UNIVERSITY OF CALIFORNIA, SAN DIEGO

EEG analysis of gait movement preparation in the normal state and in the abnormal state of Parkinson’s disease with freezing of gait

A dissertation submitted in partial satisfaction of requirements for the degree Doctor of Philosophy

in

Neurosciences

by

Priya D. Velu

Committee in charge:

Professor Virginia R. de Sa, Chair
Professor Howard Poizner, Co-Chair
Professor Eric Halgren
Professor Marta Kutas
Professor Terry Sejnowski

2014
The dissertation of Priya D. Velu is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

______________________________

______________________________

______________________________

______________________________

Co-Chair

Chair

University of California, San Diego

2014
DEDICATION

To my father, who has encouraged me throughout my life to pursue knowledge and to make positive contributions to society.
# TABLE OF CONTENTS

Signature Page ........................................................................................................ iii

Dedication ................................................................................................................ iv

Table of Contents ...................................................................................................... v

List of Figures ........................................................................................................... vii

List of Tables ........................................................................................................... x

Acknowledgements ................................................................................................... xi

Vita ............................................................................................................................ xiv

Abstract of the Dissertation ....................................................................................... xv

Chapter 1  Introduction ............................................................................................ 1
  1.1  Background ..................................................................................................... 1
      1.1.1  Locomotive systems at the cortical level
      1.1.2  EEG as an imaging modality for recording movement preparation
      1.1.3  EEG analysis using signal processing and machine learning methods
      1.1.4  Abnormal gait states: freezing of gait in Parkinson’s disease patients
  1.2  Dissertation overview .................................................................................. 22
      1.2.1  Chapter 2
      1.2.2  Chapter 3
  1.3  Significance and contributions of work ...................................................... 23
  1.4  References ................................................................................................... 25

Chapter 2  Single-trial classification of gait and point movement preparation from human EEG ................................................................. 29
  2.1  Abstract ....................................................................................................... 29
  2.2  Introduction ................................................................................................ 30
  2.3  Materials and Methods ............................................................................. 34
      2.3.1  Data collection
      2.3.2  Data analysis
      2.3.3  Rationale behind feature selection
LIST OF FIGURES

Figure 1.1. Schematic of nine-level wavelet decomposition.........................14
Figure 2.1. Experiment set-up, trial structure, and EMG trace ......................37
Figure 2.2. L walk/L point feature visualization for S4 and S8.......................51
Figure 2.3. Combined directions feature visualization for S4 and S8..............52
Figure 2.4. Scalp topographies of PCs for L walk/L point in S4....................53
Figure 2.5. Confusion matrices for MDA classification result ......................54
Figure 2.6. Different action, same direction classification using AR
and bandpower ..................................................................................55
Figure 2.7. Same action, different direction classification using AR
and bandpower ..................................................................................56
Figure 2.8. Action versus stand classification using AR
and bandpower ..................................................................................57
Figure 2.9. Different action, different direction classification using AR
and band power ..................................................................................58
Figure 2.10. Combined directions classification using AR coefficients
and bandpower ..................................................................................59
Figure 2.11. Feature space of training and test trials – bandpower ..............60
Figure 2.12. Feature space of training and test trials – AR coefficients ........61
Figure 2.13. Feature space of training and test trials – wavelets ..................62
Figure 2.14. LDA weights of training and test trials – bandpower ..............63
Figure 2.15. LDA weights of training and test trials – AR coefficients ..........64
Figure 2.16. LDA weights of training and test trials – wavelets .................65
Figure 2.17. LDA weights in feature space – bandpower .................66
Figure 2.18. LDA weights in feature space – AR coefficients .................67
Figure 2.19. LDA weights in feature space – wavelets .................68
Figure 2.20. LDA weights in PC space – bandpower .................69
Figure 2.21. LDA weights in PC space – AR coefficients .................70
Figure 2.22. LDA weights in PC space – wavelets .................71
Figure 3.1. Experiment set-up and trial structure .........................97
Figure 3.2. Number of steps taken to reach door and to finish ..........105
Figure 3.3. Amount of time taken to reach door and to finish ..........106
Figure 3.4. Histogram of step count for each block for Subject TB ..........107
Figure 3.5. Spectra and RPDC of Block A for controls .........................110
Figure 3.6. Spectra and RPDC of Block C for controls .........................111
Figure 3.7. Spectra and RPDC of Block E for controls .........................112
Figure 3.8. Spectra and RPDC of Block A for PD subjects .........................113
Figure 3.9. Spectra and RPDC of Block C for PD subjects .........................114
Figure 3.10. Spectra and RPDC of Block E for PD subjects .........................115
Figure 3.11. Spectra and RPDC of Block A-C for controls .........................116
Figure 3.12. Spectra and RPDC of Block A-E for controls .........................117
Figure 3.13. Spectra and RPDC of Blocks A-C for PD subjects .........................118
Figure 3.14. Spectra and RPDC of Block A-E for PD subjects..........................119
LIST OF TABLES

Table 2.1. Number of channels and trials for each subject and class ............39
Table 2.2. Wavelet levels and coefficients for each frequency band..............41
Table 2.3. Classification errors for different direction, same action..............47
Table 2.4. Classification errors for same direction, different action..............47
Table 2.5. Classification errors for different direction, different action..........48
Table 2.6. Classification errors for action versus standing..........................48
Table 2.7. Classification errors for combined directions...............................49
Table 3.1. Characteristics of PD subjects..................................................94
Table 3.2. Characteristics of control subjects.............................................94
ACKNOWLEDGEMENTS

First, I would like to thank my advisor Virginia de Sa for her guidance, teaching, and patience for I have not only acquired much knowledge but also learned how to remain optimistic and forge ahead with new analyses or approaches during those gray days of suboptimal results. I also appreciate the level of autonomy that Professor de Sa gave me in designing and pursuing my projects, as I found it to be very useful practice for the independence that I will inevitably face in my future research career. I would also like to thank my co-advisor, Howard Poizner, for teaching me about and directing me toward the large amount of literature on motor control in Parkinson’s disease that I could never have fathomed of in medical school and for patiently guiding me through the labyrinth of protocol involved in experimentation on patient populations.

I thank the rest of my committee, Professors Eric Halgren, Marta Kutas, and Terry Sejnowski, for being excellent role models in science who have played no small part in developing my ability to examine and question data from multiple perspectives, design good experiments, and understand how results fit into the greater context of science.

The Medical Scientist Training Program, Neurosciences Graduate Program, and School of Medicine have provided me with graduate and medical school experiences that have beyond met my expectations from when I was a starry-eyed first-year many years ago. They have also been very
generous in their flexibility in allowing me to shuttle between clinic and lab as needed while I was finishing my dissertation this past year.

I have gained much knowledge from fellow graduate(d) students Paul Hammon, Adam Koerner, Tim Mullen, and Eunho Noh, and I could not have done without the many conversations that we had about techniques, methodology, and interpretation. I have also enjoyed mentoring Matthew Valdivia, a former Cognitive Science undergraduate, as one learns more from teaching than from learning itself.

My sanity during graduate school owes large regards to my friends in the Medical Scientist Training Program and Neurosciences Graduate Program, which includes my husband Adam Calhoun who I met in graduate school.

Finally, I would like to thank my mother and father for all of their support throughout the years. I would not be where I am now without their persistent encouragement and guidance. And my sister, because siblings in the end aren’t really that bad, and can actually be quite necessary.

