essential. Another issue is synchronisation of feature extraction and classification methods. It is not possible to get good recognition results without proper combinations. The selection of methods in every signal decoding stage may depend on the types of the biosignal to be classified. In addition to this, it is also necessary to incorporate some pre-processing and fine-tuning approaches to improve the recognition rate. As observed in the literature, to analyse the biosignals for feature extraction, wavelet method is more convenient. It is also efficient to localise discriminative signal feature rather than the noise feature from the corrupted biosignals, which leads to developing an effective classifier.

So far most causal study was performed to understand the neural mechanism of the brain activity by evaluating directional influences or connectivity. To the best of our knowledge, no study has yet incorporated causal analysis for feature extraction to decode neural activity. To investigate the potentiality of causal or directional information as feature, Granger causality analysis is evaluated for neural decoding. It is expected that causal feature combined with other domain features will enhance the decoding of neural activity.

As extracted features have high dimensionality and also contain redundant features, due to these issues feature selection is important to increase interclass discriminability. Several review studies also suggested that integration of feature selection techniques significantly improve the classification accuracy and reduce the dimensionality for biosignals decoding (Bashashati et al. 2007; Gerven et al. 2009). The choice of the feature selection technique depends on the problem characteristics. Among various selection techniques, sequential methods (forward or backward feature selection), genetic algorithm and PCA got more attention for neural signal decoding, however due to simplicity sequential method is a good option for classification.

On the other hand some of the classification methods discussed earlier showed potential benefit for different biosignal classifications, particularly in developing assistive HMI. However, they have some limitations as well. According to the review of the several classification techniques, it was observed that Bayesian and SVM classifier provide better or comparable results in many cases and these classifiers are simple compared to neural networks. Ideally, performance of the classification methods needs to evaluate within the same context, i.e. using the same dataset (or same user input signal) as well as using the same feature extraction and selection methods. In our study, Bayesian and support vector machine classification methods are evaluated to identify movement related states from the extracted features of tongue movement ear pressure (TMEP) and deep brain local field potential (LFP) signals.

Chapter 4 : *Pattern Classification of Tongue Movement Action from TMEP signals*

4.1 Introduction

A wide range of research has been conducted to develop various human-machine interfaces (HMI) based on human physiological signals for hands-free communication to assist physically impaired patients (a brief summary can be found in chapter 2). Specific hands-free communication and control devices are essential for an individual who has limited mobility or severe motor dysfunctions, for example due to spinal cord injury, congenital limb deformities or arthritis. In spite of significant progress made in the development of techniques and devices for HMI systems, current products have not yet fully addressed patient-specific requirements and better interfaces between the patient and peripheral devices are still greatly needed. Recently a novel hands-free communication concept based on tongue-movement ear pressure (TMEP) signals has been introduced (Vaidyanathan et al. 2007). Users express their intention by making impulsive actions of the tongue, which create unique acoustic pressure signals within the ear canal. These pressure signals can be recorded easily using a microphone earpiece positioned non-invasively within the ear canal. The advantage of utilising the tongue is that it has an inherent capability for fine motor control, involving multiple degrees of freedom, as it has evolved to perform sophisticated motions during speech and mastication (Vaidyanathan et al. 2007; Huo et al. 2008). The system also has the additional benefits of being simple, cheap and non-invasive. For people with limited control of their limbs, paraplegia or even quadriplegia, if they still have the ability to perform their tongue movement in daily life, they can use different prescribed tongue movements to communicate with computers and control assistive devices through the sensing of bio-acoustic pressure signals.

Previously, different types of tongue movements recorded from healthy subjects relating to the controlled (intended) actions have been classified using a decision fusion

algorithm (Vaidyanathan et al. 2007). The performance of the classifier reached an average of 97% correct accuracy using time domain features and large training sets for recognising four types of tongue movement ('Up', 'Down', 'Left' and 'Right') actions. The performance of this classifier was shown to be better than or similar to three other strategies using time domain information, namely, the matched filter (86%), the parametric autoregressive (AR) Gaussian classifier (85.98%) and the nonlinear alignment classifier (96.27%). Moreover, to improve the classification performance a single channel independent component analysis (ICA) was used to isolate the critical components of TMEP signals associated with the four different tongue motions (Vaidyanathan & James 2007). This method potentially extracted features and may be more useful when a higher number of movement actions or commands are required, or the signals are contaminated with noise. However, the higher computational load makes it unsuitable for real-time applications.

It was also reported that the TMEP based control system can be applicable in developing rehabilitation devices as well as robotic applications. Specifically, an investigation involving hands-free robot tele-operation using TMEP signals demonstrated that with time domain feature and the decision fusion classifier, it achieved 97.51% accuracy for recognising tongue movements (Vaidyanathan et al. 2007a). In that study, the TMEP acquisition system was extended into a dual mode recording to monitor both tongue movements and isolated speech utterances as control commands. Feature extraction and classification of speech signals were performed using Mel Frequency Cepstral Coefficients (MFCCs) and Hidden Markov Models (HMMs) respectively, and achieved 92.77% correct recognition accuracy. It was reported that the TMEP signals relating to tongue movements are faster, quieter, and (in most cases) more intuitive to the user for directing and controlling a moving vehicle compared to speech commands. The initial analysis of dual-channel (i.e. two ears) TMEP acquisition and control operation has shown slightly better classification rate compared to a single channel. However it potentially creates problem in subject's hearing, thus only single channel recordings of TMEP signals is considered in this study.

To implement a real-time hand-free HMI system based on TMEP signals for enabling physically impaired people, several challenges need to be addressed. One significant challenge is the ability of the system to classify TMEP signals in real environments under the influence of interference, including external noise from the surrounding

environment (e.g. conversation, road noise), motion artifacts (e.g. head movements), internal noise or artifacts due to natural tongue movements (e.g. speech, mastication). Such interferences may substantially deteriorate the performance of the assistive HMI system and could potentially generate false actions or commands. It may cause significant even fatal effects to those disabled people. Also interference problems are generally challenging in any human machine interface system. Superior performance of TMEP signal classification has been achieved in the datasets collected in controlled environments (Vaidyanathan et al. 2007). Previously, no work has been particularly performed to investigate the TMEP action signal classification in adverse or challenging environments mentioned above. Nevertheless, such investigations are vitally important towards the development of real-time TMEP based assistive HMI applications. Another challenge is that only a limited number of signals are available to train and calibrate the classifier in real environments. On the other hand, the accuracy and robustness of a classification algorithm depends highly on its input and therefore efficient feature enhancement is essential, especially in noisy environments.

Part of this thesis is aimed to investigate such challenges mentioned above and develop an efficient feature enhancement and classification strategy for decoding intended TMEP actions in adverse condition with limited datasets. Particularly, to address the issue of classifying TMEP actions in the presence of external interferences (e.g. conversation, road noise), as well as the issue of a minimal size of the training dataset to provide satisfactory accuracy, the analysis and outcomes are presented in this chapter. Again to address the substantial challenge for identifying TMEP action signals in real environments under the influence of internal interference due to natural tongue movements (e.g. speech, mastication), an efficient strategy was developed and initial evaluation was performed in real-time, and is presented in chapter 5.

All of the previous analysis on TMEP signals to recognise specific actions have exclusively focused on time domain signatures of the acoustic pressure waves in the ear resulting from various tongue movements. It also gained success for differentiating tongue movement actions from one another in controlled environments. To focus on the tongue movement action classification from TMEP signals in noise, an initial analysis was performed to observe the behaviour of TMEP signal by separating its rhythmic and random noise components using wavelet decomposition with soft thresholding (Vaidyanathan et al. 2008). According to the literature review, wavelet based methods

can capture localised time-frequency information of signals and has been implemented widely in signal analysis and modelling, with significant successful application in diverse fields such as signal detection, classification, compression, noise reduction and image processing (Unser & Aldroubi 1996; Learned & Willsky 1995; Oskoei & Hu 2007; Addison et al. 2009; Xie et al. 2009; Antonini & Orlandi 2001). It is hypothesised that feature enhancement of TMEP signal using wavelet based methods will provide robust features for classification of tongue movement action in both controlled and noisy environments. To improve the classification performance in the presence of external interference, the wavelet packet transform (WPT) and stationary wavelet transform (SWT) were applied separately to extract features of TMEP signals for classification of tongue movement actions one another. Based on these extracted features, three classifiers, the Euclidian distance, Manhattan distance, and Bayesian classifiers were designed separately with very limited, small as well as large size training sets for classifying clean and low signal to noise ratio (SNR) TMEP signals. These investigations and outcomes are presented in this chapter. This study provided a foundation to select the efficient feature extraction and classification methods for subsequent analysis to design robust classification strategies to decode tongue movement actions from TMEP signals in real-time.

Section 4.2 describes the procedure of TMEP signal acquisition. Detailed analyses of signal processing and classification stages are illustrated in section 4.3. Tongue movement action classification results are presented in section 4.4. Finally section 4.5 provides a discussion of the TMEP action classification, stating the limitations and draws conclusions pertaining to the efficacy of the techniques employed.

4.2 TMEP signal acquisition

As mentioned in chapter 2 (cf. section 2.3.3), the oral cavity is connected to the ear via the Eustachian tube, which is normally in a closed position (Gelfand 2007). The tube opens reflexively by the action of different activity within the oral cavity such as swallowing, shouting, eating, drinking, coughing, smoking and/or speaking. All of these actions involve active participation of tongue movements, which create a pressure signal that travels through the Eustachian tube to the ear. This pressure signal causes a change in the air pressure or airflow within the ear canal and can be monitored by inserting a sensor into the ear canal as shown in figure 4.1 (Vaidyanathan et al. 2007). The strength

of this signal corresponds to the direction, speed and intensity of the movement of tongue which is unique to the respective action. This phenomenon guides us to use a tongue movement as control command for assistive HMI applications.

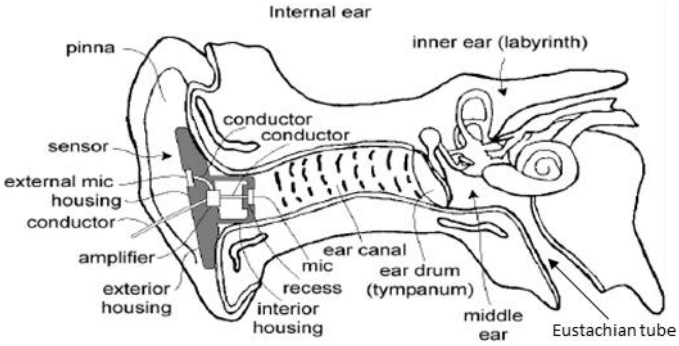


Figure 4.1: Earpiece housing and insertion into the ear to record the TMEP signal for different tongue movements. From (Vaidyanathan et al. 2006).

When a human subject moves their tongue in a certain direction, the generated air pressure changes may be characterised as sound waves or vibrations within the ear. As sound or vibration signal spreads through mediums other than air, such as bone and tissue, the intensity of the signal falls, which is inversely proportional to the distance from the source (Vaidyanathan et al. 2007). Taking this into consideration, the placement of the sensor is an important issue to clearly record the desired initiating action from the pressure signal with a detectable intensity level. With an experimental calibration the optimal placements of the sensor to capture the pressure signal was observed and identified. The sensor consists of an internal microphone and a shielded housing. An illustration of the sensor placement to record TMEP signal is shown in figure 4.1. The recording microphone sensor with housing can be custom designed for the individual to form a close and comfortable fit within the ear. Figure 4.2 (a) shows the custom designed microphone-earpiece housing and a subject comfortably wearing the microphone sensor, which is defined as the first generation sensor (Vaidyanathan et al. 2007). The problem with the first generation sensor is that it requires subject specific customisation. To avoid such problems, a second generation microphone sensor was introduced, which is suitable for use with a wider range of subjects with little or no customisation as shown in figure 4.2 (b). Both of these generations are wired for transferring the TMEP signal; however a wireless device such as one which is Bluetooth enabled, might be possible in future. The microphone earpiece is capable to recording

various forms of initiating physical movement of tongue such as clicking the tongue against portions of the mouth, touching the tongue lightly in different parts of the mouth, touching the tongue to certain parts of the mouth, or any of these combinations.

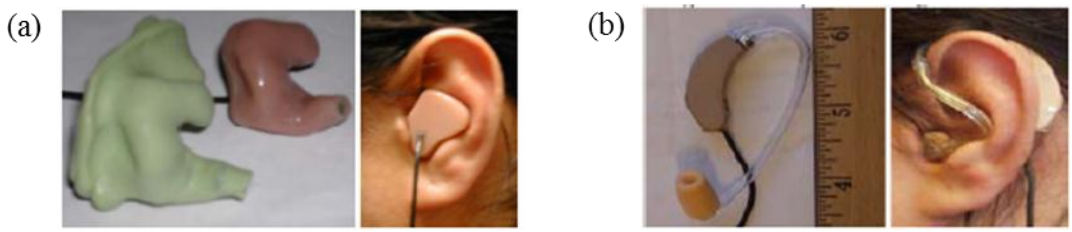


Figure 4.2: First (a) and second (b) generation microphone earpiece for TMEP signal recording. From (Vaidyanathan et al. 2006).

TMEP signals were recorded when subjects performed four types of tongue movement actions: moving the tongue from the neutral position to the top/front centre of the roof of the mouth ('Up'), touching the tongue to the bottom/front centre of the mouth ('Down'), the front/right side of the mouth ('Right'), the front/left side of the mouth ('Left'). The graphical representations of these four tongue movements (Up, Down, Left and Right) and their corresponding typical pressure signal are shown in figure 4.3. These four movements have been selected previously because they can be formed, quite easily and consistently by most individuals, and its generated pressure signal will help the user to operate assistive HMI devices such as a wheel chair.

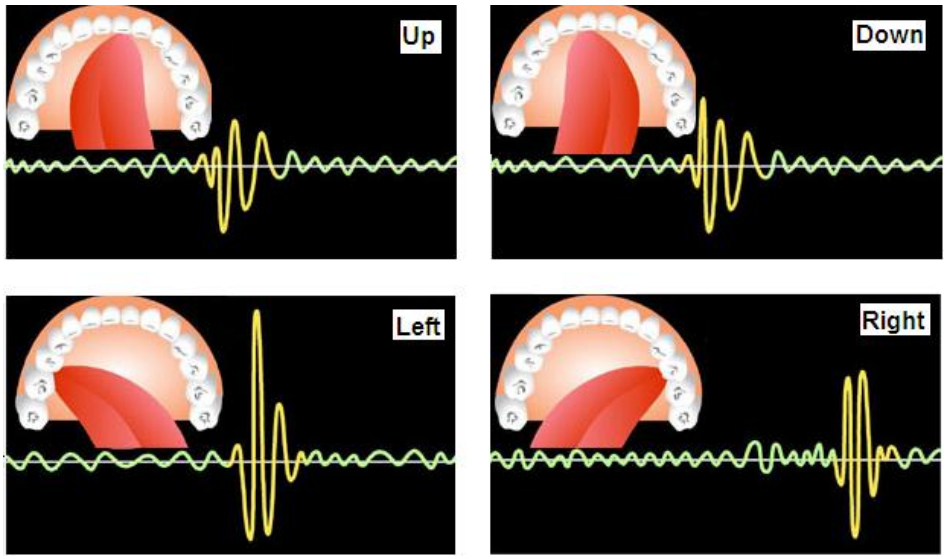


Figure 4.3: The graphical representations of four tongue initiating movements (Up, Down, Left and Right) and their corresponding pressure signals.

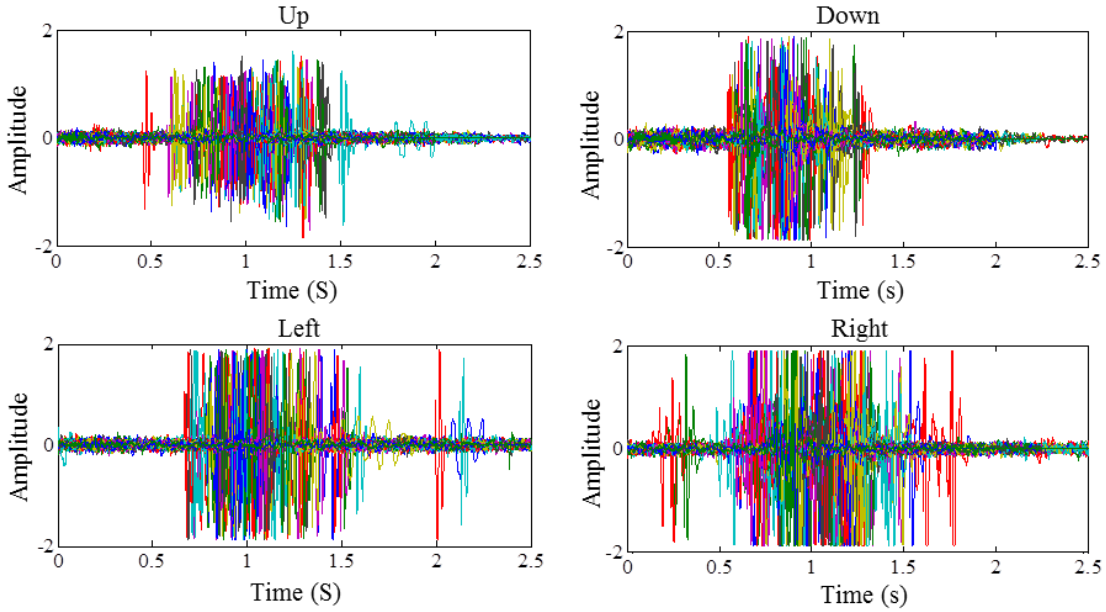


Figure 4.4: Superimposed sampled TMEP signals of subject 3 containing tongue movement actions by moving or flicking the tongue Up, Down, Left and Right directions (amplitude vs. time). Each TMEP signal has 100 repetitions.

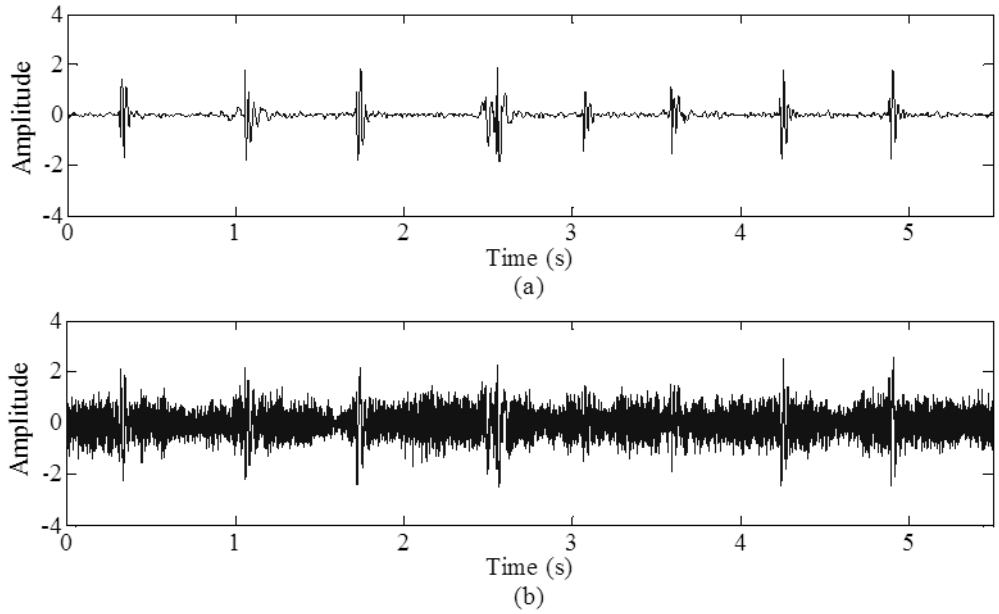


Figure 4.5: TMEP signals were recorded during different 'Up', 'Down', 'Left' and 'Right' actions (a) and contaminated signal with babble noise in 0 dB signal to noise ratio (b).

Previously, TMEP signals were recorded from five healthy subjects (age ranges from 20 to 54 years) with sampling frequency 2 kHz. Each of four movements (Up, Down, Left and Right) was repeated 100 times. For visual representation, superimposed sampled

TMEP signals of subject 3 containing tongue movement actions with 100 repetitions of each of four actions are presented in figure 4.4. From the visual analysis it was observed that the characteristics of the TMEP signal of these movements are distinct, which is also shown in figure 4.3. All of these recording tasks were performed in clean (laboratory) environments with the approval from the local research ethics committee.

As mentioned earlier, external interferences can merge with TMEP signals in real life such as the signal acquired in the street or at a party. To simulate such environments babble noise was added to the recorded TMEP signals with low SNR (0 dB) in this study and is denoted as noisy TMEP signal throughout the rest of this chapter. The continuous recorded clean TMEP signal and its noise contaminated signal at SNR 0 dB during different tongue movement actions are shown in figure 4.5.

4.3 Signal processing and classification

The objective of the TMEP signal decoding is to classify the patterns of tongue movement related to an action based on the efficient feature enhancement and classification method. The pattern classification of four tongue movement actions (‘Up’, ‘Down’, ‘Left’ and ‘Right’) from TMEP signals consisted of pre-processing, activity detection and segmentation, feature extraction, and classification. The detail flowchart of these stages is shown in figure 4.6.

4.3.1 Pre-processing of TMEP signal

Through power spectral density (PSD) analysis of TMEP signals, it was observed that the average power spectra of four tongue movement actions shows the dominant peaks at low frequency around 30 Hz with two separated peaks in some cases with signal power spread up to 200 Hz. Similarly, PSD analysis was performed for noisy TMEP signals and observed that average power spectra for four tongue movement actions show the dominant frequencies below 100 Hz with signal power distributed within 200 Hz. The average power spectra of four tongue movement actions in the TMEP signals from five subjects are shown in figure 4.7. Hence, recorded signals were low-pass filtered (Chebyshev Type I filter with zero-phase shifting) with cut-off frequency 200 Hz to remove high frequency information from the TMEP action signals. According to the PSD, it can be noted that the TMEP signal for each tongue movement action may have its own distinguishable pattern in each subject.

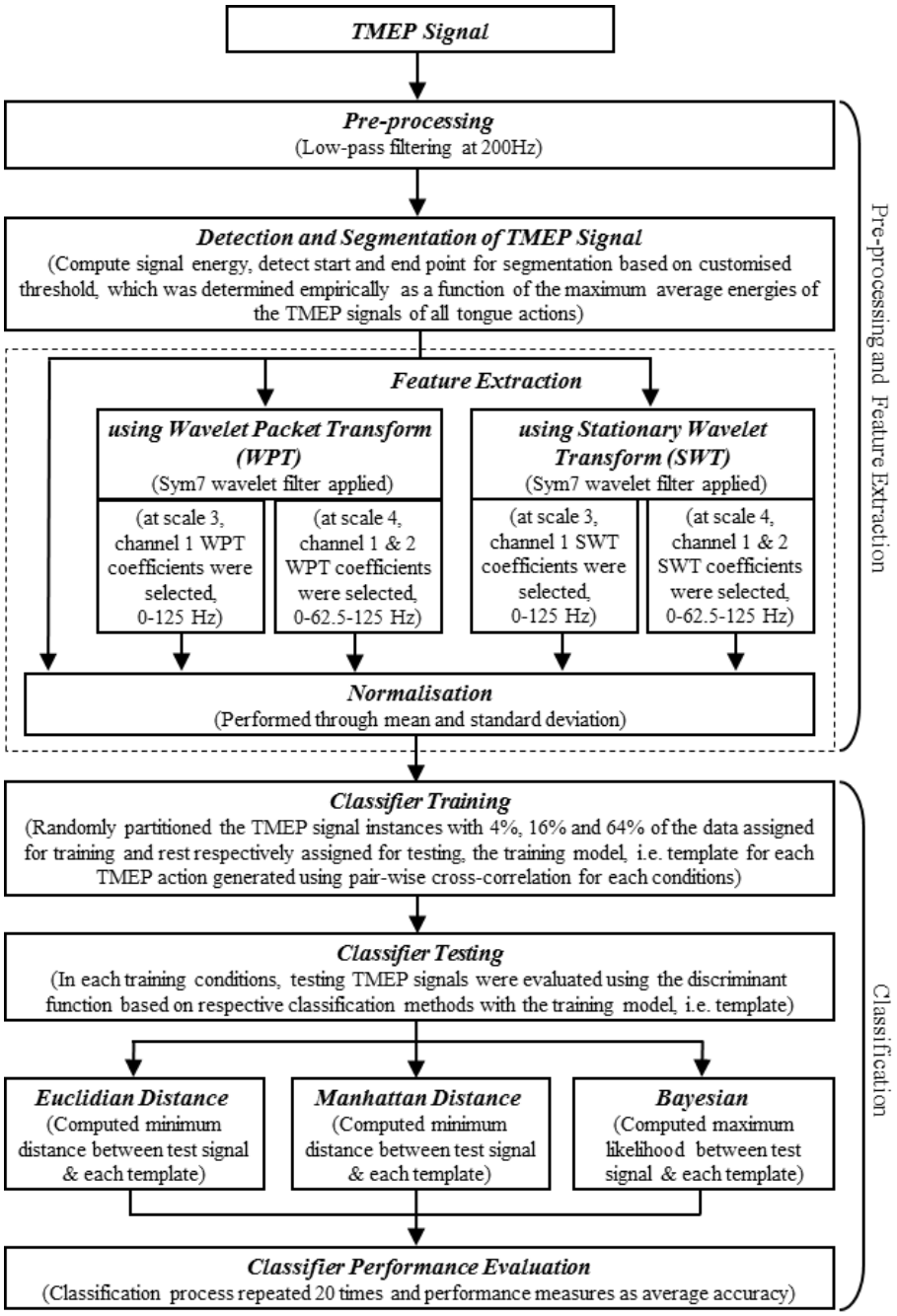


Figure 4.6: The detail flowchart of tongue movement action classification from TMEP signal.

4.3.2 Detection and segmentation of TMEP signal

The activity detection and segmentation involves monitoring continuously to identify the presence of tongue movement action in the TMEP signal recording and locate its start and end point respectively. Accurate detection and segmentation is an important issue for a feature enhancement process, which will lead the classifier to estimate the reference or template TMEP signal for each tongue movement action. From the

extensive observation of different tongue movement action related TMEP signals, it was empirically recognised that the amplitudes of ear pressure signal containing the tongue movement actions were significantly higher than the normal resting ear pressure signal. Therefore, an approach based on energy thresholding was set to detect the TMEP actions from the recorded ear pressure signal, in which it identifies the start and end point of the segment. The detection approach is similar to that used in automatic speech recognition systems by setting a threshold on the short-term energy of the incoming signal. The threshold was determined for each subject as 50% of the maximum average peak energy across training TMEP signals of all tongue movement actions. Based on this threshold, the data segment of 512 samples was extracted for further analysis. It is noted that the typical length of the TMEP signal containing tongue movement actions was 0.2 seconds (400 samples). Details of detection and segmentation approach can be found in (Vaidyanathan et al. 2007). The superimposed segmented TMEP signal for tongue movement actions ('Up', 'Down', 'Left' and 'Right') extracted from the ear pressure signals of subject 3 are shown in figure 4.8. Also a typical segmented TMEP 'Up' action is presented in figure 4.10(a). After detection and segmentation, segmented TMEP signals were selected for feature extraction.

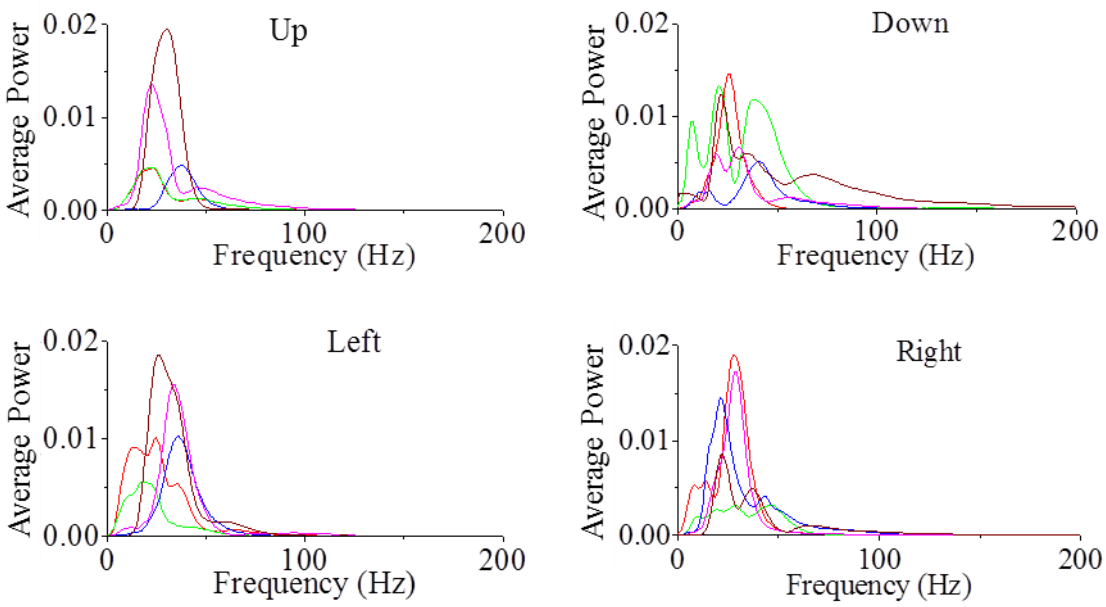


Figure 4.7: Average power spectral density of four tongue movement actions in the TMEP signals from 5 subjects.

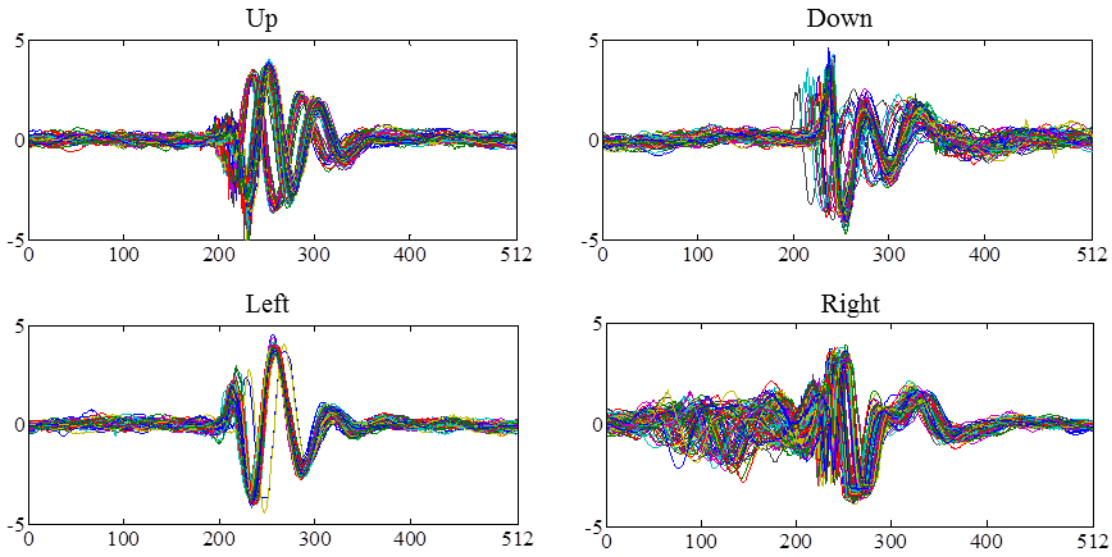


Figure 4.8: Superimposed segmented TMEP signal containing tongue movement actions ('Up', 'Down', 'Left' and 'Right') extracted from ear pressure signals of subject 3 (amplitude vs. time index @ 2 kHz). The segmentation was performed by detecting energy levels which is above the defined threshold of TMEP signal with a window size of 512.

4.3.3 Feature extraction of TMEP signal

One major focus of the pattern classification of tongue movement actions is signal feature extraction, which is the process of enhancing discriminative information from the segmented TMEP signal for classification. To increase the discriminability between different classes of TMEP signal for tongue actions, efficient feature enhancement is required. It can be possible in several domains, which includes the time, frequency, time-frequency and statistical domain discussed in chapter 3 (cf. section 3.2). Previously time domain features of TMEP signals were extensively investigated for classification. With clean TMEP signals, time domain information has showed good discrimination between the classes of action and provided better recognition accuracy. However, if TMEP signals are contaminated with noise, the time domain approach has failed to extract discriminative features for classification that results in a lower classification performance. In such situations, time-frequency analysis approach for feature enhancement can provide greater discriminative information to improve the classification performance. To enhance both time and frequency information of the TMEP signals together, a wavelet based method is investigated for feature extraction. Two wavelet methods, namely, the discrete wavelet packet transform (WPT) and

stationary wavelet transform (SWT) have been explored to extract the discriminative features from the segmented TMEP signal.

4.3.3.1 Wavelet packet transform (WPT)

The discrete WPT represents a generalisation of multi-resolution analysis to decompose a signal into sub-bands and presents both approximation and detail spaces in a binary tree (Walden 2001; Percival & Walden 2000). The wavelet packet coefficients at one level can be recursively decomposed into the coefficients at the next level using a low-pass and high-pass analysing filter. To compute the WPT coefficients of levels j=1, … , J, it filters the WPT coefficients recursively at the previous stage. Let $W_{j,p}(k)$, p=0, …, 2^j-1 represent the WPT coefficients at level j. Then the following two wavelet packet orthogonal base equations are used to compute the wavelet packet coefficients:

$$W_{j,2p}(k) = \sum_{l=0}^{L-1} h(l) W_{j-1,p}(2k+1-l \bmod N_{p-1}) \tag{4.1}$$

$$W_{j,2p+1}(k) = \sum_{l=0}^{L-1} g(l) W_{j-1,p}(2k+1-l \bmod N_{p-1}) \tag{4.2}$$

where k = 1,…, N and $N_p = N/2_p$. $h(l)$ and $g(l)$ are the impulse responses of scaling and wavelet filters, which represents low-pass and high-pass filters, respectively. They are quadrature mirror filters and have only finite non-zero filter coefficients, which results in an efficient way to compute the WPT coefficients.

4.3.3.2 Stationary wavelet transform (SWT)

The SWT is a redundant way of performing multi-resolution analysis to decompose a signal into sub-bands (Percival & Walden 2000; Walden 2001). SWT has been introduced to overcome the limitation of discrete wavelet transform to improve the time resolution of the decomposition (discussed in section 3.2.5). An SWT algorithm usually decomposes the original signal $x(k)$, k=0,1,…,N-1 at each level into its low- and high-frequency components at the next level. The two new components, each have the same length as the original signal. Unlike DWT (cf. section 3.2.5), the low-frequency (scaling) and high-frequency (wavelet) filter in SWT are rescaled as $\bar{g}(k) = g(k)/\sqrt{2}$ and $\bar{h}(k) = h(k)/\sqrt{2}$ respectively. Let $V_0(k) = x(k)$, then the SWT pyramid algorithm

computes the approximation $V_j(k)$ and detail $W_j(k)$ coefficients of level j from $V_{j-1}(k)$ as follows:

$$V_j(k) = \sum_{l=0}^{L-1} \bar{g}(l) V_{j-1}(k - 2^{j-1} l \bmod N) \qquad (4.3)$$

$$W_j(k) = \sum_{l=0}^{L-1} \bar{h}(l) V_{j-1}(k - 2^{j-1} l \bmod N) \qquad (4.4)$$

where $k = 0,1,\ldots, N\text{-}1$. The major advantage of SWT over DWT as well as WPT is the preservation of more time information of the original signal sequence at each level, which may provide more discriminative features. However, due to the high dimensionality of the decomposed signal, it may be computationally expensive for classification.

4.3.3.3 Wavelet filter selection

The efficacy of the wavelet transformation, WPT or SWT is dependent on the wavelet basis or filter. One common approach to specifying the wavelet filter is to select one with minimum reconstruction error according to an entropy cost function (Coifman & Wickerhauser 1992; Wang et al. 2004). This is considered optimal for signal compression, but may be inappropriate for signal classification. A modified algorithm was proposed to maximise the discriminant ability of the wavelet transform by using a class separability cost function (Saito & Coifman 1995). More often the wavelet filter selection is performed empirically according to the above criteria. In this TMEP signal analysis, the selection of wavelet filter was made with the criteria (1) properties of the wavelet filter and (2) a class separability based objective function for evaluation amongst all possible wavelets in the following families: Daubechies, Coiflets and Symlets (figure 4.9). These families of wavelets were considered due to their properties of (1) orthogonal transform, (2) compact support, and (3) optimal number of vanishing moments. The objective function for class separability criterion was defined as a measure of Euclidean distance between classes. Through evaluating the objective function, the optimal wavelet filter and its order was determined by comparing performance among all three wavelet families and chosen filter orders (Daubechies with order 4, 5 and 6, Coiflets with order 3, 4 and 5, and Symlets with order 5, 6, and 7), using all the available data. A Symlet wavelet filter of order seven (Sym7) was selected as it gave the maximum classification performance based on a Euclidean distance measure among the available wavelet families in both WPT and SWT. A few other

wavelets, i.e., Daubechies with order 5, Coiflet with order 4 and Symlet with order 5 also achieved comparable performance.

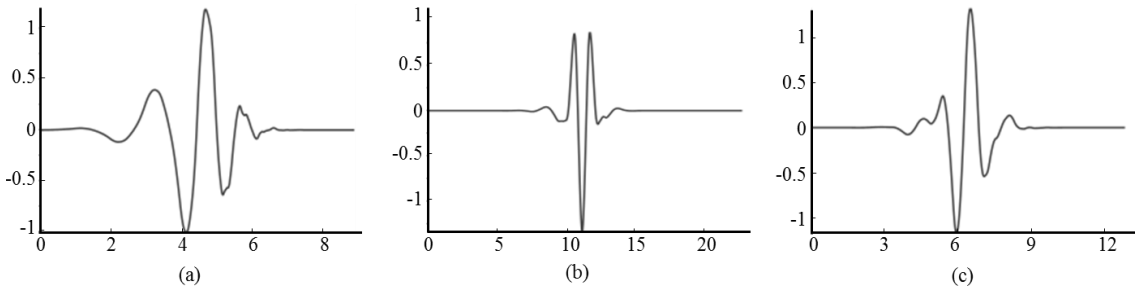


Figure 4.9: The Daubechies with order 5 (a), Coiflet with order 4 (b) and Symlet with order 7 (c) wavelet filter.

4.3.3.4 Feature extraction based on WPT and SWT

To extract features, segmented TMEP signals were decomposed into the WPT domain using a Sym7 wavelet filter. To compare the performance with improved time resolution, segmented TMEP signals were also decomposed into the SWT domain using a Sym7 wavelet filter. In both WPT and SWT transforms, the selection of the optimal decomposition level (interchangeably called scale in this thesis) for feature extraction was made by comparing different decomposition levels (2, 3, 4 and 5) based on the frequency spectrum of the TMEP signals using PSD and their capability to localise discriminative information. Finally, decomposition level 3 and 4 were selected in both WPT and SWT to compare their performance for improving classification ability of tongue movement actions in clean and noisy conditions. An illustration of the WPT (b) and SWT (c) decomposition and its coefficients of TMEP signal for 'Up' tongue movement action (a) with decomposition scale 3 are presented in figure 4.10.

According to the average power spectra of TMEP signal for tongue actions, it was observed that the majority of signal power was concentrated below 100 Hz (figure 4.7). Therefore, the coefficients of WPT at the first frequency band (0-125 Hz) of decomposition scale 3 were selected as feature for classification. Similarly the approximation coefficients of SWT with frequency band (0-125 Hz) at scale 3 were selected as feature for classification. To define the frequency bands of WPT or SWT decomposition, the term 'channel' was used to denote the each frequency band of a decomposed TMEP signal. Again for decomposition scale 4, in WPT, first (0-62.5 Hz) and second (62.51-125 Hz) frequency band coefficients, and in SWT level 4

approximation (0-62.5 Hz) and detail (62.51-125 Hz) coefficients were selected as features for classification, which covers the majority of respective time-frequency information of the tongue movement action in the TMEP signal. It is noted that for the case of decomposition scale 4, the classification process performs based on multi-channel feature (as two separate frequency band or channel coefficients selected) of the TMEP signal in both WPT and SWT. To compare the performance of time domain feature with wavelet based feature, the time information of the segmented TMEP signal was also considered for classification. Finally five normalised feature sets, time information as TIME, WPT coefficients at scale 3 as WPT-DS3, WPT coefficients at scale 4 as WPT-DS4, SWT coefficients at scale 3 as SWT-DS3, and SWT coefficients at scale 4 as SWT-DS4 from clean and noisy TMEP signals was separately evaluated with the classification method and a comparison of their performance was made in respect to their classification ability.

4.3.4 Classification of TMEP signal

4.3.4.1 Estimation of TMEP action templates

From the recording and segmentation, it was observed that TMEP signals vary from subject to subject and also their different tongue movement actions. The segmented TMEP signals (as well as its extracted feature) for different movement actions within the same subject varied in amplitudes, shapes from trial to trial due to inconsistencies in the intensity and speed of the tongue motions. Some variation might also occur from the segmentation process. In such situations, it is not possible to get a predefined standard reference TMEP signal for each tongue movement action. Therefore, signal estimation is essential for generating a reference or template signal for each action from the respective subject for classifier design. Signal averaging is a mostly used operation to estimate the template signal in this case. However, direct averaging of repeated TMEP signal of actions will result in poor, imprecise estimation because all signals are not correctly aligned in time. In this case pair-wise cross-correlation averaging is appropriate for estimation of template. A pair-wise cross-correlation based averaging produces a single template for each class of TMEP action signals. In this approach, a pair of signals is first aligned to each other using cross-correlation and then averaged across them. This is repeated for all trials of each action to generate the template.

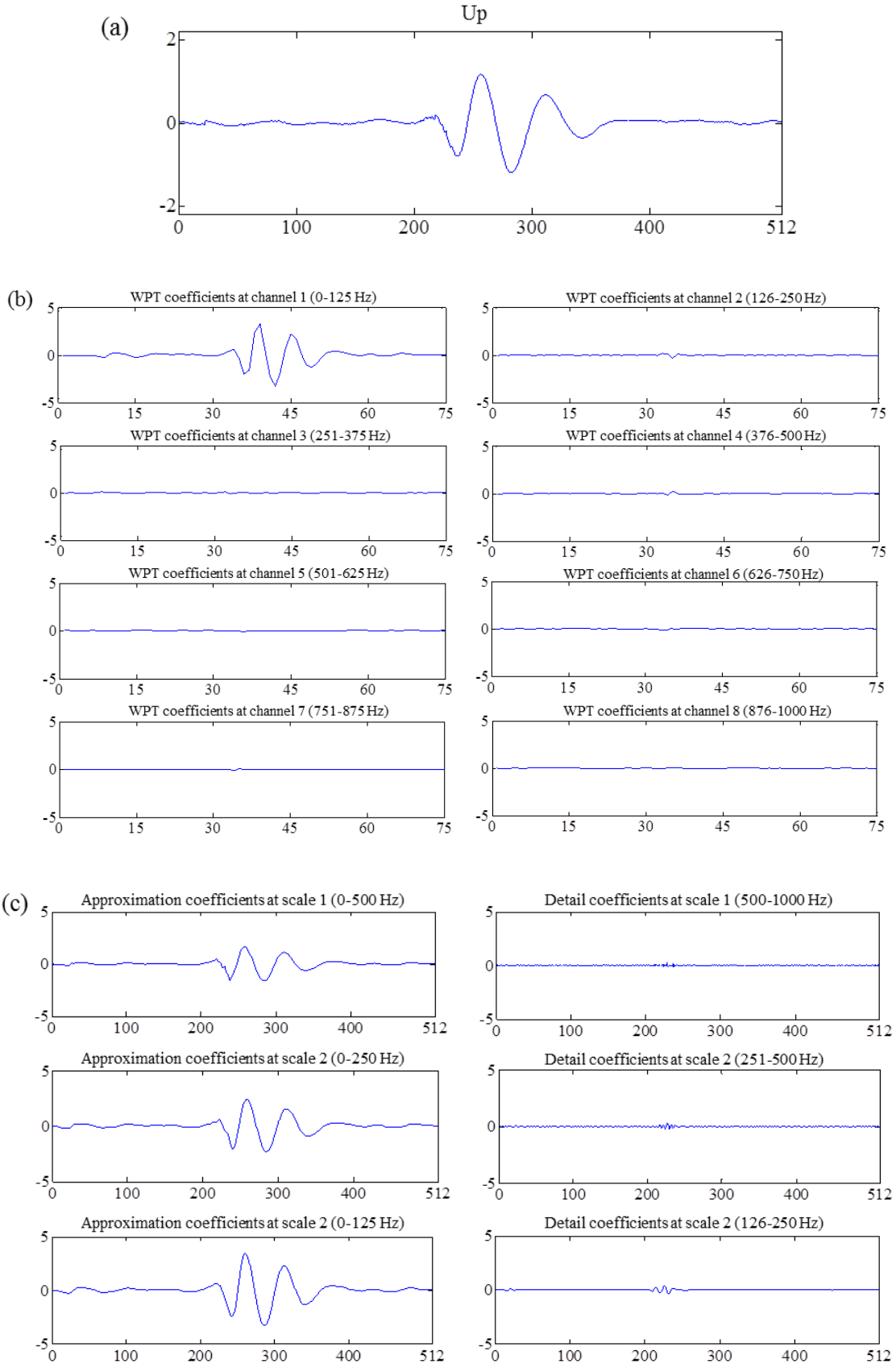


Figure 4.10: A segmented TMEP signal of 'Up' action for subject 3 (a), and its WPT (b) and SWT (c) decomposition and its coefficients using Sym7 with decomposition scale

3. The decomposed signal has 8 channel (2^3) WPT coefficients. The signal frequency distributed across WPT channel with 125 Hz for each and it started from 0-125 Hz for channel 1, 126-250 Hz for channel 2 and so on. Similarly, the SWT decomposed signal has 3 levels of approximation and detail coefficients. The signal frequency distributed as first half to the approximation (0-500 Hz) and other half to the detail (501-1000 Hz) at first level (scale), and it repeats for next level with decomposing only approximation.

