DETECTION OF MOVEMENT INTENTION ONSET FOR BRAIN-MACHINE INTERFACES

by

Steven Christopher McGie

A thesis submitted in conformity with the requirements for the degree of Master of Applied Science (M.A.Sc.)

Graduate Department of Biomedical Engineering

University of Toronto

© Copyright by Steven Christopher McGie (2009)
Abstract

The goal of the study was to use electrical signals from primary motor cortex to generate accurate predictions of the movement onset time of performed movements, for potential use in asynchronous brain-machine interface (BMI) systems. Four subjects, two with electroencephalogram and two with electrocorticogram electrodes, performed various movements while activity from their primary motor cortices was recorded. An analysis program used several criteria (change point, fractal dimension, spectral entropy, sum of differences, bandpower, bandpower integral, phase, and variance), derived from the neural recordings, to generate predictions of movement onset time, which it compared to electromyogram activity onset time, determining prediction accuracy by receiver operating characteristic curve areas. All criteria, excepting phase and change-point analysis, generated accurate predictions in some cases.
Table of Contents

Title Page.......................................................................................................................i
Abstract.........................................................................................................................ii
Table of Contents..........................................................................................................iii
List of Abbreviations......................................................................................................v
List of Equations...........................................................................................................vi
List of Figures................................................................................................................vii
List of Tables................................................................................................................viii

Chapter 1: Introduction..................................................................................................1
  1.1. Motivation.............................................................................................................1
  1.2. Classifications of Brain-Machine Interface Systems.............................................1
    1.2.1. Communication vs. Motor Restoration.........................................................1
    1.2.2. Invasive vs. Noninvasive Recording..............................................................3
    1.2.3. Performed vs. Imagined Movements............................................................4
    1.2.4. Synchronous vs. Asynchronous..................................................................4
  1.3. Techniques for Neural Signal Analysis...............................................................5
  1.4. The Present Study...............................................................................................8

Chapter 2: Objectives and Hypothesis..........................................................................9
  2.1. Objectives...........................................................................................................9
  2.2. Hypothesis..........................................................................................................9

Chapter 3: Methods......................................................................................................10
  3.1. Data Collection..................................................................................................10
    3.1.1. Overview.....................................................................................................10
    3.1.2. Subject 1....................................................................................................10
    3.1.3. Subject 2....................................................................................................12
    3.1.4. Subject 3....................................................................................................13
    3.1.5. Subject 4....................................................................................................15
  3.2. Data Analysis....................................................................................................16
    3.2.1. Overview.....................................................................................................16
    3.2.2. Automated Analysis Program: “Adapt From Raw”.....................................20
    3.2.3. Automated Analysis Program: “Preprocessing”..........................................21
    3.2.4. Automated Analysis Program: “Prepare Criteria”.......................................22
      3.2.4.1. Bandpower............................................................................................23
      3.2.4.2. Bandpower Integral...............................................................................24
      3.2.4.3. Phase....................................................................................................24
      3.2.4.4. Variance...............................................................................................25
      3.2.4.5. Sum of Differences...............................................................................26
      3.2.4.6. Fractal Dimension................................................................................26
      3.2.4.7. Spectral Entropy..................................................................................28
    3.2.5. Automated Analysis Program: “Analyze”.......................................................28
  3.3. Change-Point Analysis.........................................................................................33
  3.4. Motion-Artefact Investigation..............................................................................34

Chapter 4: Results.........................................................................................................36
  4.1. Overview............................................................................................................36
  4.2. Subject 1............................................................................................................42
  4.3. Subject 2............................................................................................................43
4.4. Subject 3........................................................................................................46
4.5. Subject 4........................................................................................................47
4.6. Motion-Artefact Investigation.................................................................48
Chapter 5: Discussion..................................................................................53
  5.1. Possible Sources of Contamination....................................................53
    5.1.1. Motion-Induced Artefact...............................................................53
    5.1.2. Eye-Blink Artefact..........................................................................57
    5.1.3. Duration of Neural Events............................................................58
    5.1.4. Temporal Misalignment.................................................................60
    5.1.5. Non-Homogenous Movements....................................................61
    5.1.6. Non-Stationarity............................................................................62
  5.2. EEG vs. ECoG Recordings.................................................................63
  5.3. Volitional vs. Cued Movements.........................................................64
  5.4. Unsuccessful Criterion Functions....................................................66
    5.4.1. Change-Point Analysis...............................................................66
    5.4.2. Phase Analysis.............................................................................66
  5.5. Effects of Parameters........................................................................67
    5.5.1. Window Length...........................................................................67
    5.5.2. Acceptable Window Around EMG Activity..................................68
    5.5.3. Preprocessing and Electrode Location.........................................69
    5.5.4. Super-Threshold and Sub-Threshold Conditions.......................70
  5.6. Imagined Movements........................................................................72
Chapter 6: Conclusions............................................................................73
References..................................................................................................74
List of Abbreviations

AUC - area under the receiver operating characteristic curve
BMI - brain-machine interface
CUSUM - cumulative sum
DAQ - data acquisition system
ECoG - electrocorticogram
EEG - electroencephalogram
EMG - electromyogram
ERD - event-related desynchronization
ERS - event-related synchronization
FES - functional electrical stimulation
FFT - fast Fourier transform
GUI - graphical user interface
M1 - primary motor cortex
PCA - principal component analysis
PMC - premotor cortex
ROC - receiver operating characteristic
SMA - supplementary motor area
SSA - singular spectrum analysis
List of Equations

Equation 1: Final number of neural signals analyzed..............................................22
Equation 2: Fractal dimension calculation.................................................................27
Equation 3: Spectral entropy calculation.................................................................28
List of Figures

Figure 1: Screenshot of the main GUI program used to analyze the current data……..17

Figure 2: Screenshot of the GUI used to establish or modify a parameter file for the analysis program.................................................................18

Figure 3: Summary of the best AUC values obtained for a given pairing of trial and analysis type.................................................................36

Figure 4: Sample ROC curve, criterion function, predictions, and true EMG for an analysis with one of the highest prediction qualities obtained in the current study (AUC = 0.863).........................................................................................38

Figure 5: Sample ROC curve, criterion function, predictions, and true EMG for an analysis with the minimum AUC value that is considered “successful” (AUC = 0.700)...........................................................................39

Figure 6: Sample ROC curve, criterion function, predictions, and true EMG for an analysis with predictions being generated at random chance (AUC = 0.500).................................40

Figure 7: Sample of the results obtained using the method proposed by Hosaka et al. (2006) for cancelling the motion-induced artefact in EEG signals.................................51
List of Tables

Table 1: Summary of the relative frequencies with which specific parameters gave rise to successful movement onset predictions……………………………………………………………41

Table 2: Summary of correlation values from the experiment investigating the possible motion-induced artefacts within Subject 3’s trials………………………………………………49
Chapter 1: Introduction

1.1. Motivation

The term “neuroprosthetics” broadly describes the division of medical technology that is concerned with restoring lost functions to individuals with neurological injury or disorder. One of the problems facing neuroprosthetics researchers is that of developing an appropriate control system to dictate how and when the system is activated. Existing systems use a variety of control systems, including push-buttons, joysticks, and electromyograph (EMG) signals or joints movements from less disabled areas (Popovic and Thrasher 2004). For patients with more severe disabilities (such as high-level spinal cord injuries), who may have difficulty generating appropriate muscle activity and joint movements, these conventional systems may be of limited use. More opaque systems, such as the push-buttons, still require small motor functions that the patient may not be able to generate. Efforts have therefore begun to develop brain-machine interfaces (BMI), in which neuroprosthetics are controlled by signals from the patient’s own brain. The present study sought to develop techniques which could contribute to a larger BMI system, and to the field of BMI research.

1.2. Classifications of Brain-Machine Interface Systems

1.2.1. Communication vs. Motor Restoration

Work to date in the field of brain-machine interface (BMI) has had a variety of different goals. Many studies have aimed at using BMI systems for communication or computer interface purposes (Birbaumer & Cohen 2007). This approach is generally the simplest to implement, since the end effector does not require any direct interaction with the user’s
body. There are a wide variety of approaches even within this aspect of the field (for example one-dimensional or two-dimensional cursor control, or single-switch control strategies) (Birbaumer & Cohen 2007).

By contrast, systems that use BMI to control prosthetic or assistive devices for motor restoration are rarer, and generally more difficult. A great deal of work has been done in a basic neurophysiological setting to try to decode the neural activity related to a person’s intention to move (for example, Georgopoulos et al. 1988). It is hoped that if this activity is understood, that it can be used as a very natural and intuitive control signal for movement-related BMI systems. In brief, this would mean that a paralyzed individual would, through their intention to move, activate their prosthetic or assistive devices in such a way as to perform the movement that they intended. The theory behind this has been applied with some success to robotic systems (for example, Schwartz et al. 2006). However, applications to neuroprosthetic systems such as functional electrical stimulation (FES) (in which electrical stimuli are delivered to nerves or muscles that do not typically receive innervation, to cause contractions) (Popovic & Thrasher 2004) have been less common, owing largely to the more complicated nature of movement production through this route. For example, only two studies to date have connected FES with BMI: one in humans (Pfurtscheller et al. 2003), and one in monkeys (Moritz et al. 2008).
1.2.2. Invasive vs. Noninvasive Recording

A further division within brain-machine interface systems being developed for prosthetic and/or assistive devices is whether or not they require invasive procedures, particularly in terms of the recording modality being used. Noninvasive recording is limited to electroencephalography (EEG). Invasive recordings can be obtained from a variety of electrode types, which run the gamut from specific, single-cell recordings to broader recordings from entire cortical areas (the latter being termed electrocorticography, or ECoG). Work involving invasive recordings tends to be limited to animal models, whereas noninvasive EEG-based studies tend to come from human subjects. There are, however, exceptions, usually when an individual has required the implantation of recording electrodes for unrelated clinical reasons (for example, Schalk et al. 2007).

To some extent, the type of recording being used informs the type of interface that is to be developed, since different modalities are better for deriving different kinds of information from the brain. For example, the majority of studies that examine the direction of an intended reaching movement are performed with fine, invasive electrodes in animal models, since this information is conveyed by relatively small areas of cortex (Schwartz 2007). By contrast, many of the single-switch communication systems being developed are based on EEG recordings from human subjects, since such a system requires only broad cortical activity, and since communication cannot be easily tested with an animal model (Birbaumer & Cohen 2007).
1.2.3. Performed vs. Imagined Movements

Many of the current studies concerning the neural control of movement are centered around having subjects actually perform certain movements while neural recordings are being obtained. This is especially true of animal studies, where there is little alternative. While these studies are highly informative, they pose a unique problem for researchers interested in developing BMI systems for individuals with high-level paralysis. A user in this target population is, by definition, unable to actually perform the movements that result from the neural activity. Additional research has therefore been conducted to address the question of how neural signals appear in either a healthy individual imagining themselves performing a movement, or a paralyzed individual attempting to move an affected body part. The general consensus of these studies has been that imagining or attempting a movement tends to activate the same brain areas as actually performing the movement, and that similar analysis methods can be used in either case (Lacourse et al. 1999, Leuthardt et al. 2004, Pfurtscheller & Neuper 2006). It should be noted, however, that imagined or attempted movements tend to produce changes of a smaller amplitude than performed movements (Lacourse et al. 1999).

1.2.4. Synchronous vs. Asynchronous

One final way of classifying BMI systems concerns the timing of their input and output signals. “Synchronous” systems are those in which the user is cued as to when they should provide the system with an input (that is, when they should attempt to modulate their neural activity). By contrast, an “Asynchronous” system is one that continually searches for an input. This type of system does not use a cue, and instead provides an
output in response to the user’s volition. At present, most systems in development are of the synchronous variety, owing to their simpler implementation (Townsend et al. 2004, Mason et al. 2006). Synchronous systems are, however, limited in their applicability, and may be more cumbersome and less intuitive for a user, especially for motor control purposes (Townsend et al. 2004). The study being proposed herein will therefore be focused on the development of asynchronous systems, in the hopes that they can provide intuitive and effective control for motor control applications.