I would like to thank the funding agencies that helped to fund my research at UCSD, including the National Institutes of Health (5T32GM007198-35 and R01-NS036449) and the National Science Foundation (CBET-0756828, IIS 1219200, SBE-0542013, SMA 1041755). This work would not have been possible without their generous contributions.
The text of Chapters 2 and 3 has already appeared in journals. Chapter 2 includes material as it appears in Velu PD, Mullen T, Noh E, Valdivia MC, Baram Y, Poizner H, de Sa VR, “Effect of visual feedback on the occipital-parietal-motor network in Parkinson’s disease patients with freezing of gait,” *Front Neurol.* 4: 209, 2013. The dissertation author was the primary author of this publication. The copyright for this publication is owned by Frontiers and used in accordance with their terms and conditions.

Chapter 3 includes material as it appears in Velu PD, de Sa VR, “Single trial classification of gait and point movement preparation from human EEG,” *Front Neurosci.* 7:84, 2014. The dissertation author was the primary author of this publication. The copyright for this publication is owned by Frontiers and used in accordance with their terms and conditions.
VITA

2005 Bachelor of Arts in Biophysics, Johns Hopkins University

2014 Doctor of Philosophy in Neurosciences, University of California San Diego

2014 Doctor of Medicine, University of California San Diego

RESEARCH

2008-2013 University of California, San Diego, Department of Neurosciences, Virginia de Sa, PhD; PhD thesis, “EEG analysis of gait movement preparation in the normal state and in the abnormal state of Parkinson’s disease with freezing of gait.”

2002-2005 Johns Hopkins University, Department of Biophysics, Bertrand García-Moreno, PhD; Honors thesis, “Thermodynamic stability of variants of staphylococcal nuclease with internal polar or ionizable groups.”

PUBLICATIONS


ABSTRACT OF THE DISSERTATION

EEG analysis of gait movement preparation in the normal state and in the abnormal state of Parkinson’s disease with freezing of gait

by

Priya D. Velu

Doctor of Philosophy in Neurosciences

University of California, San Diego, 2014

Professor Virginia R. de Sa, Chair
Professor Howard Poizner, Co-Chair

The cortical control of gait is an important aspect of locomotive function in healthy and diseased states. Here we used electroencephalography (EEG), signal processing, and machine learning methods to capture neural signals related to movement preparation of gait in healthy controls and in Parkinson’s disease patients with freezing of gait (FOG). We focused on pre-movement EEG in tasks that required natural ambulation of the subjects through the environment with the ultimate goal of application to brain computer interfaces.
First we aimed to predict the intent to start gait before the onset of any movement in normal, healthy individuals, and to distinguish this signal from different actions such as standing and pointing. Wavelets were used as features in LDA classification, which resulted in errors as low as 17% when averaged across nine subjects. The neural signal used for classification mostly consisted of contributions from slow cortical potentials (0.1-1 Hz) with additional, smaller contributions from mu (8-13 Hz) and beta (14-25 Hz) frequency bands from channels located over sensorimotor cortex.

We then examined this pre-movement period in subjects who experience freezing of gait (FOG) to determine how PD FOG patients differ from healthy controls in baseline neural activity and how these measures change with visual modulation of the environment. The spectral activity and connectivity in and between occipital, parietal, and motor regions were assessed by using EEG signals recorded at channels Oz, Pz, and Cz in a network analysis of causal information flow. The results suggest possible pathological over-binding of cortical areas important for sensory integration and motor control through increased baseline theta and beta oscillations in freezing of gait patients who respond to visual feedback.
Chapter 1

Introduction

1.1 Background

Locomotive systems at the cortical level

A literature search on the study of human motor systems at the level of the cerebral cortex reveals a large discrepancy in the number of studies that focus on movements of arms or hands and those that focus on locomotion. This may partly be an issue with experimental design as many recording methods are not amenable to movement of the subject across space, or it may be that the movements and functions of the upper extremities are considered more varied, complex, and essential to everyday life. The work in this thesis centers on cortical aspects of locomotion in the pre-movement preparatory state.

Mammals perform locomotive movements through a multi-level control system. The flow of control in basic, automatic gait starts with the selection of a motor program and release of the inhibition on the diencephalic locomotor region (DLR) and mesopontine locomotor region (MLR) by the basal ganglia. The DLR and MLR additionally integrate inputs from the optic tectum (where to go) and vestibular nuclei (go without falling) and send reticulospinal
projections to central pattern generators (CPG) in the spinal cord. During gait, online sensory input from limbs and trunk are relayed locally back to CPGs and also to higher brain centers. Of note, early experiments showed that decerebrate cats and rabbits showed the ability to locomote to search for food as long as thalamic innervation to the basal ganglia was maintained (Grillner 08).

For an extended time, the findings from such animal studies suggested that gait was essentially automatic and limited to connections between subcortical, brainstem, and spinal regions. However, the role of cortical control has become increasingly appreciated. It is now known that extensive corticostriatal and corticospinal pathways allow for more complex gait behavior. Cortical basal ganglia circuits are important for on-line corrections of already learned behaviors (Graybiel 2005). Corticospinal pathways have been shown to be vital for skilled movement, and corticostriatal pathways may be important for online change of gait in response to obstacles or new goals (Grillner 07).

Cortical contributions to gait have been elucidated through neuroimaging studies of healthy humans using electroencephalography (EEG), functional magnetic resonance imaging (fMRI), near infrared spectroscopy (NIRS), positron emission tomography (PET), and single photon emission computed tomography (SPECT) methods. These studies examined

**EEG as an imaging modality for recording movement preparation**

Several imaging modalities can capture aspects of cortical activity, albeit at different time scales and spatial resolutions. For instance, local field potentials (LFPs) and single unit recordings provide excellent temporal and spatial resolutions, electroencephalography provides comparable temporal resolution, and electrocorticography (ECoG) maintains the high temporal resolution of EEG but improves its spatial resolution by mitigating the signal distortion from scalp and skull layers by its subdural placement. All three are also direct measures of cortical activity that record the superimposition of all measurable currents at the electrode location. However, they are invasive and require surgery to access brain tissue at the intraparenchymal or subdural
level. While this is feasible and ethical in subjects who have medical requirements such as epilepsy who require such precise monitoring, the risks of infection and tissue damage currently preclude their routine use in experiments with human subjects.

Non-invasive imaging modalities such as PET and fMRI work under the assumptions that increased metabolic activity or increased cortical blood flow to specific brain regions indicates increased neural activity at those regions. These measurements occur on the scale of seconds, which is a poor temporal resolution given that neural activity occurs on the order of milliseconds (Heeger and Ress, 2002). Two non-invasive imaging modalities that can record at high time resolution are EEG and magnetoencephalography (MEG), which measure scalp electric potentials or magnetic fields produced by electrical currents produced in the brain.

Unfortunately, the spatial resolutions of these four non-invasive imaging modalities are diametric to their temporal resolutions. fMRI can resolve brain structures to 1-5mm and PET to 2-5mm, but both have temporal resolutions in the order of seconds due to the dynamics of blood flow. Both fMRI and PET can also image activity in all parts of the perfused and metabolically active brain, including subcortical structures and the brainstem. EEG has a time resolution of 1 ms but is limited to cortical activity due to the rapid loss of signal with distance, resulting in the largest contributions coming from post-
synaptic potentials (PSPs) generated at superficial layers of the cortex.

Further, all cortical activity may not be observed as the orientation of dipoles formed by PSPs may be closed field or may be positioned in a way that they cancel out to zero (Luck S, Kappenman E, 2011). Electrical potentials are also smeared together as they travel through the CSF, dura, and skull so that the signal ultimately recorded at the scalp is a sum of many different sources of cortical activity. This poor spatial resolution restricts most source interpretation of EEG data to only large brain regions on the order of 1 cm at best. In addition, the topography of the cortical surface varies from subject to subject so that an anatomical source of activity in one subject will produce different voltage patterns at the same scalp location compared to another subject. MEG has spatial resolution along the order of 1 mm since the transmission of magnetic fields are unaffected by the presence of the skull, scalp, dura, and CSF. However, only tangential, not radial, dipoles can be detected; recording requires use of large, expensive equipment including a superconducting quantum interference device that must be maintained a few degrees above absolute zero; and any tiny movement can create enough artifacts to render a recording unusable (Arbib, 2002).