The pair-wise cross-correlation averaging process can be considered as a binary tree-like structure, where a pair of input signals represents a leaf (child) node at the first level and their averaged signal represents a parent node at the second level. Using this pair-wise fashion, second level averaged signal is generated from all trials of an action. Again, applying pair-wise cross-correlation averaging to the second level averaged signal and a third level averaged signal is generated. This process is repeated towards generating a root node at level $log_2 N$, where N is the number of trials of an action. Finally aligned and averaged signal at the root node represents the estimated template for a TMEP action. The details of this approach are available in (Gupta et al. 1996; Vaidyanathan et al. 2007).

4.3.4.2 Euclidean and Manhattan distance based classifier

Euclidean or Manhattan metric is a minimum distance based classifier also called a nearest neighbour classifier (Duda et al. 2001). A nearest neighbour based classifier attempts to classify a test pattern based on the class of the "closest" training pattern(s) or template(s).

The Euclidean distance is a straight line distance between two points and is defined as the square root of the sum of squared differences between the corresponding coordinates of the points. Suppose, $x = (x_1,...,x_n)$ and $y = (y_1,...,y_n)$ are two one dimensional time series, then their distance in Euclidean space is computed as

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}. \qquad (4.5)$$

The Manhattan distance is also known as the taxicab metric. The Manhattan distance between two points is defined as sum of absolute differences between the corresponding

coordinates of the points. Suppose, $x = (x_1,...,x_n)$ and $y = (y_1,...,y_n)$ are two one dimensional time signals and their Manhattan distance is computed as

$$|x_1 - y_1| + |x_2 - y_2| + ... + |x_n - y_n| = \sum_{i=1}^{n} |x_i - y_i|. \qquad (4.6)$$

4.3.4.3 Bayesian classifier

A Bayesian classifier is based on the principle of Bayes' decision theorem, which provides a fundamental methodology for solving statistical classification problems when the probability distribution of the pattern is known. Bayes' theorem was used to convert a prior probability into a posterior probability by incorporating the evidence provided by the observed data. Assuming that the prior probability of the i^{th} class (before observation of the data) is $p(c_i)$, the evidence of the observed data is $p(x)$ and the likelihood of the belonging class is $p(x|c_i)$. According to Bayes' theorem the posterior probability $p(c_i|x)$ is formulated as

$$p(c_i|x) = \frac{p(x|c_i)\,p(c_i)}{p(x)} \qquad (4.7)$$

which aims at assigning a feature vector to the class with the highest posterior probability (Duda et al. 2001). When the prior distribution is uniform it is also called a maximum likelihood classifier. Under the assumption of a Gaussian distribution, the probability of observing data x for the given class model $M_i = (\mu_i, \sigma_i)$, can be expressed by the probability density function (univariate case)

$$p(x|M_i) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left[-\frac{1}{2}\left(\frac{x - \mu_i}{\sigma_i} \right)^2 \right] \qquad (4.8)$$

where μ_i and σ_i are mean and standard deviation. Similarly the joint probability density function (multivariate case) for the observing data sequence $\vec{x}$ with the given class model $M_i = (\vec{\mu}_i, \Sigma_i)$, can be expressed as

$$p(\vec{x}|M_i) = \frac{1}{(2\pi)^{d/2}|\Sigma_i|^{1/2}} \exp\left[-\frac{1}{2}(\vec{x} - \vec{\mu}_i)^T \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i) \right] \qquad (4.9)$$

where $\vec{\mu}_i$ is a mean vector and Σ_i is the $d \times d$ covariance matrix. Using the probability density functions for a given model, the decision of the belonging class c_i for data x_j (x or $\vec{x}$) can be expressed as

$$x_j \in c_i,\ P(c_i \left| x_j, M_i \right.) \geq P(c_l \left| x_j, M_l \right.),\ \forall l \neq i. \quad (4.10)$$

4.3.4.4 Classification of tongue movement actions

Three classifiers (Euclidean distance, Manhattan distance and Bayesian classifier with univariate Gaussian assumption) were designed and evaluated based on each extracted feature set (TIME, WPT-DS3, WPT-DS4, SWT-DS3, and SWT-DS4) for decoding tongue movement action from clean and noisy TMEP signals. In the classifier construction for each subject, all of the TMEP signal features were randomly partitioned into two mutually exclusive sets to generate a training set and a testing set for each action class. Previously, the classifier was trained with large training set (64%) and evaluated with small testing set (36%). However, the challenge in real environments is that there are only limited training data available to train the classifier and therefore a classification algorithm with high robustness is necessary. To address these aspects each of the classifiers is separately trained through the three different size training sets consisting of very limited (4%), small (16%) and large (64%) samples for classifying tongue movement action from clean and low SNR TMEP signals. Also each trained classifier respectively evaluated with rest of the data as test set consisted of 96%, 84% and 36% of available data in each action class. In the classifier training, to estimate the TMEP signal template of each action class from the respective training sets, pair-wise cross-correlation averaging was used. Four templates for four tongue movement action class were generated from respective training set to evaluate the test signal from the corresponding test sets. Figure 4.11 represents the templates of each tongue movement action class obtained through pair-wise cross-correlation for TIME, WPT-DS3 and SWT-DS3 feature set of subject 3 for clean TMEP action signals. Similarly, tongue movement action class templates of WPT-DS4 and WPT-DS4 feature set obtained for clean TMEP signal. Using the same procedure, templates of tongue movement action from noisy TMEP signal also obtained for each feature set (TIME, WPT-DS3, WPT-DS4, SWT-DS3, and SWT-DS4) based on the corresponding training set.

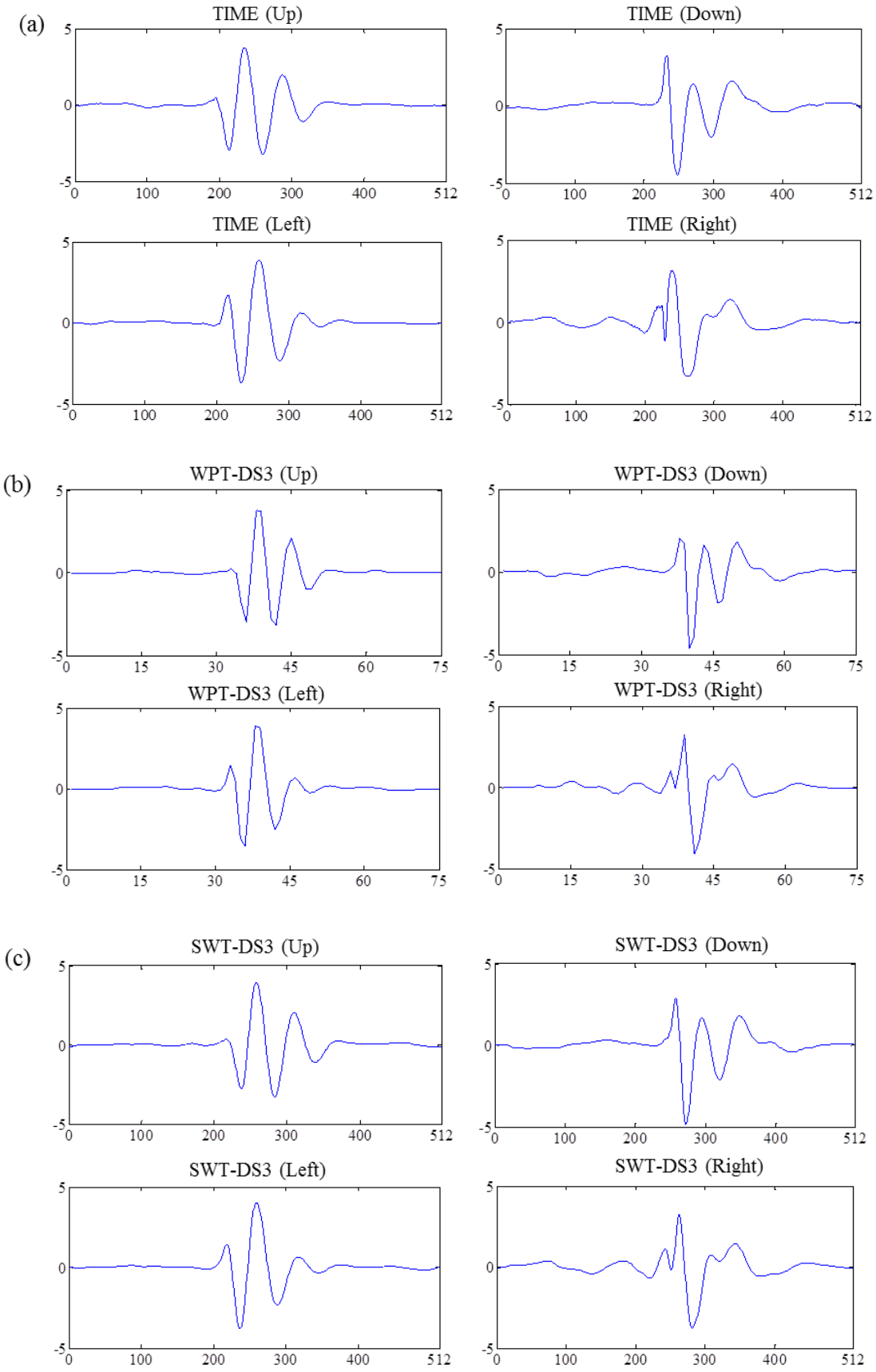


Figure 4.11: Templates of four TMEP action class obtained from TIME, WPT-DS3 and SWT-DS3 feature set of subject 3 for clean TMEP signal.

In the pattern classification phase, each segment from a test set was first aligned to the four action templates using cross-correlation. Then the pattern classification method was applied to find minimum distance using the Euclidean and Manhattan distance, or maximum likelihood probability using Bayesian classifier between the four cross-correlated signal and four templates to generate the decision for the test signal. Finally test segments were classified as an action of a respective class ('Up', 'Down', 'Left' and 'Right') using the classification procedure. The above procedure was followed for each feature set (TIME, WPT-DS3, WPT-DS4, SWT-DS3, and SWT-DS4) in clean as well as noisy conditions, with each pattern classification methods based on each of three training sets for classification. The classification performance was computed as the averaged classification accuracy - the percentage of correctly classified instances for each subject, with all the subjects' accuracies averaged and presented as mean±1SD. The averaged classification accuracy in each classification method was measured by repeating the process twenty times and each time the templates were generated from randomly partitioned corresponding training sets and evaluated on the testing set. The classification rate was then averaged. The averaged classification performance of tongue movement action was compared based on the classification methods as well as the feature extraction approach for clean and noisy TMEP signal with the respective training set.

4.4 Results

The tongue movement action classification results were obtained from five feature extraction sets, TIME, WPT-DS3, WPT-DS4, SWT-DS3 and SWT-DS4, using three classification methods, Euclidean distance, Manhattan distance and Bayesian classifier with 3 different sizes of training set, very limited (Tr-4), small (Tr-16) and large (Tr-64) for clean and noisy TMEP signals. Figure 4.12 (a, b, c) shows the averaged classification accuracy for four tongue movement actions one another in clean and noisy (0 dB SNR) conditions from the five subjects and their average in each feature set using each classifier with very limited training set (Tr-4). Similarly averaged classification accuracy in each feature set using each classifier with small (Tr-16) and large (Tr-64) training set presented in figure 4.13 (a, b, c) and 4.14 (a, b, c) respectively. It was observed that the maximum average classification accuracy among all five feature sets achieved 95.06% using Euclidean distance, 96.24% using Manhattan distance, and 97.11% using Bayesian classifier based on SWT-DS4 feature set for clean TMEP action signal from five subjects with large (64%) training set. When number of the training data segments

decreases to very limited (4%), the performance was reduced in all cases. With the SWT-DS4 feature set, the average classification accuracy achieved only 90.16% using Euclidean distance, 90.80% using Manhattan distance, and 92.30% using Bayesian classifier. However, when small (16%) training set used, the performance was improved to 94.50% using Euclidean distance, 95.17% using Manhattan distance, and 95.42% using Bayesian classifier based on the same SWT-DS4 feature set. It is obvious that large training sets in most cases provide better performance, however with one fourth of the large training set, only using 16% of the data for training and all of the classification approach achieved slightly less (1~2%) accuracy. Due to the limitation of gathering a large training dataset, a smaller size (16%) training set could be a potential solution as it provides satisfactory performance. Again for the case of noisy TMEP action signal, the classification accuracy was highly reduced (3~15%) for the time domain feature in all three classification process. Conversely, when the classification was performed based on the feature extracted from WPT and SWT coefficients, the classification accuracy was improved and its performance is similar to clean TMEP signal classification. According to the classification result shown in figure 4.12-4.14, it can be said that the wavelet based feature extraction approach is robust in noise for providing satisfactory accuracy. Comparing the wavelet based feature extraction approach for providing high accuracy, WPT-DS3 performed better (3~15%) than WPT-DS4; SWT-DS4 performed slightly better (0~2%) than SWT-DS3; and SWT-DS4 performed slightly better (1~3%) than WPT-DS3 in all classification methods for clean and noisy TMEP signal. It is also noted that in both conditions of clean and noisy TMEP signal Bayesian classifier performed better (1~13%) than Euclidean and Manhattan distance classifier for each feature set with each of three sizes of training set. The comparative result of average classification accuracy for clean and noisy TMEP action signals in three classification methods with three training set sizes based on feature sets, TIME, WPT-DS3 and SWT-DS4 are shown in figure 4.15.

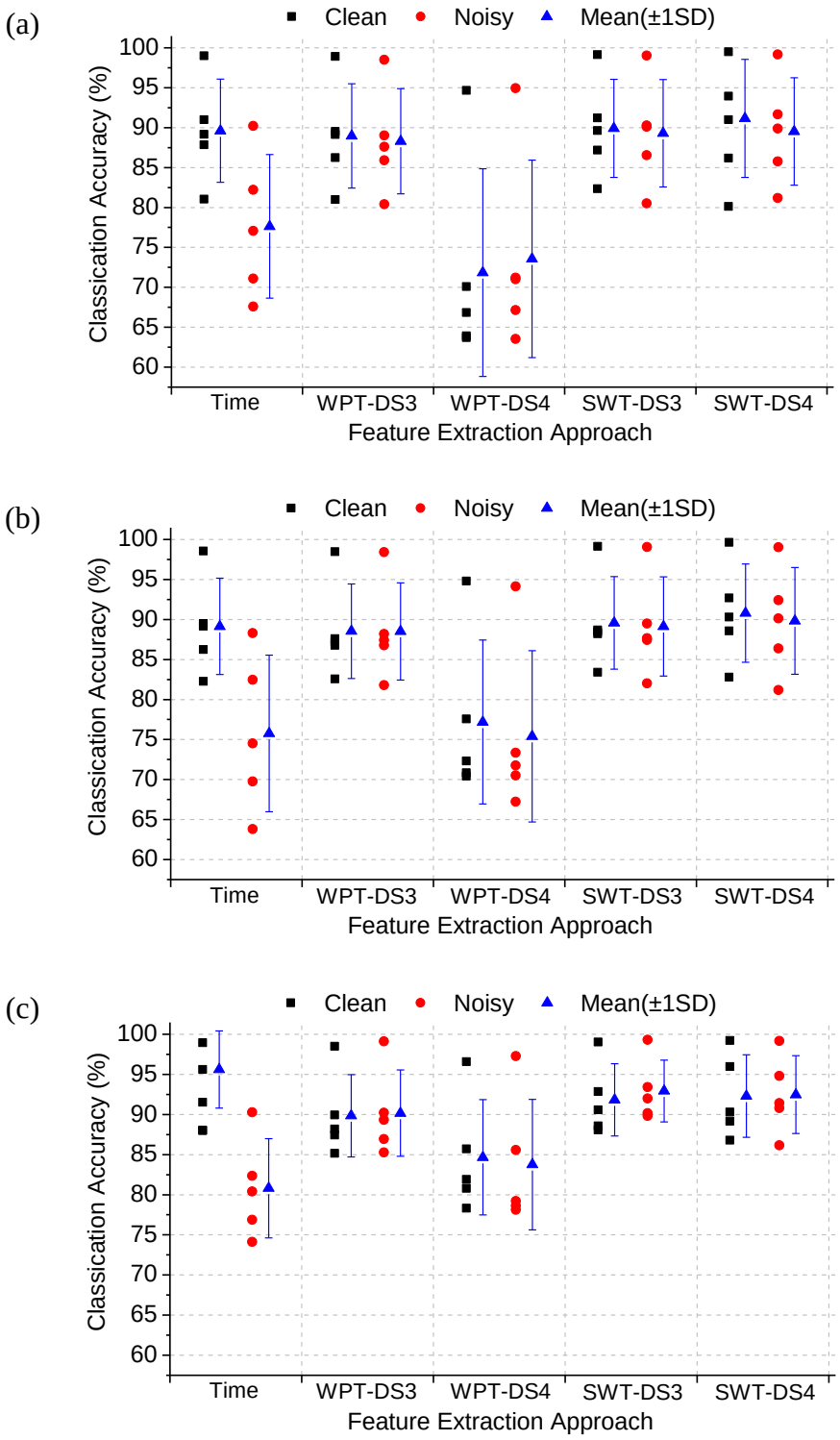


Figure 4.12: Averaged classification accuracy for four tongue movement actions each in clean and noisy (0 dB SNR) conditions from each of five subjects and their average (mean±1SD) based on five feature sets (TIME, WPT-DS3, WPT-DS4, SWT-DS3 and SWT-DS4) with Euclidean distance (a), Manhattan distance (b) and Bayesian (c) classifier for very limited training set (Tr-4).

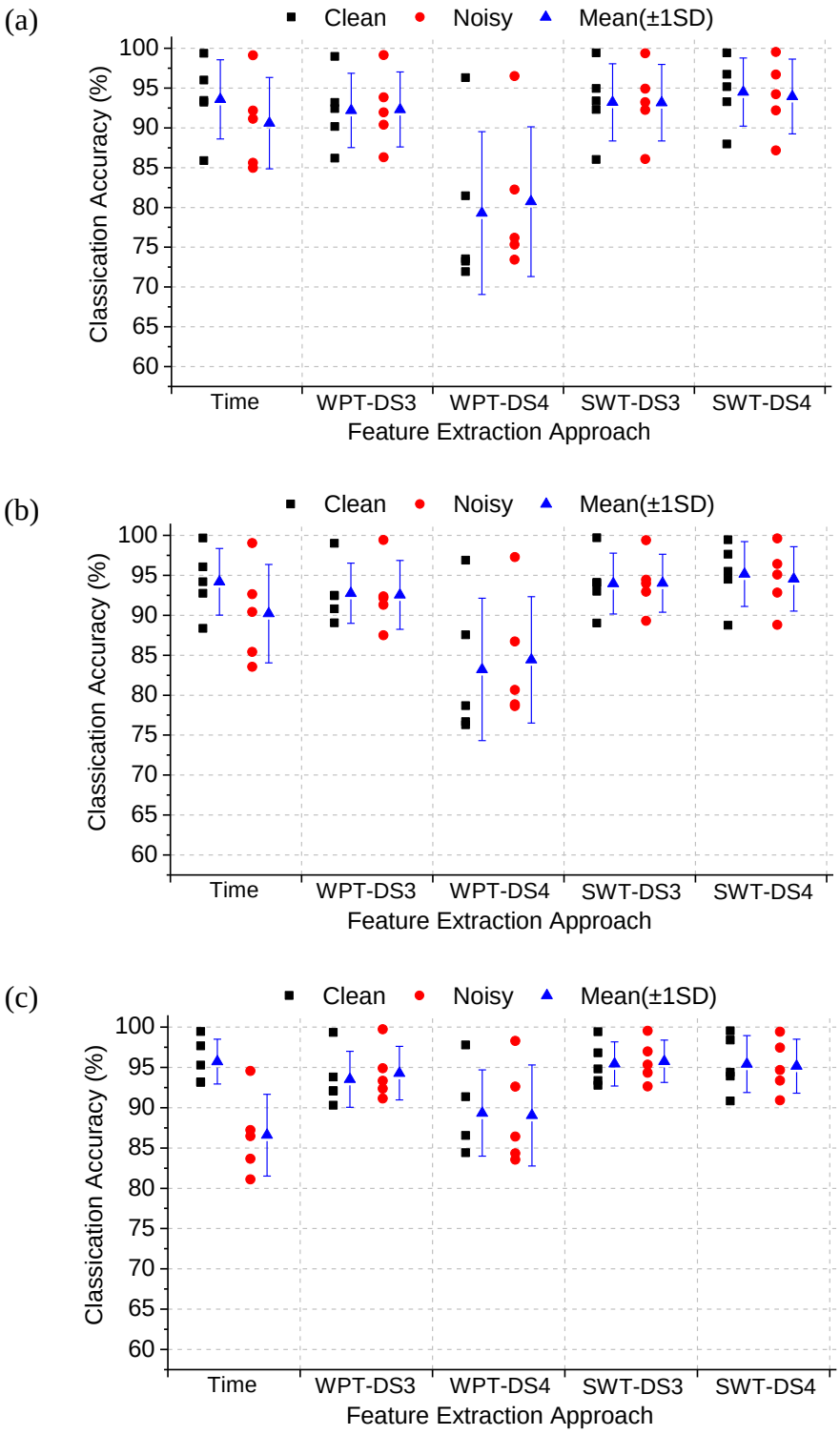


Figure 4.13: Averaged classification accuracy for four tongue movement actions each in clean and noisy (0 dB SNR) conditions from each of five subjects and their average (mean±1SD) based on five feature sets (TIME, WPT-DS3, WPT-DS4, SWT-DS3 and SWT-DS4) with Euclidean distance (a), Manhattan distance (b) and Bayesian (c) classifier for small training set (Tr-16).

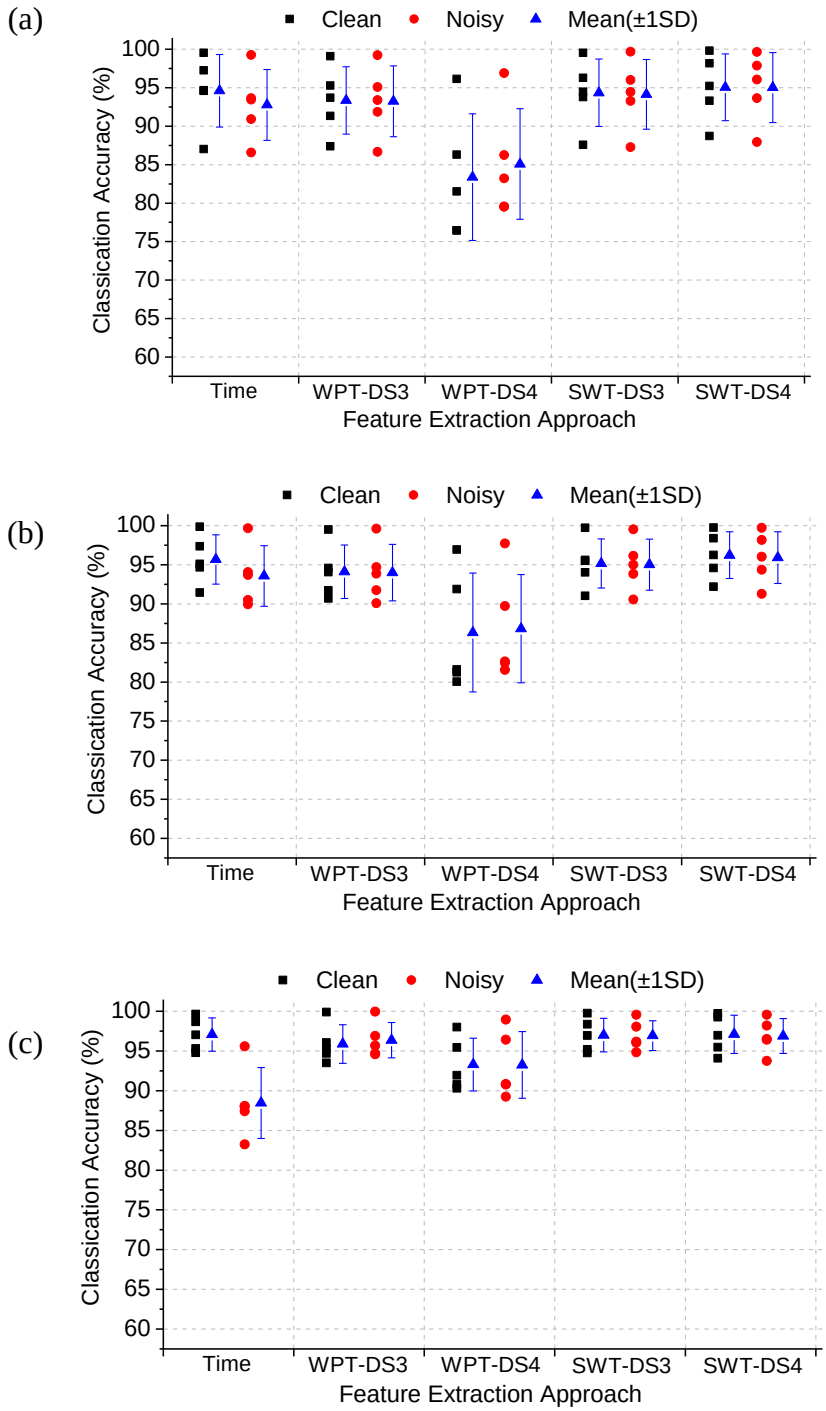


Figure 4.14: Averaged classification accuracy for four tongue movement actions each in clean and noisy (0 dB SNR) conditions from each of five subjects and their average (mean±1SD) based on five feature sets (TIME, WPT-DS3, WPT-DS4, SWT-DS3 and SWT-DS4) with Euclidean distance (a), Manhattan distance (b) and Bayesian (c) classifier for large training set (Tr-64).

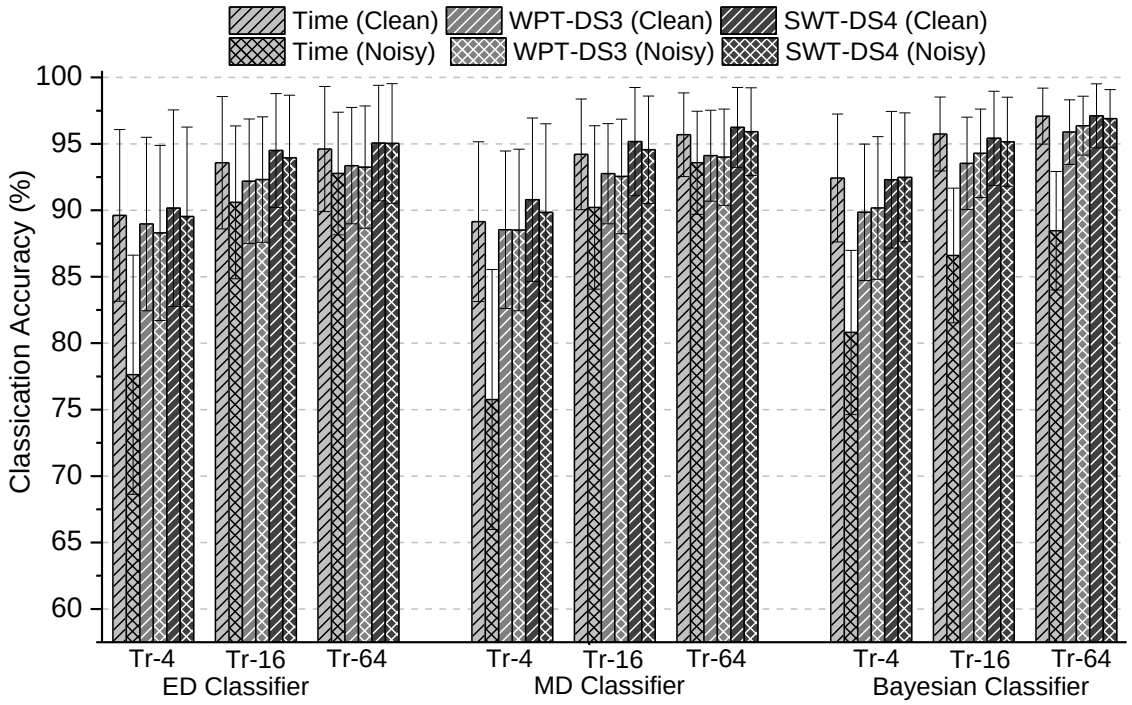


Figure 4.15: The average classification accuracy for tongue movement actions in clean and noisy TMEP signals using three classification methods (Euclidean distance, Manhattan distance and Bayesian classifier) with three training sets (Tr-4, Tr-16 and Tr-64) based on three feature sets (TIME, WPT-DS3 and SWT-DS4).

4.5 Discussion and conclusions

Tongue movement ear pressure signals have been proposed to generate controlling commands or actions for developing human machine interface systems to assist people with limited movement ability. This chapter investigated the feature extraction and classification methods for TMEP signals. Particularly analysis was performed on discrimination of tongue movement actions from one another in clean and noisy conditions. Signal features were extracted using a wavelet packet and stationary wavelet transform and classified with three different classifiers. Each classifier was trained with three different size training sets to find the optimal data size for providing satisfactory performance towards development of real life applications. A Bayesian classifier with SWT features in clean and noisy (SNR 0 dB) conditions achieved higher accuracy for recognising tongue movement actions. Slightly lower classification accuracy was achieved with the WPT features; however, this approach is computationally more efficient than the SWT approach for classification. This research provided an alternative approach for decoding TMEP actions signal in clean and noisy environments.

The WPT and SWT has been widely used for feature extraction from biosignals, such as ECG (Ubeyli 2007), myoelectric signal (MES) (Englehart et al. 2001) or EEG (Hsu & Sun 2009; Samar et al. 1999), and has provided better performance for pattern recognition over time and spectral domain features compared to the discrete wavelet transform and short-time Fourier transform. This may be owing to its capability to precisely localise the time-scale information in non-stationary signal dynamics. To improve the classification performance of TMEP signals in both clean and noisy environments, the wavelet packet transform and stationary wavelet transform were used to extract features for classifying tongue movement actions. The efficacy of the wavelet based approach, WPT or SWT is dependent on the wavelet basis or filter selection. To achieve optimal performance, the wavelet filter and decomposition scale were empirically determined by comparing several wavelet families and scales. It should be noted that some wavelet filters achieved comparable performance to the selected Sym7. The empirical method used to select the wavelet filter and scale in this study is based on classification performance of TMEP signals from limited subjects, hance it may not be feasible to consider the selected wavelet filter and scale as an optimal parameter for wider population. Therefore this selection method may need to be validated with respect to generalisation by considering larger TMEP datasets.

When time domain feature of the TMEP signal was used, the maximum average classification accuracy obtained was 97.09% through Bayesian classifier using a large training set, while with small and very limited size training set the performance obtained 95.74% and 92.43% respectively in clean conditions. However, the classification performance highly deteriorated (3~15%) for the noisy TMEP signal in all classification methods based on time domain feature sets. However, when feature extraction was performed through WPT and SWT, the recognition accuracy of the noisy TMEP signal classification increased and it achieved similar performance to the clean TMEP signal in all of the classifiers. It is noted that only babble noise was considered to simulate the noisy environment, however in real life situation there are various source of noises that can potentially contaminate TMEP signals and may threaten classification performance. The babble noise considered may not be very relevant for TMEP signals to simulate the true noise condition, given that its frequency content lies above that of the TMEP signal and may be removed in preprocessing. Therefore more extensive

analysis of various relavant noisy conditions is essentiall for the predicting performance of TMEP-based assistive communication in real-world application.

With clean and noisy TMEP signal classification, all of the wavelet based feature extraction approach provided high accuracy except WPT-DS4. The performance of the WPT-DS4 feature set for classifying TMEP signal is less (about 10%) than the other wavelet based feature sets (WPT-DS3, SWT-DS3 and SWT-DS4). The potential drawback of this feature extraction is the limitation of DWT includes lack of translation invariance and the reduction of wavelet and scaling coefficients with increase of decomposition scale (cf. section 3.2.5). So the multiple frequency band (channel 1 and 2) selection at higher scale reduces the dimension of the actual signal therefore losing some time information of the signal. This cause may be influenced in the feature space to obtain good discrimination between the different classes of action, and lead lower classification performance. However, the redundant approach, SWT minimises the loss of time information and both decomposition scales of SWT provided very good discriminative feature for classification and achieved high accuracy in all classification methods. Moreover, in some cases the SWT-DS4 feature provided slightly better accuracy than SWT-DS3 (figure 4.15). In addition, SWT-DS4 performed slightly better than WPT-DS3, however, it involves more computation expense than the WPT-DS3 due to the higher dimensionality of the feature space (75 vs. 512).

A previously designed classifier had achieved 97% recognition rate using a decision fusion algorithm with large (64%) training sets in clean conditions (Vaidyanathan et al. 2007). However, it is difficult to obtain large dataset in real life. In this analysis, to evaluate the optimal training set size, three sizes, very limited (4%), small (16%) and large (64%) were considered. Based on the classification performance in all cases of feature extraction and classification approach, it has been observed that classification performance of small and large size training set are similar and very limited size training set is lower. So, to make a trade-off between size of the training set and satisfactory performance small size training set might be an optimal choice for designing classifiers to decode TMEP signal. Given the limited or large amount of available data set, quality of the signal is essential. In some cases due to lack of appropriate experimental or instrumentation setup, the acquired signal quality may be reduced or distorted and it may degrade the overall system performance. Occasionally there was clipping distortion in the data recorded in early stage, and it may cause

performance reduction. Therefore, the overall performance may even become better if proper experimental setup is used, for instance, high sensitivity microphone, customised ear-piece.

In the current analysis, a Bayesian classifier performed better than Euclidean and Manhattan distance classifier in almost every feature extraction methods in clean and noisy condition. It should also be noted that according to methodological perspective, the discriminative functions in Bayesian classifier are computed based on posterior probability and class decision made with maximum likelihood to the belonging class. On the other hand, Euclidean and Manhattan distance classifier is distance based classifier and their decision made based minimum distance to the belonging class. According to the performance and perspective of real-time application, Bayesian classifier would be an optimal choice for TMEP signal classification. Moreover, due to the limited size of the training dataset compared to the large number of features, it was observed in the preliminary analysis that univariate Bayesian classifier performs better than multivariate classifier. The inaccurate estimation of variance matrix in multivariate case may be the reason for lower performance. With larger training dataset, it can be possible to design multivariate Bayesian classifier in futute that may provide better performance.

In summary, the notable technical contributions that were introduced in this chapter are as follows:

- A wavelet packet and stationary wavelet transformation based feature extraction approach developed for TMEP signal classification.
- The initial analysis of robustness of the feature extraction approach was evaluated with clean and noisy TMEP signal (SNR 0 dB), and achieved high accuracy.
- To overcome the limitation of using large size training data, three sizes of training dataset, very limited, small and large were evaluated and found that with small size (16%) training data, it is possible to get satisfactory performance, which is very close to the performance of the large size training set.
- Three simple classifiers, Euclidean distance, Manhattan distance and Bayesian classifier were evaluated and according to performance it was found that a Bayesian classifier would be a good choose for designing a TMEP signal classifier in applications of assistive HMI.

- Typical results from five subjects in offline mode have proved the success of the method.

Chapter 5 : ***Robust Identification of Tongue Movement Actions from Interferences***

5.1 Introduction

In the previous chapter a wavelet packet and stationary wavelet transform based feature extraction approach was presented for classifying tongue movement actions from TMEP signals. It was found that the feature extraction approach was robust to provide discriminative information from the clean as well as noisy (simulated external interferences) TMEP signals, and achieved high accuracy with the Bayesian classification method for both clean and noisy signals. In that analysis, classification performed to distinguish four tongue movement actions (Up, Down, Left and Right) from one another to generate control commands for assistive devices. However, towards development of a TMEP signal based assistive communication system for real life applications, some specific and defined challenges need to be overcome. As mentioned in the previous chapter (cf. section 4.1), one significant challenge in the TMEP signal based system is interferences, which is a generally challenging problem in any HMI system driven by human physiological signals. Moreover, the solution to control or reduce the level of interferences depends on the specific problem and predominant nature of the input signals that will drive the HMI devices. Signals like TMEP have two major sources of interferences, from the surrounding environment (e.g. conversation, road noise) can be denoted as external interferences, and others from inside the oral cavity due to natural tongue movements (e.g. speech, mastication) can be denoted as internal interferences.

External interference problems are common in various signal processing system, also extensively investigated in automatic speech recognition system to separate speech and noise (Chu et al. 2009). The influence of external interferences in the tongue movement action classification was investigated with simulated noisy TMEP signals (SNR 0 dB)

and presented in chapter 4. It was found that signal feature enhancement based on the WPT or SWT is robust and recognised tongue actions with comparable accuracy (97%) as clean TMEP signals. It was also perceived that WPT based feature extraction approach is computationally more efficient than SWT, due to the redundancy and high dimensionality of the SWT based feature.

On the other hand, the internal interference problem is more challenging than the external interferences in the TMEP signals due to the characteristics of the tongue. Tongue is naturally involved in a range of activities in our daily life, such as speech, swallowing, coughing, eating, drinking, moving the jaw. All of these activities are accomplished with natural movement of tongue. As the TMEP signal based system is intended to use tongue movements (e.g. Up, Down, Left and Right) as an action commands, these planned tongue movements are not natural, which allows them to be differentiated from the natural movements of tongue. However, during continuous recording of TMEP signals for identifying intended tongue movement actions, all of the natural tongue movements have potential to contaminate, and hence it can be detected as an action or command and lead an unwanted or false response for the device.,Such situations are not acceptable in any assistive devices for real life applications. So, all of these natural movements of tongue in the TMEP signal are characterised as internal interferences, and need to be identified and rejected towards the development of real-time TMEP signal based assistive HMI system such as a wheelchair to assist people with disability.

Previously no research work has addressed this significant challenge, therefore to improve the accuracy, reliability and robustness of a real-time assistive HMI system based on TMEP signals, this study aimed to identify intended actions related TMEP signals from a variety of adverse internal interferences as presented in this chapter. A new signal acquisition experiment was designed and acquired tongue movement actions and interferences related TMEP signals. The signal features were extracted using a WPT to capture the transient changes in the TMEP signals, and were optimally selected according to statistical distributions of the wavelet packet coefficients so as to maximise the separability between tongue movement actions and interferences. Two types of classifiers, Bayesian and support vector machine (SVM), were implemented to perform the classification between two classes of actions and interferences. Their performance was evaluated in both offline and online conditions using both subject specific and

generalised interference for training. This work has significantly improved the accuracy and robustness of both offline and real-time assistive human machine interface systems based on TMEP signals.

Section 5.2 illustrates the tongue movement actions and interferences related TMEP signal acquisition. Detailed analyses of signal processing and classification procedures are described in section 5.3. Results of tongue movement action identification from interferences in both offline and online are presented in section 5.4 and 5.5 respectively. Finally section 5.6 provides a discussion of the TMEP action identification method with stating potential limitations, and draws the conclusion of this chapter.

5.2 Experimental paradigm and signal acquisition

5.2.1 Participants

Ten healthy subjects (6 males, 4 females) ranging in age from 18- 45 years (30.7 ± 6.4; mean ± 1 SD) participated in the experiment. It is noted that within this subject group, five subjects (S6-S10) were well trained to perform the controlled TMEP actions whilst the remaining five subjects (S1-S5) had only half an hour practice directly prior to the data collection process. The experiment was approved by the ISVR Human Experimentation Safety and Ethics Committee of the University of Southampton. Participants gave their written informed consent before taking part in the study.

5.2.2 Experiment and signal recording

As discussed in chapter 4, tongue movements cause pressure changes within the ear canal, which can be detected by a microphone sensor. The microphone was inserted into the ear canal and connected to an amplifier. The pressure change was picked up by the microphone and digitised and stored in a computer similarly as previous study in (Vaidyanathan et al. 2007). The distinct movement related actions can be differentiated from signatures of the recorded ear pressure signals.

In this study the classification was performed between tongue movement actions and interference related TMEP signals. TMEP signals were recorded when subjects performed six types of tongue movement actions, four (Up, Down, Left and Right) of them already defined in the previous chapter (cf. section 4.2). The other two types of tongue movement actions are pushing and flicking, which are respectively defined as

moving the tongue to the outside of the oral cavity in a straight manner with closed lips (‘Pushing’) and flicking the tongue up and down once (‘Flicking’). TMEP signals during these six intended tongue movements were defined as controlled or intended tongue movement action related TMEP signals.

In contrast, non-controlled movement or interference related TMEP signals were collected while subjects were speaking, coughing, drinking or resting. The speech activity included utterances of words consisting of numbers from 0 to 9, and words ‘start’, ‘stop’, ‘open’, ‘close’, ‘on’ and ‘off’. The drinking activity was to drink 15 ml of water from a glass, whilst the resting activity was recorded during normal relaxation. This set of activities represents a wide range of tongue movement patterns.

Each subject was seated in a comfortable armchair with a recording microphone sensor inserted into the ear canal. A generic microphone earpiece sensor was used with a comfortable ear canal tip shown in figure 5.1. Prior to the experiment, the selection of ear (left or right) to insert the earpiece was made by the participants according to individual preference. The signals were recorded using custom made software written in Microsoft C# running on a laptop computer.

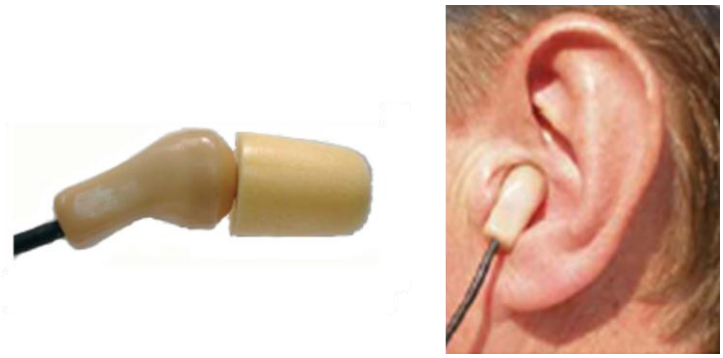


Figure 5.1: Generic microphone earpiece sensor with a comfortable ear canal tip – showing it both out of the ear (left) and inserted into the ear canal (right).

A visual cue was presented on a computer screen to instruct the subject to perform a specific tongue movement action. Subjects were instructed to move their tongue in the respective direction as much as possible, so as to perform each action correctly. The cues were represented by text as well as direction, via a moving circle on the screen. Before making each movement, the participant was instructed to always place the tongue back into its neutral position. Each action was randomly repeated every 5 seconds, to minimise the possible effects of fatigue or learning. Each controlled (six)

and non-controlled (nineteen) movement were repeated 100 and 20 times respectively. After each movement, the direction or type of movement was labelled by the subject and indexed in customised software for classification analysis. Signal acquisition was performed in a standard office/laboratory environment. Signals were sampled at 8 kHz and then digitally down-sampled to 2 kHz for further analysis. It is noted that previously TMEP signals were recorded with sampling frequency of 2 kHz, however in this experiment TMEP signals were acquired with higher sampling frequency (8 kHz) to capture high frequency content of the speech and other natural tongue activities. Generally speech recorded in the ear tends to have frequency content up to around 2.5 kHz (usually this would be higher at around 4 kHz but the frequency content is significantly dampened by the ear canal). Thus the higher sampling rate can prevent aliasing of the higher frequency information into the respective frequency band of TMEP signals. The detail experimental and data acquisition procedure are also available in Appendix B. A superimposed representation of all (six) controlled actions and three (Cough, Drink, Speech 'close') non-controlled movement or interferences related raw TMEP signal is shown in figure 5.2.