The inherent problem with the implementation of an asynchronous BMI is that it requires the system to have an “onset” time. Most of the studies in the motor-prediction BMI literature are presently focused on typifying or classifying a movement intention when it is known to occur (Townsend et al. 2004). The problem of “detecting” movement intention is therefore relatively unexplored, but those groups that have examined it have approached it from a number of different angles. Some groups, for example, treat the problem as an extension of the classification problem, looking for instances in time when the classification shifts from one category to another (Townsend et al. 2004). Other groups simply calibrate for a resting state, and look for any statistical deviation from this (Schalk et al. 2008).

1.3. Techniques for Neural Signal Analysis
In most cases, BMI systems, whether for communication, computer interaction, or movement, use the same general principle for decoding neural activity. This is the principle of event-related synchronization (ERS) and event-related desynchronization
(ERD) (Pfurtscheller & Lopes da Silva 1999). In brief, it is thought that the signals obtained through most neural modalities (such as EEG or ECoG) may be approximately represented as a sum of component sinusoidal waveforms of different frequencies. The bandpower of these component waveforms is indicative of the number of neurons under the recording electrode that are firing at the given frequency in synchrony. It is widely believed that brain areas that are inactive will have more synchrony in components that cycle at lower frequencies (for example, 8-12 Hz), whereas active areas will synchronize at higher frequencies (from 12 Hz up, although 40 Hz is a classically suggested value) (Toro et al. 1994). A decrease in a particular frequency range is termed ERD, while an increase is called ERS. BMI systems will therefore often be designed to look for ERD in lower frequency ranges, and ERS is higher frequency ranges.

It has been suggested by some researchers that ERS and ERD can be seen in ways other than frequency-domain analysis, as well. Most notably, the phases of the higher-frequency component waveforms of the different cortical areas involved in the movement intention are thought to “lock” together during ERS (Lopes da Silva 2006). Conversely, inactive areas are not expected to be phase-locked, and the relative disarray in the phases of different areas can be taken as an indication of relative inactivity. This approach to ERD and ERS is not as well-founded as investigating the bandpowers, as outlined above.

There are, however, alternative approaches to deriving information from neural signals. It has been suggested that any of a number of “complexity” measures (including but not limited to fractal dimension and spectral entropy) can provide useful information about
the signal (Liu et al 2005, Rezek & Roberts 1998). The reason for applying these methods is that activity in a given brain area is thought to increase the complexity of the waveforms recorded from it, since an active signal should contain a wider range of component waveforms, and tend to be cycling at a higher frequency (Rezek & Roberts 1998).

An additional analysis technique that holds the potential to detect onset times in neural signals is referred to as the “change-point” method. In brief, the method provides estimates of the fundamental spectral components of the signal within a sliding time window, and looks for windows in which the actual signal deviates from this predicted model (Moskvina & Zhigljavsky 2003). The estimation is produced by singular spectrum analysis (SSA) (Broomhead & King 1986). SSA works by first defining a trajectory matrix (that is, a matrix composed of the original time series and a given number of sequentially-lagged copies of itself), and then taking a certain number of the principal components of this matrix. Deviations from this model are measured as Euclidean distances between the SSA model and a delayed trajectory matrix based on the original signal (called the test matrix). Further details of the procedure can be found in Moskvina & Zhigljavsky (2003). It is worth noting that the term “change-point” has been applied to a variety of different analytical techniques, which, despite sharing the same final goal, may employ entirely different mathematical principles (for example, Taylor 2000, and Kaplan & Shishkin 2000). However, only the method proposed by Moskvina & Zhigljavsky (2003) was tested in the current study, as it was the most well-founded method available within the literature, and the simplest to reproduce.
1.4. The Present Study

The present study sought to take a simple approach to the problem of detecting the onset time of performed movements. It was based largely on “classical” methods of extracting information from neural signals (such as ERD/ERS and complexity measures), while also applying a number of novel techniques. Due to the uncertainty within the field about the best way to interpret neural signals, a wide range of analytical techniques and parameter sets were employed, and the relative accuracy of the onset time predictions they provided were be objectively assessed. It was hoped that, through this, a conclusive decision could be reached about the relative usefulness of each analytical method, and about the feasibility of using neural signals to derive predictions of movement intention onset time, for use as control signals in neuroprosthetic systems.

The current thesis is mostly focussed on the development of a robust program with which to automate analysis. This first requires the development of the specific analysis techniques which the program will employ. Several experiments worth of preliminary data has been collected and analyzed as well, although the nature of the data and its subsequent analysis make quantitative analysis complex. The present results can therefore only be considered pilot data, and a demonstration of the capability of the analysis program that has been developed to generate predictions and give an approximate measure of their accuracy.
Chapter 2: Objectives and Hypothesis

2.1. Objectives

The objectives of the current study followed one another in a step-wise manner. The first goal was to develop analysis techniques with which to derive potentially useful information from neural signals and to generate predictions of movement intention onset time. Following this, the next step was to develop analysis techniques by which the accuracy of these predictions could be quantitatively assessed. The third step was to integrate the aforementioned techniques into a robust program which received recordings of neural and muscular activity as inputs and which tested a variety of parameter combinations to determine accuracy measures for each combination as outputs. The final goal was to perform pilot experiments to assess the feasibility of the methods that had been developed, and to provide initial results to suggest which parameter sets and analytical techniques provided the most useful onset time predictions.

2.2. Hypothesis

The main hypothesis of the current study was that at least one analytical technique exists whereby a user’s intention to move his/her upper limb can be predicted in a robust, reliable manner, and with a high degree of temporal accuracy, based on EEG or ECoG recordings from primary motor cortex.
Chapter 3: Methods

3.1. Data Collection

3.1.1. Overview

The experiments took neural recordings from either electrocorticogram (ECoG) or electroencephalogram (EEG) electrodes. Two subjects participated in EEG-based experiments, and two in ECoG-based experiments, for a total of four subjects. Each individual experiment differed in terms of design, despite many commonalities that allowed for comparisons between them. The ECoG subjects gave written consent as approved by the University Health Network Research Ethics Board, and the EEG subjects gave written consent as approved by the Toronto Rehab and University of Toronto Research Ethics Boards.

3.1.2. Subject 1

Subject 1 was a 55 year old male who had four ECoG electrodes implanted over his right primary motor cortex (M1) as part of his treatment for chronic pain. The subject was seated in front of a wooden frame, which had seven large buttons attached to it (Jelly Bean Twist, Technical Solutions, Australia). These included one at the center of the frame, and one in each cardinal direction relative to the center (up, down, left, right, forward, and backward). A cue system was also displayed on a monitor in front of the subject. The subject started with his left hand touching but not pressing the center button. The cue program began by cueing the subject with one of six commands (for example, “Left”). The order in which the six commands were presented to the subjects was randomized within each set. After five seconds, a light appeared to signal that the subject
should begin his reaching movement and press the button indicated in the cue. After another second, the light disappeared. After three more seconds, the next cue appeared, which was always “Return” after a movement had been completed. The subject then moved to press the center button. The next cues given to the subject were to imagine performing the same movements that he actually did just perform, with the same timing as the performance thereof. One trial consisted of 24 cues, such that one performed and one imagined reaching movement was conducted in each of the six directions. Ten trials were performed in total, each divided as a separate recording. An additional test was conducted in which the subject was asked to play the Tennis portion of the video game Wii Sports (Nintendo of Canada, Canada) for approximately five minutes, while recordings were obtained. This game required the player to make swinging motions with the controller, and therefore served the purpose of getting them to perform large, multi-joint arm movements.

Recordings over the course of Subject 1’s experiment were of several types. Four ECoG electrodes, placed along the subject’s right M1, provided the neural signals. These signals were acquired via a digital system on a personal computer running NeuroScan software (NeuroScan Laboratories, USA), at a sampling frequency of 200 Hz and with a bandpass filter from 0.3 to 100 Hz. Two electromyogram (EMG) electrodes were placed on the subject’s left arm, one over the biceps and one over the wrist extensor, to capture information about the activity of relevant muscle groups. The EMG signals were captured with the same Neuroscan system as the ECoG signals. Three motion-tracking sensors (as part of the Polhemus Fastrak system (Polhemus, USA)) were also attached to the left arm:
one approximately 3 cm below the lateral condyle of the humerus, one approximately 4 cm below a point midway between the radial and ulnar styloid process, and one over the dorsal aspect of the distal phalanx of the index finger. Each sensor acquired a continuous recording of the sensor position in six dimensions (X, Y, Z, roll, pitch, and yaw) at a sampling frequency of approximately 40 Hz. Electrical activity from the buttons being pressed was also recorded, for timing information. The cue system was designed to send an electrical pulse to the recording system when the signal to perform a movement was given, also for timing information. A specialized mouse was used to synchronize signals from the Polhemus Fastrak with the other recordings. Clicking the button of this mouse simultaneously started the motion-tracking recordings and delivered an electrical pulse to the recording system. The electrical signals for button presses, cue commands, and mouse clicks were all interfaced to the NeuroScan system by means of a National Instruments DAQPad (National Instruments, USA) plugged into the EMG channels of the NeuroScan system.

It is worth noting that Subject 1 had previously had surgery done to fuse his left wrist with implanted screws. The left wrist was therefore incapable of bending, although the relevant musculature remained intact. Furthermore, the subject was right-handed, and movements were therefore being performed with his non-preferred limb.

3.1.3. Subject 2

Subject 2 was a 67 year old female, who had four ECoG electrodes implanted over her left M1 as part of her treatment for essential tremor. She was medicated at the time of
testing. The electrodes were confirmed to be placed over the area of M1 associated with upper extremity control, by providing stimuli to the electrode and observing the subsequent muscle contractions (100 Hz, 100 ms, monopolar, 3-10 µA) (as per Hanajima et al. 2002). In specific, this stimulation produced a hand closing motion in the subject. Recordings from the ECoG electrodes were obtained in a manner identical to that described for Subject 1. Two EMG electrodes were placed on the subject’s right arm, one to capture information from wrist flexors, and one to capture information from wrist extensors. One Polhemus Fastrak sensor was attached to the third metacarpal of the subject’s right hand. The same specialized mouse that was used in Subject 1’s experiment in order to synchronize the Fastrak recordings with the other signals was used in Subject 2’s experiment. The subject, while seated, was asked to perform repetitions of a specific motor task with her right arm, in response to two auditory cues (one warning cue and one execution cue). The motor tasks that the subject was asked to perform are as follows: resting, elbow flexion to a random angle, elbow flexion to a specified angle, reaching movements to the left and right, wrist flexion, and wrist extension. In some cases (particularly elbow flexion to a random angle and wrist flexion), the tasks were repeated for verification purposes.