The ambiguity of source in EEG can be somewhat mitigated using inverse modeling methods such as dipole fitting to locate likely sources of potentials. Methods such as dipole fitting and beam-forming attempt to
improve the spatial resolution of EEG, but these require solving the inverse problem. In the forward problem, all the parameters of the cortical sources of interest are known, and it is easy to calculate the electrical potential that will be produced at the scalp through a simple summation of the currents produced by different sources at different locations with adjustments for cortical topography and spatial dissipation of the electrical signal through the different tissues. When only the EEG signal recorded at the scalp is known, we are faced with the conundrum that infinite combinations of sources can produce that same potential. A way of avoiding the constraints of the inverse problem is by using an alternate method such as independent components analysis (ICA), which is a blind source method that attempts to separate the EEG signal from n channels into n statistically independent non-Gaussian subcomponents. Currently this method is only widely accepted for eye and muscle artifact removal. All of these methods are computationally time-consuming, and are currently impractical for use in real-time brain-computer interfaces (BCIs), which is the ultimate application of the results from this dissertation. Here we limit our location sensitivity to scalp locations of channels or scalp surface Laplacian estimates of surface potentials derived from the channel data.

Movement preparation has been extensively studied with EEG in finger movements or reach tasks with seated subjects. But it is feasible to use EEG
in a mobile setting, unlike MEG, fMRI, or PET. The system itself can be carried around as the amp is small enough to fit into a backpack and the produced signal can be transmitted through long fiber-optic cables or wirelessly to a distant computer. Motion artifact is a concern, but it can be avoided by analyzing the pre-movement period or mitigated through multiple types of artifact rejection methods. The recent proliferation of wireless EEG systems further eases the barrier to mobile imaging of neural activity.

**EEG analysis using signal processing and machine learning methods**

The EEG signal has traditionally been confined to analysis in the time-voltage domain. Though continuous recordings of EEG have been used and are still used extensively in medical applications such as monitoring of epilepsy and sleep pathology, most research applications have focused on event related potentials (ERPs) that are averaged over hundreds of trials in order to increase the signal to noise ratio. The ERPs are evaluated for features such as onset, amplitude, and duration as well as general spatial location.

Thanks to increased computing power, various signal-processing methods can be applied to EEG recordings to extract more complex information in other domains, such as time-frequency Fourier transformation, wavelet decomposition, or determination of causal information flow using autoregressive models and Granger causality. These domains tend to be high-
dimensional, but interpretation can be facilitated by pruning irrelevant dimensions with machine learning methods.

An additional advantage of these methods for processing data is that they are better at isolating useful information from single trials. Single trial or real-time interpretation of the EEG signal is particularly useful for applications in BCIs or brain state monitoring. BCIs, which were first developed to help patients with amyotrophic lateral sclerosis (ALS) who had lost most of their motor function, bypass the spinal cord to directly connect the user’s cortex to a system that performs functions, such as a computer programmed with a speller for communication, or a prosthetic limb to regain motor function.

Currently, BCIs are targeted towards many different populations, including patients with diseases such as stroke as well as healthy subjects. Examples of continuous brain state monitoring include detection of pre-ictal states in epilepsy patients, vigilance or stress levels in operators or soldiers, emotional states in psychiatric patients, and receptive attention states in memory paradigms (Zander and Kothe, 2011; Gevins A, McEvoy LK, Smith ME, Chan CS, Sam-Vargas L, Baum C, Ilan AB, 2012; Olvet and Hajcak, 2012).

Features commonly extracted from EEG signals for BCI application and used in this thesis include Fourier transformation of the signal from the time domain to the frequency domain, or coefficients from autoregressive modeling or wavelet decomposition of the time domain signal. One of the first features
used for BCIs, the desynchronization of cortical rhythms in the mu frequency band occurs before and during voluntary movement or active imagining of a movement (Pfurtscheller and Aranibar 1979). This desynchronization also has a spatial distribution focused over the primary sensorimotor area (Pfurtscheller and Neuper 1997). Other phenomenon such as a similar desynchronization of beta frequency rhythms, increased synchronization of mu rhythms in motor areas not related to the task, and the spatial topography of these rhythms may also serve as features (Pfurtscheller G, Stancák A Jr, Edinger G, 1997; McFarland DJ, Miner LA, Vaughan TM, Wolpaw JR, 2000; Pfurtscheller G, Brunner C, Schlögl A, Lopes da Silva FH, 2006.).

A computationally fast method of capturing the amplitude of the EEG signal at a particular frequency range (band power) and specified window of time is through discrete fast Fourier transform (FFT). This makes assumptions about the EEG signal that are simply not true: 1) it is infinite outside of the measured interval, 2) it is periodic, and 3) it can be optimally modeled as a sum of sine and cosine waves. It is also results in a discrete frequency spectrum with a resolution that depends on the number of available samples; a window of signal with short time length results in a periodogram with discrete frequency values spaced far apart so that resolving spectral peaks with close frequency values becomes difficult (Pardey J, Roberts S, Tarassenko L, 1995).
A similarly fast but less restrictive method of capturing these frequency
domain EEG phenomena is through non-adaptive autoregressive (AR)
spectral estimation of the signal. In contrast to FFT, the AR model is not only
descriptive but also predictive, is continuous, and does not make assumptions
about the periodicity or finitude of the EEG signal. It is a parametric method
that assumes that the signal 1) is stationary with unchanging statistical
properties such as average amplitude and frequency content, 2) is a
stochastic process, and 3) can be linearly modeled by a small number of
parameters. EEG signals can be considered locally stationary over small time
intervals, such as 1s or below, so that non-adaptive, computationally fast
approaches such as the Burg method can be used, but adaptive approaches
that continually update the AR model with new input are also available if the
stationarity of the signal is of concern (Pardey J, Roberts S, Tarassenko L,
1995). For a sequence $s_1 \ldots s_n$, where $s_n$ is the current sample, a non-adaptive
AR model can predict a value $x_n$ by summing the $p$ previous points in the
sequence after they have been multiplied by the respective coefficient $a_p$ so
that $x_n = \sum_{i=1}^{p} a_p s_{n-i}$. Thus, the forward prediction error $e_{p,n}$ between the actual
value $s_n$ and the predicted value $x_n$ is $s_n - x_n$, and the mean of the squared
prediction errors for the whole sequence is $E = \frac{1}{N} \sum_{n=1}^{N} e_{p,n}^2 = \frac{1}{N} \sum_{n=1}^{N} (s_n - x_n)^2$. The
least-squares error criterion can then be used to find values for coefficients $a_1$,
to $a_p$ by finding a best fit for the $p$th-order model of $x_n$. Burg’s method uses both forward and backward recursion to find values for coefficients, and is better for short data sequences with spectral content as used in this work (Chen, 1988).

A major caveat of AR modeling is that the experimenter must choose the order of the model. The lowest appropriate model order for a specific set of data can be found by minimizing a function that balances the monotonic decrease in prediction variance with increasing model order and the monotonic increase in the statistical bias of the estimation of AR coefficients from the true signal with increasing model order. Aikake defined the final prediction error (FPE) criterion:

$$FPE(k) = \frac{N + k + 1}{N - k - 1} \sigma(k)^2$$

where $N$ is the number of samples in a frame, $k$ is the model order, and $\sigma(k)^2$ is the variance associated with order $k$. The optimal model order as predicted by this function should minimize bias and mean square prediction error. Aikake developed a second criterion, the ‘minimum information theoretical criterion’ (AIC):

$$AIC(k) = -2 \ln(M \hat{\theta} + 2k)$$

where $ML$ is the maximum likelihood estimate of $k$ AR parameters (Schwindlein and Evans, 1989). Other criteria also exist and usually result in similar optimal model orders. Studies of BCI applications have shown that a
model order of 16 is optimal for EEG classification of motor tasks based on mu (8-13 Hz) or beta (14-25 Hz) motor rhythms (McFarland and Wolpaw, 2008).