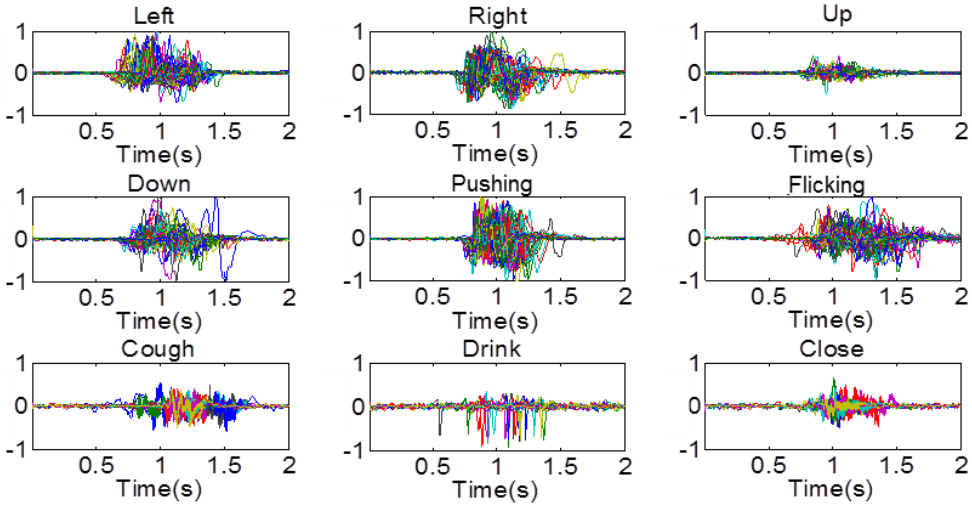


Figure 5.2: Raw TMEP signals for tongue movement actions (Left, Right, Up, Down, Pushing and Flicking) and interferences (Cough, Drink and Speech-'Close') for S5 (amplitude vs. time). The duration of each signal is 2 second and each activity repeated 100 and 20 times for movement actions and interferences activity respectively with sampling frequency 8 kHz.

5.3 Signal processing and classification

5.3.1 Support Vector Machine classification

The goal of the SVM classifier is to find the optimal separating hyperplane which is able to discriminate classes with the maximum possible margin (Gunn 1998; Duda et al. 2001). The separating hyperplane is a function, $f(x) = wx + b$, where the parameters w and b must be optimised during SVM training in order to maximise the boundaries among classes. The training procedure of SVM classification consists of solving a constrained quadratic optimisation problem. Given a training set of instance-label pairs $(x_i, y_i),\ x_i \in \mathbb{R}^n,\ y_i \in \{-1, +1\},\ i = 1, \ldots, N$, where x_i is the ith input vector with known binary target y_i, then the original SVM classifier satisfies the following conditions

$$w^{\mathrm{T}}\tau(x_i) + b \geq 1 \qquad \text{if } y_i = 1, \qquad (5.1)$$

$$w^{\mathrm{T}}\tau(x_i) + b \leq -1 \quad \text{if } y_i = -1, \qquad (5.2)$$

or equivalently

$$y_i[w^{\mathrm{T}}\tau(x_i) + b] - 1 \geq 0; \quad i = 1, \ldots, N, \quad (5.3)$$

where $\tau: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is the mapping function that maps the input space to a usually high-dimensional feature space, where the data points become linearly separable by a hyperplane defined by the pair $(w \in \mathbb{R}^m, b \in \mathbb{R})$ (Zheng et al. 2004; Fonseca et al. 2007). Then, the classification function is obtained as

$$f(x) = \text{sign}\{w^{\mathrm{T}}\tau(x_i) + b\} \qquad (5.4)$$

In order to allow for the violation in equations (5.3), slack variables ξ_i are introduced, such that (Zheng et al. 2004; Gunn 1998)

$$y_i[w^T\tau(x_i) + b] - 1 \geq 1 - \xi_i; \quad i = 1, \ldots, N, \quad \text{and } \xi_i > 0. \qquad (5.5)$$

To find a good linear separating hyperplane for classification, these slack variables, ξ_i must be minimised to obtain low errors in the training and for a better generalisation, the margin among classes must be maximised. So, combining these issues and considering the following minimisation problem (Zheng et al. 2004; Gunn 1998)

$$\min_{w,b,\xi} \quad J(w, b, \xi) = \frac{1}{2}(w^T w) + C\sum_{i=1}^{N}\xi_i \qquad (5.6)$$

$$\text{subject to } y_i[w^T\tau(x_i)+b]-1 \geq 1-\xi_i; \quad \xi_i \geq 0; \quad i=1,\ldots,N; \text{ and } C>0,$$

where C is a positive constant cost or regularisation parameter used to control the trade-off between the training error and the margin. By using Lagrange multiplier techniques, the minimisation problem of equations (5.6) leads to the following problem (Gunn 1998; Fonseca et al. 2007)

$$\max \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i\alpha_j y_i y_j K(x_i, x_j) \qquad (5.7)$$

$$\text{subject to } \sum_{i=1}^{N} \alpha_i y_i = 0; \quad 0 \leq \alpha_i \leq C; \text{ and } i=1,\ldots,N.$$

The function $K(x_i, x_j)$ is called a SVM kernel. As mentioned earlier, the kernel function is used to map the input space into a high-dimensional feature space, and that constructs an optimal separating hyper-plane in the feature space.

To obtain optimal performance of the SVM classifier, selection of a proper kernel function is essential. The optimal kernel function is dependent on the specific dataset. In general, when a dataset are linearly separable, a linear kernel is capable of providing better accuracy. However, biosignals more likely contain certain non-linear properties, which in turn may increase the difficulty to separate patterns linearly, so nonlinear kernels may work better. Common nonlinear kernels are radial basis function (RBF) and polynomial kernels. In the case of polynomial kernel, it requires heavier computation and more parameters to optimise. It may also not be suitable for situation with a limited training set and a large number of features . RBF have been considered a better option in the case of biosignal classification for following reasons, 1) it nonlinearly maps the data into a higher dimensional space unlike to linear kernel, 2) it has less hyperparameters than the polynomial kernel, 3) it is computationally less expensive and 4) it showed comparatively better performance than polynomial kernel in sevseral biosignals classification problems with limited dataset (Lotte et al. 2007; Fonseca et al. 2007). Therefore, according to nature of our dataset in this study, the RBF kernel was selected. The RBF kernel function is defined as

$$K(x_i, x_j) = \exp\left(-\eta\left\|x_i - x_j\right\|^2\right), \quad \eta > 0. \qquad (5.8)$$

The hyperparameters of a SVM classifier, i.e., the regularisation parameter C and the RBF kernel parameter η, were estimated during training to optimise the classification performance.

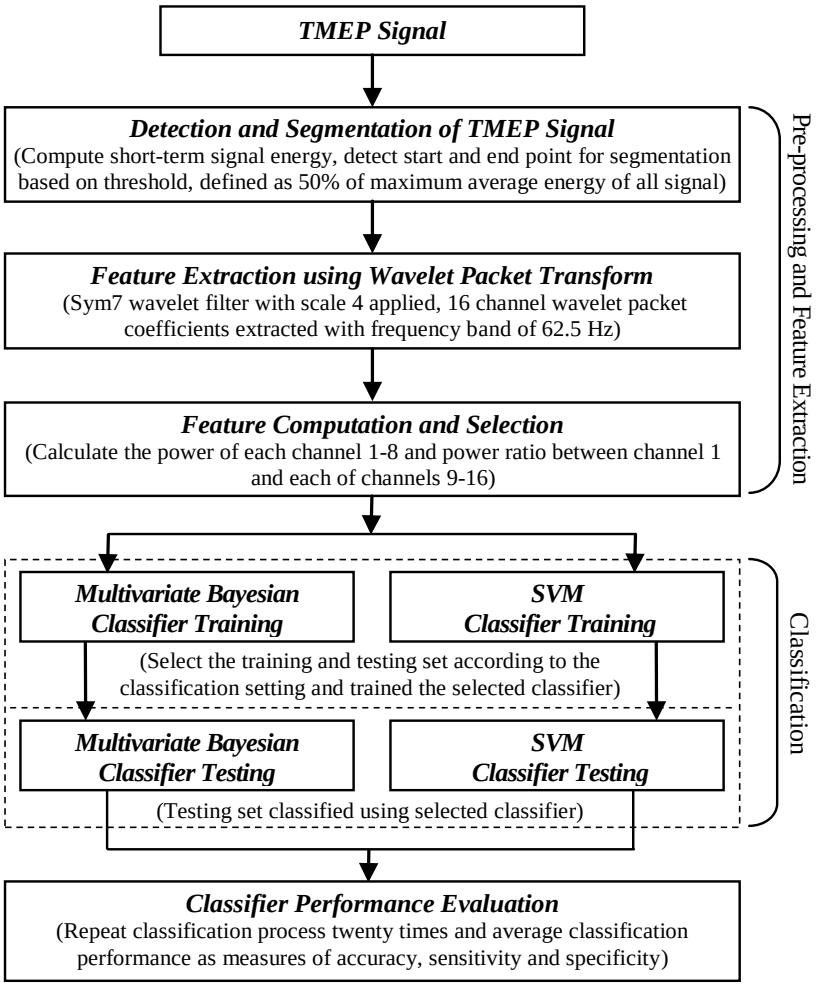


Figure 5.3: The detail flowchart for identification of tongue movement actions and interferences related TMEP signal.

5.3.2 Feature extraction and identification of action and interferences

The classification of tongue movement actions and interferences related TMEP signals consisted of signal activity detection, segmentation, feature extraction, feature selection and classification. The flowchart of these stages is shown in figure 5.3. The TMEP signals during tongue movement actions need to be detected and segmented appropriately. The detection method is similar to that in automatic speech recognition systems by setting a threshold on the short-term energy of the incoming signal. The threshold was determined for each subject as 50% of the maximum average peak energy across training TMEP signals during tongue movement commands calibration. The signal is then segmented to a section with 512 samples, which is slightly longer than the typical 0.2 second duration of the TMEP signal during a tongue movement action. Details of detection and segmentation methods are given in chapter 4 (cf. section 4.3.2).

The superimposed segmented TMEP traces of tongue movement actions and interferences have large variation in shape, duration and frequency as shown in figure 5.4.

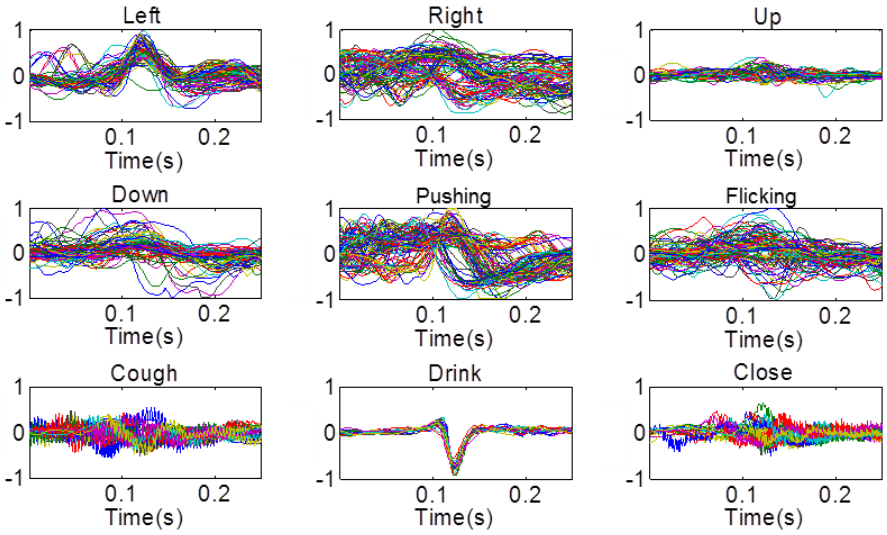


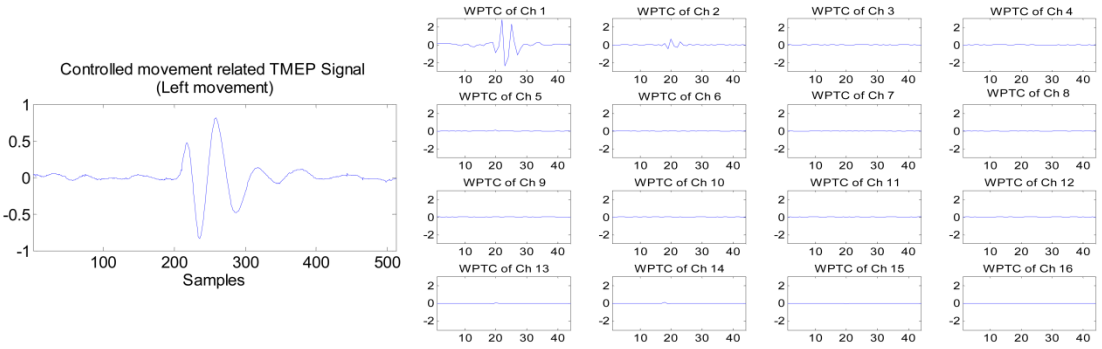
Figure 5.4: The superimposed segmented TMEP traces for tongue movement actions (Left, Right, Up, Down, Pushing and Flicking) and interferences (Cough, Drink and Speech-'Close') related signals for S5 (amplitude vs. time). The duration of each segment is 0.25 seconds and each activity repeated 100 and 20 times for movement actions and interferences respectively.

To extract features, segmented TMEP signals were decomposed into the WPT domain using a Sym7 wavelet filter at scale 4 (cf. section 4.3.3.3). The selection of the decomposition scale was made by comparing two other scales (3 and 5) based on their capability to localise discriminative information. After transformation, 16 channels of WPT coefficients at scale 4 were obtained, each with a frequency bandwidth of approximately 62.5 Hz (i.e. 0-62.5 Hz, 62.5-125 Hz and so on). An illustration of the WPT coefficients of a controlled movement or action (left) and two interferences (coughing and speech of 'close') related TMEP signal is presented in figure 5.5. It has been found that the tongue movement actions related TMEP signals have the majority of their energy located in the low frequency band (0-62.5 Hz) (cf. section 4.3.1), whereas the energy of interfering TMEP signals is distributed at low and/or high frequency. There are large variations of the energy distributions associated with interferences. Some only have signal energy at low frequency, such as drinking, and some have signal energy at low and high frequencies. Therefore all channels of the

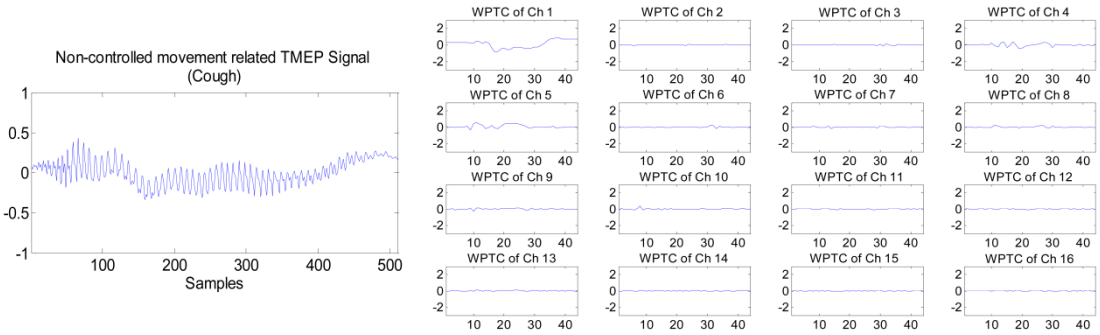
WPT coefficients were considered for feature definition. Sixteen discriminative features ($x_1,...,x_{16}$) were computed as the absolute power and low-to-high frequency power ratios. The power p_i of each WPT channel was calculated as the variance of the wavelet coefficients. The first eight features ($x_1 = p_1,...,x_8 = p_8$) were computed as the absolute power of the WPT channels 1-8 (frequency range 1-500 Hz). To isolate the discriminative information content between controlled tongue movement action and interfering TMEP signals by utilising the very low (0-62.5 Hz) and high (500 Hz or more) frequency WPT channels, the power ratios between channel 1 and each of channels 9-16 were computed as the remaining eight features ($x_9 = p_1 / p_9,...,x_{16} = p_1 / p_{16}$) for each signal. These features were determined based on maximised class separability to provide optimal classification performance after comparison with other channels and various combinations of their ratios and quantities. The distribution of the 16 features related to TMEP controlled actions and interferences for S5 is shown in figure 5.6. In this case, it was noted that the classes were almost linearly separable by features 5 to 16, while features 1 and 2 overlapped significantly, and features 3 and 4 were only partially overlapping. However feature 1 (channel 1) is important as it carries most of the power of the signal in both TMEP controlled actions and interferences as shown in figure 5.5 and 5.6(c).

Based on the extracted features, a multivariate Bayesian classifier (with Gaussian assumption) and an SVM classifier were designed to classify the controlled movement or action related signals from interferences. Each classifier was constructed in specific and generalised interference situations. In a specific interference situation, the classifier was trained and tested using each subject's specific TMEP actions and interferences related features. The training and testing datasets of the specific interference were selected randomly, with 60% from each type of signal assigned to train the classifier and the rest (40%) used to test the classifier. The training and testing data were mutually exclusive. In the generalised interference situation, the classifier was further extrapolated to be more robust to address a wide variety of interferences from all subjects. As the characteristics of the TMEP controlled actions are unique to each subject and the types of the interference are not limited to them, the classifier was constructed with subject specific TMEP controlled actions and generalised interferences related features. The training and testing datasets of the generalised interference were

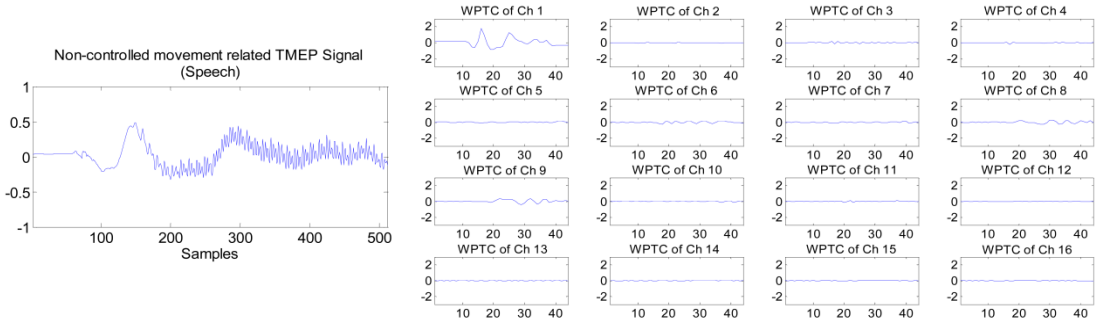
selected similarly to the specific interference for controlled actions, and using a leave-one-subject-out cross-validation procedure (i.e. data from one subject used for testing and the remaining subjects used to train the classifier) for interferences.



(a)

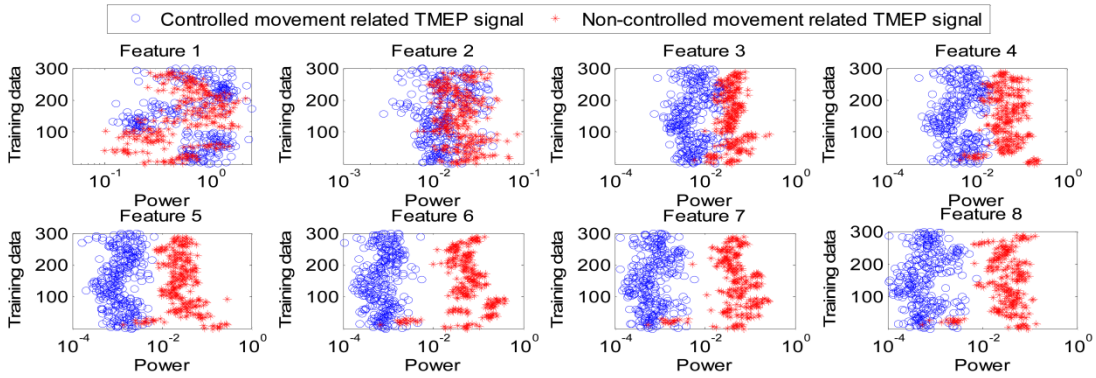


(b)

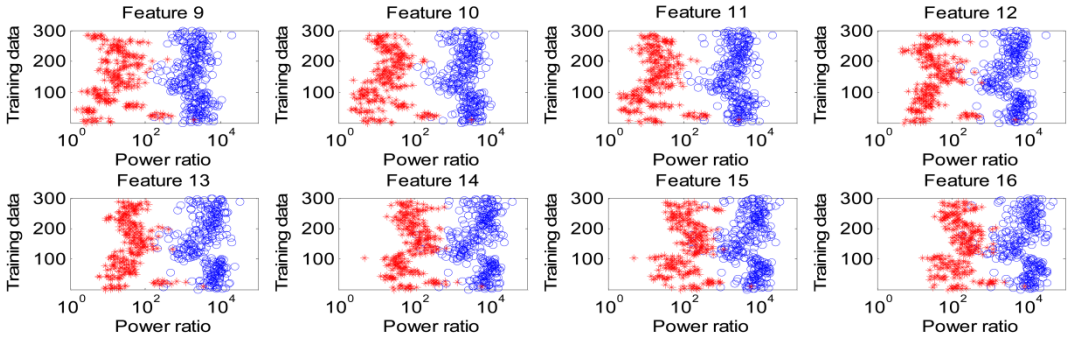


(c)

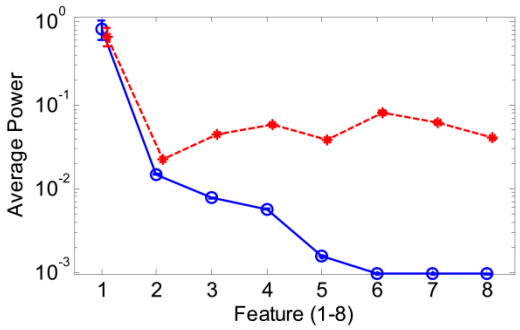
Figure 5.5: Tongue movement actions (a, moving left action) and interferences (b, cough and c, speech of 'close') related segmented TMEP signals respectively and their 16 channels wavelet packet transform coefficients at level 4 for S5.



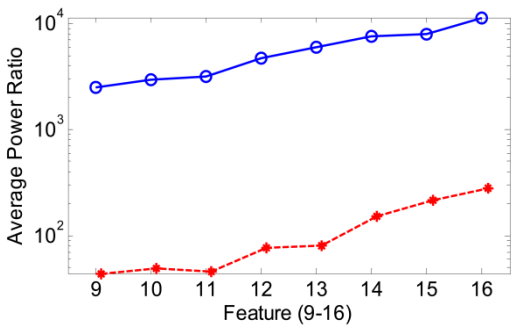
(a)



(b)



(c)



(d)

Figure 5.6: Feature distribution of controlled movements or actions (circle, blue) and non-controlled or interferences (star, red) related TMEP signals among all training dataset in S5. Feature 1-8 computed as power of each WPT channel 1-8 (a) and feature 9-16 computed as power ratio between the channel 1 and each of channel 9-16 (b). The average power for feature 1-8 (c) and average power ratio for feature 9-16 (d) of controlled movements (blue, circle) and interferences (red, star) related TMEP signal were computed and are presented on a logarithmic scale.

The multivariate discriminant functions were used to separate different classes in a Bayesian classifier (Mamun et al. 2010) (cf. section 4.3.4.3). Figure 5.7 shows the discriminant functions for classifying tongue movement actions related TMEP signals from the specific interferences situation. The discriminant function for the tongue movement action class was much higher than the interference class when the input was features of tongue movement action related TMEP signals in most cases. The opposite occurred when the input was features of interferences related TMEP signals. It indicates a separation boundary existing between the tongue movement actions and interferences related TMEP signals.

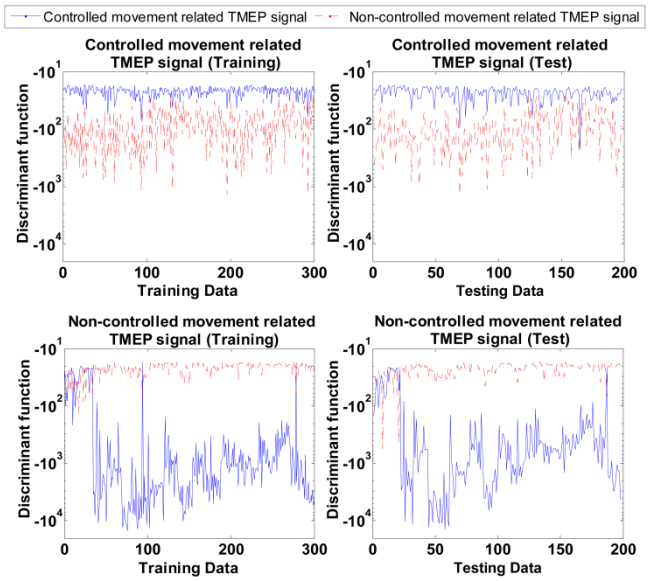


Figure 5.7: Discriminant functions of Bayesian classifier for the training and testing data to identify the state of controlled movements or actions (solid, blue) and interferences (dotted, red) related TMEP signals in S5.

The classification between the tongue movement actions and interferences related TMEP signals was further explored using an SVM classifier in both specific and generalised interference situations. The optimal selection of SVM parameters was performed through a 5-fold cross-validation procedure with the variation of kernel (η) and cost (C) parameters. A typical selection of optimal parameters for SVM classifier to identify the tongue movement actions and interferences related TMEP signals for S1-S5 shown in figure 5.8. In this way optimal parameters were obtained for each subject to optimise the SVM classifier. The SVM classifier was implemented using LIBSVM (Chang & Lin 2001). To statistically compare the performances among the

classification methods in specific and generalised interference situations, as well as trained and un-trained groups, a Student's *t*-test was performed using SPSS (Ver. 15, Chicago, Illinois).

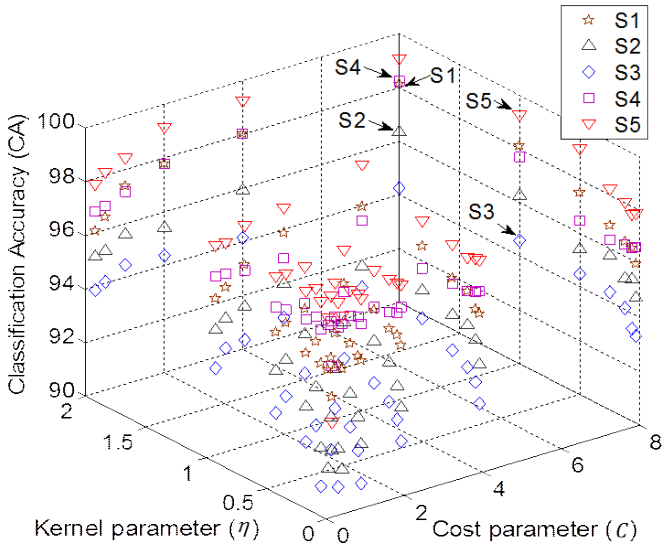


Figure 5.8: Typical selection of optimal parameters for SVM classification to identify the tongue movement actions and interference related TMEP signals for S1-S5.

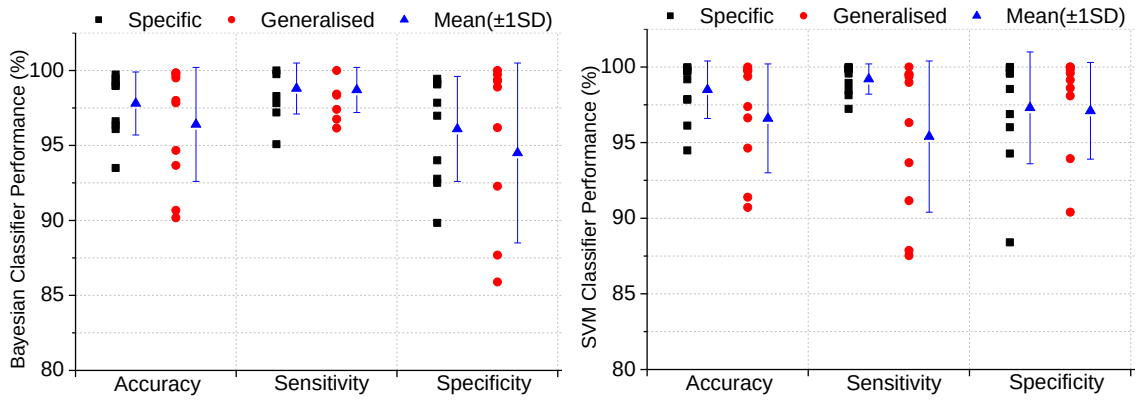
5.4 Results

The classification performance was evaluated with averaged accuracy, sensitivity and specificity, by repeating the classification process twenty times with random selection of the training and testing data. Accuracy is defined as the percentage of correctly classified instances. Sensitivity is defined as the ratio of the number of true positives classified to the number of actual total positive cases. Specificity is defined as the ratio of the number of true negative classified to the number of actual total negative cases. The Bayesian classifier and SVM classifier were evaluated on the ten subjects in both the specific and generalised situations.

The average classification accuracy of the multivariate Bayesian classifier was 97.8 ± 2.1% (mean ± 1 SD) across all 10 subjects when subject specific interferences were used (figure 5.9a). The sensitivity and specificity were 98.8 ± 1.7% and 96.1 ± 3.5%, respectively (figure 5.9a). The SVM classifier achieved slightly better performance based on the classification accuracy, sensitivity and specificity of 98.5 ± 1.9%, 99.2 ± 1.0% and 97.3 ± 3.7%, respectively (figure 5.9b). In the generalised interferences situation, the performance remained at a similar level. The accuracy, sensitivity and

specificity were 96.4 ± 3.8%, 98.7 ± 1.5% and 94.5 ± 6.0% for the Bayesian classifier (figure 5.9a), and 96.6 ± 3.6%, 95.4 ± 5.0% and 97.1 ± 3.2% for the SVM classifier (figure 5.9b), respectively. In the specific interference situation, the SVM classifier performed significantly better than Bayesian in terms of accuracy (98.5±1.9% vs. 97.8±2.1% (t(9)=-4.1, p<0.05)) and specificity (97.3±3.7% vs. 96.1±3.5% (t(9)=-2.3 p<0.05)) (figure 5.10a). In the generalised interference situation, the Bayesian classifier performed significantly better than the SVM in terms of sensitivity (98.7±1.5% vs. 95.4±5.0% (t(9)=3, p<0.05)) although the SVM performed better in terms of specificity (97.1±3.2% vs. 94.5±6.0% (t(9)=-2.8, p<0.05)) (figure 5.10b). Overall these two classifiers achieved similar levels of performance.

The effect of training was further investigated. Among 10 subjects, half (S1-S5) had a short practice before the experiment (un-trained group) and the other half (S6-S10) had intensive training to adequately make tongue movement actions (trained group). The trained group had significantly better performance than the un-trained group: accuracy 99.3±0.3% vs. 96.3±2.0% ($t(8)=-3.4$, $p<0.05$), sensitivity 100.0±0.0% vs. 97.6±1.7 ($t(8)=-3.1$, $p<0.05$) and specificity 98.8±1.01% vs. 93.4±2.9% ($t(8)=-3.9$, $p<0.05$) in the Bayesian classifier (figure 5.10c), and accuracy 99.8±0.2% vs. 97.1±1.8% ($t(8)=-3.3$, $p<0.05$), sensitivity 99.9±0.1% vs. 98.5±0.9 ($t(8)=-3.7$, $p<0.05$) and specificity 99.8±0.2% vs. 94.8±3.9% ($t(8)=-2.8$, $p<0.05$) in the SVM classifier (figure 5.10d).



(a) (b)

Figure 5.9: The average offline classification performance of accuracy, sensitivity and specificity for each subject and their average (mean±1SD) across all subjects based on specific and generalised interference situations with both Bayesian (a) and SVM (b) classifiers.

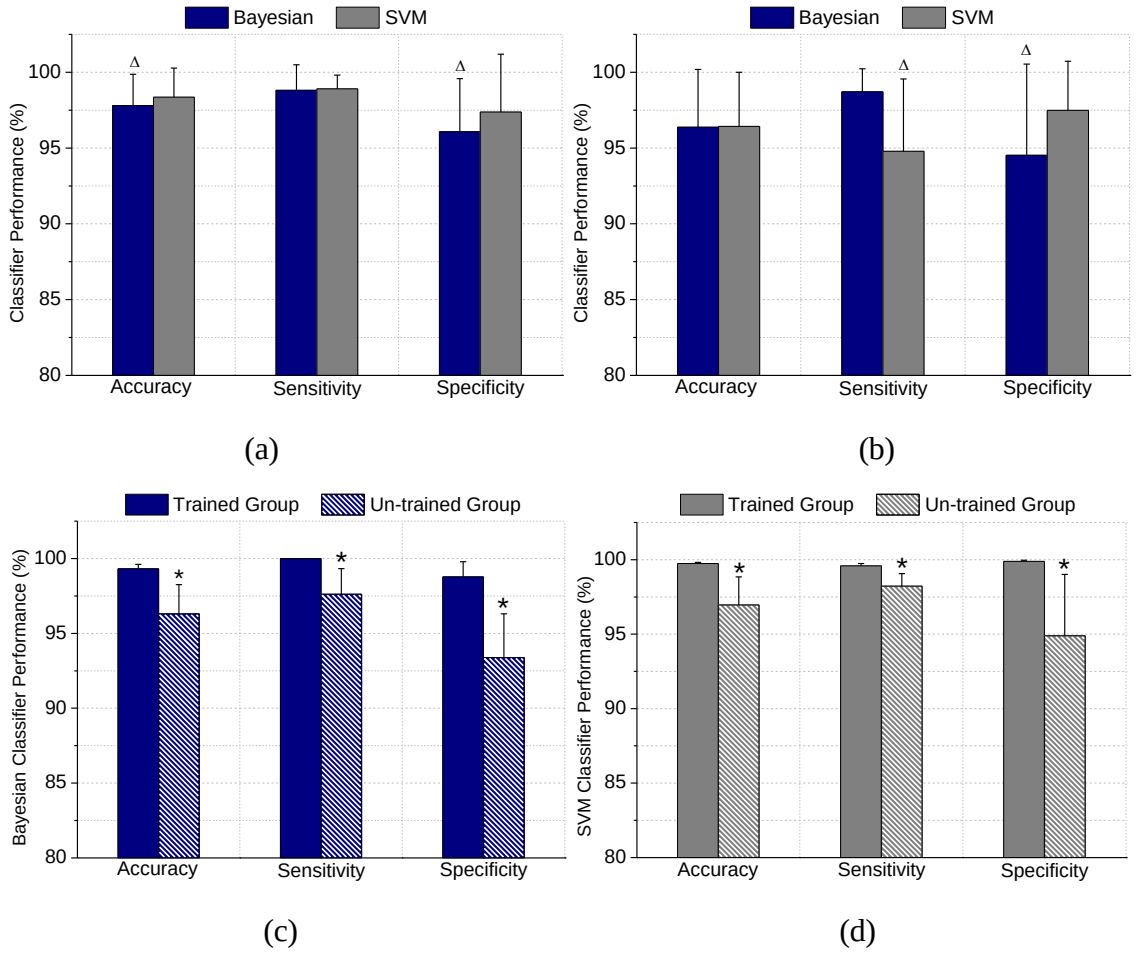


Figure 5.10: The average offline classification performance of accuracy, sensitivity and specificity across all subjects based on specific interferences (a) and generalised interference (b) situations. The training can significantly improve the performance of both Bayesian (c) and SVM classifiers (d). Δ p<0.05, Bayesian vs. SVM, and * p<0.05, trained vs. untrained group, mean ±1SD.

5.5 Real-time evaluation

In the above offline experiments, various interferences of speaking, coughing or drinking were investigated. In a real world implementation there will be a lot of other interferences, for instance, free speech, and the algorithm needs to be incorporated with the successive inter-action classification scheme (presented in chapter 4). Therefore the algorithm was further evaluated in a real-time environment. Both the Bayesian and SVM classifiers were trained with generalised interferences from three randomly selected subjects with a small training set and tested on an additional two subjects, in which one test subject had long training and the other only short training. The training set has only 120 trials of tongue movement actions (4 actions ‘up’, ‘down’, ‘right’ and

'left' with 30 trials each) and 162 trials of interferences (18 interferences activities with 3 trials each from 3 subjects). The test subjects performed tongue movement commands of 'up', 'down', 'right' and 'left' actions and interferences segmented during a 5-minutes newspaper reading, 1-minute conversation, swallowing, coughing, and drinking. The testing was carried out in a normal office environment and the signals were detected, segmented and classified in real-time. The real-time evaluation results for TMEP controlled actions and interferences classification are presented in figure 5.11. The results showed that average performance with the Bayesian classifier achieved 88.1% in accuracy, 95.0% in sensitivity and 85.9% in specificity, while the SVM achieved only 68.6% in accuracy, 97.5% in sensitivity and 59.4% in specificity. As with the offline result, the long training subject performed better than the short training subject in both classification methods. The SVM performed considerably worse than the Bayesian classifier. It may be due to the fact that the classifier parameters optimisation is only based on a small training set with large variability. The SVM is more sensitive to the size of training set than the Bayesian classifier. The small training size tends to cause the SVM classifier to over-fit to the training data and therefore have poor generalisation during testing (Gacquer et al. 2011; Lotte et al. 2007). After rejecting interferences, the movement commands were further identified and used to control a simulated wheelchair on a computer screen (Mace et al. 2010). The wheelchair was well controlled with only a few false actions. In contrast, the wheelchair went quickly out of control when no interference rejection procedure was utilised. A demonstration video of the system is available at http://www.swanglab.com/software.htm. These are the first real-time results for rejecting interferences and shown the potentiality of the system for assistive HMI applications.

5.6 Discussion and conclusions

Interference is one of the major challenges in developing human machine interfaces, including brain computer interfaces, due to its variety and uncertain sources (Bashashati et al. 2007; O'Malley 2007). In this study, robust identification of tongue movement commands from interferences was explored using Bayesian and SVM classifiers with features extracted by a WPT. The average classification accuracy for discriminating between the tongue movement actions and the interfering signals achieved 97.8% (Bayesian) and 98.5% (SVM). The classifiers were robust remaining at a similar

performance level when generalised interferences from all subjects were used. The robustness of the classification was also tested in a real-time environment. Both classifiers performed better offline, although the multivariate Bayesian classifier achieved higher accuracy than the SVM in the real-time system.

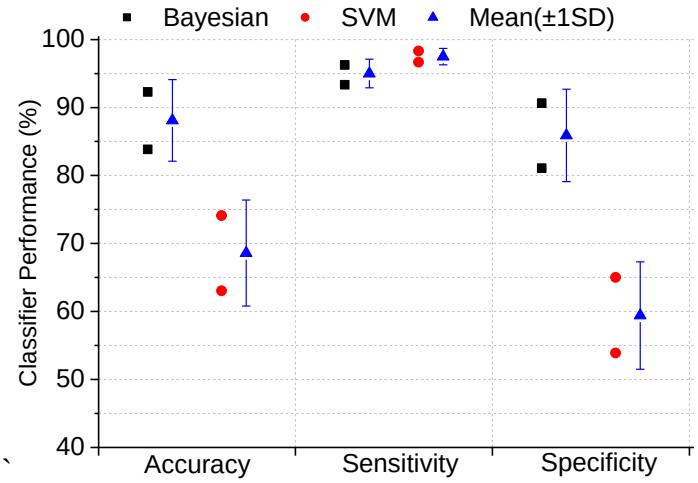


Figure 5.11: The average real-time classification performance of accuracy, sensitivity and specificity for each subject and their average (mean±1SD) in both Bayesian and SVM classifiers.

As discussed in the previous chapter, the wavelet packet transform was used to extract features for classifying tongue movement actions and achieved higher performance than time domain features in both clean and noisy environments. It is also noted that SWT based feature extraction achieved high performance for classifying tongue movement actions, and in some cases SWT features are distinguished slightly better than WPT features. However, the WPT based feature extraction approach is computationally more efficient than SWT (Addison et al. 2009; Walden 2001). This is due to the redundancy and high dimensionality of the SWT based feature. Taking care of such aspects, WPT was selected for extracting features from TMEP signals for identifying tongue movement actions and interferences. As before (cf. chapter 4), to achieve optimal performance of WPT, the decomposition scales were empirically determined by comparing other scales (3 and 5) with Sym7 wavelet filters.

Two classifiers were investigated in this study. The Bayesian classifier is simple and computationally efficient (Bashashati et al. 2007; Denison et al. 2002), while the SVM is more complex due to its optimisation characteristics (Gacquer et al. 2011; Lotte et al. 2007). The SVM performed slightly better than the Bayesian classifier during offline

classification of tongue movement actions and interferences related TMEP signals (figure 5.10a, 5.10b). Both classifiers were robust in both the subject specific and generalised interference situations. In the generalised interference situation, the classifier requires no previous information about interfering signals from the test subject. This implies that as it is trained with other subjects' signals, it does not require collecting a personalised training set for interferences. The performance in the generalised interference situation is slightly lower than in the specific interference case (figure 5.9a, 5.9b). Training for the subject associated with the execution of the controlled actions can significantly improve the classification performance (figure 5.10c, 5.10d). The classifiers were further implemented in a real-time system using generalised interference data for training. The Bayesian classifier performed much better than SVM in a real-time system environment. The performance was worse than that in an offline situation but the classifier made it possible to control a simulated wheelchair precisely in real-time. As a problem-specific application, it has been demonstrated that the Bayesian classifier with features extracted using the wavelet packet transform is suitable for a real-time system of identifying tongue movement actions from interferences, and potentially can be used within the command classification algorithm as well. The reduction of the performance in a real-time environment may be related to the wide feature variation of the TMEP signals. The implementation of a classification threshold with these results under the Bayesian classifier will reasonably improve the performance, specifically specificity, which is very important for safety critical applications.

The variance of features of interfering signals is higher than those of the controlled movement related signals. Such asymmetric distribution might be a contributing factor to the higher specificity error than the sensitivity error in the Bayesian classifier in both offline and online classification. Another possible cause of the high specificity error may be that some interfering signals have energy concentrated at very low frequencies, such as drinking, and they have comparable signatures to the actual controlled actions (figure 5.2, 5.4). Taking temporal patterns into account could potentially reduce the error and improve the overall performance.

In the current TMEP signals acquisition experiment a generic earpiece sensor of an audio microphone was used, which is capable of capturing frequency as low as 20Hz. As frequency of the controlled tongue activities are in a range of 0-100Hz, therefore

microphone with better frequency response will probably be able to capture more information of TMEP signal in lower frequency range and hence have more discriminative information to improve the classification performance. This issue will be addressed in future study. On the other hand, the placement of a generic microphone earpiece sensor varies between insertions and it may affect the waveform of the recorded TMEP signals. In this recording optimal position of the microphone earpiece was determined during the experimental setup for each subject to obtain the best quality signal. Therefore customised mould design for microphone housing may consider in the future experiment to reduce such variability. Finally it can be said that the Bayesian classifier with features extracted by the wavelet packet transform can reliably distinguish controlled tongue movement actions from the interference signals both offline and online. The rejection of various interfering signals has significantly improved the robustness of the assistive HMI based on TMEP signals and makes the real-time implementation and application in real living environments possible.

In summary, the notable technical contributions that were introduced in this chapter are as follows:

- A wavelet packet based feature extraction and selection approach developed to identify TMEP actions and interferences.
- The robustness of the identification method was evaluated with various types of interference signals in subject specific and generalised interference setting, and achieved high accuracy.
- Feasibility of robust identification of tongue movement commands from interferences was also evaluated in a real-time setting while considering a wide range of potentially interfering factors, for example free speech and the method was still able to maintain a good performance level.
- Typical results from ten subjects in offline and two subjects in real-time have proved the success of the method.
- The effect of training to perform the task was also investigated as a way of improving classification performance.