3.1.4. Subject 3

The experimental design in the EEG experiments followed a similar format to the ECoG experiments. Subject 3 was a 24 year old male with no neurological injury. The movements being performed by Subject 3 were all conducted with the right arm, and fell into several categories: reaching left, reaching right, closing the hand, flexing the elbow,
and playing the Tennis component of the video game Wii Sports. Movements were performed volitionally (that is, without a cue, and at the subject’s own pace), with the exception of the Wii Sports game (wherein the timing was partially dictated by the pace of the game). Each movement was performed 30 times, with one single continuous trial for each movement type. An exception to this was the Wii Sports trial, in which approximately 40 movements were performed (covering approximately three minutes of gameplay). Movements were not randomized, since the subsequent analysis required repetitions of the movement to be contained within a single trial. Four EEG electrodes (Grass F-E5GH gold disc electrodes, Astro-Med, Canada) were placed along the subject’s left hemisphere, in the following positions: Cz, C3, halfway between Cz and C3, and halfway between C3 and T3. These signals were amplified with a Grass IP511 AC amplifier (Astro-Med, Canada). Two EMG electrodes (Bortec BiPole Ag/AgCl electrodes, Bortec Biomedical, Canada) were placed on the right arm, one on the biceps and one on the wrist flexor, grounded with a third electrode (Kendall Medi-Trace 200 Ag/AgCl electrode, Kendall, USA) over the epicondylus lateralis (that is, the bony protrusion at the side of the elbow), and amplified with a Bortec AMT-8 amplifier (Bortec Biomedical, Canada). A single motion-tracking (Polhemus Fastrak) sensor was placed overtop of the third metacarpal of the right hand. The onset of motion-tracker recording was synchronized with the other recordings by means of the same specialized mouse used in the ECoG experiments. The amplified EEG and EMG signals, as well as the signal from the mouse clicks, were collected via a data acquisition system (DAQ) (National Instruments, USA) and recorded through a custom LabView program (National Instruments, USA) running on a designated PC.
3.1.5. Subject 4

Subject 4 was a 22 year old male with no neurological injury. The experiment with Subject 4, while similar in structure to that of Subject 3, tested different movement types, and employed different recording equipment. The movement types included the following: elbow flexion (both volitional and cued), hand opening and closing (both volitional and cued, both allowing the hand to return to a naturally relaxed position in which the fingers were still mostly curved towards the palm, and forcing the fingers into a fully opened position), random multidirectional arm movement (cued only), reaching forward to touch a pack of playing cards (both volitional and cued), reaching forward to establish a pinch grip on a pack of playing cards (both volitional and cued), rotating a pack of playing cards held in the hand by 180° using finger movements (both volitional and cued), tensing of the arm muscles without movement (both volitional and cued), writing the sentence “the quick brown fox jumps over the lazy dog”, either with pauses after each word (both volitional and cued) or with a pause at the completion of the sentence (volitional only), playing the Tennis component of the video game Wii Sports, and playing the Bowling component of the same game. The purpose of having the subject play both the Tennis and Bowling components of Wii Sports was to include situations in which the subject’s movements are completely volitional (Bowling) and cases in which the timing of the movements is at least partially guided (Tennis, in which the player’s swings must be timed for when the ball approaches their character). In cases where the movement being performed was volitional, the subject was told to perform them at their own pace, but to be consistent in their movements, and to pause in between movements.
for at least one or two seconds. In cases where the movement was cued, the cue was a colored box on a computer monitor, which stayed red (meaning to relax) for three seconds, yellow for two seconds (meaning to be ready to perform the movement), and green for one second (the shift from yellow to green indicating the time at which the subject should start performing the movement). Four EEG electrodes (Grass F-E5GH gold disc electrodes) were placed along the subject’s left hemisphere, in the same positions as those used for Subject 3 (Cz, C3, halfway between Cz and C3, and halfway between C3 and T3). Two EMG electrodes were also played on the subject’s right arm, one over top of the bicep, and one over top of the wrist flexors. These signals were acquired via a digital system on a personal computer running NeuroScan software (NeuroScan Laboratories, USA) at a sampling frequency of 2000 Hz (later down-sampled to 200 Hz to expedite analysis) and with a bandpass filter from 0.05 to 200 Hz. Three Polhemus Fastrak sensors were also attached to the subject’s arm: one approximately 3 cm below the lateral condyle of the humerus, one approximately 4 cm below a point midway between the radial and ulnar styloid process, and one over the dorsal aspect of the distal phalanx of the index finger.

3.2. Data Analysis

3.2.1. Overview

Once all of the signals from an experiment were acquired, they were run through an automated analysis program that was developed in MATLAB (MathWorks, USA). The program operated through a graphical user interface (GUI). The main GUI for the program is shown in Figure 1.
Figure 1. Screenshot of the main GUI program used to analyze the current data. The user selects a parameter file that dictates the data to be analyzed and the techniques and parameters applied to it. Analysis then begins sequentially from the step that the user specifies, such that analysis can be halted between steps and resumed at a later time.

The program divided analysis into four stages:

1. Converting the raw signals obtained from a data acquisition (DAQ) system and/or Fastrak into a MATLAB format (“Adapt From Raw”).

2. Preprocessing the signals (“Preprocessing”).
3. Preparing “criterion” signals by applying a variety of functions to windowed segments of the preprocessed signal (“Prepare Criteria”).

4. Producing predictions of movement onset and assessing their accuracy (“Analyze”).

The exact manner in which analysis was conducted was dictated by a “parameter file”, which can be established and edited by a user with a separate GUI program (which is shown in Figure 2).

Figure 2. Screenshot of the GUI used to establish or modify a parameter file for the analysis program. The parameters included are as listed in Section 3.2.1. In brief, the user is able to select the trials to analysis, the methods and frequency ranges with which to analyze them, the layout of DAQ channels included in the analysis, the windows in which
predictions might be considered ‘true’, the windows with which the signals are scanned, and the thresholds for EMG analysis.

The following is a list of the different parameters which these files dictate, the details of which will be elaborated on in the explanation of the analysis:

- Sampling frequency
- Names of all of the movement types performed
- Names of all of the analysis types included
- Number of channels in the DAQ configuration
- Specification of the DAQ configuration
  - Which channels contain neural signals
  - Which channels contain EMG
  - Which of the EMG channels correspond to which muscles
- Thresholds used for determining “true” onset times from the EMG signals
- Number of Fastrak channels
- Number of thresholds examined for each criterion function in the “Analyze” step
- Directory names for the locations of both the raw DAQ and the raw Fastrak files
- Frequency ranges included in the frequency-domain-based analyses
- Acceptable time windows around the “true” onset times in which a prediction of onset can be considered “correct”
- Window sizes used for windowing signals in the “Prepare Criteria” step.
Note that not all of the data that had been collected could be properly analyzed with this program. In particular, Subject 1’s experiment with reaching in different directions could not be properly analyzed. The reason for this was that each trial recorded for that experiment included multiple different movement types (for example, reaches in each of the six cardinal directions) as well as imagined movements. Furthermore, the subject in these experiments was not entirely consistent in his behaviour. In early trials, for example, the subject held the push-buttons down rather than pressing them, which was changed for later trials. EMG traces also indicate that certain positions required the subject to sustain muscle contraction during the waiting period, which may provide additional confounds. Due to the multiple confounds arising from this experimental design, these trials could not be properly compared to the other recorded trials, and the validity of onset accuracy predictions within them remained uncertain.

Note also that while the analysis program allowed a user to specify Fastrak inputs or to specify certain DAQ channels as accelerometer signals, the analysis for this particular study did not employ these signals. These options were included to allow flexibility in the program should it need to be modified in the future for alternate analyses.

3.2.2. Automated Analysis Program: “Adapt From Raw”

The first step in the automated analysis program was to adapt the collected data from its raw format into a MATLAB format. In the case of the DAQ, the original data collection was done through a custom LabVIEW program. In the case of the motion-tracking data, it was done through a GUI provided by Polhemus. Through these programs, the raw data,
both from the DAQ and from the Fastrak, was originally stored in an ASCII format. Files from the DAQ consisted of one column for each DAQ channel in use. The analysis program read this data, and sorted it into arrays based on the layout of the DAQ, as specified in the chosen parameter file (that is, it grouped all neural recordings into a single two-dimensional array, all EMG recordings into another, all accelerometer data into another, and all information about clicks of the customized mouse button into another). ASCII files from the Polhemus Fastrak system consisted of seven columns, which indicated the following: sensor number, X location, Y location, Z location, roll, yaw, and pitch. A custom MATLAB program was therefore employed to first read the sensor number, and then to assign the subsequent data to a particular sub-array. The result was a three-dimensional array, arranged by sensor, component (ie. X or yaw), and time point. The program also upsampled the Fastrak data by linear interpolation so that the adapted Fastrak data had the same sampling frequency as that from the DAQ. The adapted data was all saved in a MATLAB format for future use.

3.2.3. Automated Analysis Program: “Preprocessing”

The second step in the analysis was to preprocess the neural signals. Two different kinds of preprocessing were applied. The first kind was principal component analysis (PCA). The purpose of PCA was to determine separate, underlying components which may be contributing to differing extents in all of the signals it receives (in this case, the recordings from all EEG or ECoG channels). For example, a given EEG signal may be a combination of signals from different sources within the brain, and the goal of PCA would be to isolate the signals coming from the different sources. It was therefore hoped
that using PCA could provide cleaner, more meaningful signals to our analysis. The end result of PCA was a number of “principal component” signals equal to the number of electrodes originally used. The second kind of preprocessing was differential signals. This simply meant subtracting the signal of one channel from that of another channel. In addition to the preprocessed signals, the original signals were also run through analysis, without subtraction or PCA.

The final number of neural signals subsequently analyzed was therefore:

\[
2n + \frac{n!}{2 \times (n-2)!}
\]  

(1)

where \(n\) is the number of electrodes originally used. One \(n\) in the first term represents the raw signals, while the second represents the principal components. The second term is half of the possible permutations of two electrodes being selected from the full set of \(n\) electrodes (halved because half of the combinations will select two electrodes that have already been paired, but in a different order.

3.2.4. Automated Analysis Program: “Prepare Criteria”

The third step in the analysis program was the production of “criterion signals”. In all cases, this process was based around the use of a sliding time window, the signal within which was used as an input to one of seven functions. The output of each of these functions was a single data point in the criterion signal. The same analysis, applied to the next time window, produced the next data point in the criterion signal, and so on. The parameter files allowed a user to enter several different widths of time window for use in the analysis. In the analysis in question, three window lengths were tested: 0.25 s, 0.5 s,
and 1 s. These window lengths were chosen originally due to their ability to produce promising results in initial visual inspection. Longer window lengths were not chosen so as to minimize the potential for delays, should these techniques be eventually applied to a system that requires real-time operation. Shorter lengths were not chosen so as to provide each window with sufficient data with which to extract meaningful information, and to minimize the potential effects of aberrant or outlying points.

The seven analytical functions previously alluded to will be referred to as follows: bandpower, bandpower integral, phase, variance, sum of differences, fractal dimension, and spectral entropy. The parameter files allowed a user to specify which of these they would like included in the analysis, and in the case of frequency-domain-based methods (bandpower and phase), which specific frequency ranges to analyze. Criterion signals, once generated by the analysis program, were saved in MATLAB-format files, so that they did not need to be recalculated in subsequent analysis.

3.2.4.1. Bandpower

The “bandpower” criterion was obtained by first running the windowed signal through a fast Fourier transform (FFT), and deriving the amplitude of a specified frequency component, or the mean of components within a given range (Pfurtscheller & Lopes da Silva 1999). This technique is among the most common in brain-machine interface literature. Typically, it is thought that neural activity causes a decrease in lower frequencies (approximately 8-12 Hz), and an increase in higher frequencies (greater than 20 Hz) (Pfurtscheller & Lopes da Silva 1999). These events are typically referred to as
event-related desynchronization (ERD) and event-related synchronization (ERS) respectively, since they are thought to indicate changes in the way in which neurons temporally synchronize their firing activity. The frequency bands employed in the current study included both ranges that are classically considered to be relevant for ERS and ERD (1-5, 8-12, 12-20, and 36-44 Hz), and ranges that were more novel and exploratory (20-30 and 90-99 Hz).

3.2.4.2. Bandpower Integral

The “bandpower integral” technique applied the same FFT to the windowed signal, but subsequently summed the bandpowers of all derived frequency components up to 60 Hz. This specific measure has not been widely investigated, and may in fact be considered a novel approach. However, it is thought that a larger bandpower integral could indicate a signal with higher amplitude in general, or a signal in which many changes are occurring.

3.2.4.3. Phase

The “phase” technique was also centered around applying FFT to the windowed signal. Whereas the bandpower technique focused on the real components of the FFT results (which represent the amplitudes of the component sinusoids), the phase technique employed the imaginary components of the results (which represent the temporal alignment of those same components). Literature investigating phase changes is more limited than that looking at amplitudes. Much of it has focussed on the same ERS and ERD phenomena suggested in other Fourier-based studies, but takes an approach that requires different electrode locations to be included in the analysis (Lopes da Silva 2006).
It is suggested that ERS should be visible as certain high-frequency components suddenly taking on similar phases in connected areas of the cortex (Lopes da Silva 2006). That is, these areas are expected to become “phase-locked” to each other. ERD, likewise, is expected to be seen as a number of low-frequency components suddenly losing their synchrony, and taking on widely differing phase values. However, due to the limited cortical range covered by our current recordings, this technique was not implemented or investigated within the scope of this study. Instead, the current analysis simply used the raw phase values as criteria. It is still considered possible that this may be sufficient analysis, since both ERS and ERD should be accompanied by a change in phase. The same frequency ranges that were applied to the bandpower analysis were applied to the phase analysis. This analysis was complicated by the fact that the use of a sliding window causes apparent phase changes when a Fourier transform is applied. The changes caused by a sliding time window, however, may be more regular and rhythmic in nature than those caused by ERS or ERD. For example, in the case of a perfectly sinusoidal signal, the effect of the sliding window would be a phase that appears to change linearly over time, with the same period as the sinusoid itself. While the neural signals in question are much more complex than this example, the same principle applies. Within a given neural state, the cycling in the phase is expected to be at least somewhat regular, and changes in neural state should still cause rapid changes independent of that.