Wavelet decomposition is another method of capturing information about the frequency content of an EEG signal, and similar to AR, it makes fewer assumptions about the nature of the signal than FFT. It not only provides information on the spectral nature of the signal, but also provides good time resolution for higher frequencies, making it particularly good for analyzing discontinuities, trends, and patterns that exist in biological signals. The power in this technique lies in the ability to model the signal through any imaginable basis function with wavelets, or waveforms with amplitudes that begin at zero and end at zero. There are many established families of wavelets, and some experimenters have designed custom wavelets tailored to each subject using classification cross-validation schemes that choose optimal waveform properties based on their ability to correctly predict an outcome (do Nascimento and Farina, 2006).

In the continuous wavelet transform (CWT), wavelets are scaled in order to determine the spectral content of different frequencies within the signal so that for larger frequencies, the chosen wavelet is scaled to be longer and vice versa. These wavelets are shifted in time along the signal, multiplied by the signal, and integrated over the entire time length of the sample. The smaller scaled versions of the chosen wavelet, which correspond to higher
frequencies, are essentially smaller windows of analysis that naturally result in better time resolution (Chui 1997).

In the discrete wavelet transform, the EEG signal goes through multiple filtering and sub-sampling steps (Figure 1.1), with coefficients usually sampled dyadically from the CWT. For example, a signal sampled at 512 Hz and of one-second duration has a Nyquist frequency of 256 Hz and 512 data points. This signal is low-pass filtered and high-pass filtered at half of the highest frequency in the signal, or 128 Hz, and every other point is removed, effectively doubling the frequency resolution and halving the time resolution. This is repeated in a process called wavelet decomposition and is equivalent to scaling the wavelet in CWT. The wavelet function is applied during the filtering stages, resulting in approximate and detailed coefficients at each level that decrease in number as the decomposition progresses to the final level, which represents the lowest frequency analyzed. The total number of coefficients is equal to the number of sample points in the original signal. In the example, a wavelet decomposition consisting of nine levels would result in the lowest frequency ranges of 0-0.5 Hz and 0.5-1 Hz. The DWT is much faster to compute than the CWT, making it more suitable for most applications (MATLAB 2012b).
Another aspect of brain activity rich in information is the interaction between different neural processes. Multivariate autoregressive modeling of time series EEG data can provide information about how multiple neural processes progress over time. Application of Granger causality functions to these models of different neural processes can describe if and how these neural processes influence each other. Importantly, Granger causality implies true causality in that a process Granger-causes another process only if information from the upstream process improves prediction of the downstream process (Bressler and Seth, 2011). More specifically, if $X$ and $Y$ represent two
variables, the prediction of $X_{t+1}$ is improved by using both past terms of $X$ and $Y$ as opposed to only using past terms from $X$ (Bressler and Seth, 2011) if $Y$ Granger-causes $X$. Analyzing these interactions may yield otherwise unknown features for classification or monitoring that can be more easily managed in real-time.

One legacy of artificial intelligence is the rapidly growing field of machine learning. Indeed, one could argue that the most important aspect of an intelligent system is its ability to learn. Intelligent systems observe examples that provide incomplete information about a statistical phenomenon and construct a rule from these observed examples that will generalize to new, unobserved examples of the phenomenon. Machine learning uses the tools of statistics to perform this inductive inference.

Two types of learning exist: supervised and unsupervised. In supervised learning, a rule is built using labeled examples that form a training set. For example, there are $n$ samples ($x_1, x_2, \ldots, x_n$) in a training set, where $x_i$ is a vector of features. There are also $k$ labels so that each sample has a label associated with it, thus forming $k$ “groups” within the $n$ samples. Unsupervised learning does not have the luxury of labeled samples, therefore there is no training set and regularities (or irregularities) in the data must be used to separate the examples into clusters. Labels are discrete in classification (also
known as pattern recognition) problems and continuous in regression problems (Bishop 2006).

There are two main steps in a learning problem. The first is extracting informative features from the collected data, and the second is constructing an algorithm using statistical parameters derived from these features that will allow for generalization to new cases. Selecting features is arguably the most crucial step since a classification algorithm constructed using noisy, high dimensional data is not going to perform as optimally as it could both in terms of speed and error rates. Also, by extracting the most informative features possible and reducing the dimensionality of the data, simple algorithms can provide similar performance to more complex and computationally intense algorithms, which also have a greater tendency to over-fit data (Lotte F, Congedo M, Lécuyer A, Lamarch F, Arnaldi B, 2007). The work in this thesis uses linear discriminant analysis (LDA), a parametric supervised algorithm, for classification of type of movements based on features from the pre-movement EEG signal. LDA has shown similar performance to more complex classification algorithms such as neural networks and support vector machines when used in brain computer interface applications (Renfrew M, Cheng R, Daly JJ, Cavusoglu M, 2008).

As an example, many EEG based brain computer interfaces use the phenomenon of desynchronization of cortical rhythms in the mu (8-13 Hz) or
beta (14-25 Hz) bands over relevant motor cortex regions during motor imagery as a control signal (McFarland DJ, Miner LA, Vaughan TM, Wolpaw JR, 2000). One can use the EEG channels over the right and left hand regions of the motor cortex (cut down to two channels: C3 and C4) and take the power from these channels within the mu and beta frequency bands (two values for each channel) over the time period (one chunk of 500 ms) during motor imagery to produce features for the classification problem (is the person thinking of moving their right or left hand). This provides the classifier with information relevant to the task and reduces the data from 64 channels x 512 Hz x 0.5 seconds to 2 channels x 2 bandpower values per trial.

Another way to reduce dimensionality is by using the non-parametric method of principle components analysis. The goal of this method is to minimize redundancy and maximize information in data $X$ of $n \times m$ dimensions by finding a projection matrix $P_{n \times n}$ that transforms $X$ into a new $n \times m$ matrix $Y$ in which variances within dimensions are maximized and covariances between dimensions are minimized. In other words, the off-diagonal terms in the covariance matrix $Y$ should be zero so that the dimensions are maximally decorrelated, and the diagonal terms are maximized so that variances are as large as possible. This diagonalization process is performed algebraically by using eigenvector decomposition to find $n$ normalized directions in $n$-dimensional space in which the variance in $X$ is maximized, where each
solution is orthogonal to the rest. The final result is the transformation matrix \( P \), or \( n \) principle components arranged in order of decreasing rank. By retaining only a select number of principle components \(< n\) by rank order, the dimensionality of the data is reduced while the most useful information is retained. In using PCA, it is important to note that we are assuming that the data is linear, that large variances provide important information about the structure of the data while covariances represent noise, and that principle components are orthogonal to one another.

The LDA classification algorithm uses as its inputs features that are continuous and outputs a target value that is discrete (0 or 1). In other terms, it performs two-class separation of data. LDA utilizes the Bayesian framework. An example \( x \) belongs to class A if \( P(A|x) > P(B|x) \), where A and B are the two groups and A does not equal B. However, these probabilities are known. What is known is \( P(x|A) \) and \( P(x|B) \) since we have a training set of data that is labeled. Using Bayes’ Theorem: \( P(A|x) = P(x|A)P(A)/P(x) \), with the following assumptions: 1) \( P(x|A) \) and \( P(x|B) \) are multivariate normal distributions; 2) covariances of group A (\( \Sigma_A \)) and group B (\( \Sigma_B \)) are equal; and 3) \( P(A) = P(B) \) if you have the same # of trials in the training set from each group.