Chapter 6 : *Decoding Movements from Deep Brain Local Field Potentials*

6.1 Introduction

The trend of current neural interface systems, such as Brain Machine Interfaces (BMI) alongside neuroprosthetics and neuro-feedback, seek to establish bi-directional communication with the brain, for instance, recovering motor functions by externally controlling devices and directly stimulating the brain (Lebedev & Nicolelis 2006; Pasqualotto et al. 2011; Mason et al. 2007). This will not only provide tools for assisting paralysed individuals through bypassing the damaged brain region, but also assist in building bi-directional brain machine communication for instance, brain-machine-brain interfaces (BMBI) for treatment of diseases or investigation of neural circuit mechanisms. The key process of this communication interface is to decode information from neural signals and encode information as electrical impulses to stimulate specific parts of the brain or to suppress abnormal neural activity.

The pioneering work of Hans Berger (Berger, 1929) on the studies of the electroencephalography (EEG), has been fundamental to our understanding the function of the human brain (Coyle 2006). Until now there have been extensive studies performed to understand and uncover the underlying mechanisms of human brain circuits by recording neural activities, both invasively and non-invasively. As discussed in chapter 2, the BMI system got substantial attention in the last decade for decoding neural activity to control an external device for neuro-motor applications such as wheelchairs or robotic arms, to restore motor functions (Huang et al. 2009). The major component of such a BMI system is the neural decoding algorithm, which extracts the activity information from the given neural signals enabling correct operation of the devices. A commonly utilised neural signal is EEG, a non-invasive recording approach of brain activity. Due to the high level of background activity, EEG signals are highly prone to external interferences as well as it has bandwidth limitations. As such,

researchers have begun to make efforts towards analysing invasive recording of subdural electrocorticograms (ECoG) or intracortical recordings (Lal et al. 2005). These neural signal recordings provide a higher quality (higher SNR and bit rate) of information. However, it carries associated risks due to the invasive surgical procedure and thus most of the research is performed on animals rather than human subjects. In some cases it is performed on patients with different neurological diseases. On the other hand, one of the major limitations is the lack of long-term stability in recordings of neural activity that acts against the widespread use of BMI technology in human subjects (Ince et al. 2010). Deep recordings of local field potential (LFP) activity could provide an alternative solution to address these issues. Local field potentials (LFPs) are predominantly generated by excitatory synaptic potentials in the vicinity of the electrode trip that represent the sum of the synaptic activity of many neurons (Andersen et al. 2004). Accordingly these signals are assumed to have greater stability in time compared to single neuron recording, making them potentially more suitable for BMI applications (Ince et al. 2010). However, the precise relation between LFPs and the brain areas involved in motor operations has not been widely explored. Under the perspective of the BMI system, a number of publications have shown that it is possible to decode movement information from LFPs in the premotor, motor and posterior parietal cortex (Scherberger et al. 2005; Rickert et al. 2005; Andersen et al. 2004; Mehring et al. 2003; Ince et al. 2010). In addition, it was established that movement target direction related information could be effectively decoded from LFPs of the primary motor and the dorsal premotor cortex (Ince et al. 2010).

On the other hand, the target brain areas for BMI to date have extensively focused on the primary motor cortex as well as premotor and parietal cortices (Patil & Turner 2008; Bashashati et al. 2007; Ince et al. 2010). It has been found that movement related amplitude decreases (de-synchronisation) and increases (synchronisation) are available in the primary motor cortex and supplementary motor area before and during voluntary movements (Alegre et al. 2005; Huang et al. 2009; Scherer et al. 2009). The primary motor cortex is the most important brain region that controls voluntary movements. The majority of motor activity decoding studies have therefore been performed in this area. However, there are many practical reasons indicating that it would be unwise to focus exclusively on this region. The area of motor cortex that is readily accessible invasively or non-invasively on the cortical network is relatively small and there are a variety of

neurological diseases in which this area may be damaged, making it unusable (Ince et al. 2010). Therefore, investigation of the neural signals from deeper areas of the brain which are heavily involved in the motor control circuit, for instance the basal ganglia – thalamus – cortical network, could provide substantial or additional information. It is also noted that motor circuits were extensively investigated for patients with movement disorders and reported that the brain area or subcircuit such as part of the basal ganglia, subthalamic nucleus (STN) or the globus pallidus interna (GPi) are involved in motor operation (Delong & Wichmann 2007). DeLong and Winchmann (2007) presented the intrinsic circuit anatomy of the motor circuit and also stated its important role in movement operation as (Delong & Wichmann 2007)

> *"The most researched cortico-subcortical circuit is the "motor circuit" because of its importance for movement disorders. The motor circuit is composed of several subcircuits that originate from the motor cortex and several premotor areas. In a general sense, tonic output from this circuit, arising in motorportions of the GPi, internal segment of the globus pallidus and SNr, substantia nigra pars recticulata may regulate the overall amount of movement."*

Another recent study also reported the anatomical and functional connectivity of human basal ganglia, thalamus and cortex in the motor circuit and identified additional pathways involved in motor operations (Lenglet et al. 2012). It is also evident that the cortex and the basal ganglia are involved in the decision making or action selection as well as the preparation, execution and imagining of movements (Bogacz & Gurney 2007). Alegre et al. (2005) also reported that the motor cortex and both STN generates parallel changes in oscillatory activity for self–initiated movements (Alegre et al. 2005). Therefore investigation of basal ganglia neural activity for motor control may provide an alternative or supportive source, and also will advance our knowledge to develop more reliable movement decoding strategies for real life BMI applications (Kringelbach et al. 2007). It may also help us to extract more complex motor tasks or to generate complex motor commands.

The development of functional neurosurgery has provided an opportunity to record neural activity directly from the human basal ganglia. The opportunity usually arises in patients with movement or motor circuit disorders, for instance, Parkinson's disease (PD) or dystonia undergoing implantation of deep brain stimulation (DBS) into basal

ganglia to suppress or control disease symptoms (Brown & Williams 2005; Kringelbach et al. 2007). The most common target for DBS implantation is the basal ganglia nuclei, subthalamic nucleus (STN) or globus pallidus interna (GPi). The DBS microelectrodes provide an opportunity to record basal ganglia oscillatory neuronal activity as form of local field potentials. The function of the basal ganglia in motor control has been extensively studied with recording of LFPs from the DBS microelectrodes in patients with Parkinson, dystonia, bradykinesia (Alegre et al. 2005; Engel et al. 2005; Kringelbach et al. 2007). The recording happens in the few days after implantation, while the macroelectrode leads are externalised prior to connection to the neurostimulator. Post-operative recordings from the implanted DBS macroelectrodes have the advantage that it allows patients to be involved in lengthy movement paradigms (Loukas & Brown 2004).

As mentioned in chapter 2 (cf. section 2.3.5), the inherent LFP activity in STN or GPi may be broadly subdivided into three frequency bands, <8, 8–30, and >60 Hz, however, these frequency bands are likely to change due to the behavioural and disease correlation of different activities. To date, such investigations of LFP activity in the region of the STN and GPi have demonstrated prominent oscillations in the 8–30 Hz band (Brown & Williams 2005). This oscillatory activity not only shows functional connectivity with similar cortical oscillations but also shows movement-related frequency dependent de-synchronisation and synchronisation in the STN and/or GPi during externally cued and self-paced movements (clicking or continuous voluntary movements) (Brown & Williams 2005; Loukas & Brown 2004), suggesting that oscillation may be involved in the preparation of the motor response. During and after externally cued and self-initiated movements the beta frequency band (12–30 Hz) de-synchronisation dominates in the STN and/or GPi (Brown & Williams 2005; Alegre et al. 2005). This observation suggests that there is an inverse relationship between beta band synchronisation and motor action. It is also useful to note that oscillatory beta activity in the primary motor cortex behaves in a similar manner to that in the STN or GPi with respect to movements (Brown & Williams 2005; Kilner et al. 2003). Both contra- and ipsi- lateral gamma band synchronisation were also found in STN LFPs during wrist extensions (Alegre et al. 2005). Basal ganglia STN activity can be modulated by imaginary movements as well, in which subjects imagine performing a specified action or watch visual images of movements (Kuehn et al. 2006). Such

imagined movements lead to event related synchronisation and de-synchronisation which are similar in frequency and time course to that during actual voluntary movements.

Basal ganglia STN or GPi LFPs are characterised by multiple frequency dependent oscillations, which are related to motor preparation, execution, imaging and decision (Kuehn et al. 2006; Loukas & Brown 2004; Alegre et al. 2005). By incorporating these distinct oscillations as features, the onset of voluntary hand-movement prediction was investigated using a neural network approach (Loukas & Brown 2004). Important features of the STN LFPs have been evaluated using three different spectral measures, the fast Fourier transform, continuous wavelet transform and the statistical properties of wavelet spectra. The wavelet transform features coupled with a neural network, optimised using linear vector quantisation, achieved high accuracy (95% sensitivity and 77% specificity) for online movement (rest or action) prediction prior to the onset of a forthcoming movement. Human limb movements are controlled by the contralateral cerebral hemisphere (Huang et al. 2009). Kilner et al. (2003) demonstrated that significant coherence only exists between the sensorimotor cortex and contralateral hand and forearm muscles, and no significant coherence was observed between the sensorimotor cortices and any ipsilateral hand and forearm muscles (Kilner et al. 2003). Also many other neurophysiological and neuroimaging studies have explored the nature of contralateral movement control (Bai et al. 2005; Huang et al. 2009). Left and right hand movements are associated with different spatiotemporal patterns of movement related de-synchronisation and synchronisation (Bai et al. 2005). Therefore, reliably decoding movements from basal ganglia oscillatory neural activities for left and right hands will provide additional information for motor control and bilateral co-ordination (Scherer et al. 2009). The basal ganglia movement onset information, specifically from the STN or GPi incorporated with the motor cortex could potentially enhance movement decoding performance towards the development of robust BMI applications. Also the ability to predict and classify movement and its laterality in real-time may open up several experimental and therapeutic possibilities for neural interface systems, as well as for treatment of diseases or investigation of neural circuit mechanisms. Early prediction and classification of the onset of abnormal neurological events, for instance, tremor in movement disorders, will provide the possibility of adaptive feedback for

therapeutic interventions to optimise the neuromodulation effects (Kringelbach et al. 2007, Rosin et al. 2011).

This study hypothesised that neural information from the basal ganglia can be used to decode movements. Therefore it examined techniques for decoding basal ganglia LFPs related to movement and their laterality, based on left or right sided visually cued movements. Basal ganglia LFPs were recorded from the subthalamic nucleus (STN) or globus pallidus interna (GPi) through deep brain stimulation electrodes implanted in patients with Parkinson's or dystonia. Based on the recorded neural activities, the study utilised signal processing and classification methods for decoding movement information as well as developing new approaches to further improve decoding performance, and presented in this chapter. The frequency dependent components of LFPs were extracted using the wavelet packet transform. In each frequency band, signal features were extracted based on instantaneous power and neural synchronisation. The instantaneous power features of each band were computed using the Hilbert transform with defined windows during a motor response. The neural synchronisation features for movement activity were computed by analysing Granger causality between contralateral and ipsilateral LFPs in each frequency band.

To overcome the challenge of large variability of features due to limited data available, a new feature selection strategy, weighted sequential feature selection (WSFS) has been developed. It allows efficient selection of the optimal subset of features from the available extracted features for decoding. Feature selection and classification of LFPs were evaluated using Bayesian and support vector machine classifiers to sequentially recognise the occurrence of movement and whether the forthcoming movement was performed by the left or right hand. One of the innovative properties of this study is the decoding of movement laterality from basal ganglia LFPs by incorporating features based on Granger causality. The movement decoding approach developed in this study, using the WSFS strategy for selecting consistent features independent of the training set is novel and efficient. To our knowledge, this is the first study that focused on such decoding techniques. The findings of this study will enhance our understanding on the underlying processes in STN or GPi for voluntary movement control and its important implications, which may be used to develop more advanced neural interface systems, BMIs for assisting people in neuro-rehabilitation as well as for identifying better treatment of diseases.

In section 6.2, the experimental paradigm and neural signal acquisition procedure from the basal ganglia using DBS electrodes is described. The methods used for signal feature extraction and classification to decode basal ganglia neural activities are illustrated in section 6.3. Section 6.3.4 particularly provides algorithmic design and analysis procedure of the decoding of deep brain LFP activities related to movements. The detailed results of the decoding methods are presented in section 6.4. Section 6.5 provides a discussion of the decoding system in design, drawbacks and its broader applications and also the contribution of this chapter.

6.2 Experimental paradigm and signal acquisition

6.2.1 Patients

Twelve patients (age 49.6±13.9 years) with Parkinson's or dystonia disease (disease duration 14.8±10.3 years) who were selected for bilateral implantation of deep-brain stimulation electrodes in the STN or GPi participated in the study. Their clinical details are summarised in Table 6.1. All participants took part with informed consent and the experimental approval was obtained from the local research ethics committee.

Table 6.1. Summary of subject clinical and recording details

Subject	Age (year) & sex	Disease duration (years)	Diagnosis	Electrode Location	Electrode Pair Used for Analysis
1	58F	10	PD	STN	L23/R12
2	53F	3	PD	STN	L12/R12
3	59M	7	PD	STN	L01/R01
4	60M	13	PD	STN	L12/R01
5	72F	21	PD	GPi	L01/R01
6	55M	10	PD	STN	L12/R01
7	36M	14	Dystonia	GPi	L12/R12
8	53M	5	Dystonia	GPi	L01/R01
9	23M	7	Dystonia	GPi	L12/R01
10	54F	38	Dystonia	GPi	L01/R01
11	40M	25	Dystonia	GPi	L01/R01
12	32F	24	Dystonia	GPi	L12/R23

6.2.2 DBS electrode implantation

The DBS electrodes (Model 3387, Medtronic Neurological Division, Minnesota, USA) were bilaterally implanted in the STN or GPi for the treatment of the disease (Parkinson's or dystonia). The electrode has four platinum–iridium cylindrical contacts (1.27 mm diameter and 1.5 mm length) and a contact-to-contact separation of 1.5 mm. The surgical procedure has been described in detail in (Liu et al. 2002; Liu et al. 2008). Patients underwent bilateral implantation of DBS electrodes in the STN or GPi, their electrode positions were chosen where a decrease in disease symptoms occurred during intra-operative electrical stimulation of the site and confirmed by examining the post-operative MRI scan or the fused images of pre-implantation MRI with post-implantation CT. All the patients included in this study clearly had at least one pair of electrode contacts in the STN or GPi or very near to the STN or GPi. The contact pair (composed of contacts 0-1, 1-2, or 2-3) within the STN or GPi chosen for analysis demonstrated the greatest percentage of β (12-30 Hz) modulation surrounding the movement relative to the amplitude of β modulation during the baseline activity period occurring 1-2 seconds before the registration of the motor response (Liu et al. 2008). LFPs were recorded via externalised leads from the STN or GPi during the week immediately post-implantation before the pulse generator was implanted.

6.2.3 Movement paradigm

LFPs from STN or GPi and surface EMGs were recorded together during two finger-pressing tasks which were carried out in a random order with short resting periods between tasks. Patients were seated approximately 60 cm from a computer screen. Prior to each motor task, they were instructed to place their left and right index fingers on distinct keys on the left or right hand side of a standard keyboard, respectively, and to look at a 10 mm cross that was continuously displayed in the centre of the screen. A visual cue (the letter A, 8 mm in height and 7 mm in width) appeared on the screen for a period of 400 ms immediately to the left or right of the central cross, thus indicating which finger to move. The laterality and interval of cues were randomised. Subjects were instructed to press a key ipsilateral to the cue with the corresponding index finger as quickly as possible.

6.2.4 Recording

The recording of STN or GPi LFPs were made from the electrode leads which were externalised for 4-6 days post-operatively after the patients had been off medication overnight. The LFPs were recorded with a bipolar configuration from the three adjacent pairs of 4 contacts (contact pair 0-1, 1-2, or 2-3). Signal segments containing premature, absent, or erroneous responses to visual cues were excluded. In addition to noting the time of the key press as registration of the motor response, the onset of the motor response and other voluntary or involuntary movements were monitored using surface EMGs recorded from the index finger. Movements that were executed <1s before or >5s after the previous movement were excluded to ensure a limited range of inter-movement intervals and to specifically avoid rapid repetitive movements. Signals were amplified using isolated CED 1902 amplifiers (×10,000 for LFPs and x1000 for EMGs), filtered at 0.5-500 Hz and digitised using CED 1401 mark II at a rate of 2000 Hz, displayed online and saved onto a hard disk using a custom written program in Spike 2 (Cambridge Electronic Design, CED, Cambridge, UK).

6.3 Signal processing and classification

6.3.1 Feature extraction

6.3.1.1 Wavelet packet transform

As discussed in chapter 4 and 5, the wavelet packet transform (WPT) is an efficient tool for extracting information from the given signal and showed promising results for TMEP signal classification. One of the main advantages of the WPT is the ability to give improved time-frequency resolution over traditional methods, such as the short-term Fourier transform (Addison et al. 2009). In particular to analyse the neural signals, WPT is better than the standard discrete wavelet transform due to its ability to localise any specific frequency information and it also provides an adequate flexibility for partitioning the signal into functionally distinct scales (Samar et al. 1999). It is often desirable to decompose a neural signal into a finer set of band-pass signals than is given by a standard discrete wavelet transform. The WPT decomposes a signal, level by level, from the time domain to a time-scale domain with sub-bands and presents both approximation and detail spaces in a binary tree. Signals at each scale are reconstructed

from the corresponding wavelet packet coefficients to provide multi-resolution analysis (Walden 2001; Samar et al. 1999).

It was established in the literature and also in the TMEP signals classification that the efficacy of the wavelet packet transform depends on the wavelet basis selection. Usually the wavelet basis selection was performed empirically which provided optimal matches between the waveform of the signal of interest and the properties of the wavelet (cf.4.3.3.3). In this study, the WPT was computed using the discrete Meyer wavelet (figure 6.1) because it broadly matches the oscillatory characteristics of LFP activities, it is orthogonal allowing efficient localisation of scale and temporal properties of the LFPs, and also its spectrum matches the spectrum of any band-limited signal as closely as possible in a least squares sense (Samar et al. 1999). It has been demonstrated that the discrete Mayer wavelet is capable of isolating frequency components of LFPs and EEG event related potentials (ERP) (Samar et al. 1999; Loukas & Brown 2004). Also STN LFP activities of Parkinson's disease have been analysed using the continuous wavelet transform with a discrete Meyer wavelet filter and reported its effectiveness (Williams et al. 2005; Loukas & Brown 2004).

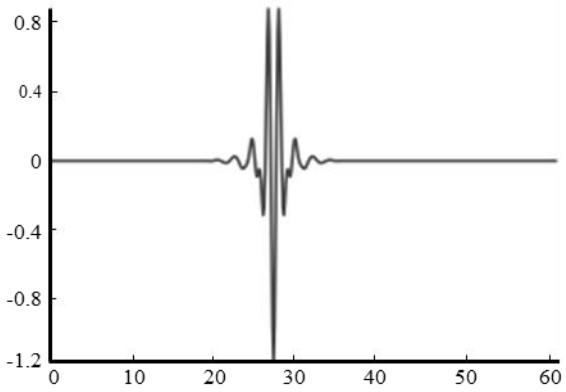


Figure 6.1: The discrete Meyer wavelet filter.

6.3.1.2 Hilbert transform

The envelope of the LFPs in each frequency band was computed by applying the Hilbert transform (HT) to the WPT reconstructed signals at each scale (Marple 1999). The Hilbert transform is a useful tool for the analysis of the oscillatory components of time-varying signals. It is used to form a complex analytic signal composed of the real narrow band time-series and the imaginary Hilbert transform of that time-series. The magnitude of the analytic signal represents the amplitude envelope of the time-series.

Theoretically this procedure creates less distortion in the estimation of the envelope than using a half-wave or full-wave rectification of the signal (Ince et al. 2010). The Hilbert transform algorithm can be summarised as: Let $x(n)$ be the reconstructed signal in one scale and $X(k)$ be the Fourier transform of $x(n)$, where n, k=1, 2, …., N. Then, $Y(k)$ is obtained from $X(k)$ by multiplying $X(k)$ with two for k=2, 3, …., N/2 and setting $X(k)$ as zero for k=N/2+2, N/2+3, …., N. The analytic signal $x_a(n)$ of $x(n)$ is computed by performing the inverse Fourier transform of $Y(k)$. $x_a(n)$ is a complex signal and can be expressed as $x_a(n)=A(n)exp(i\phi(n))$, where $A(n)$ and $\phi(n)$ are the instantaneous amplitudes (envelope) and phases of $x(n)$ respectively.

6.3.1.3 Neural synchronisation

Research in neuroscience fundamentally assumes that brain networks determine function, and it is deeply rooted in anatomical connectivity. A brain area's structural connectivity network can predict the functional response in the brain. Saygin et al. (2011) reported that anatomical connectivity patterns can predict face selectivity in the fusiform gyrus (Saygin et al. 2011). Based on the Saygin et al. (2011) work, Jbabdi and Behrens (2012) recently stated that (Jbabdi & Behrens 2012),

> *"Brain regions exhibit specialization for different functions, but such functions are constrained by anatomical connections to other brain regions. By measuring these connections, we can predict complex functional responses before the subject has even performed the task."*

These findings prove that coupling exists between structural and functional organisation in the brain. Therefore, analysis of functional coupling in oscillatory neural activity at different levels of the motor system or different functionally connected areas of the brain has led not only to a better understanding of the inherent mechanisms of movement control but also to get additional discriminative information for decoding neural activity (Wang et al. 2007). One widely used method of estimating the functional coupling between two oscillatory signals is the ordinary coherence also called magnitude-squared coherence (MSC). The MSC is a normalised cross-spectral density function that measures the strength of association and relative linearity between two stationary processes on a scale from zero to one. The coherence value indicates the strength of the coupling in the frequency domain between two signals. The conditional coupling among multiple signals may be further measured by partial coherence as

mentioned in chapter 3 (cf. section 3.3). However, these techniques based on correlation or coherence are not sufficient to describe interdependence among signals. Thus, they do not help to elucidate functional coupling or causal relationships within the system. Therefore, to fully understand information processing from the oscillatory neural activity at different levels of the motor system, directional interaction analysis to reveal causal influence or synchronisation between neural signals is essential to uncover more specific information underlying the motor activity for decoding.

The causal relations were described initially as probabilistic concepts, which is that one variable may be caused by the other if it can be better predicted by incorporating knowledge of the second one. Granger formulated the concept in terms of predictability based on the linear regression models of stochastic processes (Granger 1969). This causality was expressed as one time series is caused by the other one if its prediction error at the present time can be reduced by including the past of the second one in the model. Nowadays, Granger causality is widely used in neuroscience for analysing and identifying directional influences or synchronisations between different brain areas and neural activities (Kaminski & Liang 2005; Silchenko et al. 2010). For instance, Granger causality analysis was performed for oscillatory local field potential activity in the beta (14–30 Hz) frequency range among sensorimotor cortical recording sites during a GO/ NO–GO visual pattern discrimination task in monkeys. It was also analysed to reveal the connectively between different parts of the sensorimotor cortical network (Brovelli et al. 2004) as well as interdependence between neural and muscular activities (Wang et al. 2007).

6.3.1.3.1 Granger causality

Let $X_1(t)$ and $X_2(t)$ denote the time series from two data channels. According to Granger causality, X_2 (or X_1) causes X_1(or X_2) if the inclusion of past observations of X_2 reduces the prediction error of X_1 in a linear regression model of X_1 and X_2, as compared to a model which includes only previous observations of X_1. To illustrate the Granger causality, the temporal dynamics of $X_1(t)$ and $X_2(t)$ with length T can be described by using an autoregressive model as:

$$X_1(t) = \sum_{j=1}^{p} A_{11}(j) X_1(t-j) + e_1(t) \qquad (6.1)$$

$$X_2(t) = \sum_{j=1}^{p} A_{22}(j) X_2(t-j) + e_2(t) \qquad (6.2).$$

Similarly, incorporating both $X_1(t)$ and $\mathrm{X_2(t)}$ together in a bivariate autoregressive model as:

$$X_1(t) = \sum_{j=1}^{p} A_{11}(j)X_1(t-j) + \sum_{j=1}^{p} A_{12}(j)X_2(t-j) + e_3(t) \qquad (6.3)$$

$$X_2(t) = \sum_{j=1}^{p} A_{21}(j)X_1(t-j) + \sum_{j=1}^{p} A_{22}(j)X_2(t-j) + e_4(t) \qquad (6.4)$$

where, p is the maximum number of lagged observations included in the model (the model order, $p < T$), A is the coefficient of the model, and $e_i(t)$, $i = 1,2,3,4$ are the prediction errors with variance $var(e_i)$ for each of the time series. If the variance of prediction error $e_1(t)$ (or $e_2(t)$) is reduced ($var(e_3) < var(e_1)$ (or $var(e_4) < var(e_2)$)) by the inclusion of the X_2 (or X_1) terms in the equation 6.1 (or 6.2) as in 6.3 (or 6.4), then it is said that X_2 (or X_1) Granger causes X_1(or X_2). Assuming that X_1 and X_2 are covariance stationary (i.e. unchanging mean and variance), the magnitude of this interaction can be measured by the log ratio of the variance of prediction errors and it can be quantified as:

$$\mathrm{F_{X_2 \to X_1}} = \ln(var(e_1)/var(e_3)) \qquad (6.5)$$

If $\mathrm{F_{X_2 \to X_1}} = 0$, there is no causal influence from X_2 to X_1 and if $\mathrm{F_{X_2 \to X_1}} > 0$, there is causal influence from X_2 to X_1. Similarly, causal influence from X_1 to X_2 can be defined as:

$$\mathrm{F_{X_1 \to X_2}} = \ln(var(e_2)/var(e_4)) \qquad (6.6)$$

It assumed that the observed data can be well represented by multivariate auto regressive (MVAR) models. If the data is in the form of multiple repetitions of relatively short trials (e.g., event-related data), each trial is considered to be an independent realisation of a single statistically stationary process, such that a single MVAR model can be estimated based on the entire dataset. The estimation of MVAR model requires the inclusion of a parameter, the number of time-lags, i.e., the model order (p). Small model order can lead to a poor representation of the data, whereas large model order can lead to problems of over-fitting in model estimation. A standard means to identify the model order is to minimise a criterion that balances the variance accounted by the model, against the number of coefficients to be estimated. The most commonly used criterion is the Akaike information criterion (AIC) (Seth 2010; Wang et al. 2007). For n variables:

$$AIC(p) = \ln(\det(\Sigma)) + 2pn^2/T, \qquad (6.7)$$

where, Σ is the estimation of the prediction error covariance matrix of the bivariate autoregressive model. The model can be validated by assessing the quality of the model fitness of the prediction ratio (Wang et al. 2007), which measures how much the model can explain the variance of the signal and the percentage of the variance contributed from the model in the total variance. This provides objective criteria on whether the model is capable of characterising the system dynamics. For a perfect fit, the prediction error is zero. If the model is correct and the true parameter values are estimated properly, the prediction error would be white noise (Seth 2010). If the autocorrelation function shows pronounced patterns, such as the ripples or slow decline at low lags, it suggests model inadequacy. In cases where the model order specified by the minimal AIC is too large to permit feasible computation, or in cases where the AIC does not reach a clear minimum over the range tested, a smaller model order can be chosen on condition that the AIC shows no further substantial decreases at higher orders (Brovelli et al. 2004).

6.3.2 Feature selection

Redundant features significantly reduce the efficiency of the pattern classification process and provide poor generalisation. To avoid using redundant features and thus improving the classification process, a feature selection strategy has become essential in many signals or image classification (discussed in chapter 3, section 3.4). In particular this is true for neural signals that contain highly redundant information. Due to practical issues related to neural data acquisition techniques, lack of training, concentration, discomfort, fatigue and varied physiological or pathological conditions, it may not always be possible to collect a large amount of reliable data to alleviate the redundancy in the feature space.

Feature selection or dimensionality reduction in the extracted feature space can be achieved by eliminating the features that carry the least useful information. In this chapter, we introduce a new feature selection strategy, weighted sequential feature selection (WSFS) based on the feature ranking, sequential feature selection (SFS) and feature contributions, to efficiently select the optimal subset of features from the available features. The WSFS strategy is capable of selecting the most effective and consistent features, which is independent of changes captured in the training and testing

set. As discussed in chapter 3 (cf. 3.4), the context of pattern classification, feature selection strategies can be considered in three taxonomical categories, (1) *filter* (2) *wrapper* and (3) *embedded* approaches depending on the characteristics of the evaluation and selection criterion (i.e. feature ranking and selection) (Saeys et al. 2007). Based on the feature ranking and selection criterion used to evaluate the SFS strategy, it is considered either a *filter* or *wrapper* approach. However our new strategy, WSFS potentially overcomes the drawbacks of the SFS strategy that minimise the risk of overfitting. Therefore, according to the feature ranking and selection criterion used to evaluate the WSFS strategy, it is considered as an embedded approach (cf. 6.3.4.2).

6.3.2.1 Feature ranking

Feature ranking is a criterion that is used to rank each individual feature in the feature space. Let $A = \{a_1, a_2, \ldots, a_M\}$ be the set of M features. The function $r: A \rightarrow \mathbb{R}$ that assigns a relevant rank (value of merit) to each feature $a \in A$ based on a given criterion and orders them by its relevance. This returns an ordered list of features $a^* \in A$ that can be defined as

$$\{a^*_{j,j=1,2,\ldots,M}\} = r(\{a_{i,i=1,2,\ldots,M}\}) \qquad (6.8)$$

where , $i = 1,2, \ldots, M$ represents all features before ranking, and $j = 1,2, \ldots, M$, $r(a^*_{j-1}) \geq r(a^*_j) \geq r(a^*_{j+1})$ represents all features after ranking, ordered by its relevance (ascending or descending). By convention, it is assumed that a high score is indicative of a relevant feature, so that features are sorted in decreasing order of $r(.)$. The feature ranking criterion are defined, so to evaluate individual features, independent of the context of others with the assumption that features are independently and identically distributed. The feature ranking criteria of interclass separability based on F-score (*ISF*) and classification accuracy (*CA*) are evaluated in this project.

6.3.2.2 Interclass separability based on the F-score (*ISF*) criterion

The *ISF* is a normalised measure between the features of two classes, computed based on the F-score (Chen & Lin 2006; Duda et al. 2001). The F-score measures the discrimination between two sets of real numbers. Given training instances $X_i, i = 1, \ldots, N$, if the number of instances for the movement and rest classes is $n_{(+)}$ and $n_{(-)}$ respectively, then the ISF of the jth feature is defined as:

$$ISF(j) = \frac{\left(\overline{x}_j^{(+)} - \overline{x}_j\right)^2 + \left(\overline{x}_j^{(-)} - \overline{x}_j\right)^2}{\frac{1}{n_{(+)}-1}\sum_{i=1}^{n_{(+)}}\left(x_{i,j}^{(+)} - \overline{x}_j^{(+)}\right)^2 + \frac{1}{n_{(-)}-1}\sum_{i=1}^{n_{(-)}}\left(x_{i,j}^{(-)} - \overline{x}_j^{(-)}\right)^2} \quad (6.9)$$

where $\overline{x}_j$, $\overline{x}_j^{(+)}$, and $\overline{x}_j^{(-)}$ are the averages of the jth feature of the whole, movement and rest class datasets respectively; $x_{i,j}^{(+)}$ is the jth feature of the ith movement instance, and $x_{i,j}^{(-)}$ is the jth feature of the ith rest instance. The numerator is an indication of the separability between the movement and rest class datasets, and the denominator is indicative of the intra-class variability of a feature. The larger the ISF is, the more likely this feature is to be discriminative. Therefore, this measure can be evaluated to rank each individual feature across training instances, and then the features with a higher ISF will receive a higher rank than ones with a lower ISF.

6.3.2.3 Classification accuracy (CA) criterion

A classifier is a function that assigns a class label to a new instance ψ: $A^M \rightarrow \mathrm{C}$, where M is the number of features used for the classification. The classification accuracy (CA) is the average success rate provided by the respective classifier (ψ) given a set of test instances (N), i.e., the average number of times that ψ was able to correctly predict the class of the test instances. Let I_i be a test instance and class label pair, $I_i = (X_i, y_i)$, if $\psi(X_i) = y_i$ then X_i is correctly classified, otherwise it is misclassified. i.e. if $\psi(X_i) = y_i$ then 1 is counted else 0. So the CA can be defined as

$$CA = \frac{1}{N}\sum_{i=1}^{N}(\psi(X_i) = y_i). \quad (6.10)$$

The CA is depends on the method (ψ) that is used for classification.

The features can be ranked and ordered according to classification accuracies of univariate classifiers designed for each feature. A univariate classifier is required to developed for each feature of the set of M features, and the CA of each feature classifier is evaluated. Noticeably, the training set has to be divided further into two mutually exclusive sets: one to estimate univariate classifier parameters and the other to evaluate CA. If $CA(a_i)$ is the evaluated classification accuracy of the univariate classifier $\psi(a_i)$ for feature $a_i, i = 1,2,\dots,M$, the ranking of the feature is directly given by the $CA(a_i)$. A larger CA indicates that the feature is likely to be more discriminative. Therefore, this measure can be evaluated to rank each individual feature across training instances, and

then the features with a higher CA will receive a higher rank than ones with a lower CA. The CA is evaluated through a classification function (ψ) based on the selected classifier, Bayesian or SVM from the given training data.

6.3.2.4 Sequential feature selection (SFS)

Sequential feature selection (SFS) is a feature selection strategy that is commonly used for biosignals due its simplicity and effectiveness (Saeys et al. 2007). It can be defined as sequential forward or sequential backward feature selection. The SFS forward approach starts from the empty set, and in each iteration generates new subsets by adding a feature selected by evaluation function such as classification accuracy. It consists of a forward step which is as follows, starting from a initially empty set of features Z_0, at each forward step l add the feature $a^+ \in (A - \mathrm{Z}_{l-1})$, where $a^+ \in \{a_1^*, a_2^*, \ldots, a_M^*\}$, $a^+$is the next available feature from the set of ranked and ordered feature, such that for $\mathrm{Z}_l = \mathrm{Z}_{l-1} \cup a^+$, the probability of correct classification achieved by the classifier function ψ is maximised (Duda et al. 2001). Using this process all the ranked (high to low) features are incorporated into the feature set $\mathrm{Z}_l$until $\mathrm{Z}_l = \mathrm{Z}_M$ and, at each l, $l = 1,2, \ldots, M$ feature set is evaluated and the CA recorded. Finally the optimal subset of features is selected at l, where the classification accuracy reaches its maximum or peak, i.e.,

$$CA_{max} = \mathrm{argmax}_{l=1\ldots M}(\psi(\mathrm{Z}_l)) \qquad (6.11)$$

where CA is the classification accuracy and M is the total number of features.

Similarly, SFS backward approach starts from the complete feature set, and in each iteration generates new subsets by discarding a feature selected by evaluation function such as classification accuracy. It consists of a backward step which is as follows, starting from a initially complete set of features $\mathrm{Z}_M = A$, at each backward step l exclude the feature $a^- \in \mathrm{Z}_{l-1}$, where $a^- \in \{a_1^*, a_2^*, \ldots, a_M^*\}$, such that for $\mathrm{Z}_l = \mathrm{Z}_{l-1} - \{a^-\}$ the probability of correct classification achieved by the classifier ψ is maximised. Using this process each of the ranked (low to high) features are excluded from the feature set $\mathrm{Z}_l$until $\mathrm{Z}_l = \mathrm{Z}_0$ and, at each l, $l = M, M-1, \ldots, 1$ feature set is evaluated and the CA recorded. Similar to the forward approach, the optimal subset of features is selected at l, where the classification accuracy reaches its maximum or peak, i.e.,

$$CA_{max} = \mathrm{argmax}_{l=M,M-1,\ldots,1}(\psi(\mathrm{Z}_l)) \qquad (6.12)$$

Using SFS forward approach, the optimal feature subset is selected from the ISF or CA ranked and ordered features with maximum classification accuracy or mean *ISF* score. One of the major limitations in the SFS strategy is that it is biased towards the given training set. When the training set changes, its performance varies, which is not robust and suitable for non-stationary signal dynamics, such as those used in neural signal decoding.

6.3.2.5 Weighted sequential feature selection (WSFS)

To select the most relevant features and effectively reduces the number of redundant features, which have little or no contribution to the classification, we introduced weighted sequential feature selection (WSFS). As mentioned earlier, SFS feature selection is a greedy approach (procedure that generate local optimal choices rather than a globally optimal solution (Dash & Liu 1997)) that selects the local optimal feature set without considering global effect. Therefore it fails to provide the most relevant feature consistently with the change of training sets and also in most cases it is unable to eliminate redundant features from the feature space (Dash & Liu 1997). In this case, WSFS overcomes the limitations of SFS by selecting the most relevant features independently with changes captured in the training set. The WSFS was developed by utilising the feature ranking and the SFS approach combined with computing each feature's contributions. In WSFS approach, it divides the training set into *k* subsets for cross-validation, and evaluates each subset using SFS with ranked features for selecting suboptimal feature set. After each evaluation, feature weight is assigned based on their contribution and a total weight is computed. Finally based on the total weight, all of the features are divided into *k* overlapping subsets for further evaluation and then a subset with maximum accuracy was selected as an optimal feature set for classification. Feature weight assignment in WSFS ensures the consistency and elimination of redundancy in the selected feature subset. It is expected that this strategy will significantly improve the performance and robustness of the decoding process.

Let $\boldsymbol{S} = \{X_1, X_2, \dots, X_N\}$ be a dataset of *N* instances and $A = \{a_1, a_2, \dots, a_M\}$ be the set of *M* features. The dataset $\boldsymbol{S}$ can be divided into *k* subsets through *k*-fold cross-validation, where each subset is named as $T_1, T_2, \dots, T_k$. The subset T_p consists of n_k instances, where $n_k = \lfloor N/k \rfloor$, with each instance, $X_i = [a_1, a_2, \dots, a_M]^T$ having *M* features. During the cross-validation process, at each fold iteration, each subset T_p

was defined as a testing set, with the rest of the subsets forming a single associated disjoint training set, $D_p = \bigcup_{o=1}^{k} T_{o,}$ where, $o \neq p$ and $p = \{1, 2, \dots, k\}$. Let $\phi : A^M \rightarrow \mathbb{R}^m$ to be a function that extracts the m most significant features as ranked by a specific feature ranking function, $r(\cdot)$, and selected using a SFS strategy based on a subset $D_p \subset \boldsymbol{S}$. The cardinality of $|D_p| = \widehat{N}$, where $\widehat{N} = n_k(k-1)$, and it was created by k-fold cross-validation from $\boldsymbol{S}$**.** This process repeated q times, i.e. q repeated k-fold cross-validation, such that $D_1, D_2, \dots, D_Q \subset \boldsymbol{S}$ was created, where $Q = q \times k$. Each subset $D_i^{\widehat{N}\times M}$ with M features was evaluated during the initial feature selection procedure based on the function ϕ and obtained a resultant subset $\widehat{D}_i^{\widehat{N}\times m}$. It can be written as

$$\widehat{D}_i^{\widehat{N}\times m} = \phi\big(r(D_i^{\widehat{N}\times M})\big), \qquad (6.13)$$

where, $i = \{1,2,\cdots,Q\}$, $r(\cdot)$ is the feature ranking function, and $\phi(.)$ is SFS function with forward or backward approach. The feature ranking function can be defined by criteria such as the ranking-based ISF or CA or any other similar measure. Each subset $\widehat{D}_i$ contains the $\widehat{N}$ instances of m most significant features from subset D_i as defined by the ranking and selection function. Due to the randomness in the selection of subsets, the instances in each D_i will be different, so that the resultant features in each $\widehat{D}_i$ may be different. This generates inconsistent feature subsets and therefore estimates the classifier parameters are either inaccurate or cannot be optimised, and made the classification unreliable. To address this problem, a weight measure, ω_j associated with each feature a_j was defined to determine the consistent feature subsets independent of the instances contained within each D_i. The weight value, ρ_{ij} was assigned based on the contribution of each feature a_j in $\widehat{D}_i$ and formulated as,

$$\omega_j = \sum_{i=1}^{Q} \rho_{ij}, \text{ where } \rho_{ij} = \begin{Bmatrix} 0, & if\ a_j \notin \widehat{D}_i \\ 1, & if\ a_j \in \widehat{D}_i \end{Bmatrix} \qquad (6.14)$$

The range of the weight measure, ω_j is established in Proposition I.

Proposition I: For feature $j \in \{1,2,\cdots,M\}$, $D_i \subset \boldsymbol{S}$, where $i \in \{1,2,\cdots,Q\}$, $\rho_{ij} \in \{0,1\}$, it follows that $0 \leq \omega_j \leq Q$:

(i) $\omega_j = \sum_{i=1}^{Q}\{\rho_{ij} = 0\} = 0$

$$\text{(ii)} \qquad \omega_j = \sum_{i=1}^{Q}\{\rho_{ij} = 1\} = Q$$

Despite different training and testing sets, the feature selection method should produce consistent and robust subset of features for classification. To obtain the robust subset of features, all the features in A are grouped into Q subsets of feature based on the measured contribution of each feature a_j as dictated by weight ω_j and is therefore defined as,

$$H_i = \{a_j,\ if\ \omega_j \geq i\}, \qquad (6.15)$$

where a_j is the jth feature in A, $j = \{1,2,\cdots,M\}$, $i = \{1,2,\cdots,Q\}$, and H_i is the ith feature subset that have all the features which contributed at least i times during the Q evaluations of the function ϕ. The cardinality of H for each subset i was defined as $|H_i| = \widehat{M}_i$ and its range is established in Proposition II.

Proposition II: For feature $j \in \{1,2,\cdots,M\}$, $H_i \subset A^T$, where $i \in \{1,2,\cdots,Q\}$, $\widehat{M}_i = |H_i|$, it follows that $\widehat{M}_{min} \leq \widehat{M}_i \leq \widehat{M}_{max}$:

$$\text{(i)} \qquad \widehat{M}_i = |H_i| = \sum_{j=1}^{M}\{\omega_j = \max(\omega)\} = \widehat{M}_{min}$$

$$\text{(ii)} \qquad \widehat{M}_i = |H_i| = \sum_{j=1}^{M}\{\omega_j \geq \min(\omega)\} = \widehat{M}_{max}$$

To find the optimal feature subset for producing the most reliable and best recognition accuracy for the test set from the dataset $\boldsymbol{S}$, all the subsets H_i were evaluated and the classification accuracy was recorded as,

$$\Omega_i = \psi(H_i), \qquad (6.16)$$

where, ψ is the classification function. Finally, the optimal feature subset was selected as

$$H_{optimal} = \arg\max_{i=1\ldots Q}(\Omega_i), \qquad (6.17)$$

which provided the highest CA. This optimal subset is robust in terms of providing the best accuracy independent of changes captured in the training and testing set of $\boldsymbol{S}$.

6.3.3 Pattern classification

Based on outcomes of the previous investigation, both SVM and Bayesian classification were employed in the feature ranking, selection and classification process for deep brain LFP decoding. It was expected that both the classification methods will provide high

accuracy based on the optimally selected LFP features for decoding movement and laterality.

6.3.4 Decoding of deep brain LFPs

6.3.4.1 Pre-processing and feature extraction of deep brain LFPs

The deep brain LFP recordings were selected by excluding the segments contaminated with unintended movements based on surface EMGs. Also trials contaminated with artifacts, for which the behavioural response was incorrect, were removed using visual inspection. The LFPs were pre-processed with a low-pass Chebyshev Type I filter to remove the high frequencies. It was implemented with zero-phase shifting and a cut-off frequency of 90 Hz. A bandstop filter at 50 Hz to remove the power line noise was also implemented using a custom made adaptive filter (Wang et al. 2004). The recorded LFPs and their spectrogram (filtered at 90 Hz) from STN (a) and GPi (b) during an externally-cued left and right clicking movement tasks are presented in figure 6.2. The time of stimulus presentation and the subsequent motor response are shown using dotted and solid vertical lines, respectively. The state flow diagram for movement and its laterality decoding of deep brain (STN or GPi) LFPs and the general flowchart for decoding process of movement and laterality of LFPs presented in figure 6.3 (a) and (b) respectively.