3.2.4.4. Variance

The “variance” criterion, as its name suggests, was simply the squared standard deviation of the signal within the sliding time window. The method was implemented after visual
inspection of a particular set of signals suggested that the user’s movements coincided with an increase in the overall amplitude of the raw signal. This quality is captured in variance, since the calculation of standard deviation is centered around the difference between individual data samples and the mean of those samples. A higher-amplitude waveform is therefore expected to have a higher variance. It has been noted in the case of simple sinusoids that the frequency of oscillation seemed to have relatively little effect on the signal’s variance. The signal’s amplitude, however, had a clear and direct effect on the variance, as intended.

3.2.4.5. Sum of Differences

The “sum of differences” criterion was a computationally simple method that was intended to capture the extent to which a signal varies over time. The measure was simply the cumulative sum of the absolute differences between adjacent time points in the windowed signal. A signal that is more active is expected to show more variation over time, and will therefore have a higher criterion value. In this way, it is similar to the aforementioned “variance” criterion and the “fractal dimension” method that will be discussed below, though simpler and not as grounded in statistics. The sum of differences criterion is expected to be affected by changes in both the amplitude and the frequency of the signal.

3.2.4.6. Fractal Dimension

“Fractal dimension” was intended as a measure of signal complexity (Liu et al. 2005). Its implementation was strictly geometrical, and applied a philosophy somewhat similar to
that applied when calculating signal variance or standard deviation. The calculation was as follows:

$$FD = \frac{\ln(N - 1)}{\ln(N - 1) + \ln(d/L)}$$

(2)

Where FD is the fractal dimension, N is the number of data points included in the window, d is the maximum Euclidean distance from the first point in the window to any other point within the window, and L is the sum of the Euclidean distances between adjacent data points (Liu et al. 2005).

If changes in signal complexity (as represented by fractal dimension) are observed, they may reflect changes in the underlying neural activity. Generally, it is thought that increased activity should result in increased signal complexity, and there is considerable experimental evidence to support this. For example, Liu et al. (2005) found a positive linear correlation between grip strength and the fractal dimension of EEG signals obtained from the relevant area of M1. However, it is conceptually possible that an increase in activity could instead decrease signal complexity. This is similar to the ideas of ERS and ERD. In an active area, larger pools of neurons will be firing in synchrony. Their firing is more frequent (this being the explanation for the expected increase in complexity), but the synchrony of this firing may make the overall signal more coherent, and less like random noise (random noise itself having a comparatively high complexity). The current study therefore hoped to investigate both possibilities. In fact, it was an overarching theme of the present methodology to look not specifically for increases or
decreases in any of the derived criteria, but rather, to look for changes in them as they move away from baseline.

3.2.4.7. Spectral Entropy

“Spectral entropy” was an alternate measure of signal complexity, one which was based on a frequency spectrum, such as that generated by an FFT. The calculation following the FFT was as follows:

\[ H = \sum_{f} p_f \log(1/p_f) \]  

(3)

Where \( H \) is the spectral entropy, \( f \) is a given frequency component, and \( p_f \) is the bandpower for that frequency component. It was based on the philosophy that a more complex signal would require the involvement of a greater number of frequency components (Rezek and Roberts 1998) in order to be properly represented in the frequency domain. This method shared many similarities with the “bandpower integral” criterion outlined above. Since both fractal dimension and spectral entropy seek to represent signal complexity, the interpretation of the results from both methods should ideally be the same. However, it would not be surprising if they nonetheless produced different results in some circumstances, given their differing approaches to the problem of representing signal complexity.

3.2.5. Automated Analysis Program: “Analyze”

The fourth step in the analysis program involved the generation of predictions, and the assessment of their quality. Predictions were made by applying a threshold to each of the criterion signals that had been generated. The program tested two methods of
thresholding, one in which it predicted a movement when the criterion was below the threshold, and one in which it did the same when the criterion was above the threshold.

The assessment of prediction quality was based on receiver operating characteristic (ROC) curves. An ROC curve is a technique by which the discriminative capability of a thresholded criterion signal can be determined. This involves applying a range of different thresholds to the criterion signals and then graphing the relative true positive and false positive rates obtained with each threshold, in the form of a curve (Swets 1988). Examples of ROC curves can be seen at the tops of Figures 4-6. The quality of the criterion being employed could then be determined either by the maximum achievable distance between a point on the curve and the point representing a random situation where both the true positive rate and false positive rates are 0.5 (called d’, a distance of 0 representing random chance, and a distance of approximately 0.7071 [that is, \( \sqrt{0.5} \)], representing perfect discrimination), or by the area under the curve (abbreviated as AUC, an area of 1 representing perfect discrimination, and an area of 0.5 representing random chance) (Swets 1988). The subsequent discussion of the results is primarily based on AUC, rather than d’, since this measure gives a more thorough account of the full shape of the ROC curve, and therefore, a complete idea of the discriminative ability of the analytical method and parameter set in question. This is because AUC accounted for the results at all thresholds, whereas d’ only pertained to the results at the best threshold. The threshold which produced d’ was still, however, considered to be the best threshold, and would most likely be the threshold that would be used if these techniques were to be applied in the control of prosthetic or assistive devices. The number of thresholds
included in the ROC curve was controlled by a user via the parameter file, and the analysis program selected thresholds that covered the full range of data. In the present analysis, 102 thresholds are used for to scan across each criterion signal. These included 100 thresholds, one at each percentile of the data, one threshold less than the minimum value of the signal, and one threshold greater than the maximum value of the signal.

Unfortunately, the nature of the signals being examined and the predictions being generated made application of the ROC method difficult. First, the “true” movement onsets needed to be determined, which in the current study involved thresholding the EMG signals at levels determined in the parameter file. In the current analysis, the EMG thresholds were determined by visual inspection of the EMG signals. Kinematic parameters could alternately be used to determine movement onset, but the motion tracking system did not directly interface with the DAQ system used to record brain or EMG signals. Furthermore, its sampling frequency had been found to be variable over time, so that even if it was synchronized to the DAQ, it would lose its synchrony as the trial continued. It has been suggested that a series of accelerometers could alternately be used to determine onset time, and this possibility may be investigated further in future work. However, EMG and accelerometer signals cannot be considered substitutes for each other. For example, if a subject were to sustain a muscle contraction, such as holding their arm in the position that it ends a reaching movement with, EMG thresholding would detect this as a true movement, while accelerometer signals might not. Automated methods for thresholding EMG are available, such as change-point analysis (Vaisman et al., unpublished), or double-thresholding (Masani et al. 2009). However,
The application of such techniques would require the definition of additional parameters in the system, and often require comparison to resting periods, which were not recorded for all of the subjects examined in the current study. Furthermore, for the purposes of the current analysis, automation was not required. While visual inspection has a degree of subjectivity to it, the EMG traces recorded in the current experiments had a sufficiently high signal-to-noise ratio to allow clear distinctions between resting and active states to easily be made. Studies investigating automated methods refer to visually-based definitions of onset time when assessing the accuracy of the automated methods (Vaisman et al., unpublished), lending credence to the potential accuracy of visually-based thresholding.

The second problem in employing ROC curves for assessment of our predictions stemmed from the fact that no prediction criterion, regardless of its quality, could be expected to provide predictions that perfectly coincide with the EMG-based onsets. As one example, certain delays exist between M1 activity and muscle contraction (Rau et al. 2003). It is therefore possible that, even with an analysis method that perfectly captured the underlying neural mechanisms, the synchrony between predictions and movement onsets might be skewed. A certain acceptable time window around the “true” onset was therefore established, within which a prediction could be considered a true positive rather than a false positive. The definition of this window added parameters to the system and provided another method by which the results may be skewed one way or another, and therefore needed to be based on visual inspection and the expectation of how a system would need to perform in an eventual clinical setting. A number of different acceptable
window sizes and positions could be examined in the analysis, as dictated by the user-made parameter files. In the current analysis, three such windows were examined: one which covered 0.5 s prior to EMG activity, one which covered 0.5 s following EMG activity, and one which covered 0.25 s on either side of EMG activity.

With these concerns in mind, it is clear that the ROC-based method was not yet statistically validated in this context. Its use was therefore limited, in that it may not have been possible to derive absolute measures of prediction quality from it. It was still useful, however, for comparing one criterion to another, so that the most successful predictions could easily be identified.

No single AUC value exists for all analyses and applications at which the discriminative ability can be considered “sufficient”. Therefore, for the purposes of the present discussion, a curve area (AUC) of 0.7 or greater was considered adequate for the parameter set to be thought of as successful in producing meaningful onset time predictions. This value was chosen subjectively, based on visual inspection, as a level at which predictions can be generated without excessive numbers of false positives or false negatives. For conventional inferential statistics to be applied to the ROC results, each parameter used in the analysis would have to be included as a factor, which would result in each single AUC value representing the only value in its group. Since each group would have a sample size of one in this case, inferential statistics were not available as a tool, necessitating the use of the subjective threshold of AUC ≥ 0.7.
3.3. Change-Point Analysis

Due to certain procedural differences, change-point analysis was not conducted with the aforementioned GUI program. Rather, data was first adapted (equivalent to Step 1 of the GUI program) with a custom MATLAB script into an ASCII format suitable for the next step. The adapted data was imported into the ChangePoint program developed by Moskvina & Zhigljavsky (as outlined in Moskvina & Zhigljavsky 2003), where the change-point “detection statistics” (those being time-variant signals serving a very similar purpose to the criterion functions previously mentioned) were derived (this being equivalent to Step 3 of the GUI program). These signals were then exported from the program into MATLAB, where their usefulness as criterion functions for predicting movement onset time was assessed with a custom program that ran identical procedures to Step 4 of the GUI program. The final product of this analysis was therefore a series of ROC curves and their associated statistics, identical to those produced for the other seven analytical methods. Note that for the current study, no preprocessing (Step 2 in the GUI program) was included in our change-point analysis, and that all of the neural signals used as inputs in this analysis were therefore effectively monopolar. Note also that the parameters involved in change-point analysis were not varied, but rather kept at the default values, due to time constraints.

Two different “detection statistics” were produced by the ChangePoint program. The first was simply called the detection statistic, and represented the Euclidean distance between the SSA model and the actual signal. The second was referred to as the cumulative sum, or CUSUM statistic, and was intended to keep a “running tally” on increases in the first
statistic, and to highlight sharp jumps in its value. Further details regarding these
statistics can be found in Moskvina & Zhigljavsky (2003). Both statistics were
considered as possible criterion functions within our analysis.

3.4. Motion-Artefact Investigation

An additional experiment was also conducted to investigate the possibility of a motion-
induced artefact contaminating the recordings from the EEG system used for Subject 3.
The need for this experiment arose as a result a number of observations during collection
of data from Subject 3. One such observation was that any movement within the room,
regardless of whether the movement was performed by the subject or the experimenter,
resulted in large, visible changes in the EEG signals. Similarly, electrodes that were not
yet attached to the subject’s head, or that were placed in areas not commonly thought to
be involved in motor control, displayed large changes when any movement occurred
within the room. It was also noticed that intentional movement of the electrodes
themselves, such as when adjustments needed to be made to reattach loose electrodes,
tended to cause similar changes.