Fisher’s LDA is a variant of LDA which does not assume normal probability distributions or equal covariances in classes and reduces the
dimensionality of the data to a 1-D space in the classification process. The fundamental concept behind Fisher’s discrimination is the maximization of distance between the projected class means and minimization of variance within each class so that class overlap is minimal. For example if there are $N_1$ points in class $C_1$ and $N_2$ points in class $C_2$, the mean vectors for each class are as follows

$$m_1 = \frac{1}{N_1} \sum_{n \in C_1} x_n,$$

$$m_2 = \frac{1}{N_2} \sum_{n \in C_2} x_n.$$

Then when the data is projected onto $w$, the separation of the projected class means $m_2 - m_1 = w^T(m_2 - m_1)$, where $m_k = w^T m_k$ and $k$ is the class label, should be maximized with $w$ of unit length $\sum_i w_i^2 = 1$ so that maximization is not due to arbitrarily large values of $w$. At the same time, the within-class variance of the projected data from a class $C_k$ given by $s_k^2 = \sum_{n \in C_k} (y_n - m_k)^2$, where $y_n = w^T x_n$, should be minimized. A threshold $y_0$ can then be chosen so that a point belongs to class one if $y(x) \geq y_0$ or to class two if $y(x) \leq y_0$. A caveat is that information may be lost in the projection of multi-dimensional inputs to 1-D space, but as mentioned above, if features are appropriately chosen this may not be an issue.
**Abnormal gait states: Freezing of gait in PD patients**

Parkinson's disease (PD) results from the loss of dopaminergic neurons in the substantia nigra pars compacta. PD patients suffer from tremors, rigidity, bradykinesia/akinesia (slowing of movement or the total absence of movement), and postural instability. Current drugs that treat PD patients compensate for the loss of dopaminergic neurons by increasing dopamine levels, either through inhibiting the inactivation of dopamine or delivering extra dopamine to the area.

The motor deficits of PD patients can be further decomposed into intensive and coordinative deficits. The intensive deficits include decreased speed and amplitude of movements. These deficits are alleviated by dopamine therapy. The coordinative deficits are problems in joint coordination and the inability to combine sensory and motor information to produce complex motor sequences. These deficits may still exist even with dopamine therapies.\(^3\)\(^6\) For example, PD patients on medications show deficits in integrating proprioceptive information with visual input to coordinate an appropriate reach-to-grasp object movement (Adamovich SV, Berkinblit MB, Hening W, Sage J, Poizner H, 2001). In fact, PD patients rely heavily on visual input as opposed to other inputs to guide motion (Poizner H, Feldman AG, Levin MF, Berkinblit MB, Hening WA, Patel A, Adamovich SV, 2000; Schettino LF, Adamovich SV, Hening W, Tunik E, Sage J, Poizner H, 2006).
Freezing of gait is a debilitating phenomenon seen in a significant subset of PD patients. It is defined as the unexpected and sudden inability to start or continue walking, and frequently occurs when a patient has to walk through a doorway or other restricted space, turn a corner, or walk in a sensory heavy environment such as when crossing a busy intersection. Some cases improve with dopamine or deep brain stimulation (DBS) therapy, while other cases are resistant to these interventions.

Over the last five decades, many groups have attempted to elucidate the mechanisms underlying FOG in PD patients through behavioral and imaging experiments. The current consensus includes five categories that are not necessarily mutually exclusive: 1) abnormal gait pattern generation, 2) disruption of central drive/automaticity of gait, 3) abnormal coupling of posture and gait, 4) perceptual malfunction, and 5) frontal executive dysfunction (Heremans E, Nieuwboer A, Vercruysse S, 2013). These dysfunctional components of gait in FOG correspond to different levels in the neurological system, but there is currently no comprehensive theory that unifies all levels into a pathological neural-behavioral circuit.

Interestingly, several studies have shown that simple visual cues are effective in helping patients to overcome freezing episodes (Jiang Y and Norman KE, 2006; Nieuwboer A, 2008; Frazzita G, Maestri R, Uccellini D, Bertotti G, Abelli P, 1997). These visual cues were as simple as transverse
lines placed on the floor in front of the subject, or a laser pointer used by the subject to indicate a target location for propelling the foot towards. To date, there is extensive literature on the behavioral effects but no information on the processes occurring at the neural level that contribute to this phenomenon.

1.2 Dissertation overview

Chapter 2

This chapter, in full, is a reprint of the article as it appears in Velu PD and de Sa VR, Single-trial classification of gait and point movement preparation from human EEG, *Front. Neur.* 2013. June 11 7:84. Healthy young subjects were recruited to participate in a task that involved classifying between three possible movements (standing, walking, pointing) that use overlapping muscle groups and between three possible directions (left, front, right). Nine-level discrete wavelet decomposition that encompassed a range of 0-256 Hz was applied to the EEG signal of 750 ms in length before movement onset. PCA was used to reduce this large feature space into ten dimensions per trial for two-class regularized linear discriminant analysis (LDA) of eighteen types of movement comparisons. LDA weighting of PCA weights of channels and wavelet coefficients was visualized to determine the type of signal (spatial location and frequency range) that contributed to LDA classification. Finally,
multi-class LDA was performed to determine if there was a spatial signature in the EEG signal that contributed to left vs. forward vs. right classification.

Chapter 3

This chapter is adapted from the manuscript Velu PD, Mullen T, Noh E, Valdivia MC, Baram Y, Poizner H, and de Sa VR, Effect of visual feedback on the occipital-parietal-motor network in Parkinson’s disease patients with freezing of gait, *Frontiers in Movement Disorders* (2013, In Review). The goal of this study was to determine if 1) there is a difference between PD FOG patients and controls during the pre-movement period before visual cues are show, and 2) how real-time, closed loop visual feedback affected spectra and connectivity in and between occipital, parietal, and motor areas in PD patients with FOG that responded to visual cues compared to PD patients with FOG that did not respond to visual cues and old healthy controls. Successful response to visual feedback was determined by significant changes in behavioral performance as determined by number of steps taken and amount of time taken to complete the task as well as subjective report by the patient.

1.3 Significance and contributions of work

Chapter 2 resulted in the first published single-trial detection of movement preparation before natural gait movement using EEG in humans.
Classification against two other movements, pointing and standing, that share similar muscles to those used in ambulation was also successful, and an underlying spatial signature was discovered in the EEG signal. The LDA classifier weighted contributions from the frequency range of slow movement cortical potentials and channels in the sensorimotor area more heavily. These results support the feasibility of detection of the intent to start walking in a particular direction using non-invasive, mobile, and comparably affordable EEG technology for application towards a BCI. They also support the view that large-scale cortical neural activity is involved in preparation of gait movement.