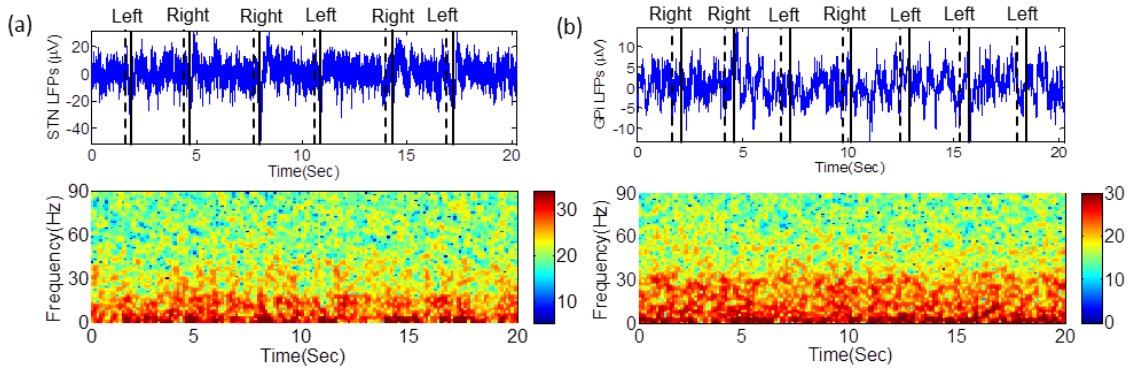


Figure 6.2: Basal ganglia LFPs recorded from STN (a) for subject 1 and GPi (b) for subject 5, and their spectrogram (filtered at 90 Hz) during an externally-cued left or right clicking movement tasks. The time of stimulus presentation and the subsequent motor response are shown using dotted and solid vertical lines, respectively.

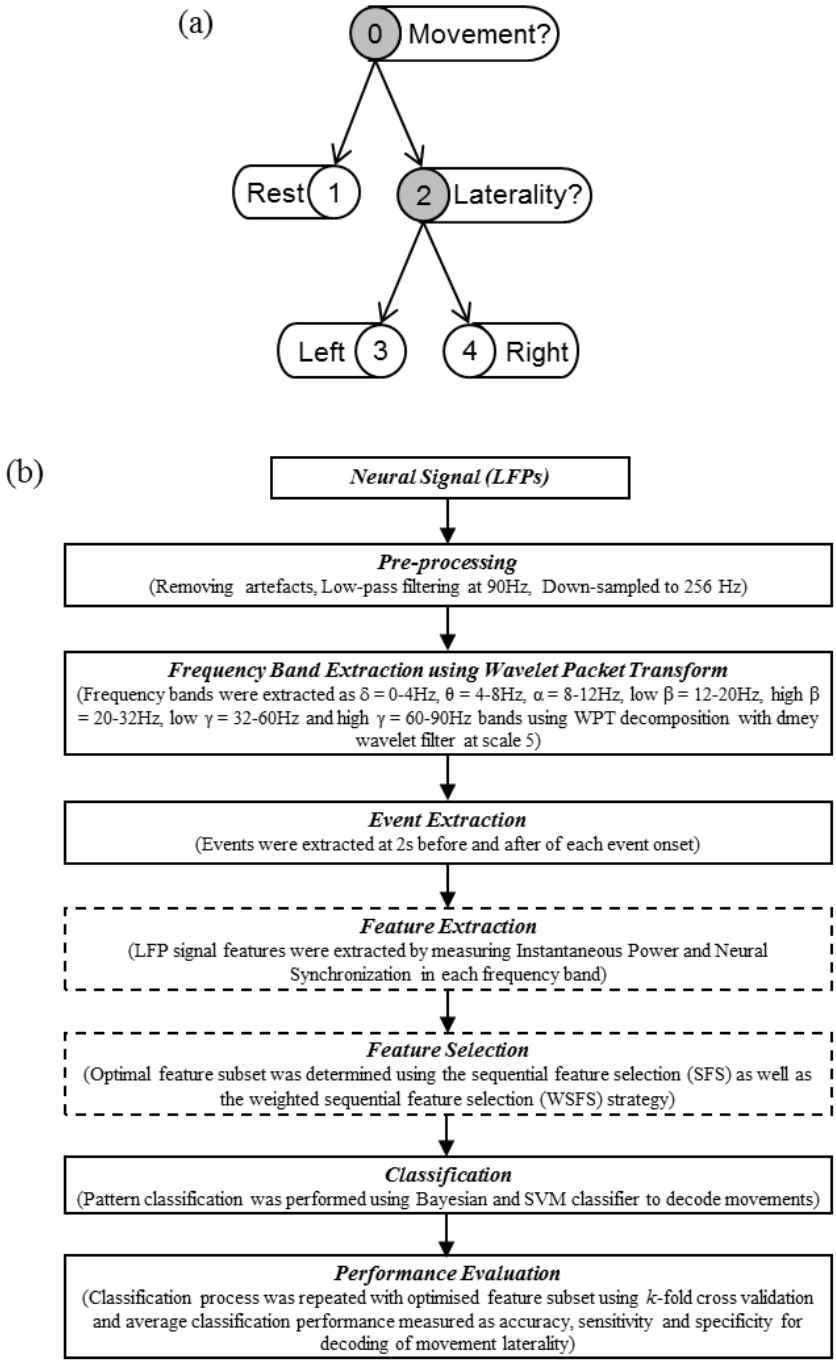


Figure 6.3: State flow diagram (a) for movement and its laterality decoding from deep brain (STN or GPi) LFPs and the general flowchart (b) for decoding process of movement and laterality of LFPs.

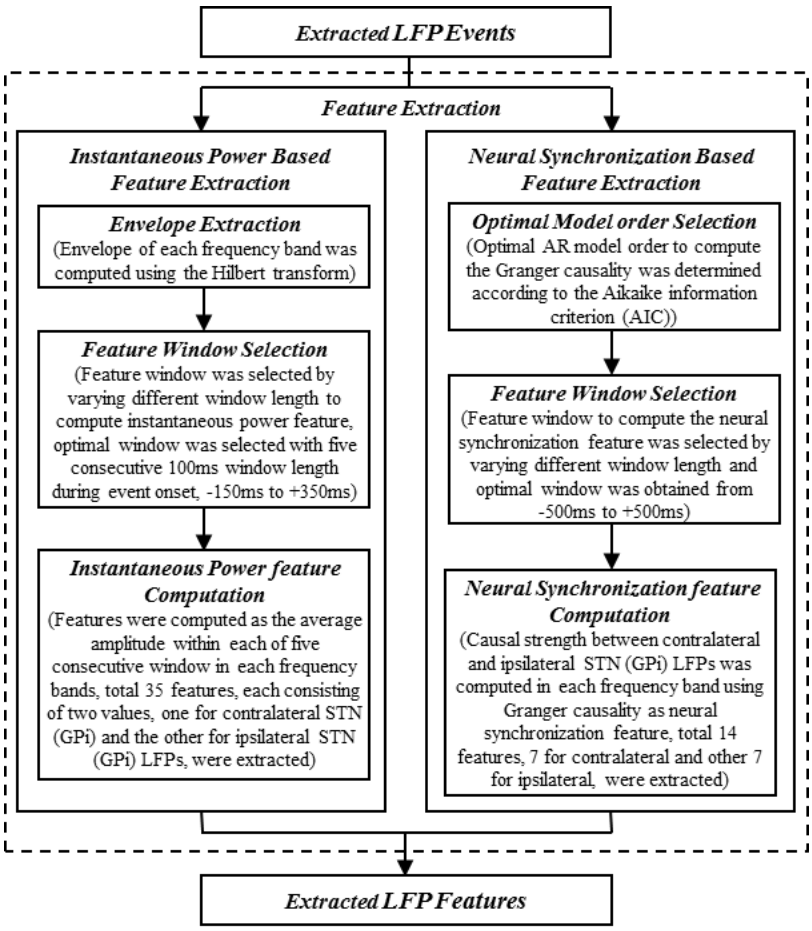


Figure 6.4: The flowchart for both instantaneous power and neural synchronisation feature extraction from the extracted LFP events.

The filtered LFPs were then re-sampled at 256 Hz and processed for feature extraction. The frequency dependent components were extracted as δ = 0-4 Hz, θ = 4-8 Hz, α = 8-12 Hz, low β = 12-20 Hz, high β = 20-32 Hz, low γ = 32-60 Hz and high γ = 60-90 Hz frequency bands using a wavelet packet transform with a discrete Meyer wavelet (dmey) at decomposition scale 5. In each component of the LFPs, the event of left and right clicking tasks were segmented as 2 s before and after at each motor response registration. Similarly the resting activity was segmented as 2 s before and 2 s after at each stimulus registration. Segmented signal features were extracted based on instantaneous power and neural synchronisation from each frequency band. The flowchart for both feature extraction approach are presented in figure 6.4. To compute the instantaneous power features, the envelope of each component was computed using the Hilbert transform. The amplitude modulations of each component during left and right clicking events over all trials recorded from left STN of subject 1 are presented in

figure 6.5. From this figure it can be seen that there seems an amplitude decrease in the β band and increases in the δ, θ, α and γ bands, most visible in the δ band. Based on the average event related de-synchronisation and synchronisation in these bands (figure 6.6), the instantaneous power features for classification were defined as the average amplitude within each of five consecutive 100-ms windows in each frequency band. The five windows for the resting condition ran from -750 to -250 ms before the stimulus and the five windows for the clicking events ran from -150 to 350 ms around the motor response (figure 6.6). In total thirty-five features (each consisting of two values, one for left STN or GPi LFPs and the other for right STN or GPi LFPs) with five from each frequency band were obtained for classification.

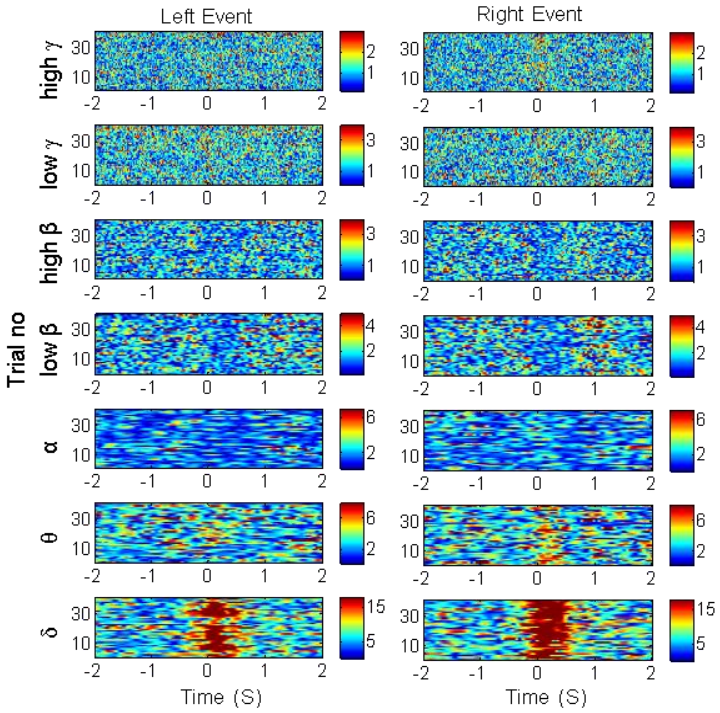


Figure 6.5: The instantaneous amplitude of deep brain left STN LFPs components were computed using Hilbert transform and all trials for subject 1 are shown for the left and right clicking events with a 4-s window centred at the time of the response.

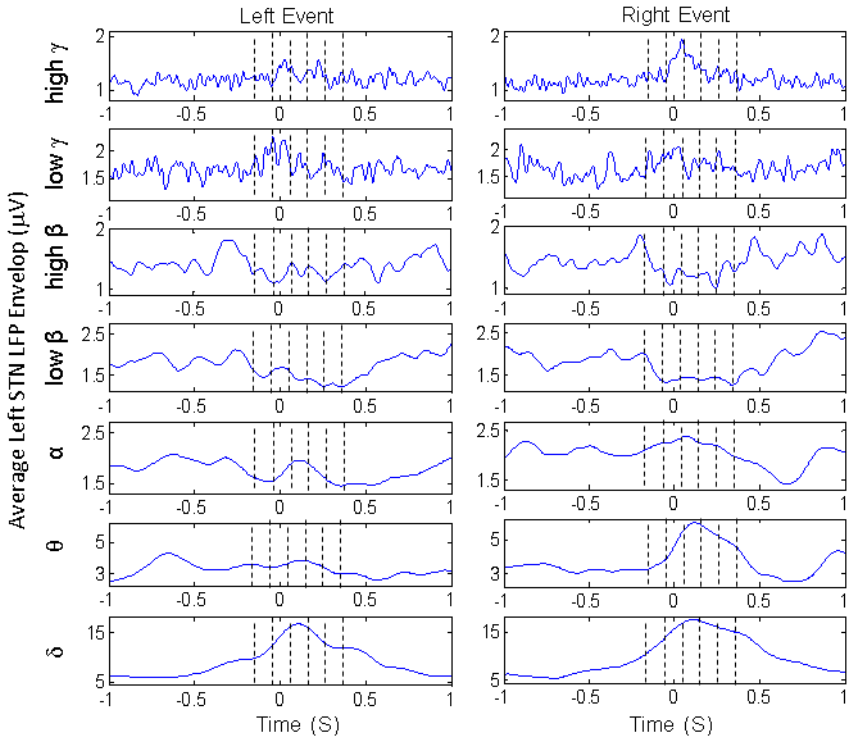
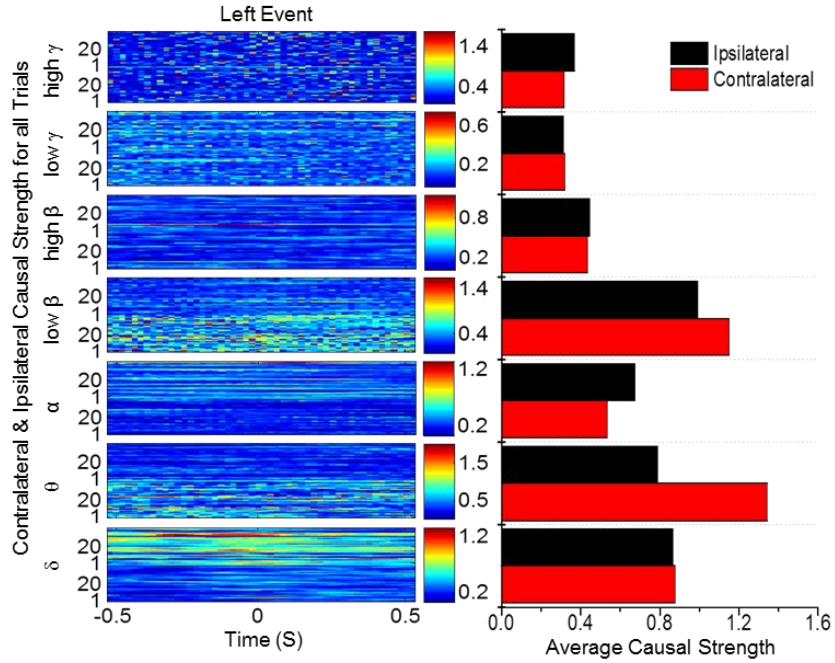


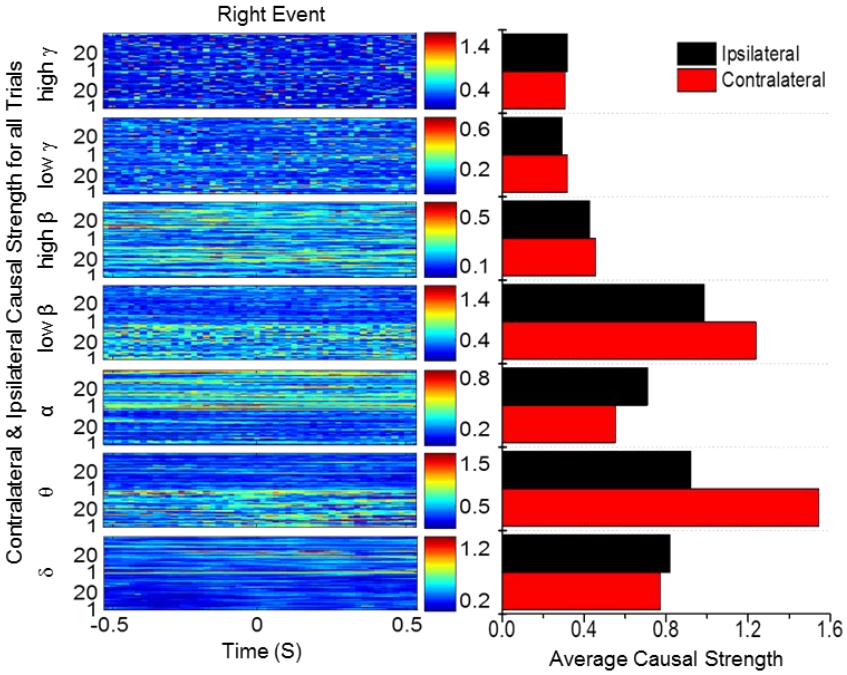
Figure 6.6: The average instantaneous amplitude of left STN LFPs in each components of subject 1 for the left and right clicking events with a 2-s window centred at the time of the response. The LFP signal features in each frequency bands were defined as average instantaneous amplitude within (W1:-150 to -50 ms), (W2:-50 to 50 ms), (W3:50 to 150 ms), (W4:150 to 250 ms) and (W5:250 to 350 ms) window around the response timing.

As it was assumed that synchronisation of the neural activities between different regions of the brain usually relates to the state or specific movements for which the signal is recorded, and will provide effective information about movements. The investigation of the dynamic changing of causal relationships between the neural signals recorded from the left and right STN or GPi for the events can provide more discriminative information to decode movement laterality. Therefore, causal strength between LFPs of left and right STN or GPi was evaluated by analysing Granger causality and denoted as the neural synchronisation feature for decoding left and right clicking events. During the causal analysis, the features were defined by computing contralateral and ipsilateral causal strength in each frequency band LFPs for each of the left and right clicking events. The contralateral and ipsilateral causal strength for left clicking events were computed as causality of right STN(GPi) → left STN(GPi) and left STN(GPi) → right STN(GPi) respectively. Similarly for right clicking events, the contralateral and

ipsilateral causal strength were computed as causality of left STN(GPi) → right STN(GPi) and right STN(GPi) → left STN(GPi) respectively. The analysis of Granger causality for each event was performed with segmented LFPs by varying analysis windows, 1) between 1 s before and after, and 2) between 500 ms before and after the onset of each motor response registration. For each analysis window, the MVAR model was estimated and the optimal order for the MVAR model was identified by locating the minimum of the AIC. However, the AIC dropped monotonically with increasing model order up to a value of 10 (5 and 15 in some cases) and then with the increase of the model order no further substantial decreases or increases of AIC was shown. Therefore average discriminability of contralateral and ipsilateral causal strength between left and right events produced from the pilot subjects and compared using model orders of 5, 8, 10, 12, 15, 20, 25 and 30, and observed that overall results varies. However, the model order 5, 10 and 25 produces more consistent and better discriminability in both analysis windows. Finally, a model order of 10 (40 ms) using a shorter window (500 ms before and after the onset) was selected as a tradeoff between sufficient discriminability and over-parameterisation (Brovelli et al. 2004). The analysed LFP data from all trials were treated as realisations of a common stochastic process, and thus were used to estimate the model coefficients for that process. Figure 6.7 presented the demonstration of contralateral and ipsilateral Granger causal strength between left and right STN for all trials and their average in all frequency bands of subject 1 for left (a) and right (b) events. Finally, fourteen neural synchronisation features were obtained for classification by evaluating contralateral and ipsilateral causal strength in each of seven frequency bands for left and right clicking events. The Granger causality analysis for extracting the neural synchronisation features was performed with the help of GCCA MATLAB toolbox (Seth 2010).



(a)



(b)

Figure 6.7: The contralateral and ipsilateral Granger causal strength between left and right STN for all trials and their average in all frequency bands of subject 1 for left (a) and right (b) events. The causal strength was defined for contralateral left event: right→left STN, right event: left→right STN, and for ipsilateral left event: left→right STN, right event: right→left STN in each frequency bands for each of left and right clicking events. The causal strength for each event was computed with in the 1-s

window (-500 to +500 ms) around the response timing in each frequency bands and defined as contralateral (red) and ipsilateral (black) neural synchronisation features.

6.3.4.2 Feature selection and classification of deep brain LFPs

During the feature extraction thirty-five instantaneous power and fourteen neural synchronisation features were obtained for decoding movement and laterality for finger clicking events. As neural synchronisation features are related to movement laterality, therefore, only instantaneous power features were used to decode the movement (event and rest). In the subsequent laterality decoding, instantaneous power and/or neural synchronisation features were used. Both the movement and laterality decoding were performed through selection of optimal feature subset and then classification. Two classifiers, Bayesian and SVM were implemented for decoding of deep brain LFPs. The classifier was trained, validated and evaluated by a *k*-fold cross-validation procedure. Cross-validation provides a relatively unbiased estimate of the generalisation capability of the algorithm by splitting the dataset into *k*-subsets and selecting each subset as the test set with the remaining sets used to train the classifier. Using cross-validation, the LFP dataset from each subject was divided into training and testing sets. The training set was used for feature selection and it was performed in two phases. Features were firstly ranked according to their individual discriminability. Then based on ordered (high to low) ranked features, a subset of features were selected by evaluating the sequential feature selection (SFS) strategy as well as our newly developed method, the weighted sequential feature selection (WSFS) strategy based on individual feature contributions in second phase. In both strategies, a sequential forward subset selection approach was used. To rank the individual features, interclass separability based on the F-score (ISF) and classification accuracy (CA) were used as a ranking criterion. The scores of ranking criteria, ISF or CA for each individual feature were computed across training instances and ordered by ranking function, $r(.)$. The detailed flowchart of both SFS and WSFS feature selection strategies to select optimal feature subset from available extracted features are presented in figure 6.8 and 6.9 respectively and detailed illustration of both strategies are also presented in section 6.3.2.4 and 6.3.2.5.

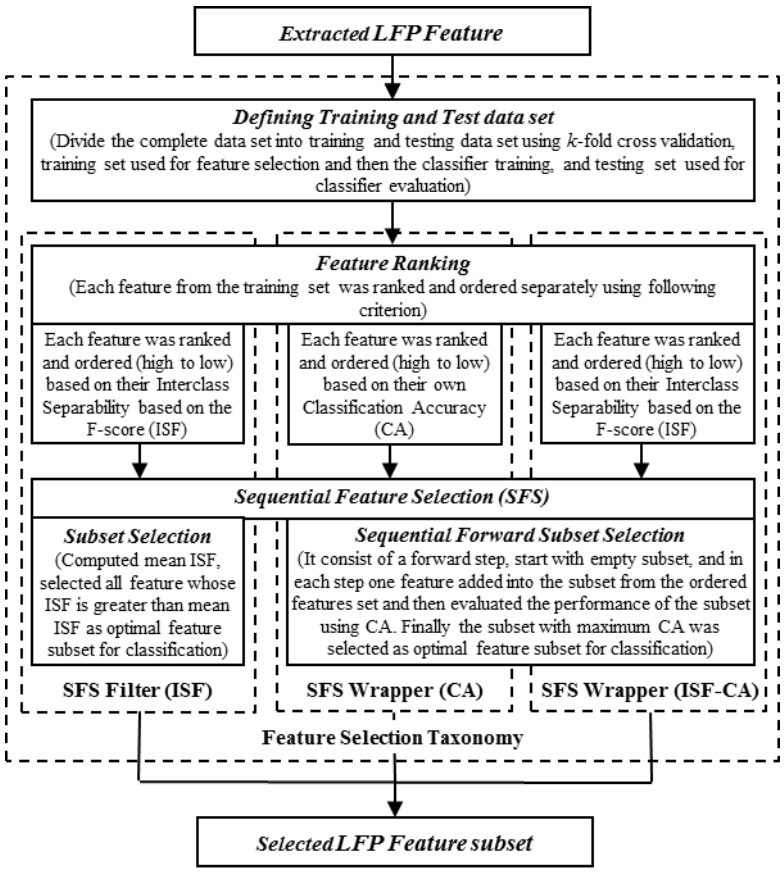


Figure 6.8: The flowchart for SFS feature selection strategy with three selection models (*SFS filter (ISF)*, *SFS wrapper (CA)* and *SFS wrapper (ISF-CA)*) designed for selecting optimal feature subset for decoding movement from deep brain LFPs.

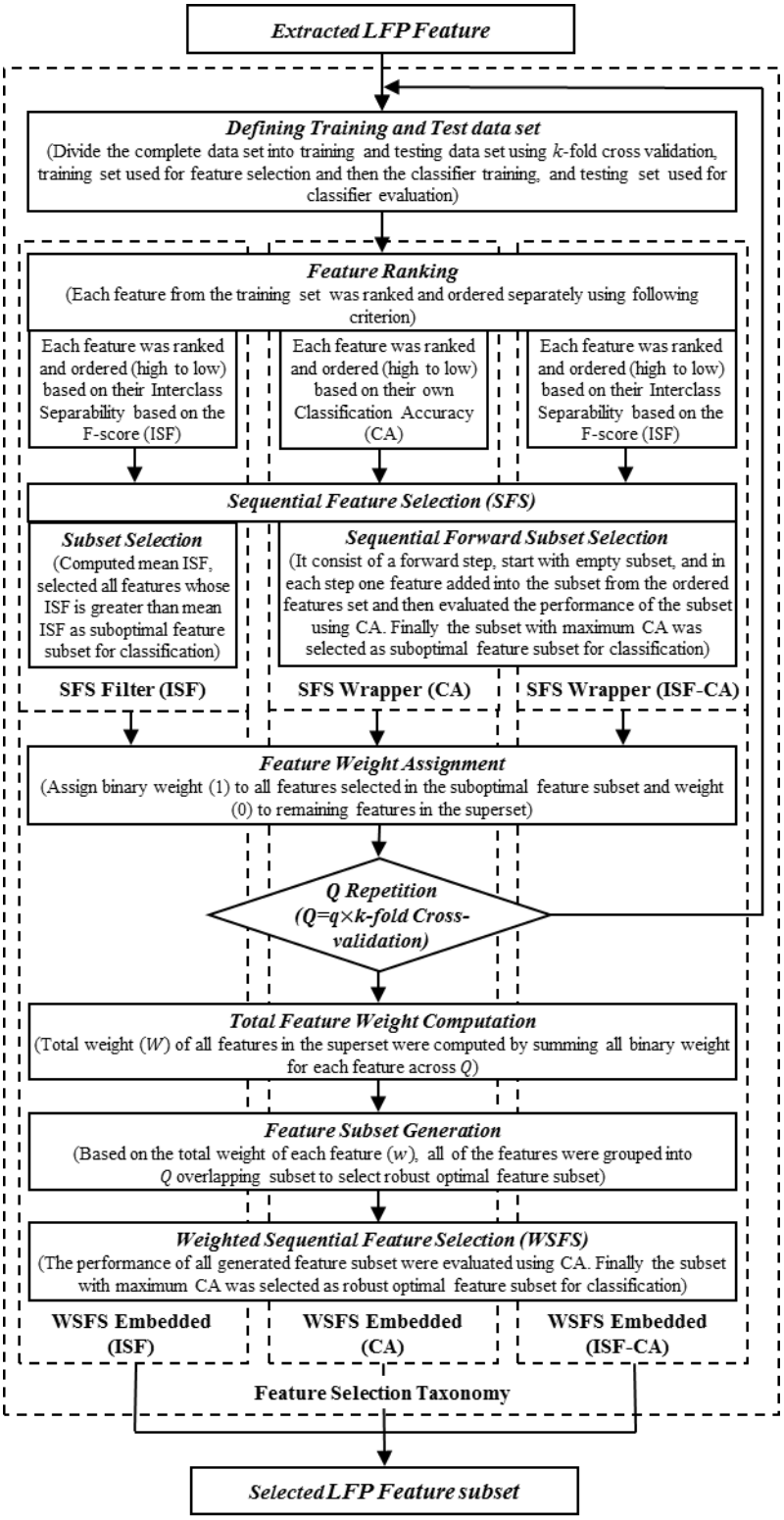


Figure 6.9: The flowchart for WSFS feature selection strategy with three selection models (*WSFS embedded (ISF)*, *SFS embedded (CA)* and *SFS embedded (ISF-CA)*) designed for selecting optimal feature subset for decoding movement from deep brain LFPs.

As discussed earlier (cf. 6.3.2), in the context of pattern classification and the construction principle of feature selection method, the SFS and WSFS strategy can be considered into three (*filter*, *wrapper* or *embedded*) taxonomical categories. Considering the taxonomical categories and feature ranking criterion (ISF or CA), the SFS and WSFS feature selection strategy for LFP decoding organised into three alternative selection models to extract the optimal feature from the feature space. Thus three selection models for SFS organised as *SFS filter (ISF)*, *SFS wrapper (CA)* and *SFS wrapper (CA-ISF)*, while three selection models for WSFS organised as *WSFS embedded (ISF)*, *WSFS embedded (CA)* and *WSFS embedded (CA-ISF)*. In both SFS and WSFS strategy, the selection models are established based on how feature ranking and selection were incorporated to generate the optimal feature subset for providing the superior decoding performance.

The *filter* approach is computationally simple and fast, and it is independent of the classification methods. As *ISF* feature ranking is a normalised measure between the features of two classes and the selection is based on its average score, which is independent of the classification methods. In other words, each feature in the feature space firstly ranked and ordered (high to low) based on their *ISF* score and then selected a subset of features whose *ISF* score is greater than the mean *ISF* score. Therefore, *ISF* feature ranking criterion incorporating subset selection based on mean *ISF* was used to generate optimal feature subset for classification modelled as *SFS filter (ISF)*. The problem of the *filter* approach is that it ignores the interaction with classifier and it considers each feature separately, thereby ignores the feature dependences. The *wrapper* approach overcomes the problem of *filter* approach through interacting with the classifier. Therefore, CA feature ranking criterion incorporating sequential forward subset selection based on CA was used to generate optimal feature subset for classification modelled as *SFS wrapper (CA)*. Similarly, *ISF* feature ranking criterion incorporating sequential forward subset selection based on *CA* to generate optimal feature subset for classification was modelled as *SFS wrapper (ISF-CA)*. The common limitation of the *wrapper* approach is that it has higher risk of overfitting, i.e. it is biased towards the training set. Particularly, as SFS wrapper approaches are mainly based on the classification function for selecting optimal feature subset, when the training dataset changes, its performance varies, which is not robust for non-stationary signal dynamics,

such as those used in neural decoding. Again *embedded* approaches try to overcome the limitation of the *wrapper* approach by addressing the risk of overfitting.

Hence, to select the robust and optimal subset of features for deep brain LFP decoding, WSFS was introduced. The WSFS is considered as an *embedded* approach and it addresses the limitation of SFS wrapper approach by integrating the individual feature contributions through computing each feature's weight with the SFS strategy. This strategy efficiently reduces the risk of overfitting and it has the ability to select the most effective features by reducing redundant features, which have least contribution to the classification, and also the selection process is independent of changes captured in the training set. Similar to the SFS *filter* and *wrapper* selection models, three WSFS *embedded* selection models were designed with integrating feature ranking (ISF or CA) criteria. The ISF feature ranking and subset selection (i.e. *SFS filter (ISF)*) incorporating WSFS strategy was used to generate an optimal subset of features for classification modelled as *WSFS embedded (ISF)*. Again CA feature ranking and sequential forward subset selection based on CA (i.e. *SFS wrapper (CA)*) incorporating WSFS strategy was used to generate an optimal subset of features for classification modelled as *WSFS embedded (CA)*. Similarly, ISF feature ranking and sequential forward subset selection based on CA (i.e. *SFS wrapper (ISF-CA)*) incorporating WSFS strategy was used to generate an optimal subset of feature for classification modelled as *WSFS embedded (ISF-CA)*. It is expected that WSFS based models will significantly improve the performance and robustness of the decoding process. In all three WSFS embedded models, the parameter Q was set to 50, i.e. five repeated 10-fold cross-validation (i.e. $Q = 5 \times 10$) was used to evaluate WSFS for optimal feature subset selection.

Based on the selected LFP feature subset obtained by each of SFS and WSFS selection models (*ISF, CA or ISF-CA*), the Bayesian and SVM classifiers were separately trained. It is noted that during the feature selection in both SFS and WSFS, the respective classifier was also used for feature ranking and subset selection. The Bayesian classifier parameters were estimated with univariate Gaussian assumption. As mentioned earlier (cf. section 4.5), due to the limited size of the training dataset compared to the large number of features, univariate classifier performed better than multivariate case in this study. The SVM classifier parameters, C and η were optimised to produce the highest classification rate with the selected features using a 10-fold cross-validation procedure

during the training stage. The decoding performance of the optimised Bayesian and SVM classifiers were evaluated through hundred repeated 10-fold cross-validation and the classification rate was calculated as an average. The performance of both classification methods in all SFS and WSFS selection models was evaluated and compared for decoding of voluntary movement and its laterality. It is noted that the decoding performances of movement laterality based on instantaneous power, neural synchronisation, and combining both features were also evaluated separately and compared for providing the better decoding performance based on the selected feature subset. To statistically compare the performances among the feature space (instantaneous power, neural synchronisation and combining both together), the selection models (ISF, CA and ISF-CA), the feature selection strategies (SFS and WSFS) as well as the classifiers (Bayesian and SVM) across all subjects, a repeated measure ANOVA was performed using SPSS (Ver. 15, Chicago, Illinois).

6.4 Results

In this study the decoding of voluntary movement activities was divided into two stages, namely classification between resting and voluntary movement, and then laterality classification between left and right hand finger clicking movements (figure 6.3(a)). Deep brain LFP data from twelve subjects (six PD and six dystonia subjects; deep brain LFP recorded from STN for five subjects and GPi for seven subjects; (cf. table 6.1)) were used for decoding the movement and its laterality. For movement and rest state classification only the instantaneous power based features were extracted for feature selection and classification. However for subsequent laterality, left and right, two types of features, instantaneous power and neural synchronisation were extracted. From the extracted features, feature selection and classification was evaluated for each type of feature separately as well as combining both features together to improve performance. The decoding performance was evaluated as averaged accuracy, sensitivity and specificity. Accuracy is defined as the percentage of correctly classified instances. Sensitivity is defined as the ratio of the number of true positives classified to the number of actual total positive cases. Specificity is defined as the ratio of the number of true negative classified to the number of actual total negative cases. The decoding performance of movement and laterality were evaluated for each subject separately and all subjects' performance was averaged and presented as mean ± 1 SD. To identify the overall influence of the features for movement and laterality decoding in SFS and

WSFS feature selections, all selected features were recorded and averaged across all subjects in both classifiers (Bayesian and SVM) and all selection models (*ISF*, *CA*, *ISF-CA*).

6.4.1 Demonstrative evaluation of SFS and WSFS

The demonstrative evaluation of both SFS and WSFS feature selection strategies is presented in figure 6.10. Here subject #1 was used to decode the movement laterality based on both instantaneous power and neural synchronisation features. The average classification accuracy (*CA*) of each individual feature ranges from 34.9 to 83.2% in all 49 features, which was used to rank each feature and presented in figure 6.10(a). Based on the ordered ranked feature, *SFS wrapper (CA)* model was evaluated and the superior performance achieved 93.1% average decoding accuracy using the first twenty-one ranked features, shown in figure 6.10(b,c). In the WSFS, ranked features were repeatedly evaluated using SFS for $Q = 50$ times and the contribution of each feature was recorded. The total contribution (weight) of each feature is shown in figure 6.10(d). Based on the total feature weight, Q subsets of features were generated and each subset was evaluated to find the most effective subset for providing better accuracy to decode left and right movements presented in figure 6.10(e). It was observed that using *WSFS embedded (CA)* model, the subset with only fourteen features provided highest accuracy 98.0% (figure 6.10(e,f)). The evaluation of this demonstration subject was performed with Bayesian classifier. The neural synchronisation features, contralateral θ, α and low β, ipsilateral δ and low β bands, and instantaneous power features of δ, θ, α, β and high γ bands during and immediately after clicking event were selected in SFS as shown in figure 6.10(c). While in WSFS, the features from the neural synchronisation, contralateral θ and α, ipsilateral δ and low β bands, and instantaneous power of δ, θ, α, β and high γ bands (α at W1(−150 -−50 ms), θ at W3(+50-+150 ms), high γ at W1-W3(−150-+150 ms), and δ, θ, β and high γ at W5(+250-+350 ms)) before, during and immediately after clicking onset were selected to provided higher accuracy as shown in figure 6.10(f). From here, it is observed that the WSFS selected a lower number of feature (20 vs. 14) and also most effective feature subset to provide better decoding accuracy than the SFS.

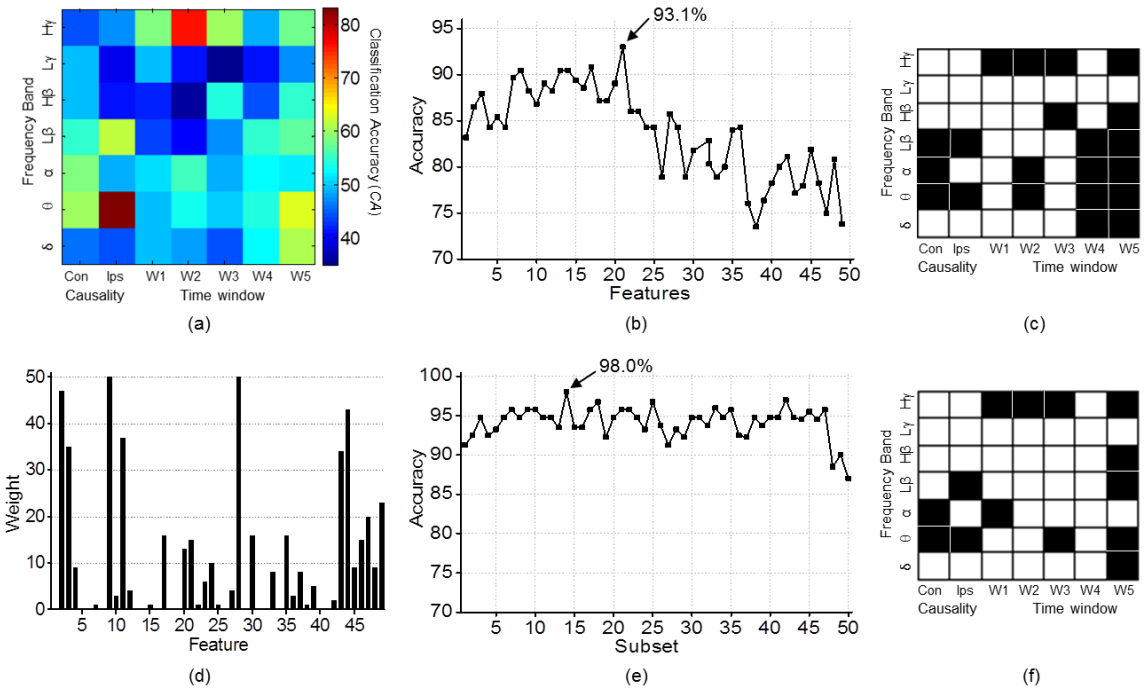


Figure 6.10: The feature ranking and feature selection strategy using SFS (*SFS Wrapper (CA)*) and WSFS (*WSFS Embedded (CA)*) for decoding of left and right events for subject #1 based on combined neural synchronisation (Con, Ips) and instantaneous power (W1 to W5) features. Each feature was ranked according to their individual classification accuracy (*CA*). All of the ranked neural synchronisation and instantaneous power features are shown in (a). Based on the ordered ranked feature, the best feature subset was selected using the SFS strategy with *SFS Wrapper (CA)* model for decoding left and right events, and presented in (b). Only 21 features were selected out of 49 features to provide the highest accuracy. The selected feature in *SFS Wrapper (CA)* model is presented in (c). In WSFS strategy with *WSFS Embedded (CA)* model, each feature obtained a weight based on their contribution shown in (d). Based on the feature weight, fifty ($Q = 50$) subset of features were generated and each subset was evaluated to obtain the optimal and robust subset for providing optimum performance to classify left and right events, presented in (e). The subset 14 contains only 14 features and provided the highest accuracy. The selected feature in *WSFS Embedded (CA)* model is presented in (f). The SFS strategy selected 21 features while WSFS strategy selected 14 features out of 49 features. With fewer feature WSFS providing better accuracy than the SFS strategy.

6.4.2 Movement decoding from deep brain LFPs

The decoding performance of movement (event vs. rest) was evaluated for all subjects by selecting the optimal feature subset using both SFS and WSFS strategies through Bayesian and SVM classifier. The performance (accuracy, sensitivity and specificity) of movement decoding for each subject (STN or GPi) and their average (mean ± 1 SD) in *SFS Filter (ISF)* and *WSFS Embedded (ISF)* (a), *SFS Wrapper (CA)* and *WSFS Embedded (CA)* (b), and *SFS Wrapper (ISF-CA)* and *WSFS Embedded (ISF-CA)* (c) selection models combined with Bayesian and SVM classifier are presented in figure 6.11. All methods have achieved high decoding performance with average accuracy, sensitivity and specificity larger than 98%. SVM with *WSFS Embedded (ISF-CA)* achieved the superior performance overall and the accuracy, sensitivity and specificity is above 98% in all subjects and the intra-subject variability is least. It is observed that there was least performance variation among classification methods (Bayesian and SVM), feature selection strategies (SFS and WSFS) as well as selection models (ISF, CA and ISF-CA). The extracted features of the movement and rest activity for all subjects have very good discrimination and provided excellent decoding performances.

The number of features selected for movement decoding across all subjects and their average in both classifiers and both feature selection strategies with *ISF* (a), *CA* (b) and *ISF-CA* (c) selection models are presented in figure 6.12. As shown in the figure, during movement decoding process only 15±2.3 and 13±0.6 features on average selected out of 35 features in all SFS and WSFS (ISF, CA and ISF-CA) selection models across all subjects in Bayesian classifier. However, the SVM classifier selected slightly less number of features, only 12±1.6 and 10±2.3 features on average in all SFS and WSFS selection models and achieved similar performance. Both SFS and WSFS strategies in all selection models performed very well, although the WSFS selected slightly less number of features.

Most of the features for movement decoding were selected from α, β and γ bands during and immediately after clicking. Features from high γ band consistently selected in most of the cases, and low β and low γ band selected in some cases during and immediately after the event onset. The overall influence of the features in SFS (a) and WSFS (b) feature selection strategies for movement decoding is presented in figure 6.13. From

this figure, it is revealed that both SFS and WSFS influenced by similar subset of features for movement identification.

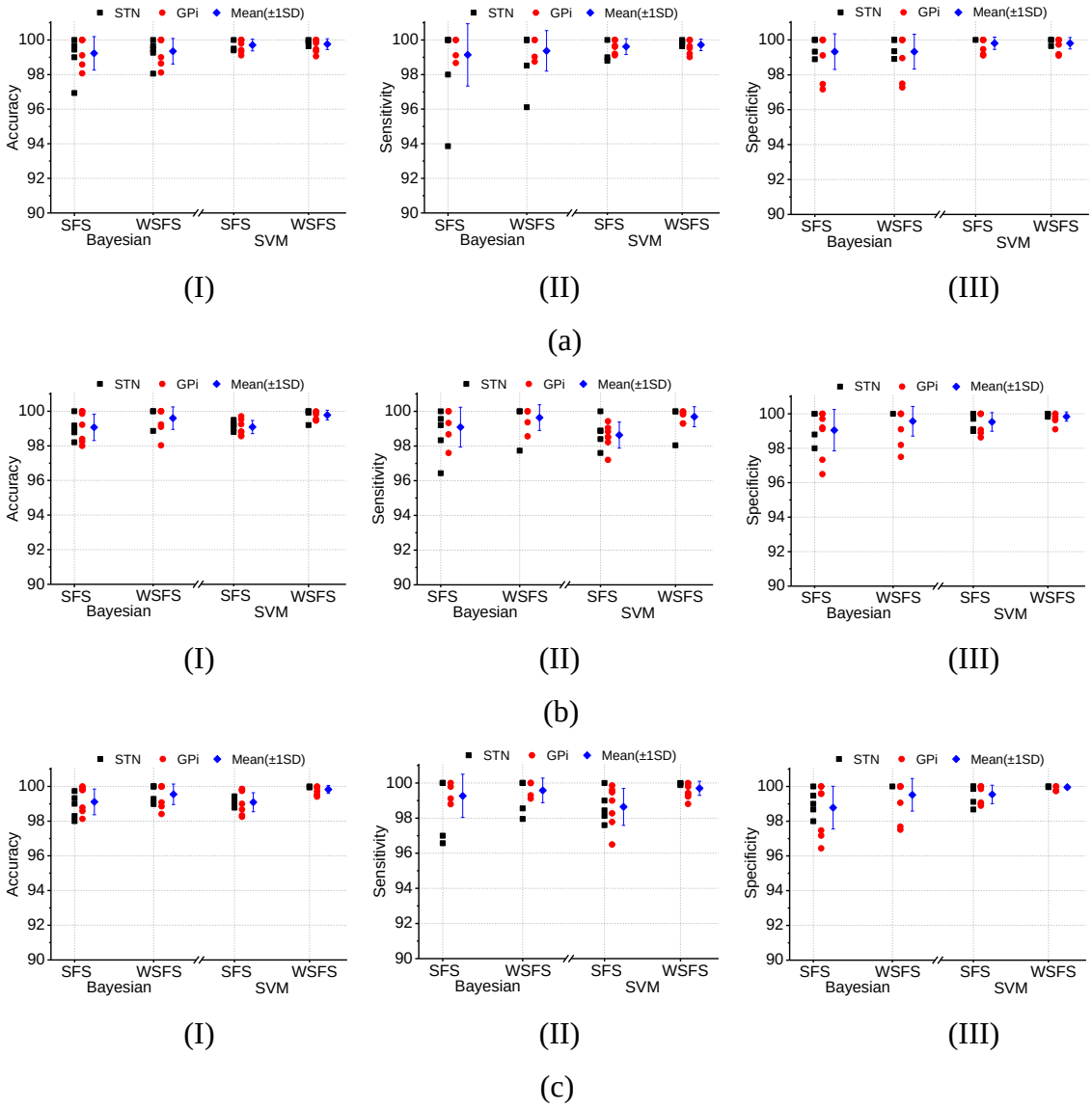
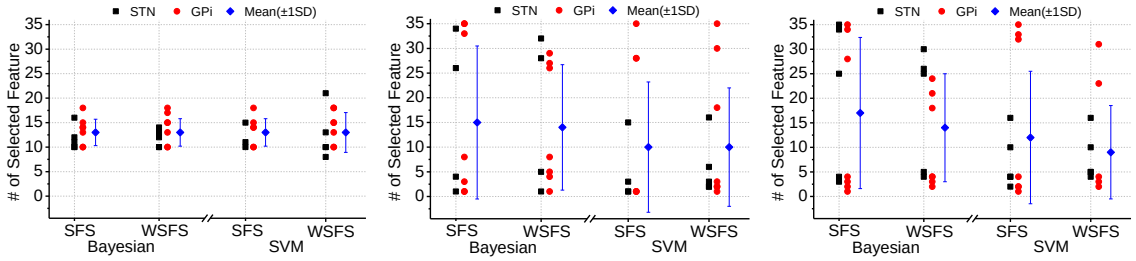


Figure 6.11: The decoding accuracy (I), sensitivity (II) and specificity (III) of movement and rest state for each subject (STN or GPi) and their average (mean ± 1 SD) using the SFS and WSFS feature selection strategies with selection model *SFS Filter (ISF)* and *WSFS Embedded (ISF*) (a), *SFS Wrapper (CA)* and *WSFS Embedded (CA)* (b), and *SFS Wrapper (ISF-CA)* and *WSFS Embedded (ISF-CA)* (c) during Bayesian and SVM classifications.