A subject (a 24 year old male without neurological injury) had one EEG electrode placed
over his left C3, and a second electrode clipped onto his left ear lobe. It was thought that
the electrode clipped onto the subject’s ear lobe should receive the same level of motion-
induced artefact, while being devoid of any actual neural activity. A third electrode was
placed, with conductive gel, onto the surface of the table in front of the subject. Three
accelerometers (Dytran, USA) were also attached to the top of the subject’s head on a
cube, such that acceleration in all three dimensions was captured. Another set of three accelerometers was also placed on the table top, alongside the third electrode.

Accelerometer and EEG data was recorded through the same DAQ system used in Subject 3’s experiment. The subject then performed the following movement tasks: taking deep breaths, making deliberate eye blinks, right elbow flexion, swinging the right foot, opening and closing the right hand, handwriting, jogging lightly on the spot, jumping lightly, making reaching movements to the left and the right with the right arm, speaking, standing and sitting, stomping the left and right feet while seated, turning a pack of cards over in the right hand, waving the right arm and playing the Tennis portion of Wii Sports. Two trials were also performed wherein another person in the room with the subject performed movement tasks (stomping either foot and waving either arm). These results were analyzed by determining $r^2$ values between the different electrodes, and between the electrodes and the accelerometer signals.

In addition to this analysis, an implementation was attempted of the artefact-cancellation method proposed by Hosaka et al. (2006). In brief, this method involved subtracting a scaled and delayed copy of one of the accelerometer signals from one of the EEG signals, with the scaling factor and delay time optimized to maximize correlation between the two signals. The underlying principle of this method was the suggestion that if the artefact is caused by movement, it should resemble the accelerometer signal.
Results

4.1. Overview

Results of the ROC-based analysis are summarized in Figure 3. Specifically, this graph presents the maximum AUC value obtained by applying a given analysis technique to a given trial of movement.

Figure 3. Summary of the best AUC values obtained for a given pairing of trial and analysis type. Warmer colors (towards red) indicate more accurate predictions, whereas cooler colors (towards blue) indicate poorer predictions. An AUC value of 0.5 indicates random chance, while an AUC value of 1 indicates perfect prediction. Notable results
include poor prediction quality for change-point and phase analysis, poor prediction quality for most of Subject 3’s trials, and more accurate predictions in Subject 4’s trials for cued movements than volitional movements. Additional results, and details of which parameters produced the AUC values presented herein, are discussed in the Results (4.2-4.5) and Discussion (5.2-5.5) sections.

Note that while AUC was chosen as the primary indicator of performance, the minimal distance (d’) was noted to correlate very highly with AUC (p < 0.0001), and the results presented would likely be very similar had this measure been used instead. The curve area presented may be from any channel (monopolar or differential) or principal component, and may be assessed with any window length and acceptable time window around the “true” EMG event. The predictions on which the curve is based may also be based on either increases or decreases in the criterion signals. For methods in which frequency ranges needed to be specified (those being the “bandpower” and “phase” methods), each frequency band investigated is presented as a separate analysis technique.

Visual inspection of the criteria deemed most successful by the ROC-based analysis (that is, those with AUC $\geq 0.7$) confirms that predictions can be made that are of high accuracy, both in terms of maximizing true positives and minimizing false positives and false negatives. Samples of the predictions generated at different AUC values are shown in Figures 4-6.
Figure 4. Sample ROC curve, criterion function, predictions, and true EMG for an analysis with one of the highest prediction qualities obtained in the current study (Area Under Curve (AUC) = 0.863). Predictions largely agreed with true onsets seen in the EMG recordings. These results are from Subject 2’s first trial of elbow flexion to a random angle, using increases in the sum of differences in monopolar ECoG channel 1 as an indication of possible movement onset. The window length in this case is 1 s, and the acceptable window surrounding EMG activity is 0.5 s pre-onset.
Figure 5. Sample ROC curve, criterion function, predictions, and true EMG for an analysis with the minimum AUC value that is considered “successful” (Area Under Curve (AUC) = 0.700). Predictions largely agreed with true onsets seen in the EMG recordings, though both false negatives and false positives are visible. These results are from Subject 2’s first trial of elbow flexion to a random angle, using increases in the bandpower integral in monopolar ECoG channel 2 as an indication of possible movement onset. The window length in this case is 0.25 s, and the acceptable window surrounding EMG activity is from 0.25 s pre-onset to 0.25 s post-onset. Note that, for the purposes of this figure, the threshold was increased above the value that gave the minimum d’ value, to make results more clearly visible. This does not, however, affect the AUC value or the discriminative power of the analysis technique being used.
Figure 6. Sample ROC curve, criterion function, predictions, and true EMG for an analysis with predictions being generated at random chance (Area Under Curve (AUC) = 0.500). Predictions had no particular correlation to EMG activity. These results are from Subject 2’s first trial of elbow flexion to a random angle, using increases in the fractal dimension in monopolar ECoG channel 3 as an indication of possible movement onset. The window length in this case is 0.25 s, and the acceptable window surrounding EMG activity is 0.5 s post-onset.

The results outlined below do not specify whether the maximum AUC values are obtained by trying to detect increases or decreases in the criterion signals (that is, whether predictions are generated by a super-threshold or a sub-threshold criterion value). The
reason for this is that which thresholding condition produces the best results tends to be highly variable, as will be elaborated on in the Discussion section.

A summary of the parameter values which produced predictions with AUC \(\geq 0.7\) is provided in Table 1.

<table>
<thead>
<tr>
<th>Successful trials (AUC (\geq 0.7))</th>
<th>ECoG</th>
<th>EEG (Wii Tennis only)</th>
<th>EEG (Subject 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window length</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.25 s</td>
<td>3.5%</td>
<td>26.8%</td>
<td>22.0%</td>
</tr>
<tr>
<td>0.5 s</td>
<td>31.5%</td>
<td>33.8%</td>
<td>33.6%</td>
</tr>
<tr>
<td>1 s</td>
<td>65.0%</td>
<td>39.4%</td>
<td>44.4%</td>
</tr>
<tr>
<td>Acceptable window</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5 s pre-EMG</td>
<td>5.2%</td>
<td>21.1%</td>
<td>58.3%</td>
</tr>
<tr>
<td>0.25 s pre- or post-EMG</td>
<td>34.5%</td>
<td>31.7%</td>
<td>39.3%</td>
</tr>
<tr>
<td>0.5 s post-EMG</td>
<td>60.3%</td>
<td>47.2%</td>
<td>2.4%</td>
</tr>
<tr>
<td>Preprocessing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monopolar</td>
<td>5.3%</td>
<td>28.5%</td>
<td>31.9%</td>
</tr>
<tr>
<td>Differential</td>
<td>64.9%</td>
<td>46.5%</td>
<td>48.2%</td>
</tr>
<tr>
<td>PCA</td>
<td>29.8%</td>
<td>25.0%</td>
<td>20.0%</td>
</tr>
<tr>
<td>Monopolar channels</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.0%</td>
<td>42.0%</td>
<td>36.2%</td>
</tr>
<tr>
<td>2</td>
<td>0.0%</td>
<td>28.9%</td>
<td>39.4%</td>
</tr>
<tr>
<td>3</td>
<td>66.7%</td>
<td>11.8%</td>
<td>24.5%</td>
</tr>
<tr>
<td>4</td>
<td>33.3%</td>
<td>17.4%</td>
<td>20.0%</td>
</tr>
<tr>
<td>Differential signals</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 &amp; 2</td>
<td>2.7%</td>
<td>3.7%</td>
<td>16.9%</td>
</tr>
<tr>
<td>2 &amp; 3</td>
<td>59.5%</td>
<td>16.0%</td>
<td>2.1%</td>
</tr>
<tr>
<td>3 &amp; 4</td>
<td>0.0%</td>
<td>20.6%</td>
<td>16.9%</td>
</tr>
<tr>
<td>1 &amp; 3</td>
<td>37.8%</td>
<td>16.8%</td>
<td>14.1%</td>
</tr>
<tr>
<td>2 &amp; 4</td>
<td>0.0%</td>
<td>21.2%</td>
<td>26.1%</td>
</tr>
<tr>
<td>1 &amp; 4</td>
<td>0.0%</td>
<td>21.7%</td>
<td>23.9%</td>
</tr>
<tr>
<td>Principal components</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First</td>
<td>5.9%</td>
<td>22.0%</td>
<td>57.6%</td>
</tr>
<tr>
<td>Second</td>
<td>0.0%</td>
<td>38.4%</td>
<td>1.7%</td>
</tr>
<tr>
<td>Third</td>
<td>58.8%</td>
<td>29.4%</td>
<td>40.7%</td>
</tr>
<tr>
<td>Fourth</td>
<td>35.3%</td>
<td>10.2%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Table 1. Summary of the relative frequencies with which specific parameters gave rise to successful movement onset predictions. Red cells indicate channels in which at least one of the electrodes being used was thought to be invalid, and should have been generating only random noise. Blue cells indicate cases in which the signal may have been
contaminated with large motion-induced artefacts. Trends in these results are discussed in the Results (4.2-4.5) and Discussion (5.5) sections.

4.2. Subject 1

In the case of Subject 1, most of the data acquired could not be properly analyzed with the ROC-based method. The reason for this was that each of the ten trials performed within the reaching experiment contained six movements, each in a different direction. The automated ROC analysis, however, was designed to examine trials in which the subject repeated a single movement many times. The presence of different movement types (that is, reaching movements in different directions) may therefore have provided a significant confound. Reaches in different directions would be expected to have produced somewhat different neural signals. This may have resulted in predictions that were successful for some reaching directions, but not others. In such a case, the AUC value that the ROC analysis presented would represent an average of the prediction accuracies for all different directions, rather than for a single movement type. It was due to this significant confound that the reaching data for Subject 1 was not analyzed.

The only other trial conducted with Subject 1 was a round of Wii Sports Tennis. Several different analysis techniques were found to be of potential use. These methods were: fractal dimension (AUC = 0.711), spectral entropy (AUC = 0.755), bandpower from 1-5 Hz, 8-12 and 12-20 Hz (AUC = 0.705, 0.714 and 0.797, respectively), time-frequency integral (AUC = 0.750), and variance (AUC = 0.767). Of the results with AUC \( \geq 0.7 \), a tendency was noticed towards better results in cases with longer window lengths (3.5%
had a window length of 0.25 s, 31.5% had a window length of 0.5 s, and the remaining 65.0% had a window length of 1 s). A tendency was also noticed towards better results in cases in which the acceptable window was mostly post-EMG-onset (5.2% had predictions within 0.5 s pre-onset, 34.5% had predictions within 0.25 s pre- or post-onset, and 60.3% had predictions within 0.5 s post-onset). Differential signals were the most common among cases with AUC $\geq 0.7$ (64.9%), principal components the second most common (29.8%), and monopolar signals the rarest (5.3%). Lower principal components tended to provide most of the significant results (5.9% of principal component results came from the first principal component, 0% from the second, 58.8% from the third, and 35.3% from the fourth).

### 4.3. Subject 2

Generally speaking, results from ECoG Subject 2 were of roughly the same quality as those generated from Subject 1. A trend towards better results in cases with longer window lengths was seen, similar to that noticed in Subject 1’s results, though not as pronounced (26.8% had a window length of 0.25 s, 33.8% had a window length of 0.5 s, and the remaining 39.4% had a window length of 1 s). The tendency towards better results in cases with post-EMG-onset acceptable windows that was noted in Subject 1 was also observed in a less pronounced manner here (21.1% had predictions within 0.5 s pre-onset, 31.7% had predictions within 0.25 s pre- or post-onset, and 47.2% had predictions within 0.5 s post-onset). Contrary to the results from Subject 1, no particular preference was noted for monopolar, differential, or principal component signals (which comprised 28.5%, 46.5%, and 25.0%, of the successful results respectively, with the
expected values in a purely random case being 28.6%, 42.8%, and 28.6%, since there are 4 monopolar channels, 6 differential channels, and 4 principal components). Furthermore, of the principal component results, no large preference was noted towards a particular component (22.0%, 38.4%, 29.4%, and 10.2% of PCA-based results came from the first, second, third, and fourth principal components, respectively).