Chapter 3 resulted in the first analysis of EEG in Parkinson’s disease patients with FOG before, during, and after visual feedback in a real ambulatory setting. The resulting neural dynamics are consistent with recently published data exhibiting increased theta activity in Parkinson’s disease patients versus healthy controls and increased theta activity during states of imbalance compared to normal ambulation. Additionally, all PD FOG patients exhibited a peak in the theta frequency range in connectivity from the occipital area to parietal and sensorimotor areas before visual cues were presented. This measure has potential for use as a diagnostic tool in managing Parkinson’s disease.
1.4 References


Chapter 2
Single-trial classification of gait and point movement preparation from human EEG

2.1 Abstract

Neuroimaging studies provide evidence of cortical involvement immediately before and during gait and during gait-related behaviors such as stepping in place or motor imagery of gait. Here we attempt to perform single-trial classification of gait intent from another movement plan (point intent) or from standing in place. Subjects walked naturally from a starting position to a designated ending position, pointed at a designated position from the starting position, or remained standing at the starting position. The 700 ms of recorded electroencephalography (EEG) before movement onset was used for single-trial classification of trials based on action type and direction (left walk, forward walk, right walk, left point, right point, and stand) as well as action type regardless of direction (stand, walk, point). Classification using regularized LDA was performed on a principal components analysis (PCA) reduced feature space composed of coefficients from levels 1 to 9 of a discrete wavelet decomposition using the Daubechies 4 wavelet. We achieved significant
classification for all conditions, with errors as low as 17% when averaged across nine subjects. LDA and PCA highly weighted frequency ranges that included movement related potentials (MRPs), with smaller contributions from frequency ranges that included mu and beta idle motor rhythms. Additionally, error patterns suggested a spatial structure to the EEG signal. Future applications of the cortical gait intent signal may include an additional dimension of control for prosthetics, preemptive corrective feedback for gait disturbances, or human computer interfaces (HCI).

2.2 Introduction

The detection of locomotive intent has potential as a control signal in brain-computer interfaces, for mitigating complications in movement disorders, or for use in environments with human computer interfaces (HCI). Detection of this signal is valuable for spinal cord injury patients and other users of lower limb prosthetics or artificial exoskeletons. It may also prove useful for alleviating abnormal states that prevent proper gait coordination such as during freezing of gait in Parkinson's disease by providing preemptive visual cues that help prevent freezing episodes (Hanakawa T, Katsumi Y, Hidemo F, Honda M, Hayashi T, 1999; Jiang and Norman 2006; Nieuwboer 2008). Other uses could be in HCI systems that anticipate user movement. Truly useful and
versatile applications will require differentiation between many possible types of movements as well as the directionality of the movements. Here we explore the possibility of electroencephalography (EEG) to differentiate the intent to produce a locomotive movement from another movement as well as the spatial direction of the movement.

EEG is an ideal modality to capture locomotive intent. Its non-invasive nature is attractive to users, and the high temporal resolution, which is lacking in other non-invasive methods such as functional magnetic resonance imaging (fMRI), is suitable for capturing the cortical dynamics of gait planning and production. EEG systems are also inexpensive compared to other imaging technologies and can be worn comfortably by the subject even during large-scale movements such as locomotion. There are many groups competing to produce commercial dry electrode EEG systems, including ones that wirelessly connect to cell phones (Matthews R, Turner PJ, McDonald NJ, Ermolaev K, Manus T, McShelby RA, Steindorf M, 2008; Yasui 2009; Wang YT, Wang Y, Jung TP, 2011).

Locomotive movement through an environment is a crucial survival trait of many organisms. Vertebrates perform locomotive movements through a multi-level control system involving the cortex, brainstem, cerebellum, and spinal cord. This system integrates visual, vestibular, and proprioceptive information from the environment to perform coordinated movement of joints.
and muscles. Direct pathways from the cortex to the spinal cord have been shown to be vital for skilled movement, and pathways from the cortex to the basal ganglia may be important for online change of gait in response to obstacles or new goals (Graybiel 2005; Grillner S, Wallen P, Saitoh K, Kozlov A, Robertson B, 2008). Recently, intracortical recordings from hundreds of neurons in monkey motor cortex were successfully decoded to 3D coordinates of leg joints during treadmill walking of varied speed and direction (Fitzsimmons NA, Lebedev MA, Peikon ID, Nicolelis MAL, 2009). A similar experiment conducted with humans and using EEG decoded the linear and angular kinematics of the ankle, knee and hip joints during treadmill walking (Presacco A, Goodman R, Forrester L, Contreras-Vidal JL, 2011).

Neuroimaging studies in humans provide further evidence that the cortex is active during gait preparation and gait production. These studies attempted to isolate the neural correlates of bipedal gait production in humans by measuring brain activity during visual observation and mental motor imagery of gait and related motor activities with fMRI and positron emission tomography scanning (PET) (Malouin F, Richards CL, Jackson PL, Dumas F, Doyon J, 2003; Jahn K, Deutschlander A, Stephan T, Strupp M, Wiesmann M, Brandt T, 2004; Bakker M, De Lange FP, Helmich RC, Scherringa R, Bloem BR, Toni I, 2008; Iseki K, Hanakawa T, Shinozaki J, Nankaku M, Fukuyama H, 2008; Wang CH, Wai YY, Kuo BC, 2008; Wang JJ, Wai YY, Weng YH, Ng KK,

This study aims to detect EEG based cortical activity that is related to gait preparation before the voluntary motor production of gait. To our knowledge, we are the first to attempt to classify the intent to walk from single trial EEG data recorded before the onset of a natural gait movement in which the subject walks in real space from a starting position to a target position. We also classify the intent to walk from another motor plan, the intent to point. The classification of reach preparation before movement onset from EEG has been well established (Hammon PS, Makeig S, Poizner H, Todorov E, de Sa VR, 2008; Wang and Makeig, 2009; Lew E, Chavarriaga R, Silvoni S, Millán J, 2012).

The classification process here uses information from all available frequency ranges and channels to allow for individual variations in cortical
topography and dynamics. By examining which channels and which frequency ranges are most weighted by the top principal components and the LDA classifier, we characterize the nature of the signal used in classification. Finally, we attempt classification of pre-movement EEG between different target positions located in spatially distinct areas either to the left, right or in front of the subject.

We used a feature space that could capture mu (8–13 Hz) and beta (14–25 Hz) frequency band desynchronization over central motor and premotor areas (PMAs) as well as slow movement related potentials (MRPs) within single trials. Wavelets, especially the Debauchies (db) family, have been extensively used for EEG classification. The intended motor plan was predicted by classification of a principal components analysis (PCA) reduced feature space using regularized linear discriminant analysis (rLDA).

2.3 Materials and Methods

Data collection

Nine healthy, right-handed subjects (18–27 years old, 2 females) participated in the experiment. All subjects read and signed informed consent forms that were approved by the UCSD Human Research Protections Office. Subjects were naive and untrained in the task and no feedback was given that
could cue subjects to modulate their cortical signals to produce better features over time.

Continuous EEG was recorded from 64 Ag/AgCl electrodes positioned on a BioSemi nylon head cap according to the 10–20 International System. The signal was amplified with fixed gain BioSemi ActiveTwo amplifiers, band-passed from 0.2 to 100 Hz, and digitized at 512 Hz with 24-bit resolution. The independent software package DataRiver was used to read and record EEG signals as well as to integrate EEG signals with events from the Stim2007 stimulus presentation software (Delorme A, Mullen T, Kothe C, Akalin Acar Z, Bigdely-Shamlo N, Vankov A, Makeig S, 2011; Delorme A, Kothe C, Vankov A, Bigdely-Shamlo N, Oostenveldt R, Zander T, Makeig S, 2012; Vankov A, Bigdely-Shamlo N, Makeig S, 2010). Two EOG electrodes were placed to record eye movements (one on the right outer canthus and one below the right eye). Right and left mastoid electrodes were averaged off-line to serve as reference. To minimize movement artifacts, subjects were encouraged to remain as still as possible and to look at a fixation cross on the wall until they heard the go cue. Two electrodes were placed on the anterior tibialis muscle (the first muscle to activate in gait) on each leg to detect premature muscle contraction during trials (Mann RA, Hagy JL, White V, Liddell D, 1979). As only eight EMG electrodes were available, the muscle activity of the arm could not be monitored through electrophysiological methods; the experimenter noted
and discarded any trials with premature pointing or arm movements unrelated to the task.