(a) (b) (c)

Figure 6.12: Number of features selected for decoding movement (event vs. rest) state for each subject (STN or GPi) and their average (mean ± 1 SD) in SFS and WSFS (with, a: *SFS Filter (ISF)* and *WSFS Embedded (ISF)*; b: *SFS Wrapper (CA)* and *WSFS Embedded (CA)*; c: *SFS Wrapper (ISF-CA)* and *WSFS Embedded (ISF-CA)*) feature selection strategies combined with Bayesian and SVM classifiers.

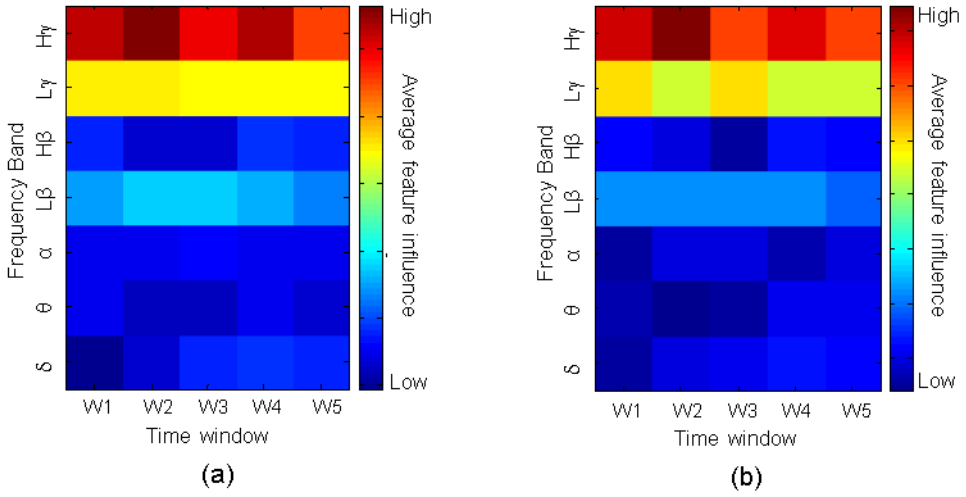


Figure 6.13: Overall influence of the features in SFS (a) and WSFS (b) feature selection strategies for movement (movement vs. rest) decoding. Overall feature influence in SFS and WSFS was computed as an average of all selected features across all subjects in both classifiers (Bayesian and SVM) and all selection models (*ISF*, *CA*, *ISF-CA*).

6.4.3 Movement laterality decoding based on instantaneous power features

After the movement decoding, movement related events were passed to the laterality decoder. Based on instantaneous power features, the movement laterality (left vs. right) decoding performance with average accuracy, sensitivity and specificity has achieved more than 61% in all feature selection models with both Bayesian and SVM classifiers. SVM classifier with *WSFS Embedded (CA)* feature selection model achieved the highest performance overall and the average accuracy, sensitivity and specificity reached to

79.8 ± 2.8%, 75.5 ± 4.3% and 82.0 ± 4.3% respectively in all subjects. SVM classifier with other WSFS selection models (*ISF* and *ISF-CA*) also achieved similar performance. The lowest performance was achieved by the *SFS filter (ISF)* with Bayesian classifier and the average accuracy, sensitivity and specificity reached to 64.6 ± 2.7%, 61.1 ± 3.5% and 67.9 ± 4.3% respectively in all subjects. The details result (accuracy, sensitivity and specificity) of movement laterality decoding based on instantaneous power features for all subjects and their average in both feature selection strategies with *ISF* (a), *CA* (b) and *ISF-CA* (c) selection models, and both classifiers are presented in figure 6.14.

It was revealed that with the same type of feature (instantaneous power), the performance for movement decoding scored much higher (~20%) than the performance for laterality decoding. With optimal selection of features, SVM classifier performed better than the Bayesian classifier for movement laterality decoding. In both classifiers, feature selection models using WSFS provided higher accuracy, sensitivity and specificity than the SFS, and maintained least intra-model (*ISF*, *CA* and *ISF-CA*) performance variability in both SFS and WSFS.

During the movement laterality decoding, average 24±7.2 and 21±4.7 features were selected out of 35 features in both Bayesian and SVM classifiers respectively with SFS strategy in all selection models. On the other hand, all of the selection models in WSFS have selected less number of features (only 14±0.9 and 15±2.3) compared to SFS strategy in both classifiers, and provided better performance. The number of features selected for movement laterality decoding across all subjects and their average in both classifiers and both feature selection strategies with *ISF* (a), *CA* (b) and *ISF-CA* (c) selection models are presented in figure 6.15. From this figure, it is observed that in SFS strategy almost every feature was selected in all subjects except *ISF* selection model. However, less than half of the available features were selected in WSFS in all selection models.

It is observed that features from δ (W4: +150-+250 ms), α (W5: +250-+350 ms), low β (W2: −50-+50 ms) and γ (W5) band were consistently selected in most cases and all other features were not consistently selected in SFS feature selection for classification. On the other hand, with the WSFS feature selection, most of the features were consistently selected from θ (W3: +50-+150 ms), α (W4), low β (W2, W3), high β (W2,

W5) and γ (W5) band during and immediately after the movement onset. Overall features from θ, β and high γ band during and immediately after event onset across all subjects were mostly influenced in the decoding process. The overall influence of the instantaneous power features in SFS (a) and WSFS (b) feature selection strategies for laterality decoding is presented in figure 6.16.

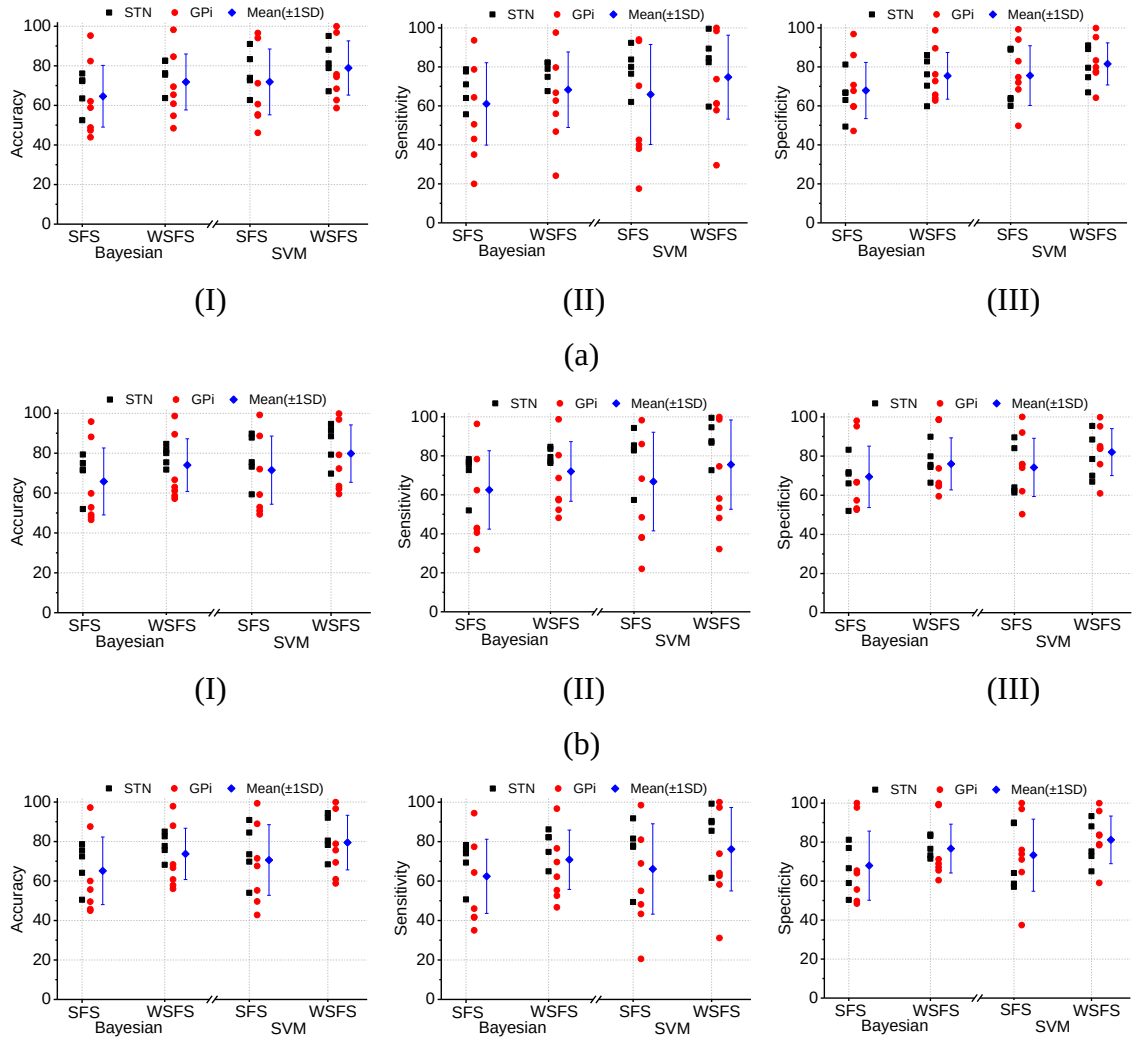
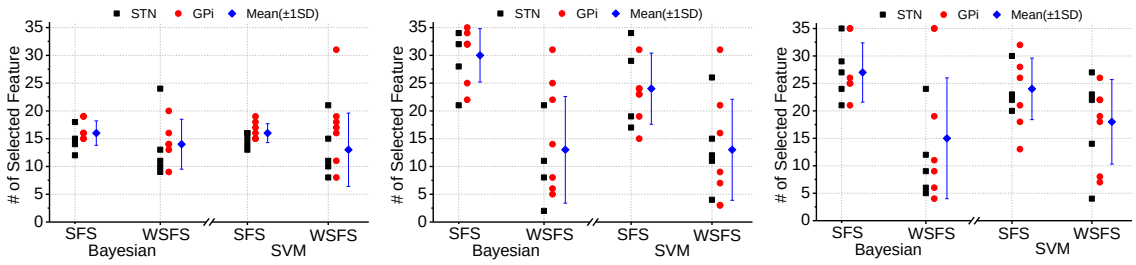


Figure 6.14: The decoding accuracy (I), sensitivity (II) and specificity (III) of movement laterality (left vs. right) based on instantaneous power feature for each subject (STN or GPi) and their average (mean ± 1 SD) using the SFS and WSFS feature selection strategies with selection model *SFS Filter (ISF)* and *WSFS Embedded (ISF)* (a), *SFS Wrapper (CA)* and *WSFS Embedded (CA)* (b), and *SFS Wrapper (ISF-CA)* and *WSFS Embedded (ISF-CA)* (c) during Bayesian and SVM classifications.



(a) (b) (c)

Figure 6.15: Number of features selected for decoding movement laterality (left vs. right) based on instantaneous power feature for each subject (STN or GPi) and their average (mean ± 1 SD) in SFS and WSFS (with, a: *SFS Filter (ISF)* and *WSFS Embedded (ISF)*; b: *SFS Wrapper (CA)* and *WSFS Embedded (CA)*; c: *SFS Wrapper (ISF-CA)* and *WSFS Embedded (ISF-CA)*) feature selection strategies combined with Bayesian and SVM classifiers.

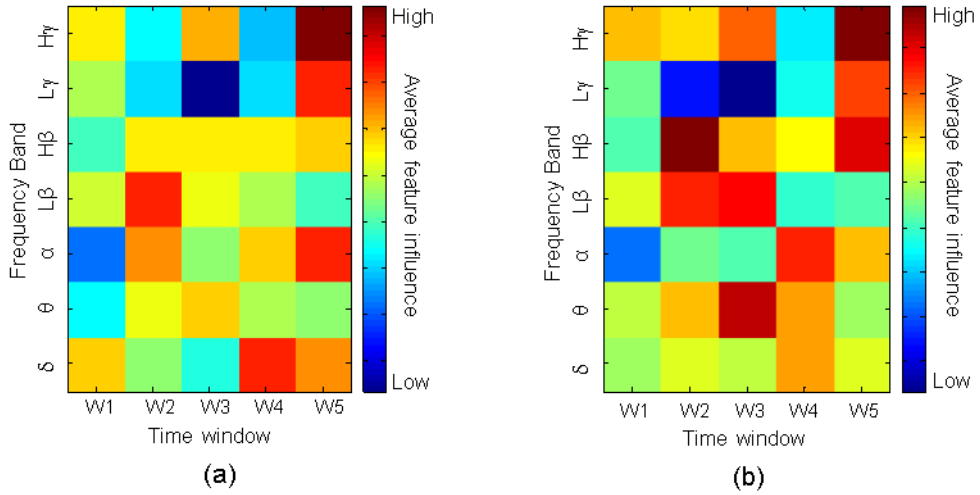


Figure 6.16: Overall influence of the instantaneous power features in SFS (a) and WSFS (b) feature selection strategies for movement laterality (left vs. right) decoding. Overall feature influence in SFS and WSFS was computed as an average of all selected features across all subjects in both classifiers (Bayesian and SVM) and all selection models (*ISF*, *CA*, *ISF-CA*).

6.4.4 Movement laterality decoding based on neural synchronisation features

When movement laterality decoding was performed using only neural synchronisation features, the performance slightly deteriorated. With optimal feature selection and classification, the average movement laterality (left vs. right) decoding performance (accuracy, sensitivity and specificity) has obtained more than 58% in all feature selection models based on neural synchronisation features. Superior performance (accuracy 77.2 ± 2.7%, sensitivity 70.7 ± 5.0% and specificity 81.7 ± 3.9%) achieved with SVM classifier incorporating *WSFS Embedded (CA)* feature selection model. SVM with all WSFS feature selection models provided similar performance, and observed minimum intra-subject variability only in *WSFS Embedded (ISF-CA)* feature selection model. The detailed result (accuracy, sensitivity and specificity) of movement laterality decoding based on neural synchronisation feature for all subjects and their average in both feature selection strategies with *ISF* (a), *CA* (b) and *ISF-CA* (c) selection model, and both classifiers are presented in figure 6.17.

As before (decoding using instantaneous power feature), it was observed that SVM classifier performed better than the Bayesian classifier in both the feature selection strategies for laterality decoding based on neural synchronisation features. However, Bayesian classifier maintained minimum intra-subject variability compared to SVM classifier in all feature selection models. Maximum decoding performance using Bayesian classifier was achieved with *WSFS Embedded (CA)* feature selection model. WSFS feature selection always provided better performance than SFS in both classifiers.

During SFS feature selection, almost all contralateral and ipsilateral neural synchronisation features were selected across all subjects in both classifiers except ISF selection model. However, only half of the features from 14 neural synchronisation features were selected on average during WSFS strategy in both classifiers. The number of features selected for movement laterality decoding based on neural synchronisation across all subjects and their average in both feature selection strategies with ISF (a), CA (b) and ISF-CA (c) selection models, and both classifiers are presented in figure 6.18.

In both SFS and WSFS feature selection, contralateral neural synchronisation features were strongly influential and more consistently selected for classification. In some cases ipsilateral neural synchronisation features were also selected. Overall contralateral α

band feature was highly contributed in both feature selections. The overall influence of the neural synchronisation features in SFS (a) and WSFS (b) feature selection strategies for laterality decoding are presented in figure 6.19.

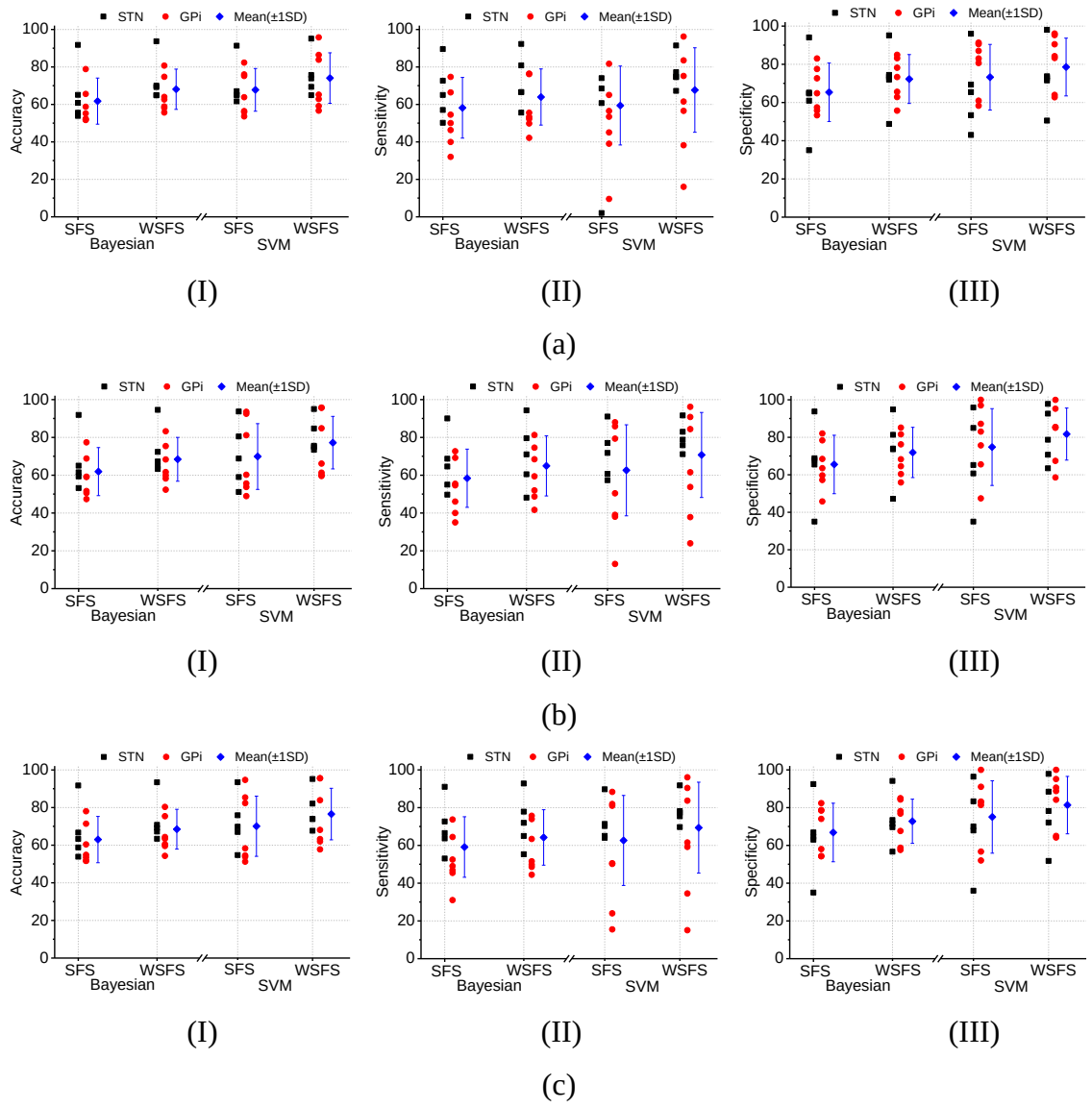


Figure 6.17: The decoding accuracy (I), sensitivity (II) and specificity (III) of movement laterality (left vs. right) based on neural synchronisation feature for each subject (STN or GPi) and their average (mean ± 1 SD) using the SFS and WSFS feature selection strategies with selection model *SFS Filter (ISF)* and *WSFS Embedded (ISF)* (a), *SFS Wrapper (CA)* and *WSFS Embedded (CA)* (b), and *SFS Wrapper (ISF-CA)* and *WSFS Embedded (ISF-CA)* (c) during Bayesian and SVM classifications.

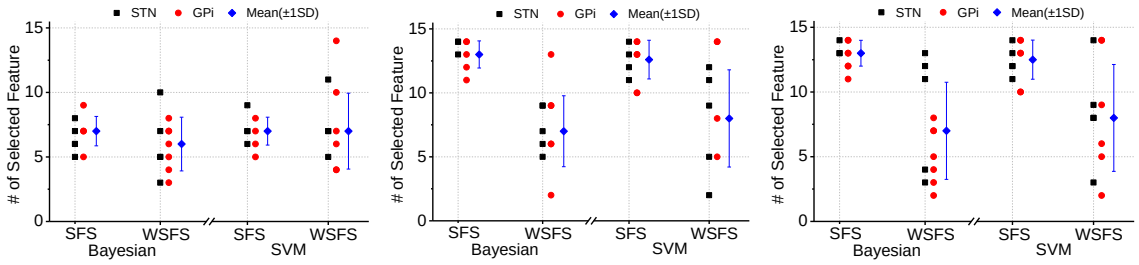


Figure 6.18: Number of features selected for decoding movement laterality (left vs. right) based on neural synchronisation feature for each subject (STN or GPi) and their average (mean ± 1 SD) in SFS and WSFS (with, a: *SFS Filter (ISF)* and *WSFS Embedded (ISF)*; b: *SFS Wrapper (CA)* and *WSFS Embedded (CA)*; c: *SFS Wrapper (ISF-CA)* and *WSFS Embedded (ISF-CA)*) feature selection strategies combined with Bayesian and SVM classifiers.

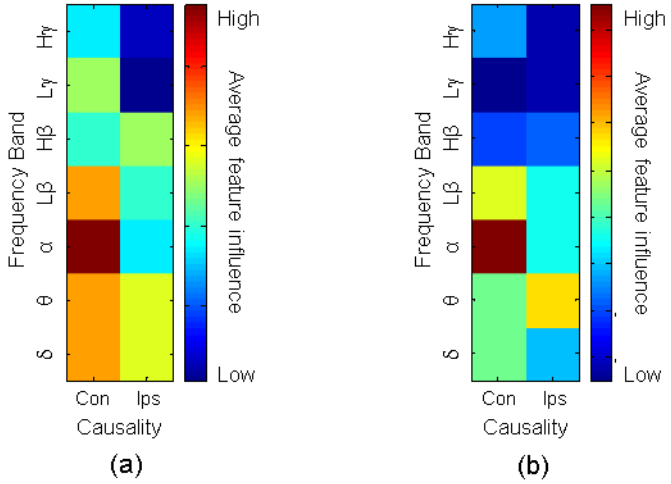


Figure 6.19: Overall influence of the neural synchronisation features in SFS (a) and WSFS (b) feature selection strategies for movement laterality (left vs. right) decoding. Overall neural synchronisation feature influence in SFS and WSFS computed as an average of all selected features across all subjects in both classifiers (Bayesian and SVM) and all selection models (*ISF*, *CA*, *ISF-CA*).

6.4.5 Movement laterality decoding based on instantaneous power and neural synchronisation features

When both the instantaneous power and neural synchronisation features combined together and were incorporated into the feature selection and classification, the decoding performance was improved. All methods have achieved good classification performance

with average accuracy, sensitivity and specificity larger than 68%. As before, SVM with WSFS achieved the superior performance overall and the accuracy, sensitivity and specificity is above 77% in all feature selection models. Bayesian classifier with all WSFS feature selection models also provided improved decoding performance and achieved more than 71% recognition accuracy, sensitivity and specificity. The performance (accuracy, sensitivity and specificity) of movement laterality decoding based on combined instantaneous power and neural synchronisation features for all subjects and their average in both SFS and WSFS strategies with *ISF* (a), *CA* (b) and *ISF-CA* (c) selection model, and both Bayesian and SVM classifiers are presented in figure 6.20.

Maximum average decoding performance (accuracy 82.5 ± 2.5%, sensitivity 78.8 ± 4.1% and specificity 85.0 ± 3.6%) achieved with SVM classifier incorporating *WSFS Embedded (ISF-CA)* feature selection model across all subjects. SVM classifier performed better than the Bayesian classifier in all feature selection models. With WSFS, Bayesian classifier maintained less performance variability across all subjects. It is observed that WSFS feature selection strategy with all selection models achieved higher performance than SFS across all subjects.

It was found that in both Bayesian and SVM classifiers, on average 33±10.1 and 30±7.7 features were respectively selected out of 49 (35 instantaneous power and 14 neural synchronisation) features across all subjects in all SFS feature selection models. On the other hand, all of the selection models in WSFS selected a lower number of features, only one third (average 17±1.1 and 16±2.8) of all available features and only half compare to SFS strategy in both classifiers, and provided much better performance. The number of features selected for movement laterality decoding based on combined instantaneous power and neural synchronisation features across all subjects and their averages in both SFS and WSFS strategies with *ISF* (a), *CA* (b) and *ISF-CA* (c) selection models in both classifiers are presented in figure 6.21.

In the selected feature subset, neural synchronisation features were more dominant and consistently selected than the instantaneous power features in WSFS feature selection. The opposite scenario was observed in SFS feature selection, with most features selected from instantaneous power. The overall influence of the instantaneous power and neural synchronisation features in SFS (a) and WSFS (b) feature selection strategies

for laterality decoding is presented in figure 6.22. The features from neural synchronisation, contralateral α, and instantaneous power, δ, θ, β bends just before and during onset, and δ, α, high β and γ bands immediately after onset were selected in SFS strategy. Again in WSFS, features from neural synchronisation, contralateral δ, θ, α, low β, and ipsilateral δ, θ, α, and instantaneous power, β bands just before and during onset, and high β and high γ bands immediately after onset were selected across all subjects. Beside this, δ and low γ band features after onset were also selected in some cases. Overall contralateral θ, α and low β features from neural synchronisation and β band features from instantaneous power were strongly influential for movement laterality decoding.

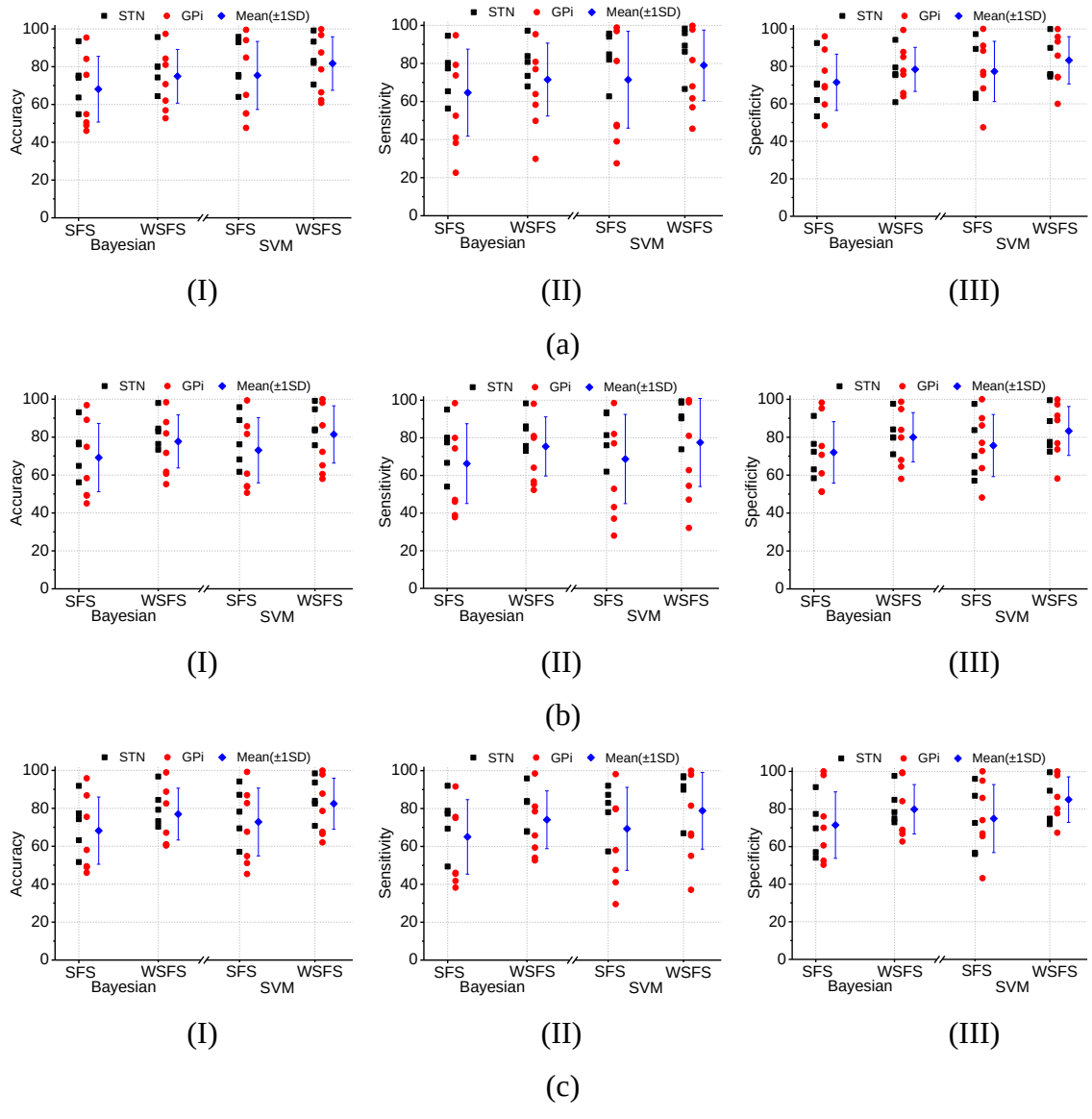


Figure 6.20: The decoding accuracy (I), sensitivity (II) and specificity (III) of movement laterality (left vs. right) based on both instantaneous power and neural synchronisation

feature together for each subject (STN or GPi) and their average (mean ± 1 SD) using the SFS and WSFS feature selection strategies with selection model *SFS Filter (ISF)* and *WSFS Embedded (ISF)* (a), *SFS Wrapper (CA)* and *WSFS Embedded (CA)* (b), and *SFS Wrapper (ISF-CA)* and *WSFS Embedded (ISF-CA)* (c) during Bayesian and SVM classifications.

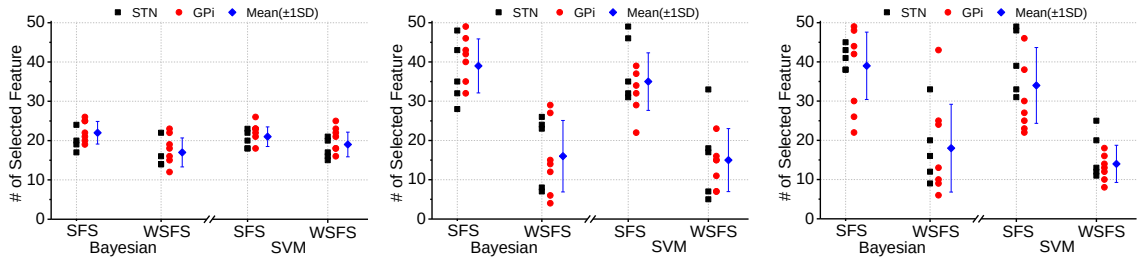


Figure 6.21: Number of features selected for decoding movement laterality (left vs. right) based on both instantaneous power and neural synchronisation features for each subject (STN or GPi) and their average (mean ± 1 SD) in SFS and WSFS (with, a: *SFS Filter (ISF)* and *WSFS Embedded (ISF)*; b: *SFS Wrapper (CA)* and *WSFS Embedded (CA)*; c: *SFS Wrapper (ISF-CA)* and *WSFS Embedded (ISF-CA)* models) feature selection strategies combined with Bayesian and SVM classifiers.

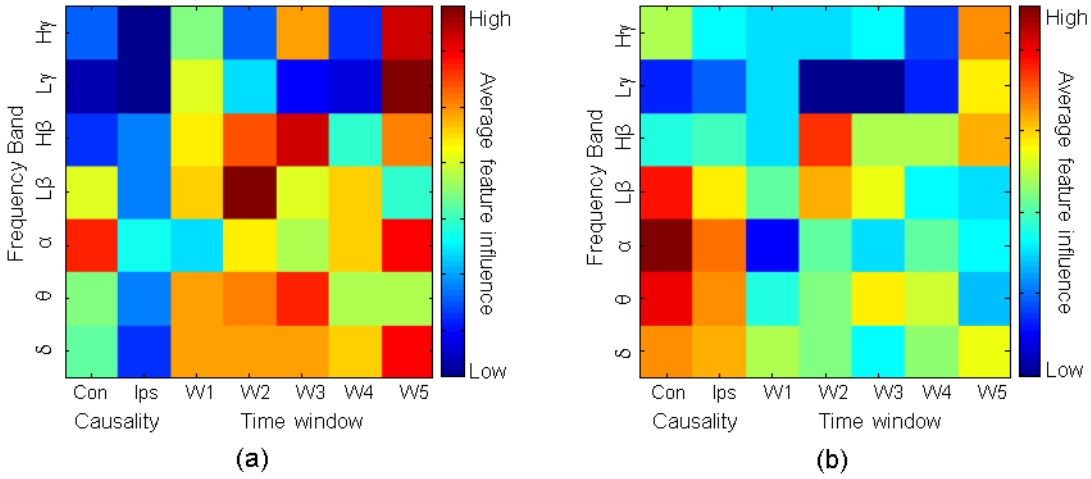


Figure 6.22: Overall influence of the instantaneous power and neural synchronisation features in SFS (a) and WSFS (b) feature selection strategies for movement laterality (left vs. right) decoding. Overall instantaneous power and neural synchronisation feature influence in SFS and WSFS was computed as an average of all selected features across

all subjects in both classifiers (Bayesian and SVM) and all selection models (*ISF*, *CA*, *ISF-CA*).

6.4.6 Decoding performance comparison

The decoding performance (accuracy) of movements (event vs. rest) was compared between the classifiers (Bayesian and SVM), between the feature selection strategies (SFS and WSFS) in all selection models (*ISF*, *CA*, *ISF-CA*), and found that there was no significant difference (figure 6.23). The extracted features for movement and rest activities have very good discrimination for all subjects. Based on the extracted features, all of the approaches provided very high classification performance.

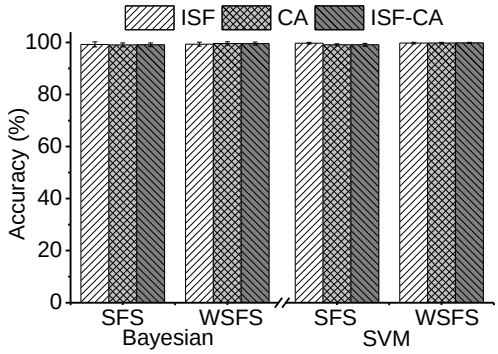


Figure 6.23: The decoding performance comparison for movement (event vs. rest) based on average (mean ± 1 SD) accuracy in SFS and WSFS (with *SFS Filter (ISF)* and *WSFS Embedded (ISF)*, *SFS Wrapper (CA)* and *WSFS Embedded (CA)*, *SFS Wrapper (ISF-CA)* and *WSFS Embedded (ISF-CA)* selection models) feature selection strategies during Bayesian and SVM classifications.

For subsequent movement laterality decoding the performance was compared within the feature space (instantaneous power, neural synchronisation and combining both together), within the selection models (ISF, CA and ISF-CA), between the feature selection strategies (SFS and WSFS) as well as between the classifiers (Bayesian and SVM) across all subjects. The details of these performance comparisons are presented in figure 6.24.

When different feature selection models are compared, it was observed that there was no significant performance difference within ISF, CA and ISF-CA selection models. All of the models achieved similar accuracy with their respective feature selection and classification methods.

When feature selection strategies are compared for movement laterality decoding, it was found that WSFS feature selection strategy performed significantly better than SFS in all cases. SFS strategy is less robust for selecting effective features and provided lower accuracy. However, WSFS achieved about 8~10% more accuracy than SFS. It is also noted that in almost all situations, WSFS strategy selected fewer (about half) of features compare to SFS and achieved significantly better accuracy with the Bayesian and SVM classifiers.

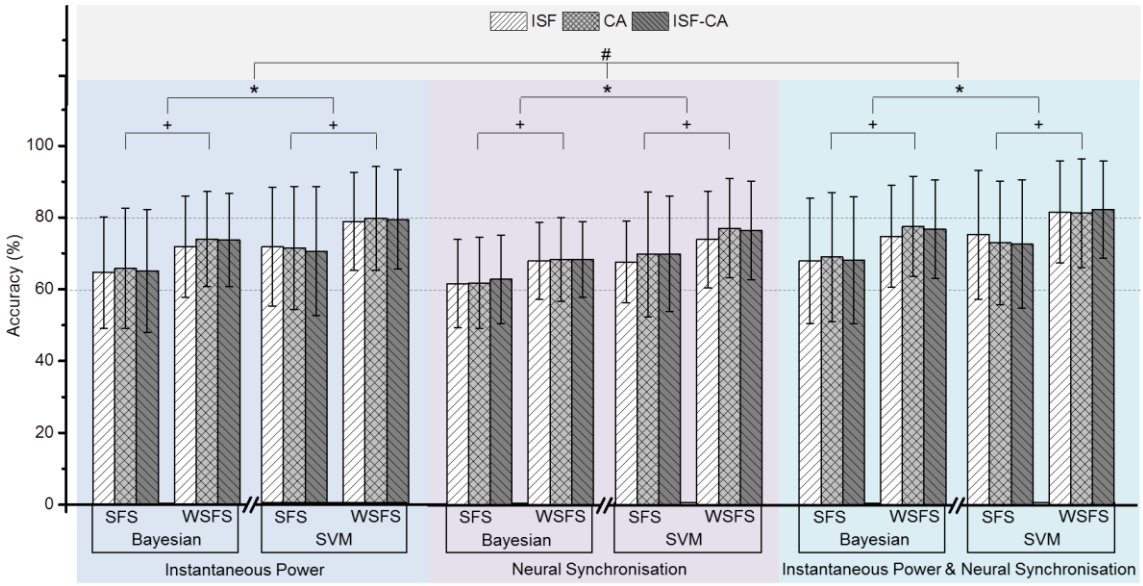


Figure 6.24: The decoding performance comparison for movement laterality (left vs. right) based on average (mean ± 1 SD) accuracy in SFS and WSFS (with *SFS Filter (ISF)* and *WSFS Embedded (ISF)*, *SFS Wrapper (CA)* and *WSFS Embedded (CA)*, *SFS Wrapper (ISF-CA)* and *WSFS Embedded (ISF-CA)* selection models) feature selection strategies during Bayesian and SVM classifications. **+** $p<0.05$, SFS vs. WSFS, * $p<0.05$, Bayesian vs. SVM, # $p<=0.05$, both instantaneous power & neural synchronisation vs. instantaneous power, neural synchronisation feature.

When the performance is compared between the classifiers, it was observed that the SVM classifier significantly outperformed the Bayesian classifier for all evaluations (instantaneous power, neural synchronisation and their combination) in the feature space. Also SVM selected fewer features than the Bayesian classifier. However, Bayesian classifier always maintained less intra-subject performance variability than SVM.

When types of features are compared within the feature space, it was revealed that combined instantaneous power & neural synchronisation features performed moderately better than the other two evaluations, i.e. when they evaluated separately. Combined instantaneous power & neural synchronisation features were improved the performance of movement laterality decoding and achieved superior decoding accuracy with WSFS feature selection and SVM classification. Finally the maximum accuracy (82.5%) was achieved in SVM classifier with *WSFS embedded (ISF-CA)* selection model based on combined instantaneous power and neural synchronisation features for decoding movement laterality.

6.5 Discussion and conclusions

Deep brain stimulation (DBS) offers a unique opportunity to sense and control human brain circuits. The function of the basal ganglia and the thalamus in motor control has been studied with invasive recordings of local field potentials (LFPs) from the DBS stimulation macro-electrodes in patients with Parkinson's or other neurological diseases. In this study recorded LFPs from basal ganglia (STN or GPi) were investigated to decode voluntary movements and their laterality. Two types of feature, instantaneous power and neural synchronisation were extracted from LFPs. With efficient feature selection and parameter optimisation during training, both Bayesian and support vector machine classifiers achieved high decoding accuracy to recognise movements (event vs. rest), and good accuracy to recognise movement laterality (left or right sided movements) using either SFS and WSFS feature selection strategies with *ISF*, *CA* and *ISF-CA* selection models. Movement decoding based on instantaneous power features achieved high performance; however using a similar set of features, the movement laterality decoder provided less accuracy. To improve the performance of laterality decoding neural synchronisation features were extracted by analysing Granger causality. When both features were applied together, the performance of movement laterality decoding was improved in both classifiers. The SVM classifier outperformed the Bayesian classifier in all cases. Again, during optimal feature selection, WSFS performed significantly better than SFS in both classifiers. Overall the SVM classifier with the WSFS feature selection strategy achieved the highest performance for decoding movement (99.8%) and its laterality (82.5%).

The wavelet transform was previously used to extract features for predicting movements from deep brain STN LFPs and found to be superior for information enhancement compared to the FFT and spectral measures (Loukas & Brown 2004). It was also reported in the literature that wavelet transformation methods are better tools for information enhancement from neuro-electric signals than conventional methods (Samar et al. 1999). Particularly within the wavelet methods, WPT is an efficient tool for precisely localising time-frequency information from non-stationary neural signals. It has also established in our TMEP signal classification work that WPT is an efficient way to extract signal information (cf. chapters 4 and 5). Based on the literature and our knowledge, WPT was selected for frequency band envelop extraction from deep brain LFPs incorporating the Hilbert transform. To achieve optimal performance, the wavelet decomposition scale was empirically determined according to the standard frequency bands for neural signals. Based on the extracted frequency band envelope, the computation of instantaneous power features was performed by averaging amplitude in defined windows. To select the optimal window, it was evaluated by varying window length (100-ms and 50-ms) as well as varying the positions from -500 ms to +500 ms. The best discriminative features to achieve maximum decoding accuracy were observed when features came from five consecutive 100-ms windows in each frequency band from -150 ms to 350 ms around the motor response. It was also found that in some cases 50-ms window features provided comparable discrimination. The laterality decoding performance was improved and achieved accuracy of 82.5%, when neural synchronisation features were incorporated (figure 6.24).

In the current analysis, neural synchronisation features were computed by evaluating the causal strength between LFPs of left and right STN or GPi using Granger causality. This can be named as a network or long range neural synchronisation. The features were defined as contralateral and ipsilateral neural synchronisation in each frequency band for both left and right events, and it was found that contralateral features are more dominant than ipsilateral features and strongly contributed to the classification during movement laterality decoding (figure 6.19).