For two of the trials performed, no analytical technique was able to produce successful predictions (those being the second trial of elbow flexion to a random angle, and the only trial of elbow flexion to a fixed angle, AUC ≤ 0.676 and AUC ≤ 0.660, respectively). By contrast, all other trials, including the first trial of elbow flexion to a random angle, had at least one analytical technique that was able to produce successful predictions (AUC ≥ 0.754). As per the results of Subject 1, however, phase measures, regardless of the frequency range employed, did not provide successful predictions for any trials (AUC ≤ 0.685). Change point measures were also largely ineffective, with the exception of the first trial of elbow flexion to a random angle (AUC = 0.778 for the aforementioned elbow flexion trial, AUC ≤ 0.676 for all other trials). Within trials involving only the wrist (one trial of extension and two of flexion), successes were limited, but visible, and no considerable differences were noted between the three trials. The most successful results in the wrist-only trials were all for bandpower analysis (in the 1-5 Hz range for extension and the first flexion trial, and in the 20-30 Hz range for the second flexion trial, AUC = 0.754, 0.782, and 0.784, respectively). Other bandpowers were still useful, however (from 1-5 Hz in the second flexion trial [AUC = 0.767], from 8-12 Hz in the first flexion trial [AUC = 0.736], and from 12-20 Hz and 20-30 Hz in the extension trial [AUC =
Aside from bandpower, useful results were also obtained through the sum of differences (extension trial, AUC = 0.718) and variance (extension trial and second flexion trial, AUC = 0.700, 0.716). More successful predictions were obtained with a broader range of analysis techniques in the reaching trials, with reaches to the left providing slightly clearer results. Successful results were obtained from fractal dimension (left only, AUC = 0.749), spectral entropy (left and right, AUC = 0.813, 0.781), the sum of differences (left and right, AUC = 0.826, 0.818), bandpower from 1-5 Hz (left and right, AUC = 0.799, 0.721), bandpower from 8-12 Hz (left only, AUC = 0.762), bandpower from 12-20 Hz (left and right, AUC = 0.835, 0.789), bandpower from 20-30 Hz (left and right, AUC = 0.879, 0.835), bandpower from 36-44 Hz (left and right, AUC = 0.708, 0.711), the bandpower integral (left and right, AUC = 0.824, 0.789), and variance (left and right, AUC = 0.837, 0.810). Results from the first trial of elbow flexion to a random angle were generally of similar quality to those of the reaching trials. Specifically, successful predictions were obtained from change point analysis (AUC = 0.778, as previously mentioned), spectral entropy (AUC = 0.832), the sum of differences (AUC = 0.863), bandpower for all of the frequency ranges examined (AUC = 0.781, 0.777, 0.782, 0.798, 0.852, 0.760, for 1-5, 8-12, 12-20, 20-30, 36-44, and 90-99 Hz, respectively), the bandpower integral (AUC = 0.858), and variance (AUC = 0.830). This was contrasted by the results of the second such trial, in which no successful predictions were generated by any analytical technique (AUC ≤ 0.676).
4.4. Subject 3

In the case of Subject 3, no analysis techniques were found to be successful (based on the criterion of AUC ≥ 0.7) for the elbow flexion, hand opening/closing, reaching left, or reaching right trials (AUC ≤ 0.649). Many techniques were, however, successful in the Wii Sports Tennis trial. These methods are: fractal dimension (AUC = 0.700), spectral entropy (AUC = 0.805), bandpower in all frequency ranges investigated (AUC = 0.770, 0.787, 0.835, 0.738, and 0.711 for 1-5, 8-12, 12-20, 2-30, and 36-44 Hz respectively), the bandpower integral (AUC = 0.798), and variance (AUC = 0.793). As will be subsequently discussed, the successful results may be more related to a movement-related artefact rather than genuine neural activity, and must be cautiously interpreted. A slight, but unclear trend was noticed within the significant results towards better results in cases with longer window lengths (22.0% had a window length of 0.25 s, 33.6% had a window length of 0.5 s, and the remaining 44.4% had a window length of 1 s). A tendency was also noticed towards better results in cases in which the acceptable window is pre-EMG-onset (58.3% had predictions within 0.5 s pre-onset, 39.3% had predictions within 0.25 s pre- or post-onset, and 2.4% had predictions within 0.5 s post-onset). No clear preference was noticed for one type of signal or another (of the significant results, 31.9% were from monopolar channels, 48.1% were from differential signals, and 20.0% were from principal components, with expected values of 28.6%, 42.8%, and 28.6% respectively in a purely random case).
4.5. Subject 4

Results from EEG Subject 4 by and large showed more successful predictions than those for Subject 3, though not as successful as those for Subject 2. Among the successful predictions, a strong preference was shown for longer window lengths (8.7% had a window length of 0.25 s, 33.5% had a window length of 0.5 s, and the remaining 57.8% had a window length of 1 s). Any trend in preferred acceptable window around EMG events was less clear (38.5% had predictions within 0.5 s pre-onset, 37.5% had predictions within 0.25 s pre- or post-onset, and 24.0% had predictions within 0.5 s post-onset). The majority of the successful results were based on principal components (50.4%, with an expected value of 28.6% in a purely random case), although differential signals also commonly provided accurate predictions (38.4%, with an expected value of 42.8% in a purely random case). The remaining 11.3% of the successful predictions were from monpolar signals (with an expected value of 28.6% in a purely random case). Among the results from principal components, the majority came from lower components (7.2%, 7.6%, 53.1%, 32.1% of these results came from the first, second, third and fourth principal components, respectively).

Change point measures were still ineffective in all of Subject 4’s trials (AUC ≤ 0.634), as was the bandpower from 90-99 Hz (AUC ≤ 0.676). Fractal dimension was only successful in the Wii Bowling trial (AUC = 0.752). Surprisingly, several phase measures showed successful predictions in two trials: cued arm tensing (1-5, 36-44, and 90-99 Hz, AUC = 0.758, 0.700, and 0.707 respectively), and cued handwriting of single words (1-5 and 8-12 Hz, AUC = 0.759, 0.705 respectively). The most significant result of the
remaining data is the fact that while successful predictions were relatively common in trials where the movement onset was cued (with some exceptions), no successful predictions occurred in trials where the movement onset was volitional (AUC ≤ 0.682). Given this, the results of the Wii Sports trials are unexpected, since the Tennis trial, which was intended to represent cued movements, provided no successful predictions (AUC ≤ 0.644), while the Bowling trial, which was intended to represent volitional movements, generated successful predictions with multiple methods (AUC ≤ 0.793). The other exceptions to the general success of cued trials were random arm movements (in which no successful predictions were generated, AUC ≤ 0.697), and the handwriting of single words (in which the only successful predictions were generated by the two aforementioned phase measures). Among the successful cued trials, accurate predictions were usually generated by multiple analytical techniques. Bandpowers between either 12-20 or 20-30 Hz were the most frequently successful, and generated the highest AUC values, but bandpowers from 1-5 and 8-12 Hz, spectral entropy, sum of differences, time-frequency integral, and variance all provided successful predictions in multiple trials.

4.6. Motion-Artefact Investigation

The most significant results of the experiment investigating the possibility of a motion-induced artefact in Subject 3’s recordings are summarized in Table 2.
Table 2. Summary of correlation values from the experiment investigating the possible motion-induced artefacts within Subject 3’s trials. High correlation values were noted between an EEG electrode that was thought to be recording genuine neural activity (C₃), and an electrode that was clipped onto the subject’s ear lobe (and should therefore only have been receiving artefacts). The high correlation values, especially in more vigorous motor tasks, indicated large motion artefacts. Maximum correlation values between any pairing of an EEG signal and an accelerometer signal were lower, but still showed highest values during more vigorous movements.

In general, correlation coefficients were high between the electrode placed over C₃ and the electrode placed on the ear lobe. This was especially true for movement types that required larger movements on the subject’s part. For example, near-perfect correlations were obtained for jogging, jumping, sitting, standing, stomping, and swinging of both the
arms and legs \((r^2 \geq 0.998)\), while the lowest correlations were observed while the subject was taking deep breaths or blinking \((r^2 \leq 0.294)\). High correlations were also observed between these two electrodes even when the movements were being performed by another person in the room with the subject \((r^2 \geq 0.623)\). Correlation coefficients between the electrode placed on the table top and the other two electrodes (not shown in Table 2) were considerably lower, though still visible in the case of some of the more vigorous movements (for example, stomping feet \((r^2 \geq 0.665)\), jumping \((r^2 \geq 0.390)\), and jogging \((r^2 \geq 0.282)\)). Correlations between the accelerometer signals and the EEG signals were comparatively very low \((r^2 \leq 0.106)\). However, visual inspection confirmed that the increases in EEG signal amplitude tended to occur at the same time as spikes in the accelerometer signals. Higher correlations were still visible when larger movements were required on the subject’s part, similar to the results seen in the EEG-to-EEG comparisons. For example, the lowest correlations were observed in the cases of speaking and handwriting \((r^2 \leq 0.001)\), while the highest were observed during jogging, jumping, and playing Wii Sports Tennis \((r^2 \geq 0.040)\).

Visual inspection suggested that the method proposed by Hosaka et al. (2006) for removing the motion-induced artefact was largely ineffective. An example of these results is presented in Figure 10.
Figure 7. Sample of the results obtained using the method proposed by Hosaka et al. (2006) for cancelling the motion-induced artefact in EEG signals. The subject was playing the Tennis portion of Wii Tennis as a motor task. The first trace is from an EEG electrode clipped onto the subject’s ear lobe, and should therefore represent only motion artefact. The second trace is from one of the three accelerometers attached to the subject’s head – in this case, the one aligned to detect front-to-back motion. The third trace is the result of the correction, following optimization. Ideally, this last trace should be a flat line, indicating cancellation of the motion artefact. Note that even though cancellation is unsuccessful, spikes in the EEG and accelerometer signals seemed to be occurring at the same time. Similar results were seen for all other movement tasks and accelerometers.

This was largely supported by the relatively low maximum $r^2$ values obtained between the original EEG signals and the scaled and delayed accelerometer signals, even after optimizing for both scaling factor and delay time. For example, the highest $r^2$ values when investigating the electrode attached to the ear lobe (which should consist only of motion-induced artefact and some noise), were found when the user was jumping ($r^2 =$
0.638) and when they were sitting and standing ($r^2 = 0.212$). Unsurprisingly given the types of movement involved, these highest correlations were all obtained from the vertically-oriented accelerometer.
Chapter 5: Discussion

5.1. Possible Sources of Contamination

5.1.1. Motion-Induced Artefact

The results of the experiment investigating the possibility of a motion-induced artefact strongly suggested that such an artefact existed within Subject 3’s signals. This was of particular concern given that the entire purpose of our experiments was to provide predictions of the time of movement onset time based on changes in the neural signal. The artefact would only be generated when the user performed a movement, and would be detected by most of our analytical techniques. The system would therefore be able to provide high quality predictions of the movement onset time. However, these predictions would not be genuinely based on any neural activity, and as such, could not be considered part of a genuine BMI system.

The motion-induced artefact is an acknowledged, but poorly-understood phenomenon within the literature. Deformation or stretching of the skin, in particular, is thought to cause changes in potential (Ödman 1982). One of the most prevalent theories to account for this is that deformation alters the resistance properties of the skin, by shifting the “barrier layer” of dead cells on the surface of the skin (that being the layer most responsible for the impedance value of skin) (Ödman 1982, Webster 1984). An alternate explanation is that the subject’s movement causes movements in the ions themselves at either the electrode-gel interface or the gel-skin interface (Ödman and Öberg 1982, Searle and Kirkup 2000).
While the motion artefact did allow for the generation of highly accurate predictions of movement onset time, these predictions could not be considered useful for any actual clinical application. The reason for this is that a user who would require a BMI system would have to have a high level of paralysis, and would therefore not be able to perform the movements needed to generate the artefact, in the same way that they would be unable to use a more conventional control strategy such as a mouse or push-button. The observation that motion-induced artefacts could also be generated by the movements of nearby individuals in fact suggests that a system which includes motion-induced artefacts would generate many false predictions of movement onset.