Cues were given in the form of auditory stimuli played through two speakers located behind the subject. A trial consisted of a command cue spoken by a computer-generated voice (walk front, walk right, walk left, point right, point left, stand still) followed by a delayed go cue (indicated by an auditory tone of 1 s in duration). The command cues were 1 s in length, and an interval of 1 s occurred after the end of the command cue and before the sound of the go cue. For point and stand trials the end of the trial was indicated by an auditory tone 2 s after the go cue while for walk trials the same tone was sounded 4 s after the go cue (Figure 1b). The experiment consisted of 60 trials for each condition, or a total of 360 trials, presented in pseudorandom order in six blocks of 60 trials with 2 min of rest between blocks, with the exception of Subject 1, who had 80 trials per condition. We reduced the number to 60 trials per condition for subjects 2–9 as the time to complete the experiment was prohibitively long. After the go cue sounded, the subject either walked forward five feet to designated spots on the floor to the right, left, or in front of standing position; pointed at designated objects on the right and left ends of a table placed five feet in front of the subject, or remained standing still with eyes focused on the fixation cross at eye level on the wall directly in front of the subject (Figure 1). The subject carried the EEG amplifier
and battery in a specially designed backpack for the duration of the experiment.

**Figure 2.1.** (A) Experiment set-up. Subject stands with arms at side and fixates on cross bar until the go cue sounds to perform one of six actions: walk left, front or right to target marked on floor, point left or right at an object on the table, or stand still. (B) Trial structure with the red bar indicating time range of data used in classification. (C) EMG from right anterior tibialis muscle of one subject. Trials in which EMG indicated movement onset prior to 200 ms were omitted. The red bar indicates the time range of data used in classification.

**Data analysis**

All data were analyzed offline. For artifact removal, data were high pass filtered above 1 Hz to remove slow cortical potentials and galvanic skin potentials. The experimenter first visually inspected the data for removal of noisy channels, epochs with artifacts, and epochs with incorrect responses or premature leg movement (Figure 1c). These data was then further cleaned using EEGLAB automatic artifact rejection functions that removed channels and epochs that had kurtosis measures 5 standard deviations from the mean.
Kurtosis value (Delorme and Makeig, 2004; Delorme A, Mullen T, Kothe C, Akalin Acar Z, Bigdely-Shamlo N, Vankov A, Makeig S, 2011). Kurtosis is the fourth moment measure of a probability distribution, and large positive kurtosis values indicate increased peaky shape whereas large negative kurtosis values indicate abnormally flat shape in the distribution. In EEG, these may represent undesirable artifacts in the data.

The above channels and time points were noted and then excluded from original raw, unfiltered data for subsequent classification analyses. The number of trials remaining for each class and the number of channels remaining for each subject after artifact rejection are shown in Table 1.

Filtering of data provided to the classifier was explicitly avoided so as to avoid any misrepresentation or distortion of the signal. High pass and acausal filters are commonly used in EEG research to remove noise, but these are poor choices for predictive classification as they produce temporal smearing of the signal such that values at future time points in which actual movement occurs may be used in the calculation of values at time points of interest for movement prediction. Causal filters avoid this problem, but still cause distortions in the original signal.

For classification, EEG data were separated into epochs starting 500 ms before and ending 200 ms after the onset of the go cue (but before the onset of gait) of each trial. We used this time range in order to capture the
MRP signal from the motor cortex as well as any pre-movement related mu or beta desynchronization. Epochs were then processed to extract wavelet coefficients as the features for the classifier.

Table 2.1. The number of channels and the number of trials for each class remaining after artifact elimination are shown for each subject. The final column lists the adjusted cut-off for binary classification significance based on Wald intervals modified for small sample sizes determined by the class with the lowest number of trials.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Channels</th>
<th>L. walk</th>
<th>R. walk</th>
<th>F. walk</th>
<th>L. point</th>
<th>R. point</th>
<th>Stand</th>
<th>Cut-off (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>63</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>72</td>
<td>41</td>
</tr>
<tr>
<td>S2</td>
<td>59</td>
<td>44</td>
<td>44</td>
<td>46</td>
<td>44</td>
<td>44</td>
<td>46</td>
<td>40</td>
</tr>
<tr>
<td>S3</td>
<td>59</td>
<td>49</td>
<td>49</td>
<td>49</td>
<td>49</td>
<td>49</td>
<td>50</td>
<td>40</td>
</tr>
<tr>
<td>S4</td>
<td>51</td>
<td>45</td>
<td>45</td>
<td>52</td>
<td>45</td>
<td>45</td>
<td>49</td>
<td>40</td>
</tr>
<tr>
<td>S5</td>
<td>62</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>48</td>
<td>58</td>
<td>40</td>
</tr>
<tr>
<td>S6</td>
<td>58</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>51</td>
<td>39</td>
</tr>
<tr>
<td>S7</td>
<td>50</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>27</td>
<td>29</td>
<td>37</td>
</tr>
<tr>
<td>S8</td>
<td>50</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>48</td>
<td>40</td>
</tr>
<tr>
<td>S9</td>
<td>59</td>
<td>47</td>
<td>47</td>
<td>47</td>
<td>47</td>
<td>47</td>
<td>47</td>
<td>40</td>
</tr>
</tbody>
</table>

Rationale behind feature selection

Wavelets provide high-resolution frequency information at low frequencies and high-resolution time information at higher frequencies. Many biological signals consist of slow oscillating background activity with rapid onset of change in activity. The temporal dynamics of such frequency perturbations may be better captured by wavelets than other methods.

Wavelet decomposition was performed on all available channels within a subject. This yielded a high dimensional feature space with \( d = n \) channels \( \times \) 358 coefficients per trial. To facilitate classification by LDA, this high dimensionality was reduced by PCA to ten dimensions per trial (Lan T, Erdogmus D, Black L, Van Santen J, 2010). We used all 10 principal components for classification since only components with the highest variance
were not ensured to be informative for classification (Lugger K, Flotzinger D, Schloegl A, Pregenzer M, Pfurtscheller G, 1998). Thus, we let the classifier decide which components were most informative by assigning those components higher weights during the training process.

**Feature parameters**

We calculated the coefficients from level 9 discrete wavelet decompositions of the data using Daubechies wavelet 4 (db4) as the mother wavelet with periodization padding. The db family of wavelets has been shown to be better than biorthogonal wavelets, autoregressive filtering, and mismatched filtering for extracting movement-related information from EEG signals (Renfrew M, Cheng R, Daly JJ, Cavusoglu M, 2008). The db4 wavelet has been used extensively to analyze EEG signals (Subasi, 2006) and has performed as well as wavelets customized to individual subjects' EEG training data (do Nascimento and Farina, 2006; Farina D, do Nascimento OF, Lucas MF, Doncarli C, 2007) in a task predicting torque direction in foot movement.

Since the EEG signal was sampled at 512 Hz, the Nyquist frequency was 256 Hz. The level 1 decomposition thus included frequency information from 128 to 256 Hz, the level 2 decomposition from 64 to 128 Hz, and so on (Table 2). Detail coefficients from levels 1 to 9 encompassed available frequency signals from 0.5 to 56 Hz, and the approximate coefficients for level
9 had signals from 0 to 0.5 Hz. This range included both the mu and beta frequency sub-bands associated with idle motor rhythms and slow cortical potentials such as the MRP.

Table 2.2. Wavelet levels and coefficients with corresponding frequency ranges.