Evaluation of other neural synchronisation features, such as local synchronisation (LFPs between the electrode channel within the STN or GPi), cross modality (cross frequency band, phase, power-phase) synchronisation may provide more discriminative features to improve the decoding performance. In the current analysis, only one channel out of

three from each DBS electrode was selected for decoding. Increasing the number of channels could enhance information about the neural activity related to the events. However, there is a limitation in electrode placement to get good recording from all available channels. On the other hand, neural synchronisation features were evaluated by analysing bivariate Granger causality. Evaluation of spectral Granger causality of LFPs in different frequency bands may enhance more directional information and might improve the performance of decoding movement laterality (Wang et al. 2007; Seth 2010). In the present decoder, causal strength was obtained based on segmented LFPs with fixed MVAR model order across all subjects. However, dynamic selection of MVAR model order for each subject based on AIC criterion for causal strength evaluation from continuous LFPs could enhance more discriminative information (Seth 2010; Brovelli et al. 2004). It is expected that a subject specific MVAR model may provide better performance compared to a fixed model.

In the SFS feature selection, SFS filter and wrapper approaches were evaluated based on mean ISF score and cumulative CA respectively to select a subset of features according to their rank. To rank each individual feature, ISF score or CA ranking criterion were used. By introducing feature weight based on the effective contribution of each feature during repeated evaluation of SFS with classification accuracy criterion, a new WSFS strategy was developed and used to select a consistent and optimal feature subset. WSFS is considered an embedded feature selection approach. It performed better than the SFS in all ISF, CA and ISF-CA selection models through selecting the most discriminative features, and avoiding redundant features for movement and laterality decoding using either Bayesian or SVM classifier. In all selection models, WSFS selected fewer features compared to SFS and achieved better performance. This was due to its ability to remove the least contributing features that affect the classifier generalisation or make poor estimation of classifier parameters. Only in the ISF selection model, the number of selected features in SFS is very close to WSFS. This is because the initial selection was performed based on mean ISF score and it is independent of classification method. As SFS is a greedy search approach, in CA and ISF-CA selection models, it selected a large number of features that included redundant features and poorly estimated the classifier parameters. Due to this, SFS selection based classifier provided lower accuracy, sensitivity and specificity compare to WSFS. It is also noted that the WSFS strategy can select optimal subset of feature from high

dimensional, small sample size datasets, and its ability in this aspect is already validated by evaluating classification performance of subjects, whose LFP dataset was very small (25 trials) and has large number of features (84 dimensions), and achieved high accuracy (SVM with WSFS (84.0%) vs. SFS (68.2%)). To validate the WSFS strategy in a broader perspective, further evaluation is needed such as evaluation of the feature selection strategy in standard high dimensional datasets as well as other feature ranking criteria, also comparing its performance with other state of art feature selection strategies.

Both classifiers, Bayesian and SVM performed very well in movement and laterality decoding. SVM outperformed the Bayesian classifier, however the complexity of the SVM is very high and also it has some drawbacks such as estimation of optimal parameters, model over-fit (Lotte et al. 2007; Mamun et al. 2012; Gacquer et al. 2011). On the other hand, a Bayesian classifier is simple and can generalise with a small training set of data, and provide reasonable accuracy (Mamun et al. 2012; Gacquer et al. 2011). It is also efficient for real-time application (Mamun et al. 2012; Alzoubi et al. 2008). The SVM is more sensitive to the size of training set than the Bayesian classifier. With small training size an SVM classifier can over-fit to the training data and therefore may provide poor generalisation during real-time testing (Gacquer et al. 2011; Alzoubi et al. 2008). According to the trade-off between the decoding performance and complexity, SVM is the better choice during offline evaluation, however to design an online decoder for classifying movement activity, a Bayesian classifier would be preferred due to its simplicity.

From figure 6.5 and 6.6, it was observed that the δ band amplitude was quite high during the event onset compared to α, θ and γ bands, nonetheless the contribution of δ band feature in instantaneous power domain was not as strong as other frequency band features for decoding movement laterality across all subjects. This is because it didn't carry much discriminative information for decoding left and right movements. In the neural synchronisation feature, the contralateral α band feature strongly contributed to laterality decoding (figure 6.19). For movement and rest state decoding the high γ feature was strongly dominant and features from α, low β and low γ during onset (-150 ms to +150 ms) were also selected in most cases (figure 6.13). In the movement laterality (left vs. right) decoding based on both instantaneous power and neural synchronisation features, most of the features contributed from neural synchronisation

with contralateral θ, α and low β band features were more consistent (figure 6.22). Ipsilateral α and low β also contributed in some cases. Most of the instantaneous power features were selected from θ, β, and γ bands with strong contribution from high β during and immediately after event onset. It is noted that the overall influence of the frequency band features observed in this movement decoding study also broadly reflected the previous finding of movement related frequency bands synchronisation in basal ganglia (Brown & Williams 2005; Loukas & Brown 2004; Alegre et al. 2005).

High decoding performance was achieved across all subjects during movement classification, however for subsequent laterality decoding all subjects did not perform as well as movement classification. Several issues accounted for these performance degradations, (1) unbalanced number of trials between the class dataset, i.e. one class has large number of trials and other has only few, (2) unbalanced variability within the class dataset, i.e. the variance within the class is very large across trials, (3) unbalanced feature variation, i.e. magnitude of one feature is very high and another feature is very low, the low magnitude feature may be ignored during feature selection, however it may contain significant information contributing to the classification. Other issues like high magnitude variation across frequency bands as well as subject's ability to concentrate and perform action according to stimulus, and subject's disease severity and ability to perform actions also affected the decoding performance. On the other hand, in most of the subjects, the number of trials for classification was less than the number of features, the average number of trials for each class across all subjects was 58±23.6 with minimum of 25 trials and maximum of 113 trials, but the features have 84 dimensions. To estimate the classification model accurately for robustness, a larger number of trials is needed (Lotte et al. 2007; Jain & Duin 2000). One or more of the above issues existed in the LFP dataset in several subjects and which prevented obtaining better decoding performance. However, WSFS feature selection strategy overcame the issues of high dimensional, small sample size datasets and improved the decoding performance. The effective solutions of one or more drawbacks are still under research in machine learning and further investigation is needed to overcome these issues related to problem specific applications (Golugula et al. 2011; Imam et al. 2006). The problem of unbalancedness in the dataset also prevented the selection of consistent features across the subjects, and WSFS strategy enhanced selection of consistent feature based on maximum contribution and provided better decoding accuracy.

One notable observation is that during the movement laterality decoding with most of the subjects, the detection of right hand clicking (average specificity, 85.0 ± 3.6%) task performed better than left (average sensitivity, 78.8 ± 4.1%). It is assumed that most of the subjects participated in the experiment were right handed, and they are very familiar or well trained to use their right hand compare to left hand in daily life. This could be the reason for getting high specificity than sensitivity. It was also shown in the literature and also found in TMEP signal classification (cf. chapter 5) that a well trained activity response is more stable and detectable than a less trained one. Thus, for the case of right clicking, based on the stimulus, subjects effectively planned and quickly acted to the given stimuli and accordingly neural reaction of the STN or GPi was synchronised. But effective planning and performing actions for less trained activities (left clicking) was slow and therefore the synchronised recorded LFPs were less consistent and have nominal information. However, training could overcome such limitations.

Currently the information of the patients disease conditions such as diseases severity, handedness are not accessible, therefore it is not possible to perform the correlational analysis of the decoding performance with disease conditions. However based on the available demographic data (cf. table 6.1), it was observed that patients with longer disease duration (more than 25 years) achieved lower decoding accuracy, and this may reflect the influence of disease progression on the decoding performance. In some cases patients with older age (more than 53 years) also achieved lower decoding accuracy. According to disease types, patients with PD achieved much higher laterality decoding accuracy on average compare to patients with dystonia (87.6 ± 10.5% vs. 76.8 ± 14.3%). Similarly, LFP activity recorded through DBS electrodes from STN achieved higher average accuracy than GPi (85.5 ± 10.3% vs. 79.8 ± 15.3%). Finally, as mentioned earlier it was observed that movement laterality decoding for right hand clicking on average achieved higher accuracy than left hand clicking. Even though we do not have handedness information of patients, with the prevalence of right-handedness across the human populations, it can be considered that right handed patients were better trained to perform an action by their right hand.

The primary motor cortex is the most important brain area for the control of voluntary movements. Decoding of motor activity based on neural signals recorded from primary motor cortex is currently regarded as the gold-standard for neural interface applications. The majority of motor decoding studies were performed in this area (Bashashati et al.

2007; Ince et al. 2010). However, for practical reasons it would be unwise to focus exclusively on this part. Acquisition of neural signals from other areas of the brain that are involved in motor operation could provide substantial or additional information, which may advance our knowledge to develop a more reliable decoder for real life applications (Pasqualotto et al. 2011). The investigation of STN or GPi LFP decoding is one such evaluation and it has been found that basal ganglia LFPs carry considerable movement related information (Brown & Williams 2005). The decoding methods developed in this study performed well to recognise movement and rest states and subsequent movement laterality (left vs. right) based on the extracted information from deep brain LFPs recorded through DBS electrodes. In future, such decoding methods integrated into DBS-driven neural interfaces should help to develop an advanced BMI system, which might be useful for patients in several pathological states. For example, it might help patients with spinal cord injuries in learning to control their limbs or prostheses (Gasson et al. 2005; Kringelbach et al. 2007). The prospect of the research in this area was also manifested by Kringelbach et at. (2007) in their publication "Translational principles of deep brain stimulation" (Kringelbach et al. 2007) and they stated that

> *"DBS-driven brain–computer interfaces might in the future modulate brain activity in order to help individuals in vegetative and minimally conscious states."*

In this study, LFPs from each subject was considered as a single session recording for analysis, although some subjects had two-sessions of recording. To evaluate the stability in decoding of basal ganglia LFPs for motor activities, future work is needed to record LFPs in a number of sessions and decoding should be performed by testing data from one session and training with other sessions (Ince et al. 2010). Investigation of decoding performance in other conditions such as self-paced and imagined movements from basal ganglia LFPs will advance our understanding and facilitate its application towards development of neural interface system for BMI. Nevertheless, the investigation to predict and recognise voluntary and involuntary (such as Parkinson tremor) movements from deep brain LFPs may enable future development of intelligent devices (for instance, demand driven deep brain stimulations) for advancing therapeutic applications.

In summary, the contributions to knowledge introduced in this chapter for decoding of movements from deep brain LFPs are as follows:

- Movement activities of voluntary movements recorded from basal ganglia STN or GPi were successfully decoded with high accuracy. These will enhance our understanding of human brain circuits related to motor operation and also facilitate developing advanced neural interface system for brain machine communications.
- Extraction of neural synchronisation features was introduced by analysing Granger causality, which strongly contributed in decoding and improved the performance when incorporated with instantaneous power features.
- A new feature selection strategy, WSFS was developed to select optimal and consistent feature subsets and the decoding performance was significantly improved in both movement and laterality classifications.
- Both SFS and WSFS feature selection strategies were also evaluated in three selection models (ISF, CA and ISF-CA) and it has been found that the WSFS is robust to eliminate features that carry least information for classification.
- Typical results from twelve subjects were evaluated and achieved high accuracy, which proved the success of the decoding method.

Chapter 7 : *Conclusion and Future Work*

7.1 Summary

This project aims to design signal decoding algorithms for identification of movement related states from biosignals in order to improve the robustness, accuracy and reliability of human machine interface (HMI) and brain machine interface (BMI) systems. The principle of HMI and BMI systems is to translate a user's intention through signal processing and pattern recognition algorithms into action commands to operate devices. Among various types of biosignals investigated in the scientific literature, TMEP signals related to tongue motion have promising potential for people who are not able to move their limbs to communicate with their environment. However, a person with severe movement impairment such as locked-in syndrome has limited options to communicate with their environment and one potential option is neural signal related to motor activity. To address the above conditions, HMI or BMI might be a possible technology to achieve such communication goal. In both TMEP and neural signals based systems, the translation algorithm plays a vital role to decode user intention for communication. To achieve the goal of this project, we have developed robust biosignal translation algorithms by integrating feature enhancement and classification methods to efficiently decode movement related states from TMEP signals and deep brain LFPs.

Real-time translation of TMEP signals requires two steps, first it needs to segment TMEP signals to identify the control tongue movements related signal from the background signals and distinguish these from natural tongue movement signals such as coughing, drinking etc. The second step is to decode motor commands from TMEP segments identified as control tongue movements, for instance, left, right, up or down actions. This thesis addressed both of these issues and developed feature extraction methods by utilising wavelet packet transform to localise discriminative information in the time-frequency domain for classification. The classification is performed using

Bayesian method, a simple and linear classifier, and support vector machine, a complex and nonlinear classifier. To test the performance of tongue movement action identification from interferences, a dataset from ten subjects was collected and evaluated. The identification methods achieved high accuracy for identifying tongue actions and interferences in offline and online situations.

When considering the neural signal translation to decode the voluntary movements from deep brain LFPs, the problems are more challenging than the TMEP signal translation due to the complex characteristics of the neural activity. The recorded neural activities of LFPs are highly non-stationary and have high variability. They also contain highly redundant neural information. On the other hand, due to less concentration and training, the discriminative information may be insufficient. Another challenging aspect is the size of the dataset, which is usually limited. Considering the inherent properties of the neural signals, a neural decoding algorithm has been developed in this thesis to subsequently decode movement and its laterality from deep brain LFPs. The decoding algorithm consists of three stages, i.e. feature extraction, feature selection and classification. In the feature extraction, frequency dependent components of LFPs were extracted using the wavelet packet transform, and then features at each frequency bands were extracted as instantaneous power and neural synchronisation measure using Hilbert transform (with computing event related amplitude desynchronisation and synchronisation) and Granger causality respectively. It was found that not all of extracted features contained high discriminative information for classification, and some of them had little discriminative information, which degraded the classification performance. To efficiently utilise the most discriminative features in the feature space, a new feature selection strategy, weighted sequential feature selection (WSFS) was developed in this thesis. The WSFS strategy was evaluated to select optimal feature subset of deep brain LFPs. It significantly outperformed the existing sequential feature selection (SFS) methods in respect of decoding accuracy. Finally based on the optimal feature subset, the classification of movement and laterality was implemented using Bayesian and support vector machine classifiers. The decoding algorithms were evaluated in twelve subjects and showed high accuracy for movement and rest classification, and good accuracy for laterality (left and right) classification. It was also observed that integration of WSFS strategy for feature selection significantly improved the decoding performance in both movement and laterality.

The signal transformation, feature extraction and selection, and classification algorithms for TMEP and deep brain LFP signals decoding can be applicable to other biosignals, for instance, EEG or EMG based HMI or BMI. Particularly, the feature reduction and selection method of WSFS can be useful for high dimensional, small sample size datasets.

7.2 Discussion

Generally biosignal decoding problem is challenging due to the properties of non-stationarity as well as inconsistency that creates large intra-class and small inter-class variability in the feature space. In many cases, limited or small size datasets are available to design and evaluate the biosignals decoding algorithm. Again artifacts or noise in a real environment makes the decoding problem more complex. On the other hand, each type of biosignal has its own unique characteristics that need to be adapted for designing robust decoding algorithms for a specific biosignal, such as TMEP or neural signals. According to the characteristics of neural signals, the recorded neural activities such as LFPs are more complex (highly non-stationary) in nature and it reflects the integration of multiple complex brain functions and provides minimal information about the underlying neural activities. Due to this complexity, neural signal decoding problems are more challenging to extract information, such as movement related activity. It requires large amount of feature extraction from single or multi-channel recording to decode neural activities. In the feature space, extracted features are highly redundant and provide less discriminative information. On the other hand, the unbalanced properties in the dataset as well as in the feature space (cf. section 6.5) reduce the robustness of classification.

During the design of biosignal decoding algorithms for TMEP and deep brain LFPs, we addressed the challenging issues for improving accuracy and robustness to identify movement related states. The properties of TMEP signals are less complex than the LFPs and the developed TMEP decoding algorithm achieved high accuracy (cf. chapter 4 and 5). On the contrary, the developed LFP decoding algorithm achieved less accuracy (cf. chapter 6). There were several issues mentioned above that were observed in the LFPs data and affected the decoding performance, which include complexity in the signal characteristics, limited size of the dataset, unbalanced properties in the dataset, large dimension of features with redundancy in the feature space, and also quality of the data as

well as the patients disease severity and ability to respond or act to the stimulus accordingly.

7.2.1 Sample size and feature dimension

Particularly in LFP decoding, the size of the dataset is limited and also unbalanced compare to TMEP dataset. TMEP dataset contains comparatively a large number of trials (100) for each action class and the amount of data across all subjects is similar (400 trials for tongue movement actions and 360 trials for different tongue movement interferences), also the datasets were collected from healthy subjects with age range 18-45 years (30.7±6.4; mean±1 SD). On the contrary, LFPs datasets were collected from Parkinson's and dystonia patients with age range 23-72 years (49.6±13.9) and disease duration 3-38 years (14.8±10.3). The collected LFPs datasets are limited in size (average 115±43.6 trials with a minimum of 56 and a maximum of 202 across all subjects), and the number of trials in each class (left or right) is unbalanced in most of the subjects. The average number of trials for each class of LFPs data across all subjects is 58±23.6 with a minimum of 25 trials and a maximum of 113 trials, and the average difference between the classes across all subjects is 14.2%±19.0 trials with a minimum of 1.2% and a maximum of 57.6%.

In the feature space, the extracted feature dimension in the LFPs is higher than the TMEP signals. The extracted features for TMEP action classification have 75 dimensions and TMEP action identification from interference have 16 dimensions. In both TMEP signal decoding tasks, the average size of datasets for each class is higher than the feature dimensions. Conversely, the feature dimension for LFP decoding is large, as well as it being greater than the size of the training dataset in many subjects (84 dimensions in the feature space and an average dataset size for each class is 58±23.6). Generally to design the generalised and robust classifier to reduce the classification error, the recommended training set size per class is more than (at least, five to ten times) the dimensionality of the features (Raudys & Jain 1991; Lotte et al. 2007). Unfortunately it is not possible to obtain such large dataset to train the classifier for biosignals decoding in many cases.

To define the training and testing sets, repeated hold-out was used for TMEP and cross-validation was used for LFP decoding. For small or limited size dataset, cross-validation is a good approach for selecting training and testing set to estimate the generalisation error. To estimate the generalisation error for classification, both approaches require

repeated evaluation during offline processing and which is are not suitable for online evaluation.

7.2.2 Feature extraction

Previously, investigated TMEP actions classification extensively used time domain features and reported good performance. However, when signals are contaminated with noise, time domain feature fail to maintain good performance (cf. section 4.4). Therefore robust feature extractions such as a wavelet based approach can overcome such limitations (Addison et al. 2009; Samar et al. 1999). This thesis has already demonstrated how efficient wavelet based techniques are able to enhance movement information from biosignals. The wavelet based approach is able to successfully extract the TMEP signals feature, and based on the extracted features, all classification methods achieved high performance for both tongue movement action identification and classifications.

The wavelet based approach also efficiently isolated the different frequency dependent components from LFPs and based on each frequency band, instantaneous power and neural synchronisation features were computed. Instantaneous power features from each frequency band provided very good discrimination for decoding the movement and rest LFP activities. On the other hand, for movement laterality, left vs. right decoding instantaneous power features have not maintained good discrimination between the classes, hence achieved poor performance (Mamun et al. 2011). However, when neural synchronisation features (computed by analysing Granger causality) were combined with instantaneous power features, the performance was improved (cf. section 6.4.6). It can be noted that during the instantaneous power feature computation in each frequency band, window averaging was used. Due to this averaging, precise local time-frequency information was reduced. Instead of averaging, the wavelet packet transform (WPT) coefficient could be selected as a feature that might increase the interclass separability. Consequently direct computation of WPT coefficient as neural features will makes the classification problem more difficult due to the curse-of-dimensionality (Gupta et al. 2010). In the extraction of neural synchronisation features, bivariate Granger causality provided good discrimination. Nonetheless, it is the first evaluation for decoding LFPs based on such features; evaluation of spectral Granger causality could enhance more discriminative neural information to provide higher inter-class separability (Wang et al. 2007; Seth 2010).

In the LFP investigation, most of the cases it is considered that the very low frequency band (δ) does not carry event related information and hence this was usually filtered out during frequency band pre-pressing. However, a recent study reported that the δ band carries significant information about events (Ince et al. 2010). This evidence was partially observed in our analysis (as we extracted δ band features) and some cases δ band features in both instantaneous power and neural synchronisation contributed during classification (cf. section 6.4).

7.2.3 Classification

In both the TMEP and LFP signal decoding problems, the supervised classifiers, Bayesian and SVM were evaluated. Bayesian is a generative classification method to compute the likelihood of each class and choose the maximum likelihood to classify a feature vector. In contrast, SVM is a discriminative classification method that computes the hyperplane to classify a feature vector (Gacquer et al. 2011; Lotte et al. 2007). SVM is considered as a regularised classifier which has good generalisation capabilities and also is robust with respect to outliers.

Both the Bayesian and SVM classifiers performed well for identifying movement related states from TMEP and LFP signals. During the offline evaluation, SVM outperformed the Bayesian classifier, however the complexity of the SVM is very high and also has some drawback such as selection of kernel function, estimation of optimal parameters, model over-fit (Lotte et al. 2007; Mamun et al. 2012). On the other hand, the Bayesian classifier is simple, computationally efficient and can generalise with a small training set of data. It has also provided reasonable accuracy for real-time identification of tongue movement actions from varied interference setting (Mamun et al. 2012). The SVM is more sensitive to the size of training set than the Bayesian classifier. With small training size, in some cases the SVM classifier can over-fit to the training data and therefore provides poor generalisation during real-time testing (Lotte et al. 2007; Mamun et al. 2012). On the other hand, the classifier that works well offline is not necessarily the classifier that performs better online, although such comparisons are difficult due to the lack of training data for the classifier model optimisation and the different testing sets in online evaluation.

One advantage of Bayesian classifier is that the outcome of the classification is a probabilistic value rather than the true or false decision. This implies that by setting

classification thresholds (or analysing receiver operating characteristic (ROC)) a false positive rate could be minimised. Also a computationally inexpensive classifier simplifies the classification process. Considering the problem-specific applications, the Bayesian classifier could be a good option for classification of biosignals in a real-time system. On the other hand, SVM is a non-probabilistic classifier and it is good for dealing with binary classification problem. For multiclass, it breaks up the problem into several binary classifications (one class versus others), which is computationally expensive (Ubeyli 2007). The selection of an appropriate SVM kernel function seems to be a trial-and-error process. One would not know the appropriateness of a kernel function and performance of the SVM until it has been evaluated on a respective dataset. Again selection of optimal or near optimal values for SVM parameters (the regularisation parameter C and the RBF kernel parameter η in our case) is very critical in order to have a properly trained SVM. It is also an experimentation process by changing different values until to have the best result (i.e. accuracy). However, overcoming the drawbacks of SVM classifications could provide better solutions for decoding biosignals especially neural signals.

7.2.4 Curse-of-dimensionality

As mentioned earlier (cf. section 3.4), the curse-of-dimensionality is an issue for classifier generalisation to decode biosignals (Raudys & Jain 1991; Lotte et al. 2007). Due to the curse-of-dimensionality, both Bayesian and SVM classifiers were less robust in some cases and did not achieve high performance, especially during LFP decoding (Mamun et al. 2011). Nevertheless, feature selection or dimensionality reduction method can eliminate the redundant features from the feature space. In many cases integration of feature selection (such as sequential feature selection) or dimensionality reduction (such as PCA) with the classification methods can overcome the curse-of-dimensionality and improve the robustness of the decoding, as well as reducing the complexity of the classification problem (Gupta et al. 2010; Bashashati et al. 2007). The sequential feature selection (SFS) is the most commonly used for biosignals feature selection due to its simplicity (Saeys et al. 2007). The major limitation in the SFS strategy is that it is biased towards the given training set. When the training set changes, its performance varies, which is not robust for high dimensional neural signal features. To reduce the dimensionality from the feature space, PCA is widely investigated. PCA is a linear transformation and can be used for dimensionality reduction by keeping lower-order

principal components, but it is not optimised for class separability (Bashashati et al. 2007). Also it may be less robust for larger dimensional and small sample size datasets.

Given the problem of small datasets with a large dimension of features as well as unbalanced dataset in LFPs, the integration of a feature selection approach significantly improved the LFP decoding performance in most cases. Practically, our newly proposed feature selection methods, weighted sequential feature selection (WSFS) efficiently eliminated the redundant or least discriminative features from the feature space and significantly improved the decoding performance for identifying movement laterality from LFPs (cf. section 6.3.2.5 and 6.4.6). It also significantly outperformed the existing SFS method by selecting fewer features (in most cases it only selected half the number). The only limitation of WSFS is that it requires selection of an optimal value for the parameter Q. However, with 5 or 10 repeated 10-fold cross-validation ($Q = 5 \times 10$, in our case, or $Q = 10 \times 10$) could provide an effective performance. In future, statistical analysis may be incorporated with WSFS to select the optimal feature subset from the available feature subset generated based on the feature weight. This will support to find out the best subset that is significantly better than all available optimal candidate subsets according to the accuracy.

7.2.5 Overall decoding performance

In both the TMEP and LFP decoding algorithms, all of the subjects have not achieved high accuracy, especially in the LFP decoding. In the TMEP action classification, with a small training set, high classification accuracy (90~100%) has been achieved across all subjects in both clean and noisy environments. Previously, similar TMEP action classification accuracy was achieved using the same dataset (Vaidyanathan et al. 2007). Again for discriminating tongue movement actions in various challenging interference conditions, the decoding algorithm also achieved high accuracy in subject specific (93~100%) and generalised (91~100%) interferences during offline and online evaluations (84~93%) across all subjects. There are several reasons that have contributed for achieving high accuracy in real environments, these include less intra-class variance in the tongue movement actions due to a high level of training and concentration to perform the movement actions (all are healthy subjects); optimal extraction and selection of time-frequency features using wavelet transform that increased the inter-class discrimination; as well as classification methods to some extent. However, un-trained subjects accuracy was significantly less compared to

trained subjects, as well as in some cases interferences that contain similar time-frequency information (such as drinking), as the TMEP actions thus increased the false positive rate (Mamun et al. 2012). Evaluation of other interference conditions (such as real-life external interferences) and addressing the drawback of this study could help to improve the robustness of the TMEP decoding algorithm. Moreover investigation of real (i.e. disabled) user's dataset will provide the feasibility to develop TMEP based HMI system for real-life applications.

In the LFP decoding, high accuracy (98~100%) has been achieved for recognising movement and rest states across all subjects. In the extracted LFPs segments, event related synchronisation or de-synchronisation were observed in each frequency band and highly in the δ, β and γ bands. Accordingly, extracted features had very good separability between movement (event) and rest activities. Previously, Loukas & Brown (2004) reported similar results (95% sensitivity and 77% specificity) for predicting movement onset from STN LFPs (Loukas & Brown 2004). During subsequent laterality (left and right) decoding, remarkably less accuracy (60~100%) was achieved, in some subjects we achieved high accuracy (~100%) and in others achieved low accuracy (60% or more). There are several issues that are exhibited in the LFP dataset and imposed significantly on the laterality decoding performance. One notable issue in the LFP decoding is that the datasets were collected from patients, who were more or less severely affected by neurological diseases (Parkinson's or dystonia) (cf. section 6.2.1, table 1). Patients with essential tremor, Parkinsonian tremor or dystonia, who cannot be treated effectively by medication, may be the candidates for DBS implantation. Due to the disease severity, age, limited or no training as well as lack of concentration and control to perform the tasks, extremely high variability was observed in the LFP datasets. On the other hand, due to the dynamic similarities of LFPs recorded from left and right STN or GPi for both (left and right) finger clicking tasks, it may be difficult to enhance high inter-class separability. Again with a limited and unbalanced dataset, as well as higher redundancy in the feature space made the decoding problem more difficult. However, considering all of the limitations, the performance of the proposed LFP decoding algorithm is extremely encouraging and showed that movement identification is possible. A number of studies in the literature also provide evidence that LFPs are more stable than the single unit activity and able to provide good neuronal information related to motor activity (Ince et al. 2010).

Deep brain LFPs have been extensively investigated in a clinical perspective to understand neural mechanisms or to improve the treatment of neuro-generative diseases (Kringelbach et al. 2007). To the best of our knowledge, this is the first study that explores the deep brain LFP decoding problem; therefore it is difficult to compare the performance with the results from the existing literature. Nevertheless, outcomes of this decoding analysis suggest that deep brain LFPs could be an alternative or supportive source for developing neural interfaces. Having said this, to develop real-life neural interface applications based on deep brain LFPs, the movement decoding algorithms need further investigation in other conditions (self-paced and imagined movements). Furthermore online evaluation of the algorithms is essential to identify its feasibility. In future studies, there are a number of issues that need to be taken into consideration, such as unbalanced characteristics in the dataset, characterising the quality of the data based on patients' disease severity and their ability to perform an action. Addressing such issues may further improve robustness and the decoding performance, and also lead to design of real-life neural interface systems for BMI applications.

7.3 Conclusion and future work

This thesis investigated the decoding ability for recognising pattern of movement related states from tongue movement ear pressure signals and deep brain local field potentials. The developed algorithms utilised time frequency methods for feature extraction and selected the subset of features to increase inter-class separability for classification. The classification results show that the decoding algorithms are capable of identifying different movement states from the respective biosignals (TMEP or LFPs). This study not only helps us to decode movement related states from biosignals, but also it will advance the possibility to develop an HMI or BMI system based on the TMEP or LFPs, to improve the quality of living for disabled people as well as for rehabilitations.

Alongside the neural decoding algorithm, this thesis introduced a new feature selection strategy, weighted sequential feature selection (WSFS) to select optimal feature subset for pattern classification. It is expected that the WSFS strategy is capable of selecting effective features with reducing redundant (least discriminative) features from the feature space to improve classification accuracy and complexity, which are very important for high dimensional, small sample size datasets.

In future, the research carried out in this thesis can be extended in the following directions.

Firstly, to improve the pattern recognition ability for high dimensional, small sample size datasets such as neural data, the proposed feature selection and dimensionality reduction strategy, WSFS was successful for deep brain LFPs decoding. However, to evaluate the robustness of the proposed strategy, further investigations on other biosignals are essential and also need comparative evaluation with similar existing feature selection strategies using benchmark dataset (Gupta et al. 2010; Saeys et al. 2007). The WSFS strategy can also be extended to investigate very high dimensional biomedical datasets such as gene and protein studies to improve the classification performance (Saeys et al. 2007).

Secondly, the neural decoding algorithms developed in this study are successfully used for recognising voluntary upper limb movements from deep brain (STN or GPi) LFPs with high accuracy. However, investigation of other movement paradigms such as self-paced, imagining as well as other movements such as lower limb movements can advance our understanding on the underlying processes of basal ganglia deep brain LFPs for voluntary movement control and its important implications for the development of neural interface systems (Kringelbach et al. 2007). These investigations will lead us towards development of alternative brain machine interface based on deep brain LFPs for real life application. Particularly, it might be potentially useful in place of primary motor cortex (M1) if the M1 is damaged or in combination with primary motor cortex and basal ganglia activities for enhanced movement identification towards the development of neuro-prosthetic applications for disables (Patil & Turner 2008; Ince et al. 2010).

Finally, future investigation of involuntary movement identification of deep brain LFPs from basal ganglia could advance our understanding towards development of bidirectional (*read-out* and *write-in*) communications such as brain machine brain interface (BMBI) for advanced therapeutic interventions (Engel et al. 2005; Kringelbach et al. 2007). Particularly early prediction and decoding of involuntary movement such as tremor incorporating with deep brain stimulators could lead us to develop closed loop or demand driven deep brain stimulation for improving the treatment of several neurological

diseases (Rosin et al. 2011; Santos et al. 2011; Gasson et al. 2005; Kringelbach et al. 2007).

All of these future investigations based on this thesis work may open up several possibilities to produce novel knowledge contribution and new inventions in the field of neural signal processing and pattern identification for developing potential neural interface system or brain machine communications and also it will create substantial impact in health care of our society.

Appendix A: Key Publications from this Research Work

Journal Papers:

1. *K. A. Mamun*, M. Mace, M. E. Lutman, R. Vaidyanathan and S. Wang, "Weighted Sequential Feature Selection (WSFS): an efficient strategy to facilitate high dimensional small sample size biosignal classification," *In Preparation, to be submitted in IEEE Transactions on Neural Systems and Rehabilitation Engineering.*
2. M. Mace, *K. A. Mamun*, L. Gupta, S. Wang and R. Vaidyanathan, "Real-time implementation of tongue-movement communication system for assistive human-machine interface," *Ready to submit in Journal of Pattern Recognition*.
3. *K. A. Mamun*, M. Mace, R. Vaidyanathan, M. E. Lutman, J. Stein, X. Liu, T. Aziz, and S. Wang, "Decoding of movements and laterality from human basal ganglia local field potentials based on neural synchronisation," *Ready to submit in Journal of Neural Engineering*.
4. *K. A. Mamun*, M. Mace, L. Gupta, C. A. Verschuur, M. E. Lutman, M. Stokes, R. Vaidyanathan, S. Wang, "Robust real-time identification of tongue movement commands from interferences," *Neurocomputing*, Elsevier, vol. 80, pp-83-92, doi: 10.1016/j.neucom.2011.09.018, March 2012.

Journal Abstracts:

1. *K. A. Mamun*, M. Mace, R. Craig, M. E. Lutman, R. Vaidyanathan, and S. Wang, "Tongue movement ear pressure signal classification using wavelet packet transform," *International Journal of Audiology*, pp. 701, vol. 49, no. 9, Sep. 2010.

Conference Papers:

1. *K. A. Mamun*, M. Mace, R. Vaidyanathan, M. E. Lutman, J. Stein, X. Liu, T. Aziz, and S. Wang, "A robust strategy for decoding movements from deep brain local field potentials to facilitate brain machine interfaces," *4th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*, Roma, Italy, Jun. 24-28, 2012, (Accepted).

2. M. Mace, *K. A. Mamun,* Shouyan Wang, Lalit Gupta and Ravi Vaidyanathan, "Ensemble classification for robust discrimination of multi-channel, multi-class tongue-movement ear pressure signal," *Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2011),* Boston, USA, Aug. 30-Sep. 3, 2011.
3. *K. A. Mamun,* R. Vaidyanathan, M. E. Lutman, J. Stein, X. Liu , T. Aziz, and S. Wang, "Decoding movement and laterality from local field potentials in the subthalamic nucleus," *Proceedings of the 5th International IEEE EMBS Conference on Neural Engineering*, Cancun, Mexico, Apr. 27- May 1, 2011.
4. *K. A. Mamun,* M. Banik, M. Mace, M. E. Lutman, R. Vaidyanathan and S. Wang, "Multi-layer neural network classification of tongue movement ear pressure signal for human machine interface," *Proceedings of the 13th International Conference on Computer And Information Technology (ICCIT 2010)*, Dhaka, Bangladesh, Dec. 23-25, 2010.
5. M. Mace, *K. A. Mamun,* Ravi Vaidyanathan, Shouyan Wang and Lalit Gupta, "Preliminary real-time results in the classification of tongue-movement ear pressure signals," *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipei, Taiwan, Oct. 18-22, 2010.
6. *K. A. Mamun,* M. Mace, M. E. Lutman, R. Vaidyanathan, and S. Wang, "SVM classification of tongue movement ear pressure signals for human machine interface," *INSPIRE 2010*, University College London (UCL), London, UK, Sep. 6-8, 2010.
7. M. Mace, *K. A. Mamun,* S. Wang, L. Gupta and R. Vaidyanathan, "Human-machine interface for tele-robotic operation using tongue movement ear pressure (TMEP) signals," *Proceedings of the 11th Conference Towards Autonomous Robotic Systems*, Plymouth, UK, Aug. 31- Sep. 2, 2010.
8. *K. A. Mamun,* M. Mace, M. E. Lutman, R. Vaidyanathan, L. Gupta and S. Wang, "Multivariate Bayesian classification of tongue movement ear pressure signals based on the wavelet packet transform," *Proceedings of the 2010 IEEE Int. Workshop on Machine Learning for Signal Processing*, Finland, Aug. 29- Sep. 1, 2010.
9. *K. A. Mamun*, M. Mace, M. E. Lutman, R. Vaidyanathan, and S. Wang, "Bayesian classification of tongue movement based on wavelet packet transformation," *INSPIRE 2009*, Imperial College London, London, UK, September 2009.
10. *K. A. Mamun,* M. Mace, M. E. Lutman, R. Vaidyanathan, and S. Wang, "Pattern classification of tongue movement ear pressure signal based on wavelet packet feature extraction," *Proceedings of the 5th UK & RI Postgraduate Conference in Biomedical Engineering & Medical Physics*, pp. 33-34, Magdalen College, Oxford University, Oxford, UK, July 2009.

Workshop or Conference (Extended) Abstracts:

1. *K. A. Mamun*, M. Mace, M. E. Lutman, R. Vaidyanathan, J. Stein, X. Liu, T. Aziz, and S. Wang, "Decoding movement and laterality from human subthalamic local field potentials for neuro-prosthetic applications," *2011 International UKIERI Workshop on the Fusion of Brain-Computer Interface and Assistive Robotics*, Londonderry, Northern Ireland, UK, Jul. 7- 8, 2011.
2. *K. A. Mamun*, M. Mace, M. E. Lutman, R. Vaidyanathan, J. Stein, X. Liu, T. Aziz, and S. Wang, "Decoding movements from human subthalamic local field potentials based on neural synchronization," *9th Annual Southampton Neurosciences Group (SoNG) Meeting*, University of Southampton, Southampton, UK, Sep. 22, 2011.
3. *K. A. Mamun*, M. E. Lutman, R. Vaidyanathan, J. Stein, X. Liu, T. Aziz, and S. Wang, "Recognition of voluntary movement from human subthalamic activity for brain computer interface," *SET for BRITAIN 2011*, The House of Commons, UK, March 14, 2011.
4. *K. A. Mamun*, M. E. Lutman, R. Vaidyanathan, J. Stein, X. Liu, T. Aziz, and S. Wang, "Recognition of Voluntary Movement from Human Subthalamic Activity for Brain Machine Interface," *Multidisciplinary Postgraduate Research Showcase*, University of Southampton, UK, March *31, 2011*.
5. *K. A. Mamun*, M. E. Lutman, R. Vaidyanathan, J. Stein, X. Liu, T. Aziz, and S. Wang, "Recognition of the laterality of voluntary movement from subthalamic activity," *8th Annual Southampton Neurosciences Group (SoNG) Meeting*, University of Southampton, Southampton, UK, Sep. 23, 2010.
6. *K. A. Mamun*, M. E. Lutman, R. Vaidyanathan, and S. Wang, "Tongue movement: A novel concept for assistive HMI," *FESM Postgraduate Research Showcase 2010*, University of Southampton, UK, May 6, 2010.
7. *K. A. Mamun*, M. Mace, M. E. Lutman, R. Vaidyanathan, and S. Wang, "State identification of tongue movement signals," *Mathematical Neuroscience 2010,* Edinburgh, UK, Apr. 19-21, 2010.

Appendix B: Experimental Design for TMEP Actions and Interferences Signal Acquisitions

Title of Experiment

Human-Machine Interface based on Tongue Movement Ear Pressure (TMEP) signal for Rehabilitation Systems

B.1 Overview

The goal of this research is to design and develop a Human-Machine Interface (HMI) which enables patients with quadriplegia, severe arthritis, spinal cord injuries, limited movement due to stroke, or other conditions causing limited or painful hand/arm movement to interface with their environment by controlling all manner of peripheral equipment, ranging from mechanical or robotic assist devices, lights, television, prosthetic aids to computers. To achieve the goal of this HMI, an unobtrusive method has been considered which includes monitoring and detection of changes in air pressure within the ear canal due to specific tongue movement, and subsequently generating a control instructions corresponding to that movement in real-time. Various movements of the tongue within the oral cavity create unique, traceable pressure changes in the ear, which can be measured by inserting a simple microphone sensor (earpiece) into the ear canal. Later, recorded Tongue Movement Ear Pressure (TMEP) signal will be analysed to generate action commands for the rehabilitation or assistive system. The initial requirement of this system needs to create a datasets of TMEP signal that includes movement of tongue in different direction of each subject. The aim of this experiment is to create such dataset with the participation of 20 subjects.

B.2 Procedure of the experiment

In the experiment, each subject will sit in a comfortable armchair with relaxed environment and recording microphone sensors (earpiece) will be inserted into both

ears. The TMEP signals will be recorded by our own software running on a desktop/laptop computer. The software has been written in Matlab (R2007b) to record and pre-process the collected signals. The data collection procedure has two stages: 1) subject preparation, which includes practicing different movement of tongue actions; 2) collection of TMEP signal for training and test the system. The recording session for each subject lasted typically about 2 hours including 30 minutes for practice. During the subject preparation stage, subjects will be instructed to move their tongue in a respective direction as much as possible to perform each action correctly. Actions are to move the tongue in one direction of left, right, up, down or flicking the tongue up and down once with putting no time interval as single click and moving the tongue straight to outside of the oral cavity (mouth) with closed lips as double click. The instruction for tongue movement will be given by running a program in the computer which will guide the subject. That is the tongue movement task is cued by visual stimulus on the computer monitor. Figure B.1 shows the typical representation of visual stimulus action window. It will guide the subject by representing the text of each action as well as direction to move the tongue through the circle moving in respective direction. The configuration of the visual action window consists of outer area and inner area with both height and weight of 21 cm and 20 cm respectively and the diameter of the circle is 5 mm. Initially the circle remains in the middle of the window and it will move in one direction according to the action and again back to the origin for the next action. The circle will move and back with random selection of action into left, right, up, down, up and down, front and back direction for tongue action of left, right, up, down, single click and double click respectively. Subjects need to place their tongue at normal position when the circle is in the middle position of the window and the subject needs to move or flick the tongue according to a visual cue. The circle will move in one direction with three different speeds 0.25 m/s, 0.5 m/s and 1 m/s for creating different variation in tongue movement. Subject will be instructed to follow the movement of the circle and try to move and back their tongue with direction and speed of the circle represented in the action window i.e. synchronisation of tongue with the cue. Visual stimulus action window will setup in front of subject. Maximum duration to complete an action by the subject will take less than 2s. The visual instruction for each action will be repeated after every 5s. In TMEP signal collection stage, the subjects will be given a random sequence of actions (100 repetitions for each) through visual action window to follow and perform it accordingly.

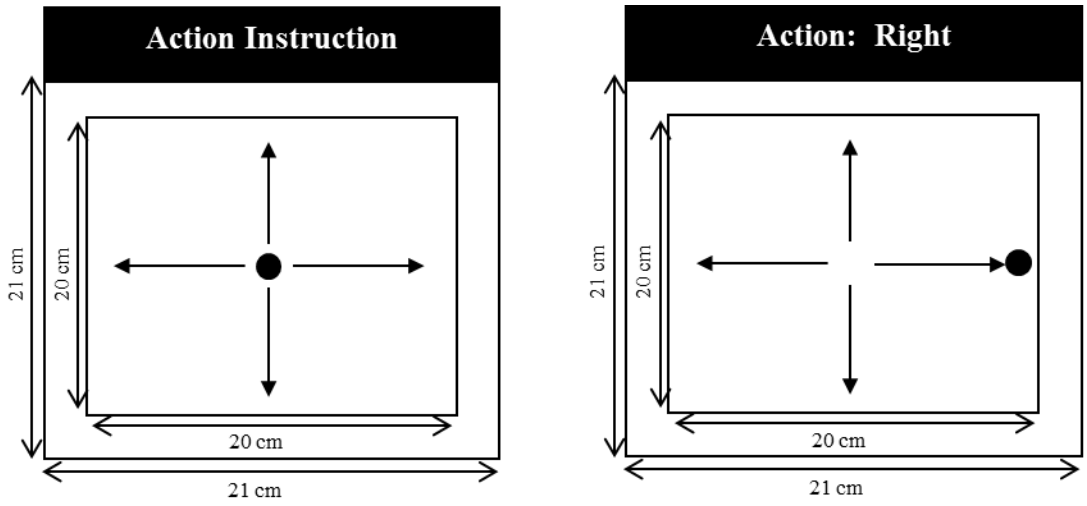


Figure B.1: Visual stimulus action window for representations of active instruction to perform by tongue.