The relative inability of the method proposed by Hosaka et al. (2006) to mitigate the motion-induced artefact was understandable, given the qualitatively different appearance and nature of the EEG and accelerometer signals, and the very low initial correlation values between these two types of signals.

Other methods of reducing or cancelling the motion-induced artefacts have not yet been implemented at the time of writing. One approach is simply to limit the focus of our study to only include ECoG recordings. This, however, greatly limits the sample size from which we are able to obtain data. It may be possible to reduce artefacts simply by using other equipment. More sophisticated electrodes may be less susceptible to motion-induced artefacts than the relatively simple electrodes used for Subject 3’s experiment (for example, those suggested by Ödman and Öberg (1982) and Webster (1984)). The use of an EEG cap or stronger conductive gel may allow the electrodes to be held more
firmly in place. Since our group does not personally have access to any other equipment at the present time, this may need to be accomplished through collaboration. The system used to record from Subject 4, while similar in many respects to that used for Subject 3, seemed to have a greatly decreased motion artefact. However, the artefact was still somewhat present, as confirmed by visual inspection of trials in which the subject tilted their head side to side. The presence or absence of the motion artefact could not be confirmed in ECoG trials, since it was not known of at the time that the ECoG experiments were run, and no trials were therefore performed to examine the possibility of it. However, it would not be expected that these trials should include motion artefacts, since the artefact is thought to arise from changes in the interface between the skin and the electrode. While it is possible that changes may be present in the interface between the cortex and the electrode, this would seem less likely, and has not been thoroughly investigated in the literature.

Traditional filtering is not likely to be able to remove the motion artefacts, since they do not occur with a fixed frequency, but rather, in relation to specific events. It has been suggested that adaptive filtering (as outlined in Sennels et al. 1997) may be able to remove the artefacts, as it has been shown to be able to remove event-related artefacts from EMG signals. However, it is presently considered doubtful that it would be able to remove motion-related artefacts from our neural recordings. There are several reasons for this. One is that adaptive filtering assumes that there are no significant correlations between the signal and either the background noise or the reference signal being used to keep track of the artefact (for our purposes, this would most likely be an electrode
recording from the subject’s earlobe, since that area should receive the same head
movement as the other electrodes, but none of the neural activity). These assumptions,
however, are violated by the basic fact that changes in all of these signals were caused by
the same events, namely the subject’s movements. As explained in the Results section,
high statistical correlations were visible between the neural recordings and the earlobe
recordings during most types of movements. An additional concern is that the signal,
noise, and reference signals are also assumed to be stationary. Again, the very fact that
the artefacts were occurring, and that they were event-related, suggests that this
assumption is false. A final concern is that the calculations of tap-weights for the
adaptive filter may require the artefacts to be roughly periodic, since one of the
parameters is set to the frequency of the stimulation causing it (Sennels et al. 1997).
Since our artefact was not periodic, this may highly complicate the calculations of tap-
weights. It was for these reasons that time had not yet been taken to test or implement
adaptive filtering in the current study.

The presence of the motion artefact is of vital importance when considering the results
observed within Subject 3’s trial. It was noted that during most of the movements
performed in that trial (elbow flexion, hand opening/closing, reaching left, and reaching
right), no analysis technique or parameter set was able to produce predictions of
sufficient accuracy. By contrast, when the subject was playing Wii Sports Tennis, many
techniques were able to produce meaningful predictions. The most probable explanation
for this is that Wii Tennis, being the fastest and most vigorous of the movements,
produced more motion artefact than the other movements. Therefore, for the reasons outlined above, these successful results could not be considered truly valid.

5.1.2. Eye-Blink Artefact

The motion-induced artefacts were not the only possible sources of contamination within our signals. Another source was the eye-blink artefacts. These are artefacts caused when the large electrical bursts used by the muscles in the eyes to cause blinks are detected by EEG electrodes, and are among the most commonly-detected artefacts within EEG signals (Fatourechi et al. 2008). Among the most common methods for removing these artefacts are principal component analysis and independent component analysis, with artefacts often being isolated in the first few components (Hyvärinen and Oja 2000). Therefore, in the cases where principal component analysis was applied, we could expect that the first few principal components would consist mostly of artefacts, which, having no relation to the movement tasks, would have produced low-quality predictions. In differential signals, the subtraction may have caused a cancellation of some of the artefact, though the differing proximity of each electrode to the eyes may have caused uneven representation in the different channels, preventing complete cancellation. Monopolar EEG signals were still likely to have contamination from eye-blink artefacts. This was confirmed by visual inspection of a trial with Subject 4 in which the only motor task was to provide deliberate blinks. Each blink in that trial resulted in a large spike in the activity of all EEG channels. Since most of the analytical techniques being applied were intended to detect increases in electrical activity or simply changes in the fundamental quality of the signal, the most likely result of such an artefact would have been the generation of a
false positive. This was not the only possible result, however. For example, the eye-blink may have happened to coincide with the actual movement, resulting in an apparent increase in the prediction quality. Measures such as change-point analysis, sum of differences, or variance, would have been expected to be more susceptible to this contamination than measures such as high-frequency bandpower. In general, however, it is simply best to keep the eye-blink artefacts in mind as a possible source of variations in prediction quality in the EEG recordings. Some investigators have also found eye-blink artefacts in ECoG recordings (Ball et al. 2009). However, these artefacts were largely localized to more frontal areas of the cortex, and were considerably smaller, relative to the background noise, than the artefacts found in EEG recordings. Their influence may therefore still have been present in our signals, but would have been less pronounced.

5.1.3. Duration of Neural Events

It is considered possible that information within the M1 pertaining to movement onset may be presented in one of two ways. The first (which will be called the “change” option) is for the neural events related to movement onset to only persist for as long as the onset itself. In this case, the neural event can be used directly as a measure of movement “onset”, as it is more closely related to changes in muscle activity than the activity itself. The second (which will be called the “duration” option) is for the neural event to persist for as long as the movement is sustained. In this case, the event is more closely related to muscle activity, rather than to changes therein. This is of concern due to the fact that the program used to analyze the signals would only have shown high accuracy predictions if the neural event was of the “duration” type, persisting for the extent of the movement. At
present, a neural event of the “change” type would have resulted in a relatively low AUC value, due to a high number of false negative counts. An entirely different accuracy measure would therefore be needed to properly assess the success of predictions for neural events of the “change” type. The relevant literature does not fully conclude that neural events within M1 conform to one type of activation or another. Vuckovic and Sepulveda (2008) found that ERS and ERD events up to 60 Hz were evenly distributed throughout the 3 s period that their subjects’ wrist muscle contractions were sustained for. Cassim et al. (2000), however, found that changes in mu and beta range bandpower returned to baseline within 4-5 s of sustained wrist muscle contraction. Crone et al. (1998a and 1998b) conducted a more wide and varied set of experiments, and found that the temporal evolution of ERS and ERD events (including the extent to which they return to baseline during a sustained contraction) varies by subject, frequency band, and motor task. This corroborates with initial visual inspection during the present experiments, in which it was noted that the events visible in some tasks could return to baseline at different rates, depending on the task being performed. Given the preference of the current analysis method for neural events of the “duration” type, it is possible that some of the negative results presented herein are simply due to events of the “change” type providing the ROC analysis with a high number of false negatives. The additional analysis that would be required to properly detect such events has not yet been investigated or implemented at the time of writing.
5.1.4. Temporal Misalignment

An issue that arises in the production of criterion signals is that of temporal alignment. The problem emerges from the fact that each point on a criterion function is derived from multiple points within a sliding time window. This results in a criterion signal that is slightly shorter than the original neural recording on which it is based. Since the neural signals and the EMG signals were recorded simultaneously, this may have caused a misalignment between the two. At the time of writing, the implemented solution has been to simply to align the point on the criterion function with the center of the sliding time window. This effectively means that the EMG signals were being cropped, with an equal number of samples being removed from the start and end of the signal until it is the same length as the criterion function.

The problem becomes more complex when multiple criterion functions, which may themselves be of slightly different lengths, are combined to generate a single set of predictions. In this case, the two criterion functions were first aligned in the manner described above, and then the combined predictions generated from them were likewise aligned with the EMG signals.

Larger window lengths will tend to cause larger discrepancies between the criterion and EMG signals. Whether one method of aligning the two signals is more correct from a physiological or practical perspective, or can provide more accurate predictions, has not yet been established. Centering the signals was therefore chosen as a compromise.
between the different possibilities, and as the closest approximation to a situation in which no such alignment is necessary.

5.1.5. Non-Homogenous Movements

A consideration which must enter into the interpretation of prediction accuracy results is the fact that many of the movement types employed were not strictly homogeneous. Rather, they consisted of both the intended movement and the return to a base position, both of which were included in the subsequent analysis. For example, in an elbow flexion, the subject would first have had to flex, and then had to return to a more extended position to prepare for the next flexion. Since the determination of true onset time was based on EMG signals, both the original movement and the return movement would have been considered as genuine onsets by the analysis system. Since both of these movements should hypothetically be brought about by neural signals passing through the motor cortex, this is not strictly incorrect. The concern is that different movement types, even if they were within roughly the same category (for example, both the elbow flexion and the relative extension thereafter were movements that involved the same muscle groups in roughly the same order of magnitude), may have presented different signals. One illustration of this is the fact that many studies which record from smaller neural populations show that reaching movements will cause differing neural changes depending on which direction the subject is reaching (for example, Georgopoulos et al. 1988). It is possible, however, that these kinds of differences may not be visible in EEG and ECoG recordings, since signals obtained in those cases may contain contributions from neural populations which prefer different reaching directions. The same principle may apply to
other movement types in which both original movements and return movements were present – that is, EEG and ECoG may not have had the resolution necessary to discern between the two movements. While it has been suggested that ECoG may be used to predict the direction of reaching movements (Schalk et al. 2007), the generalizability of this data is limited. Visual inspection of the present data and the predictions generated by the accompanying analysis did not conclusively resolve the matter. However, the heterogeneity of the multiple movements included in a given trial may still be taken as a possible source of variation within the prediction accuracy results.

5.1.6. Non-Stationarity

A question can be raised about the correctness of applying Fourier-based signal processing techniques to our EEG signals. The reason for this concern is that Fourier analysis requires the assumption that the signal being analyzed is stationary (that is, that its mean and variance are invariant over time). It is known, in general, that neural signals are not stationary, and that they are in fact changed over time by events, such as the movement onset-related changes that our methods are searching for. The fact that windowed variance was being examined as a possible criterion for detecting neural events further underscores the violation of this assumption. It can be argued, however, that the violation of this assumption does not compromise the validity of the predictions being generated through Fourier-based methods. The applications to which the Fourier-based results were being put were not intended to make any fundamental, underlying statements about the true nature of the signals. The results were only meant to be put into practical application. If an event caused the assumption of stationarity to be violated, and
if this violation caused a change in the Fourier results, then these changes could still have been used as a marker of the onset of the neural event. This is particularly true given that the Fourier analysis was only ever applied in our analysis to a sliding time window isolated from the rest of the signal. Time windows in which there was no event would likely not have included a violation of stationarity, whereas windows on the edge of such an event would have, and may have appeared different as a result. An alternate argument that could be made is simply that the vast majority of literature on the topic of motor commands uses Fourier-based methods to investigate ERS and ERD. Doing so in the current study therefore allows us to compare our results to the established literature, and allows us to investigate the potential validity of continuing to use these methods. This question could be largely circumvented by employing an alternate method of spectral analysis, such as wavelet analysis. The added complexity of this method, however, has prevented it from being implemented in our study at the time of writing.

5.2. EEG vs. ECoG Recordings

If Subject 3’s Wii Tennis results were not considered valid, then the most successful predictions generated within that subject’s trials would have been from the hand opening/closing trial, as analyzed with spectral entropy (AUC = 0.649). Based on this low ROC curve area value, it can be suggested that these recordings may not have been sufficient for producing meaningful predictions of movement onset time. However, the results from Subject 4 suggested that this should not be generalized to all EEG recordings. Note, however, that the AUC values for Subject 4 still tended to be lower than those
obtained in any of the ECoG trials. This is to be expected, since EEG signals should contain less information and have a lower signal-to-noise ratio than ECoG signals.