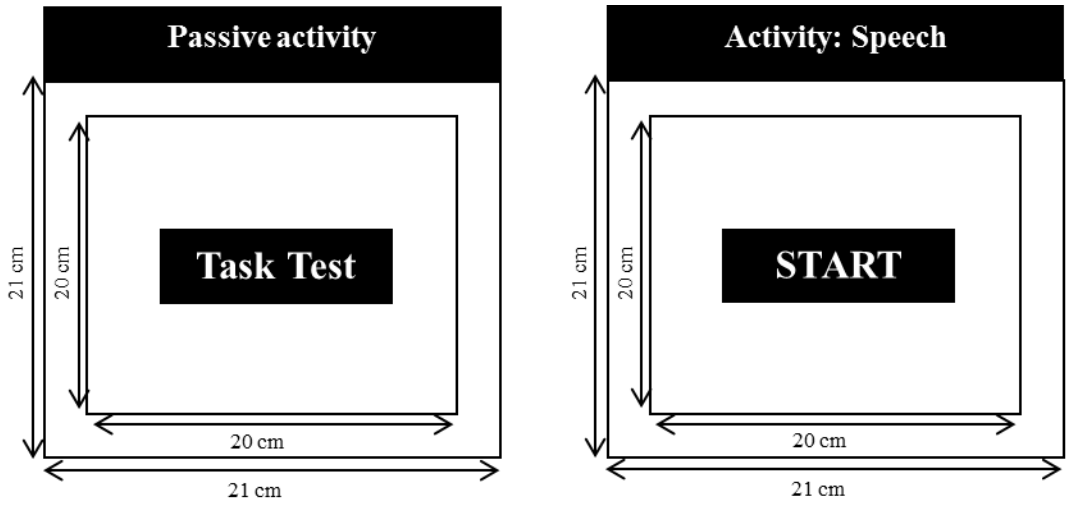


Figure B.2: Visual stimulus window for representations of passive activity instruction to perform by tongue.

To discriminate TMEP action and natural activity of tongue movements happens in our everyday life, which includes movements of tongue during talking (speech), coughing, drinking, and resting; considered as passive activity, these activities will also be recorded. To record these activities a visual as well as a verbal instruction will also be given to the subject to perform the action. For the passive activity, visual window will show the respective activity for guiding the subject. For speech, visual window will represent each task (text view) which is numbered from 0 to 9 and words of ‘start’, ‘stop’, ‘on’, ‘off’, ‘open’, ‘close’. Every task repeats after every three seconds in a random order with 20 repetitions for each task. The subjects job is to speak the

respective task presented in the visual window. Coughing is a sudden and often repetitively occurring defense reflex which helps to clear the large breathing passages from excess secretions, irritants, foreign particles and microbes. For coughing, visual window will represent a text 'Coughing' and it will repeat after every ten seconds with 20 repetitions. In the case of drinking water, the subject needs to drink 15 ml of water from a glass. Visual window also represent a text 'Drink 15 ml water' to guide the subject and it will repeat after every ten seconds with 20 repetitions. Resting is the normal relaxation activity which regularly takes every human in their life. Here resting is defined as no activity in whole body of the subject and also no movement of tongue for 5 minutes. The visual window also show the text 'Start 5 minute rest' for the starting of the resting task and after 5 minutes it also show the text 'Stop 5 minute rest' for finishing the task. A typical visual window for passive activity is shown in figure B.2. All of these natural TMEP signals may characterise as passive action or noise while command TMEP signals are characterised as active action. Time interval between two active movements is an impartment issue to discriminate active and passive action. The experiment will collect two types of TMEP signal, 1. Command signals and 2. Natural signals, which will be characterised as noise for command signal. All experiment will perform in a standard office/ laboratory environment. All the recorded signals will be sampled with a sampling frequency 8000 Hz.

B.3 Participants of the experiment

20 healthy subjects will be voluntarily recruited from the university students, staff and friends. Subjects age range 18-45 years consisting of male and female. Experimenter will ask each subject to know that they are suffering from any neurological, tongue, oral cavity or hearing disorders or not. It will ensure that they did not suffer from any of the above mentioned disorders previously (subjective test). Subjects do not have any previous experience of such experiment. All subjects to be fully briefed, their understanding confirmed and informed consent to be received before undergoing the experiment. The experiment consisting of four sessions, each subject will participate in all sessions. Duration of each session is 30 minutes and the experiment lasts no longer than 2 hours, after each session there is a 10-15 minutes break. Extra breaks to be granted on participant request. There will be a payment of £10 for each subject to attend the experiment. Table B.1 shows the overall organisation of the sessions that each subject will go through and list of risk involve with the subject presented in table B.2.

Table B.1. Session organization for each subject

Session	Activity
Session 1: Practice	1. Explanation of the procedure. Informed consent to be obtained. 2. Setting up the microphone sensor (earpiece) with comfortable adjustment. It also follows the standard clinical practice. 3. Subject will practice different movement (left, right, up, down, single click and double click) of tongue (including speed and distance) according to visual instruction presented in the display.
Session 2: TMEP signal recording (active action)	Recording of TMEP signal by inserting sensor into both ears (2 channels). Subject will be instructed through visualisation of each movement in the display to perform movement of his/ her tongue accordingly. Each movement (actions) has 100 repetitions to be carried out. The interval between the consecutive movements is 5 second. Total time to finish the task will approximately 30 minutes. The recording condition is standard office/ laboratory environment. The recording of active actions is left, right, up, down of tongue and flicking the tongue up and down once with putting no time interval as single click and moving the tongue straight to outside of the oral cavity (mouth) with closed lips as double click.
Session 3: TMEP signal recording (passive action)	The recording setup of passive action is as same as active action. The passive action consists of 4 actions. 1. Coughing: A sudden and often repetitively occurring defense reflex which helps to clear the large breathing passages from excess secretions, irritants, foreign particles and microbes. 20 repetition of coughing will record from each subject. It will take approximately 4 minutes, ten seconds for each repetition. 2. Drinking water: Activity of water drinking will record 20 times, each time subject will requested to drink 15ml of water. It will take approximately 4 minutes, ten seconds for each repetition. 3. Resting: Resting is defined as no activity in whole body of the subject and also no movement of tongue for 5 minutes. Subject will perform resting activity and this will record for 5 minutes. 4. Speech: It includes counting number (0-9) and words of ‘start’, ‘stop’, ‘on’, ‘off’, ‘open’, ‘close’. Each word will record 20 times and time for each word three seconds. Total time will take for this approximately 17 minutes. Total time to finish this session will take approximately 30 minutes.
Session 4:	If required to repeat any task or session.

Table B.2: List of risk related to experiment

	Nature of Risks	*Control Measures for the Risks*
1.	Infections in the ears.	Use of separate ear tips for each subject.
2.	Subject may feel tired.	Providing break/ resting after each session; or at any time during the experiment subjects can request for break or resting.
3.	Middle ears may be damaged.	Experiment will follow the standard clinical procedure.

B.4 Equipment

1. Sensor (Earpiece/microphone) specification.

2. Comply canal tips

3. Computer

In the TMEP signal acquisition, a microphone earpiece sensor with a comfortable canal tips is shown in figure B.3, which will be inserted into both ears to record two TMEP signals.

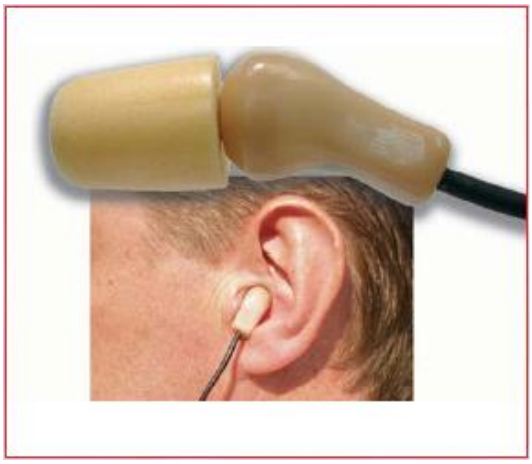


Figure B.3: Microphone earpiece sensor with a comfortable canal tip.

References

Addison, P., Walker, J., & Guido, R. (2009). Time--frequency analysis of biosignals. *IEEE engineering in medicine and biology magazine*, *28*(5), 14-29.

Alegre, M., Alonso-Frech, F., Rodriguez-Oroz, M. C., Guridi, J., Zamarbide, I., Valencia, M., Manrique, M., et al. (2005). Movement-related changes in oscillatory activity in the human subthalamic nucleus: ipsilateral vs. contralateral movements. *European Journal of Neuroscience*, *22*(9), 2315-2324.

AlZoubi, O., Koprinska, I., & Calvo, R. A. (2008). Classification of Brain-Computer Interface Data. *In the 7th Australasian Data Mining Conference (AusDM 08)*, Adelaide, Australasian, 123-131.

Andersen, R. A., Musallam, S., & Pesaran, B. (2004). Selecting the signals for a brain-machine interface. *Current Opinion in Neurobiology*, *14*(6), 720-726.

Antonini, G., & Orlandi, A. (2001). Wavelet packet-based EMI signal processing and source identification. *IEEE Transactions on Electromagnetic Compatibility*, *43*(2), 140-148.

Bai, O, Mari, Z., Vorbach, S., & Hallett, M. (2005). Asymmetric spatiotemporal patterns of event-related desynchronization preceding voluntary sequential finger movements: a high-resolution EEG study. *Clinical Neurophysiology*, *116*(5), 1213-1221.

Bai, O, Lin, P., Vorbach, S., Li, J., Furlani, S., & Hallett, M. (2007). Exploration of computational methods for classification of movement intention during human voluntary movement from single trial EEG. *Clinical Neurophysiology*, *118*(12), 2637-2655.

Barea, R., Boquete, L., Mazo, M., & Lopez, E. (2002). System for assisted mobility using eye movements based on electrooculography. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *10*(4), 209-218.

Bashashati, A., Fatourechi, M., Ward, R. K., & Birch, G. E. (2007). A survey of signal processing algorithms in brain-computer interfaces based on electrical brain signals. *Journal of neural engineering*, *4*(2), R32-57.

Bernasconi, C., & König, P. (1999). On the directionality of cortical interactions studied by structural analysis of electrophysiological recordings. *Biological Cybernetics*, *81*(3), 199-210.

Betke, M., Gips, J., & Fleming, P. (2002). The camera mouse: Visual tracking of body features to provide computer access for people with severe disabilities. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *10*(1), 1-10.

Bilmes, J., Malkin, J., Xiao, L., Harada, S., Kilanski, K., Kirchhoffi, K., Wright, R., et al. (2006). The Vocal Joystick. *In the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2006)*, Toulouse, France, 625-628.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.

Bogacz, R., & Gurney, K. (2007). The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural Computation*, *19*(2), 442-477.

Brault, M. W. (2008). *Americans With Disabilities: 2005*. U.S. CENSUS BUREAU: http://www.census.gov/prod/2008pubs/p70-117.pdf

Brovelli, A., Ding, M. Z., Ledberg, A., Chen, Y. H., Nakamura, R., & Bressler, S. L. (2004). Beta oscillations in a large-scale sensorimotor cortical network: Directional influences revealed by Granger causality. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(26), 9849-9854.

Brown, P., & Williams, D. (2005). Basal ganglia local field potential activity: Character and functional significance in the human. *Clinical Neurophysiology*, *116*(11), 2510-2519.

Bulling, A., Ward, J. A., Gellersen, H., & Tröster, G. (2011). Eye Movement Analysis for Activity Recognition Using Electrooculography. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *33*(4), 741-53.

Cannan, J., & Hu, H. (2011). *Human-Machine Interaction (HMI): A Survey*. Technical Report CES-508, School of Computer Science & Electronic Engineering, University of Essex, UK.

Center, N. S. C. I. S. (2008). *SPINAL CORD INJURY Facts and Figures at a Glance*. National Spinal Cord Injury Statistical Center, The University of Alabama, Birmingham, Alabama, USA. http://images.main.uab.edu/spinalcord/pdffiles/Facts08.pdf

Chang, C.-C., & Lin, C.-J. (2001). LIBSVM: a library for support vector machines. http://www.csie.ntu.edu.tw/~cjlin/libsvm.

Chen, Y.-wei, & Lin, C.-jen. (2006). Combining SVMs with Various Feature Selection Strategies. *Strategies*, *324*(1), 1-10.

Chu, S., Narayanan, S., & Kuo, C. C. J. (2009). Environmental Sound Recognition With Time–Frequency Audio Features. *IEEE Transactions on Audio Speech and Language Processing*, *17*(6), 1142-1158.

Coifman, R. R., & Wickerhauser, M. V. (1992). Entropy-Based Algorithms for Best Basis Selection. *IEEE Transactions on Information Theory*, *38*(2), 713-718.

Collinger, J. L., Dicianno, B. E., Weber, D. J., Cui, X. T., Wang, W., Brienza, D. M., & Boninger, M. L. (2011). Integrating rehabilitation engineering technology with biologics. *PM&R, Journal of American Academy of Physical Medicine and Rehabilitation, 3*(6), S148-S157.

Cook, A. M., & Hussey, S. M. (2002). *Assistive technologies: Principles and practice* (2nd edition). St. Louis (MO): Mosby.

Coyle, D. H. (2006). *Intelligent Preprocessing and Feature Extraction Techniques for a Brain-Computer Interface, PhD Thesis*. Intelligent Systems Engineering Laboratory, School of Computing and Intelligent Systems, Faculty of Engineering, University of Ulster, UK.

Dash, M., & Liu, H. (1997). Feature Selection for Classification. *Intelligent Data Analysis*, *1*, 131-156.

Delong, M. R., & Wichmann, T. (2007). Circuits and circuit disorders of the basal ganglia. *Archives of Neurology*, *64*(1), 20-24.

Denison, D. G. T., Holmes, C. C., Mallick, B. K., & Smith, A. F. M. (2002). *Bayesian methods for nonlinear classification and regression*. Wiley series in probability and statistics.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern Classification* (2nd edition). Wiley-Interscience.

Engel, A. K., Moll, C. K. E., Fried, I., & Ojemann, G. A. (2005). Invasive recordings from the human brain: clinical insights and beyond. *Nature Reviews Neuroscience*, *6*(1), 35-47.

Englehart, K., Hudgins, B., & Parker, P. A. (2001). A wavelet-based continuous classification scheme for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, *48*(3), 302-311.

Englehart, K., Hudgins, B., Parker, P. A., & Stevenson, M. (1999). Classification of the myoelectric signal using time-frequency based representations. *Medical Engineering & Physics*, *21*(6-7), 431-438.

Fatourechi, M., Birch, G. E., & Ward, R. K. (2007). Application of a hybrid wavelet feature selection method in the design of a self-paced brain interface system. *Journal of NeuroEngineering and Rehabilitation, 4*(11), 1-13.

Ferreira, A., Celeste, W. C., Cheein, F. A., Bastos, T. F., Sarcinelli, M., & Carelli, R. (2008). Human-machine interfaces based on EMG and EEG applied to robotic systems. *Journal of NeuroEngineering and Rehabilitation, 5(10), 1-15*.

Fonseca, E. S., Guido, R. C., Scalassara, P. R., Maciel, C. D., & Pereira, J. C. (2007). Wavelet time-frequency analysis and least squares support vector machines for the identification of voice disorders. *Computers in biology and medicine*, *37*(4), 571-578.

Gacquer, D., Delcroix, V., Delmtte, F., & Piechowiak, S. (2011). Comparative study of supervised classification algorithms for the detection of atmospheric pollution. *Engineering Applications of Artificial Intelligence*. *24*(6), 1070-1083.

Ganguly, K., Secundo, L., Ranade, G., Orsborn, A., Chang, E. F., Dimitrov, D. F., Wallis, J. D., et al. (2009). Cortical representation of ipsilateral arm movements in monkey and man. *Journal of Neuroscience*, *29*(41), 12948 -12956.

Gasson, M., Wang, S., Aziz, T., Stein, J., & Warwick, K. (2005). Towards a demand driven deep-brain stimulator for the treatment of movement disorders. *In the 3rd IEE International Seminar on Medical Applications of Signal Processing*, London, 83-86.

Gelfand, S. A. (2007). *Hearing: An introduction to psychological and physiological acoustics* (4th edition). Informa healthcare, USA.

Gerven, M. V., Farquhar, J., Schaefer, R., Vlek, R., Geuze, J., Nijholt, A., Ramsey, N., et al. (2009). The brain-computer interface cycle. *Journal of Neural Engineering*, *6*(4), 1-10.

Geweke, J. (1982). Measurement of Linear-Dependence and Feedback Between Multiple Time-Series. *Journal of the American Statistical Association*, *77*(378), 304-313.

Glassman, E. L. (2005). A wavelet-like filter based on neuron action potentials for analysis of human scalp electroencephalographs. *IEEE Transactions on Biomedical Engineering*, *52*(11), 1851-1862.

Golugula, A., Lee, G., & Madabhushi, A. (2011). Evaluating feature selection strategies for high dimensional, small sample size datasets. *In the 33rd Annual International Conference of the IEEE EMBS*. Boston, USA, 949-952.

Gourévitch, B., Bouquin-Jeannès, R. L., & Faucon, G. (2006). Linear and nonlinear causality between signals: methods, examples and neurophysiological applications. *Biological Cybernetics*, *95*(4), 349-369.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross spectral methods. *Econometrica*. *37*, 424-438.

Guerrero-Mosquera, C., Verleysen, M., & Vazquez, A. N. (2010). EEG feature selection using mutual information and support vector machine: A comparative analysis. *In the 32nd Annual International Conference of the IEEE EMBS.* Buenos Aires, Argentina, 4946-4949.

Gunn, S. R. (1998). *Support Vector Machines for Classification and Regression*. Technical Report, ECS, University of Southampton, UK.

Gupta, L., Kota, S., Murali, S., Molfese, D. L., & Vaidyanathan, R. (2010). A Feature Ranking Strategy to Facilitate Multivariate Signal Classification. *IEEE Transactions on Systems Man and Cybernetics: Part C-Applications and Reviews*. *40*(1), 98–108.

Gupta, L., Molfese, D. L., Tammana, R., & Simos, P. G. (1996). Nonlinear alignment and averaging for estimating the evoked potential. *IEEE Transactions on Biomedical Engineering*, *43*(4), 348-356.

Guyon, I. (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research*, *3*, 1157-1182.

Güneş, S., Polat, K., & Yosunkaya, Ş. (2010). Multi-class F-score feature selection approach to classification of obstructive sleep apnea syndrome. *Expert Systems with Applications*, *37*(2), 998-1004.

Herman, P., Prasad, G., McGinnity, T. M., & Coyle, D. (2008). Comparative Analysis of Spectral Approaches to Feature Extraction for EEG-Based Motor Imagery Classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *16*(4), 317-326.

Hoffmann, U. (2007). *Bayesian Machine Learning Applied in a Brain-Computer Interface for Disabled users, PhD Thesis*. EPFL, Switzerland.

Hsu, W.-Y., & Sun, Y.-N. (2009). EEG-based motor imagery analysis using weighted wavelet transform features. *Journal of Neuroscience Methods*, *176*(2), 310-8.

Huang, D., Lin, P., Fei, D.-Y., Chen, X., & Bai, O. (2009). Decoding human motor activity from EEG single trials for a discrete two-dimensional cursor control. *Journal of Neural Engineering*, *6*(4), 1-12.

Huo, X. L., Wang, J., & Ghovanloo, M. (2008). Introduction and preliminary evaluation of the Tongue Drive System: Wireless tongue-operated assistive technology for people with little or no upper-limb function. *Journal of Rehabilitation Research and Development*, *45*(6), 921-930.

Hutchinson, T. E., White, K. P., Martin, W. N., Reichert, K. C., & Frey, L. A. (1989). Human-Computer Interaction Using Eye-Gaze Input. *IEEE Transactions on Systems Man and Cybernetics*, *19*(6), 1527-1534.

Imam, T., Ting, K. M., & Kamruzzaman, J. (2006). z-SVM : An SVM for Improved Classification of Imbalanced Data, *Lecture Notes in Computer Science 4304: Advances in Artificial Intelligence*, Springer, 264-273.

Ince, N. F., Gupta, R., Arica, S., Tewfik, A. H., Ashe, J., & Pellizzer, G. (2010b). High Accuracy Decoding of Movement Target Direction in Non-Human Primates Based on Common Spatial Patterns of Local Field Potentials. *Plos One*, *5*(12), 1-11.

Jain, A. K., & Duin, P. W. (2000). Statistical pattern recognition: a review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*(1), 4-37.

Jbabdi, S., & Behrens, T. E. J. (2012). Specialization: the connections have it. *Nature Neuroscience*, *15*(2), 171-172.

Jerbi, K., Vidal, J. R., Mattout, J., Maby, E., Lecaignard, F., Ossandon, T., Hamamé, C. M., et al. (2011). Inferring hand movement kinematics from MEG, EEG and intracranial EEG: From brain-machine interfaces to motor rehabilitation. *IRBM*, *32*(1), 8-18. Elsevier Masson SAS.

Jolliffe, I. T. (2002). *Principal Component Analysis* (Second Edition). Springer.

Ju-Jang, L., Kap-Ho, S., Changmok, O., & Bien, Z. (2007). Development of a future Intelligent Sweet Home for the disabled. *Artificial Life and Robotics*, *11*(1), 8-12.

Kaminski, M., & Liang, H. (2005). Causal influence: advances in neurosignal analysis. *Critical Reviews in Biomedical Engineering*, *33*(4), 347-430.

Kamiński, M, Ding, M., Truccolo, W. A., & Bressler, S. L. (2001). Evaluating causal relations in neural systems : Granger causality, directed transfer function and statistical assessment of significance, *Biological cybernetics*, *85*, 145-157.

Kennedy, P., Andreasen, D., Ehirim, P., King, B., Kirby, T., Mao, H., & Moore, M. (2004). Using human extra-cortical local field potentials to control a switch. *Journal of Neural Engineering*, *1*(2), 72-77.

Kilner, J. M., Salenius, S., Baker, S. N., Jackson, A., Hari, R., & Lemon, R. N. (2003). Task-dependent modulations of cortical oscillatory activity in human subjects during a bimanual precision grip task. *NeuroImage*, *18*(1), 67-73.

Kiymik, M. K., Guler, I., Dizibuyuk, A., & Akin, M. (2005). Comparison of STFT and wavelet transform methods in determining epileptic seizure activity in EEG signals for real-time application. *Computers in Biology and Medicine*, *35*(7), 603-616.

Kohlmorgen, J., & Blankertz, B. (2004). Bayesian classification of single-trial event-related potentials in EEG. *International Journal of Bifurcation and Chaos*, *14*(2), 719-726.

Korzeniewska, a. (2003). Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method. *Journal of Neuroscience Methods*, *125*(1-2), 195-207.

Kringelbach, M. L., Jenkinson, N., Owen, S. L. F., & Aziz, T. Z. (2007). Translational principles of deep brain stimulation. *Nature Reviews Neuroscience*, *8*(8), 623-635.

Krumin, M., & Shoham, S. (2010). Multivariate autoregressive modeling and granger causality analysis of multiple spike trains. *Computational Intelligence and Neuroscience*, *2010*, 752428, 1-9.

Kuehn, A. A., Doyle, L., Pogosyan, A., Yarrow, K., Kupsch, A., Schneider, G. H., Hariz, M. I., et al. (2006). Modulation of beta oscillations in the subthalamic area during motor imagery in Parkinson's disease. *Brain*, *129*(3), 695-706.

Kumar, V., Rahman, T., & Krovi, V. (1997). Assistive Devices For People With Motor Disabilities. *Wiley Encyclopaedia of Electrical and Electronics Engineering*. Wiley.

Kühn, A. a, Williams, D., Kupsch, A., Limousin, P., Hariz, M., Schneider, G.-H., Yarrow, K., et al. (2004). Event-related beta desynchronization in human subthalamic nucleus correlates with motor performance. *Brain*, *127*(4), 735-46.

Lal, T. N., Hinterberger, T., Widman, G., Schr, M., Hill, J., Rosenstiel, W., Elger, C. E., et al. (2005). Methods Towards Invasive Human Brain Computer Interfaces. *Advances in Neural Information Processing Systems*, *17*, 737-744.

Lau, C., & O'Leary, S. (1993). Comparison of Computer Interface Devices for Persons With Severe Physical Disabilities. *The American Journal of Occupational Therapy*, *47*, 1022-1030.

Law, C. K. H., Leung, M. Y. Y., Xu, Y., & Tso, S. K. (2002). A cap as interface for wheelchair control. *In the IEEE/RSJ International Conference on Intelligent Robots and Systems*, EPFL, Switzerland, 1439-1444.

Learned, R. E., & Willsky, A. S. (1995). A Wavelet Packet Approach to Transient Signal Classification. *Applied and Computational Harmonic Analysis*, *2*(3), 265-278.

Lebedev, M. A., & Nicolelis, M. A. L. (2006). Brain-machine interfaces: past, present and future. *Trends in Neurosciences*, *29*(9), 536-546.

Lee, G., Rodriguez, C., & Madabhushi, A. (2008). Investigating the efficacy of nonlinear dimensionality reduction schemes in classifying gene and protein expression studies. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *5*(3), 368-384.

Lenglet, C., Abosch, A., Yacoub, E., De Martino, F., Sapiro, G., & Harel, N. (2012). Comprehensive in vivo Mapping of the Human Basal Ganglia and Thalamic Connectome in Individuals Using 7T MRI. *PLoS One*, *7*(1), e29153.

Leuthardt, E. C., Schalk, G., Roland, J., Rouse, A., & Moran, D. W. (2009). Evolution of brain-computer interfaces: going beyond classic motor physiology. *Neurosurgical Focus*, *27*(1): E4.

Levan, P., Urrestarazu, E., & Gotman, J. (2006). A system for automatic artifact removal in ictal scalp EEG based on independent component analysis and Bayesian classification. *Clinical Neurophysiology*, *117*(4), 912-927.

Liu, H., & Yu, L. (2005). Toward integrating feature selection algorithms for classification and clustering. *IEEE Transactions on Knowledge and Data Engineering*, *17*(4), 491-502.

Liu, X., Ford-Dunn, H. L., Hayward, G. N., Nandi, D., Miall, R. C., Aziz, T. Z., & Stein, J. F. (2002). The oscillatory activity in the Parkinsonian subthalamic nucleus investigated using the macro-electrodes for deep brain stimulation. *Clinical Neurophysiology*, *113*(11), 1667-1672.

Liu, X., Wang, S., Yianni, J., Nandi, D., Bain, P. G., Gregory, R., Stein, J. F., et al. (2008). The sensory and motor representation of synchronized oscillations in the globus pallidus in patients with primary dystonia. *Brain*, *131*(6), 1562-1573.

LoPresti, E. F., Bodine, C., & Lewis, C. (2008). Assistive technology for cognition. *IEEE Engineering in Medicine and Biology Magazine*, *27*(2), 29-39.

Looney, C. G. (1997). *Pattern Recognition Using Neural Networks*. Oxford University Press.

Lotte, F., Congedo, M., Lecuyer, A., Lamarche, F., & Arnaldi, B. (2007). A review of classification algorithms for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, *4*(2), 1-13.

Loukas, C., & Brown, P. (2004). Online prediction of self-paced hand-movements from subthalamic activity using neural networks in Parkinson's disease. *Journal of Neuroscience Methods*, *137*(2), 193-205.

O'Malley, M. K. (2007). Principles of Human-machine Interfaces and Interactions. *In Life Science Automation: Fundamentals and Applications*, 101-125, Artech House Publishers.

Mace, M, Mamun, K. A., Vaidyanathan, R., Wang, S., & Gupta, L. (2010). Real-Time Implementation of a Non-Invasive Tongue-Based Human-Robot Interface. *In the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Taipei, Taiwan, 5486-5491

Mace, M., Mamun, K. A., Wang, S., Gupta, L., & Vaidyanathan, R. (2011). Ensemble classification for robust discrimination of multi-channel, multi-class tongue-movement ear pressure signals. *In the 33rd Annual International Conference of the IEEE EMBS*, Boston, USA, 1733-1736.

Mahmoud, S. S., Hussain, Z. M., Cosic, I., & Fang, Q. (2006). Time-frequency analysis of normal and abnormal biological signals. *Biomedical Signal Processing and Control*, *1*(1), 33-43.

Mamun, K. A., Mace, M., Gupta, L., Lutman, M. E., Stokes, M., Vaidyanathan, V., & Wang, S. (2012). Robust real-time identification of tongue movement commands from interferences. *Neurocomputing*, *80*, 83-92.

Mamun, K. A., Vaidyanathan, R., Lutman, M. E., Stein, J., Liu, X., Aziz, T., & Wang, S. (2011). Decoding Movement and Laterality from Local Field Potentials in the Subthalamic Nucleus. *In the 5th International IEEE/EMBS Conference on Neural Engineering (NER 2011)*, Cancun, Mexico, 128-131.

Mamun, K. A., Mace, M., Lutmen, M. E., Vaidyanathan, R., Gupta, L., & Wang, S. (2010). Multivariate Bayesian classification of tongue movement ear pressure signals based on the wavelet packet transform. *In the 2010 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Kittila, Finland, 208-213.

Marple, S. L. (1999). Computing the discrete-time "analytic" signal via FFT. *IEEE Transactions on Signal Processing*, *47*(9), 2600-2603.

Marzullo, T. C., Dudley, J. R., Miller, C. R., Trejo, L., & Kipke, D. R. (2005). Spikes, Local Field Potentials, and Electrocorticogram Characterization during Motor Learning in Rats for Brain Machine Interface Tasks. *In the 27th Annual International Conference of the IEEE EMBS*, Shanghai, China, 429-431.

Mason, S. G., Bashashati, A., Fatourechi, M., Navarro, K. F., & Birch, G. E. (2007). A comprehensive survey of brain interface technology designs. *Annals of Biomedical Engineering*, *35*(2), 137-169.

Mehring, C., Rickert, J., Vaadia, E., de Oliveira, S. C., Aertsen, A., & Rotter, S. (2003). Inference of hand movements from local field potentials in monkey motor cortex. *Nature Neuroscience*, *6*(12), 1253-1254.

Micera, S., Citi, L., Rigosa, J., Carpaneto, J., Raspopovic, S., Pino, G. D., Rossini, L., et al. (2010). Decoding Information From Neural Signals Recorded Using Intraneural Electrodes: Toward the Development of a Neurocontrolled Hand Prosthesis. *Proceedings of the IEEE*, *98*(3), 407-417.

Muller, K. R., Anderson, C. W., & Birch, G. E. (2003). Linear and nonlinear methods for brain-computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *11*(2), 165-169.

Muller, K., Krauledat, M., Dornhege, G., Curio, G., & Blankertz, B. (2004). Machine Learning Techniques for Brain-Computer Interfaces. *Journal of Biomed Tech*, *49*(1), 11-22.

Musallam, S., Corneil, B. D., Greger, B., Scherberger, H., & Andersen, R. A. (2004). Cognitive control signals for neural prosthetics. *Science*, *305*(5681), 258-262.

Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain Computer Interfaces, a Review. *Sensors*, *12*(2), 1211-1279.

Ojakangas, C. L., Shaikhouni, A., Friehs, G. M., Caplan, A. H., Serruya, M. D., Saleh, M., Morris, D. S., et al. (2006). Decoding movement intent from human premotor cortex neurons for neural prosthetic applications. *Journal of Clinical Neurophysiology*, *23*(6), 577-584.

Oskoei, M. A., & Hu, H. (2007). Myoelectric control systems-A survey. *Biomedical Signal Processing and Control*, *2*(4), 275-294.

Pasqualotto, E., Federici, S., & Belardinelli, M. O. (2011). Toward functioning and usable brain–computer interfaces (BCIs): A literature review. *Disability and Rehabilitation Assistive Technology*, 7(2), 89-103.

Patil, P. G., & Turner, D. A. (2008). The Development of Brain-Machine Interface Neuroprosthetic Devices. *Neuroprosthetics*, *5*(1), 137-146.

Percival, D. B., & Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge University Press.

Peterson, D. A., Knight, J. N., Kirby, M. J., Anderson, C. W., & Thaut, M. H. (2005). Feature Selection and Blind Source Separation in an EEG-Based Brain-Computer Interface. *EURASIP Journal on Advances in Signal Processing*, *2005*(19), 3128-3140.

Pires, G., Nunes, U., & Castelo-Branco, M. (2007). Single-Trial EEG Classification of Movement Related Potential. *In the IEEE 10th International Conference on Rehabilitation Robotics*, Noordwijk, Netherlands, 569-574.

Quinquis, A. (1998). A few practical applications of wavelet packets. *Digital Signal Processing*, *8*(1), 49-60.

Quiroga, R. Q., & Garcia, H. (2003). Single-trial event-related potentials with Wavelet Denoising. *Clinical Neurophysiology*, *114*(2), 376-390.

Rangayyan, R. M. (2002). *Biomedical Signal Analysis - A Case-Study Approach*. *Signals*. John Wiley & Sons.

Raudys, S. J., & Jain, A. K. (1991). Small sample size effects in statistical pattern recognition: Recommendations for practitioners. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *13*(3), 252-264.

Rezaei, S., Tavakolian, K., Nasrabadi, A. M., & Setarehdan, S. K. (2006). Different classification techniques considering brain computer interface applications. *Journal of Neural Engineering*, *3*(2), 139-144.

Rickert, J., Oliveira, S. C. D., Vaadia, E., Aertsen, A., Rotter, S., & Mehring, C. (2005). Encoding of movement direction in different frequency ranges of motor cortical local field potentials. *Journal of Neuroscience*, *25*(39), 8815-8824.

Roebroeck, A., Formisano, E., & Goebel, R. (2005). Mapping directed influence over the brain using Granger causality and fMRI. *NeuroImage*, *25*(1), 230-242.

Rosin, B., Slovik, M., Mitelman, R., Rivlin-Etzion, M., Haber, S. N., Israel, Z., Vaadia, E., et al. (2011). Closed-loop deep brain stimulation is superior in ameliorating parkinsonism. *Neuron*, *72*(2), 370-84.

Rui, Z., McAllister, G., Scotney, B., McClean, S., & Houston, G. (2006). Combining wavelet analysis and Bayesian networks for the classification of auditory brainstem response. *IEEE Transactions on Information Technology in Biomedicine*, *10*(3), 458-467.

Saeys, Y., Inza, I., & Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics*, *23*(19), 2507-17.

Saito, N., & Coifman, R. (1995). Local discriminant bases and their applications. *Journal of Mathematical Imaging and Vision*, *5*(4), 337-358.

Samar, V. J., Bopardikar, a, Rao, R., & Swartz, K. (1999). Wavelet analysis of neuroelectric waveforms: a conceptual tutorial. *Brain and Language*, *66*(1), 7-60.

Santos, F. J., Costa, R. M., & Tecuapetla, F. (2011). Stimulation on demand: closing the loop on deep brain stimulation. *Neuron*, *72*(2), 197-8.

Sato, J. R., Junior, E. A., Takahashi, D. Y., Felix, M. M., Brammer, M. J., & Morettin, P. A. (2006). A method to produce evolving functional connectivity maps during the course of an fMRI experiment using wavelet-based time-varying Granger causality. *NeuroImage*, *31*(1), 187-196.

Saygin, Z. M., Osher, D. E., Koldewyn, K., Reynolds, G., Gabrieli, J. D. E., & Saxe, R. R. (2011). Anatomical connectivity patterns predict face selectivity in the fusiform gyrus. *Nature Neuroscience*, *15*(2), 321-327.

Schalk, G., & Mellinger, J. (2010). *A Practical Guide to Brain–Computer Interfacing with BCI2000*. Springer London.

Scherberger, H., Jarvis, M. R., & Andersen, R. A. (2005). Cortical Local Field Potential Encodes Movement Intentions in the Posterior Parietal Cortex. *Neuron*, *46*(2), 347-354.

Scherer, R., Zanos, S. P., Miller, K. J., Rao, R. P. N., & Ojemann, J. G. (2009). Classification of contralateral and ipsilateral finger movements for electrocorticographic brain-computer interfaces. *Neurosurgical Focus*, *27*(1).

Seth, A. K. (2007). Granger Causality. *Scholarpedia*, *2*(7), 1667.

Seth, A. K. (2010). A MATLAB toolbox for Granger causal connectivity analysis. *Journal of Neuroscience Methods*, *186*(2), 262-273.

Shen, K.-Q., Ong, C.-J., Li, X.-P., Hui, Z., & Wilder-Smith, E. P. V. (2007). A feature selection method for multilevel mental fatigue EEG classification. *IEEE Transactions on Biomedical Engineering*, *54*(7), 1231-1237.

Shiliang, S., Changshui, Z., & Naijiang, L. (2005). On the on-line learning algorithms for EEG signal classification in brain computer interfaces. *Lecture Notes in Artificial Intelligence*, Springer, *3614*, 638-647.

Shin, K., & Hammond, J. K. (2008). *Fundamental of Signal Processing for Sound and Vibration Engineers*. Wiley publications.

Silchenko, A. N., Adamchic, I., Pawelczyk, N., Hauptmann, C., Maarouf, M., Sturm, V., & Tass, P. a. (2010). Data-driven approach to the estimation of connectivity and time delays in the coupling of interacting neuronal subsystems. *Journal of Neuroscience Methods*, *191*(1), 32-44.

Simpson, R. C., LoPresti, E. F., & Cooper, R. A. (2008). How many people would benefit from a smart wheelchair? *The Journal of Rehabilitation Research and Development*, *45*(1), 53-72.

Spadotto, A. A., Gatto, A. R., Guido, R. C., Montagnoli, A. N., Cola, P. C., Pereira, J. C., & Schelp, A. O. (2009). Classification of normal swallowing and oropharyngeal dysphagia using wavelet. *Applied Mathematics and Computation*, *207*(1), 75-82.

Stone, J. V. (2005). Independent Component Analysis, *Encyclopedia of Statistics in Behavioral Science*, 907-912, Jhon Wiley & Sons.

Stroke, N. I. of N. D. and. (2008). *Spinal cord injury: Hope through research*. National Institute of Neurological Disorders and Stroke, Bethesda (MD): National Institutes of Health. http://www.ninds.nih.gov/disorders/sci/detail_sci.htm

Struijk, L. N. S. A. (2006). An inductive tongue computer interface for control of computers and assistive devices. *IEEE Transactions on Biomedical Engineering*, *53*(12), 2594-2597.

Takami, O., Irie, N., Kang, C., Ishimatsu, T., & Ochiai, T. (1996). Computer interface to use head movement for handicapped people. *In the 1996 IEEE Tencon - Digital Signal Processing Applications Proceedings*, Perth, Western Australia, 468-472.

Thulasidas, M., Guan, C., & Wu, J. K. (2006). Robust classification of EEG signal for brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *14*(1), 24-29.

Tibble, M. (2004). *User's guide to disability estimates and definitions*. Department for Work and Pensions, UK. http://www.dwp.gov.uk/asd/asd5/ih2003-2004/IH128userguide.pdf

Ubeyli, E. D. (2007). ECG beats classification using multiclass support vector machines with error correcting output codes. *Digital Signal Processing*, *17*(3), 675-684.

Unser, M., & Aldroubi, A. (1996). A review of wavelets in biomedical applications. *Proceedings of the IEEE*, *84*(4), 626-638.

Vaidyanathan, R., Fargues, M., Gupta, L., Kota, S., Dong, L., & West, J. (2006). A dual-mode human-machine interface for robotic control based on acoustic sensitivity of the aural cavity. *In the 1st IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics*, Pisa, Italy.

Vaidyanathan, R., Chung, B., Gupta, L., Kook, H., Kota, S., & West, J. D. (2007). Tongue-movement communication and control concept for hands-free human-machine interfaces. *IEEE Transactions on Systems Man and Cybernetics Part A-Systems and Humans*, *37*(4), 533-546.

Vaidyanathan, R., Fargues, M. P., Serdar Kurcan, R., Gupta, L., Kota, S., Quinn, R. D., & Lin, D. (2007a). A Dual Mode Human-Robot Teleoperation Interface Based on Airflow in the Aural Cavity. *The International Journal of Robotics Research*, *26*(11-12), 1205-1223.

Vaidyanathan, R., & James, C. J. (2007). Independent component analysis for extraction of critical features from tongue movement ear pressure signals. *In the 29th Annual International Conference of the IEEE EMBS*, Lyon, France, 5481-5484.

Vaidyanathan, R., Wang, S., & Gupta, L. (2008). A Wavelet Denoising Approach for Signal Action Isolation in the Ear Canal. *In the 30th Annual International Conference of the IEEE EMBS*, Vancouver, Canada, 2677-2680.

Vapnik, V. N. (1999). An overview of statistical learning theory. *IEEE Transactions on Neural Networks*, *10*(5), 988-999.

Vaughan, T. M., Heetderks, W. J., Trejo, L. J., Rymer, W. Z., Weinrich, M., Moore, M. M., Kubler, A., et al. (2003). Brain-computer interface technology: A review of the second international meeting. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *11*(2), 94-109.

Walden, A. (2001). Wavelet analysis of discrete time series. *In the 3rd European Congress of Mathematics*, Barcelona, Spain, 627-641.

Wang, L., Zhou, N., & Chu, F. (2008). A General Wrapper Approach to Selection of Class-Dependent Features. *IEEE Transactions on Neural Networks*, *19*(7), 1267-1278.

Wang, S., Chen, Y., Ding, M., Feng, J., Stein, J. F., Aziz, T. Z., & Liu, X. (2007). Revealing the dynamic causal interdependence between neural and muscular signals in Parkinsonian tremor. *Journal of the Franklin Institute*, *344*(3-4), 180-195.

Wang, S, Liu, X. G., Yianni, J., Aziz, T. Z., & Stein, J. F. (2004). Extracting burst and tonic components from surface electromyograms in dystonia using adaptive wavelet shrinkage. *Journal of Neuroscience Methods*, *139*(2), 177-184.

Wang, Suogang. (2009). *Enhancing Brain-computer Interfacing through Advanced Independent Component Analysis Techniques*, *PhD Thesis*. Institute of sound and vibration research, University of Southampton, UK.

Wei, L., & Hu, H. (2011). *Towards Multimodal Human-Machine Interface for Hands-free Control : A survey*, Technical Report CES-510, School of Computer Science & Electronic Engineering, University of Essex, UK.

Widrow, B., & Stearns, S. D. (1985). *Adaptive Signal Processing*, Prentice-Hill Inc, USA.

Williams, D., Kühn, A., Kupsch, A., Tijssen, M., van Bruggen, G., Speelman, H., Hotton, G., et al. (2005). The relationship between oscillatory activity and motor reaction time in the parkinsonian subthalamic nucleus. *The European Journal of Neuroscience*, *21*(1), 249-58.

Wolpaw, J. R., Birbaumer, N., Heetderks, W. J., McFarland, D. J., Peckham, P. H., Schalk, G., Donchin, E., Quatrano, L. A., Robinson, C. J., & Vaughan, T. M. (2000). Brain-computer interface technology: A review of the first international meeting. *IEEE Transactions on Rehabilitation Engineering*, *8*(2), 164-173.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, *113*(6), 767-791.

Xie, H. B., Zheng, Y. P., & Guo, J. Y. (2009). Classification of the mechanomyogram signal using a wavelet packet transform and singular value decomposition for multifunction prosthesis control. *Physiological Measurement*, *30*(5), 441-457.

Yom-Tov, E., & Inbar, G. F. (2002). Feature selection for the classification of movements from single movement-related potentials. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *10*(3), 170-177.

Yu-Luen, C., Fuk-Tan, T., Chang, W., May-Keun, W., Ying-Ying, S., & Te-Son, K. (1999). The new design of an infrared-controlled human-computer interface for the disabled. *IEEE Transactions on Rehabilitation Engineering, 7*(4), 474-481.

Zabinsky, Z. B. (2009). *Random Search Algorithms*, Wiley Encyclopedia of Operations Research and Management Science.

Zandi, A. S., Javidan, M., Dumont, G. A., & Tafreshi, R. (2010). Automated real-time epileptic seizure detection in scalp EEG recordings using an algorithm based on wavelet packet transform. *IEEE Transactions on Biomedical Engineering*, *57*(7), 1639-1651.

Zhai, S., & Salem, C. (1997). An isometric tongue pointing device. *In the SIGCHI Conference on Human Factors in Computing Systems*. Georgia. New York, 22-27.

Zhao, Z., Morstatter, F., Sharma, S., Alelyani, S., Anand, A., & Liu, H. (2010). Advancing Feature Selection Research. *ASU Feature Selection Repository Arizona State University*, 1-28.

Zheng, S., Liu, J., & Tian, J. W. (2004). A new efficient SVM-based edge detection method. *Pattern Recognition Letters, 25*(10), 1143-1154.

Zou, C., Denby, K. J., & Feng, J. (2009). Granger causality vs. dynamic Bayesian network inference: a comparative study. *BMC Bioinformatics, 10*(122), 1-17.