5.3. Volitional vs. Cued Movements

As mentioned in the Results section, during Subject 4’s trials, no successful predictions were generated while the subject was performing movements on a volitional basis, but successful predictions were relatively common when the subject’s movements were cued. This is of particular concern since the aim of the present study is to develop techniques for an asynchronous BMI system, in which all of the subject’s command signals would be delivered on a volitional basis. However, the most likely reason for this discrepancy in the results is that the subject’s movements were slightly quicker and more distinct in the cued condition (as confirmed by visual examination of the motion-tracking signals). In particular, the subject was quicker to return to a ‘ready’ position in order to prepare for the next movement. The subject verbally reported after the experiment that during the volitional movements, he was not consciously considering his movements, which largely corroborated with the suggestions of the motion-tracking recordings. The suggestion is that the subject may have put more “intent” into each movement, thereby resulting in a more pronounced signal. If this is the case, then an asynchronous BMI should still be a viable possibility, since a subject should be able to spontaneously generate the necessary level of intention with or without a cue.

It is known that, prior to reaching the M1, motor commands generally come from one of two different streams. Movements that are “internally-guided” (i.e. volitional) arrive by
way of the supplementary motor area (SMA), while those that are “externally-guided” (i.e. cued) arrive through the premotor cortex (PMC) (Geyer et al. 2000). While these areas were not being targeted by our recording electrodes, it was considered possible that contamination may still have been occurring from these areas, or that the M1 signals may simply have appeared different depending on which area it receives inputs from. It was for this reason that both cued and volitional movements were examined in Subject 4’s experiment. It was originally anticipated that any discrepancies between the two conditions would be caused by this. Given the location and alignment of the EEG electrodes, however, the most likely result of contamination from these areas would have been an increased signal from SMA during volitional movements, seen at the electrode sites closer to Cz. The electrodes did not come as near to PMC, and direct contamination of the signals from that area was considered less likely. Given that the results showed the opposite of this, it is unlikely that our electrodes were receiving direct contamination from more frontal cortical areas. It cannot yet be ruled out that signals from SMA and signals from PMC may simply have different appearances once they arrive in M1. However, given the differences in the motion-tracking signals and the subject’s own report, it seems more likely that the difference observed between cued and volitional conditions simply resulted from the cued condition providing a more focussed, intense movement. This is a factor that will need to be taken into consideration for future experiments.
5.4. Unsuccessful Criterion Functions

5.4.1. Change-Point Analysis

No comprehensive explanation is presently available for the general failure of change-point analysis to produce meaningful predictions of movement onset time. One possibility stems from the fact that, due to procedural differences (as outlined in the Methods section), no preprocessing was done on the raw signals before applying change-point analysis to them, and all of the signals were therefore effectively monopolar. While there is little evidence from other analysis methods to suggest that preprocessing greatly affected the accuracy of the eventual predictions, it could be possible that the absence of this variety places change-point at a disadvantage when compared to the other methods. Similarly, the parameters within change-point analysis were not fully optimized, and only their default values were used. Brief investigation of this possibility, however, suggested that changing these parameters would not have drastically changed the criterion signals generated by the analysis, or the prediction accuracy obtained by using them in our ROC analysis. It therefore seems unlikely that parameter optimization would have allowed for greatly improved prediction quality. It should be noted that the current results are in accordance with other attempts to apply change-point analysis to neural signals (Ritov et al. 2002, Vaisman et al. unpublished).

5.4.2. Phase Analysis

Phase analysis was noted to be largely ineffective in predicting movement onset in most subjects and movement types. This may be because our implementation of phase analysis did not perfectly coincide with the measures being used in current literature (in that we
were not looking at phase-locking between different electrode locations, but rather, the specific phases within single electrodes). It may also have been partly due to the fact that the use of a sliding window will itself cause apparent phase changes if phase is assessed by means of a Fourier transform. If the phase method is to be pursued in future work, revisions will need to be made, either to investigate phase-locking between different electrode locations, or to use instantaneous measures of phase, such as a Hilbert transform.

5.5. Effects of Parameters

5.5.1. Window Length

By investigating which parameters were applied in cases which gave rise to successful predictions (as defined by AUC $\geq 0.7$), it may be possible to determine general trends and thereby suggest how best to optimize the accuracy of our predictions. The clearest such trend was that successful results became more common as the window length increased. While this was visible in all subjects, it was strongest in Subjects 1 and 4. It is possible that window lengths of greater than 1 s could have provided even more accurate predictions than those seen here, since no decrease in the frequency of successful predictions was noted within the range of window lengths examined. Future analysis should therefore focus on a higher range of window lengths. One possible reason for this effect is that shorter window lengths may have allowed outlier or aberrant data points to have a more prominent effect on points in the criterion signals. This would mean that shorter window lengths would have produced noisier criterion signals, while longer window lengths would have produced smoother curves. At the time of writing, however,
this possibility had not yet been confirmed. Note that in a practical, real-time application of these techniques, longer window lengths may introduce longer delays into the system, due to both increases in the processing time (as more data is sent to the computational functions at each time point), and the inclusion of data from farther in the past in the decision-making process. The current experimental results, since they did not in any way simulate the conditions of a real-time application, could not be used to conclude how much of a factor this could be, and future studies may need to take this unexplored balance into consideration.

5.5.2. Acceptable Window Around EMG Activity

The effect of the acceptable window surrounding EMG activity was unclear from the current data. Results from ECoG recordings tended to suggest that the neural events in question were most visible after EMG activity, while EEG results seemed to suggest that the period before EMG activity was ideal. The effects of this parameter were further complicated by the fact that this window was applied to every data point. Therefore, a point in the middle of a burst of EMG activity would have been examined by all of the different acceptable windows. The distinctions between the conditions therefore only truly had an effect at the beginning or end of a burst of EMG activity, where time points outside that EMG burst may have been included. For this reason, it would be incorrect to use the current analysis to make statements about the genuine timing of the neural events in question. Rather, the parameter should be considered from a purely practical standpoint.
5.5.3. Preprocessing and Electrode Location

The results pertaining to the use of different preprocessing techniques and different channels varied widely from subject to subject. For example, Subject 1 showed a strong preference for differential signals, Subject 4 showed stronger results from principal components, and Subject 2 showed little preference for any type of preprocessing. Within principal component analysis, Subjects 1 and 4 showed a strong preference for lower principal components, while Subject 2 showed broader preferences. It was therefore difficult to draw any definitive conclusions about which techniques or channels were able to provide the most meaningful predictions. This, in and of itself, was contrary to the expected results. For example, it would be expected that in ECoG, where the signals could not be properly referenced (as opposed to the EEG cases, in which the use of the earlobe as a reference provided a proper comparison), one would expect that differential signals would have been more meaningful. The lack of this proper reference may not have hindered prediction accuracy is cases where the criterion under examination focuses on deviation from the mean, rather than the absolute value (for example, sum of differences or variance). In any of the subjects, one would also expect that earlier principal components (excepting the first, which may contain large signals such as eye-blink artefacts), would have contained more relevant information, as lower components are meant to contain things like fundamental noise. That this was not the case may suggest that the signal-to-noise ratio in our recordings was low enough that the noise forms a more prominent component of the signal than the information relevant to movement prediction.
Within the results from monopolar EEG recordings, there was insufficient evidence to suggest whether or not the specific location of the electrode had an effect on whether or not successful predictions were generated (ECoG recordings could not be used for this purpose, since their exact location was not always certain). For example, it would be predicted based on the known structure of the M1 that in trials where the task primarily required fine manipulation with the fingers, the relevant signals (and therefore, the more accurate predictions) should have been coming from more lateral areas, whereas tasks involving shoulder or elbow movements should have been coming from more medial areas. The limited results, however, did not provide sufficient evidence for or against this hypothesis. Part of the reason for this may be that, visually, many of the channels seemed to share common information. An increase in one channel often occurred at the same time as increases in the other channels. This sharing of common information was, in fact, the fundamental basis of principal component analysis, and it was therefore not entirely unexpected. It did, however, make approximate localization of the signal sources, as above, difficult. Channels tended to show more independence in ECoG than in EEG.

5.5.4. Super-Threshold and Sub-Threshold Conditions

Most of the BMI literature provides specific suggestion about whether a given criterion is expected to increase or decrease as a result of neural activity. This stands in contrast to the present results, in which no such clear pattern was apparent for any given criterion. It was noted that in many cases, accurate predictions may have been generated with one thresholding condition (that is, whether predictions are generated by super-threshold or sub-threshold criterion values) when a monopolar signal was analyzed, and another when
differential signals and/or principal components were examined. Since preprocessing of the original signals can have such a drastic effect on the thresholding required for accurate predictions, and since, in some cases, monopolar signals may not have generated accurate predictions, it could be conclusively stated whether, on a fundamental level, movement-related neural activity should be expected to cause increases or decreases in any of the criterion functions examined.

An alternate explanation for the lack of clarity in this matter is based on the large number of neurons whose activities were included in the final signal of a given electrode. For example, if the neurons being recorded from were directionally-tuned (as suggested by Georgopoulos et al. 1988), then an EEG or ECoG electrode may have recorded from neurons that prefered reaching movements in many different directions. Some of these neurons may have increased their firing rate, and some may have decreased them, for any given reaching movement, and the overall effect on the recording would have been due to the balance between them all. This may therefore have resulted in an increase or a decrease in the signal’s overall power or complexity, depending on the balance of neurons involved. Another way of considering this is to assume that M1 activity is closely linked to downstream muscle activity. An area that is classically considered to control elbow movements may therefore have neurons relating to both biceps and triceps activity (that is, elbow flexors and elbow extensors). If the area being recorded from happened to have a greater representation of extensors than flexors, the area may in fact have shown decreased activity when the subject flexed their elbow (note, however, that this need not be related to reciprocal inhibition, which is most commonly represented at
the spinal cord level). In either of these two cases, the wide range of neurons being recorded from contributes to the possibility that either increases or decreases in any of the criterion functions may be indicative of neural activity.

5.6. Imagined Movements

A single trial of imagined reaching movements was included in Subject 4’s experiment. However, at the time of writing, this trial has not been fully analyzed. The main reason for this is the difficulty in obtaining “true” onsets in trials where the task requires imagined movements. At present, these onsets were defined by thresholding the EMG signals. An imagined movement, however, produces no such signals, and as such, an alternate method would need to be used to quantify the accuracy of any onset predictions produced by the analysis. Since the only other trial in which imagined movements were performed could not be properly analyzed (the reaching frame tests from Subject 3, as outlined in the Methods section), no comment could be made on the usefulness of the analytical methods in question for detecting the onset time of imagined movements. This limited the broader applicability of these results to BMI applications, since a user in the target population (that is, a user with high-level paralysis) would be unable to fully perform the movements assessed in our experiments. It is reasonable to suggest that, since imagined or attempted movements generally produce the same kinds of neural signals that performed movements do, albeit at a lower amplitude (Lacourse et al. 1999), imagined or attempted movements should produce similar results to those discussed herein, though perhaps with a reduced accuracy. This, however, currently remains a matter of speculation.
Chapter 6: Conclusions

In conclusion, it can be suggested that neural signals from primary motor cortex, as measured through EEG and ECoG, contained sufficient information for accurate predictions to be made about the timing of many different types of movements. Fractal dimension, spectral entropy, sum of differences, most bandpowers up to 40 Hz, the bandpower integral, and variance were all suggested as appropriate methods for deriving this information, while the current implementations of change-point analysis and phase were not considered as suitable. However, it was also shown that the accuracy of these predictions was highly dependent upon a number of different parameters, in ways that were not always consistent. Parameter sets that provided ideal results for one subject and movement type may not have provided similar results for another movement or another subject. It is therefore suggested that a wide range of parameter combinations would need to be tested for each case if the techniques developed herein are to be successfully applied to an asynchronous BMI system. However, as originally stated, the nature of the data and its subsequent analysis were such that these results should only be considered as pilot data, and a demonstration that the analysis program that has been developed herein is capable of generating predictions of movement onset time and providing approximations of the accuracy thereof.
References