<table>
<thead>
<tr>
<th>Level</th>
<th>Frequency (Hz)</th>
<th>Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>0-0.5</td>
<td>1 (approximate)</td>
</tr>
<tr>
<td>9</td>
<td>0.5-1</td>
<td>2 (detail)</td>
</tr>
<tr>
<td>8</td>
<td>1-2</td>
<td>3, 4</td>
</tr>
<tr>
<td>7</td>
<td>2-4</td>
<td>5-7</td>
</tr>
<tr>
<td>6</td>
<td>4-8</td>
<td>8-13</td>
</tr>
<tr>
<td>5</td>
<td>8-16</td>
<td>14-25</td>
</tr>
<tr>
<td>4</td>
<td>16-32</td>
<td>26-48</td>
</tr>
<tr>
<td>3</td>
<td>32-64</td>
<td>49-93</td>
</tr>
<tr>
<td>2</td>
<td>64-128</td>
<td>94-183</td>
</tr>
<tr>
<td>1</td>
<td>128-256</td>
<td>184-362</td>
</tr>
</tbody>
</table>

LDA classification

Features were extracted for six classes of trials: walk left, walk right, walk front, point left, point right, and stand. The classes were paired into 15 different binary classification problems, which can be separated into four different categories: same action with different directions, different actions with same direction, different actions with different directions, and actions vs. standing. Classes were also collapsed over directions to test the ability to classify walk/stand, point/stand, and walk/point.

A regularized LDA classifier (with optimized regularization parameter k) was trained on 10 features per trial using a 10-fold cross validation scheme.
Each class was assigned an equal number of trials and these trials were then randomly shuffled and partitioned into 10 sets, with nine used for training and one used for testing. Training and testing was done 10 times with a different set used each time as the test set. Thus, test data were completely separate from the training data. This complete process was repeated 9 additional times, and the prediction errors from the resulting 100 test cross validation folds were averaged and reported as the final results.

In situations with few trials compared to features, we do not have enough data to accurately estimate the covariance matrices used by LDA. In the absence of adequate information, a spherical covariance matrix is often assumed. Regularization provides a way to smoothly interpolate (using regularization constant $k$) between the sample covariance matrix and spherical covariance matrix. Within each fold, the optimal value for the regularization constant $k$ was chosen from 10 values covering the interval $[0.09, 0.5]$ based on the value with the highest performance in an inner 4-fold cross validation scheme that used training data only.

Classifier performance was evaluated based on the number of trials used in classification. Chance level in a binary classification problem is not exactly 50%, but 50% with a confidence interval for a given $p$-value depending on the number of trials. The Wald interval is a normal approximation of the binomial confidence interval. As we had a small number of trials, we calculated
Wald intervals with adjustments for a small sample size by adding four dummy observations, or two for each type (Agresti and Caffo, 2000; Müller-Putz GR, Scherer R, Brunner C, Leeb R, Pfurtscheller G, 2008). These intervals were then used to determine if the classifier performed significantly above chance or not.

**Location and frequency source analysis using PCA results**

To better understand the nature of the signal used in classification and to ensure that classification was not based on eye or muscle movements, the final weights (all 10 PC weights multiplied by their respective LDA weights) were visualized with respect to space (channel scalp topography) and frequency band (wavelet coefficients.) The topographies of the final weights were plotted at each frequency band as well as averaged over all frequency bands to pinpoint the channels that contributed most to classification. The absolute values of the final weights were plotted for each wavelet coefficient of each channel to visualize which region of the frequency spectrum was most informative for successful classification. Scalp topographies from each of the 10 PC components averaged over all 100-folds and all frequencies from the subject with the best performance were plotted in order of the PC with the highest contribution to LDA to the PC with the lowest to determine if PCs with higher eigenvalues contained the most pertinent information for classification.
Contribution from EOG and peripheral scalp channels

As an additional measure to assess the contribution of eye and muscle movements to the classification, LDA classification was performed using only EOG channels and those channels that would most be influenced by eye movements or muscle movements: Fp1, AF7, F7, FT7, T7, TP7, P9, P7, PO7, O1, Iz, Oz, O2, PO8, P8, P10, TP8, T8, FT8, F8, AF8, Fp2, and Fpz. These channels were located on the periphery of the EEG cap and were close to the origin of eye movements and muscles, and thus probably most strongly reflected the activity from these artifacts when compared to the other scalp channels.

Subjects with greater than chance performance based on the average of 10 runs of 10-fold CV from a specified LDA classification using this subset of channels were eliminated from the reported grand averages over subjects for that classification.

Error patterns

To determine if there was underlying spatial structure in the EEG signal, multiple discriminant analysis (MDA) using the one class vs. rest scheme was applied to the following classification problems: (1) left walk, front walk, right walk (three classes); (2) left point, front walk, right point (three classes); and (3) left walk, left point, right walk, right point (four classes). To visualize
classification performance, trials from all cross validation folds were grouped based on true class and predicted class. For example, in a three-class problem, the groups were arranged into a $3 \times 3$ confusion matrix with row labels corresponding to true classes and column labels corresponding to predicted classes. Cells were normalized by dividing by the total number of trials in the row class, so that the value of one cell was the fraction of trials that the classifier predicted to be in the class as defined by the column location of the cell.

2.4 Results

LDA classification

Individual subject errors and errors averaged over all nine subjects are shown in Tables 3a-3d for the 15 classification problems that consider direction as well as action. Mean and individual subject errors are shown in Table 4 for the three classification problems that collapse trials across directions and only consider differences in actions. Classification was successful if it yielded an error lower than the calculated threshold that was based on the number of trials used (Table 1).

Only 9/162 classifications had significant performance using EOG and peripheral channels: L walk/F walk and R point/stand in subject 6; L point/R point in subject 9; walk/stand in subjects 1, 2, 4, 7; and point/stand in subjects
1 and 4. Within these nine classifications, the performance of EOG and peripheral channels was either poorer (8/9) or similar (1/9) to that of all scalp channels.

Left tailed one-sample $t$-tests using subject errors as samples, $\alpha = 0.05$ and a mean set of 50% resulted in $p$-values below 0.05 for all 18 binary classifications. One-Way ANOVA testing suggested significant differences in means between the 15 classification problems that accounted for both direction and action ($p = 0.035$) and between the three classification problems that accounted for action only ($p = 0.019$). Post-hoc multiple comparisons test using Tukey's honestly different significance criterion both with and without assuming that the data followed a normal distribution revealed that only the mean error for walk vs. stand was significantly lower than the mean error for walk vs. point.
Table 2.3. Classification errors for different direction, same action. Above chance performances in the classification using only EOG+peripheral scalp channels are shown in parentheses next to the reported performances that used only scalp channels. These values were not included in the calculation of the mean, and are crossed out.

<table>
<thead>
<tr>
<th>Subject</th>
<th>L walk/ R walk</th>
<th>L walk/ F walk</th>
<th>F walk/ R walk</th>
<th>L point/ R point</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.07</td>
<td>0.09</td>
<td>0.14</td>
<td>0.15</td>
</tr>
<tr>
<td>S2</td>
<td>0.37</td>
<td>0.42</td>
<td>0.41</td>
<td>0.36</td>
</tr>
<tr>
<td>S3</td>
<td>0.12</td>
<td>0.08</td>
<td>0.23</td>
<td>0.20</td>
</tr>
<tr>
<td>S4</td>
<td>0.15</td>
<td>0.01</td>
<td>0.23</td>
<td>0.27</td>
</tr>
<tr>
<td>S5</td>
<td>0.06</td>
<td>0.03</td>
<td>0.20</td>
<td>0.12</td>
</tr>
<tr>
<td>S6</td>
<td>0.32</td>
<td>(0.37 (0.36))</td>
<td>0.46</td>
<td>0.41</td>
</tr>
<tr>
<td>S7</td>
<td>0.31</td>
<td>0.47</td>
<td>0.48</td>
<td>0.51</td>
</tr>
<tr>
<td>S8</td>
<td>0.07</td>
<td>0.08</td>
<td>0.14</td>
<td>0.11</td>
</tr>
<tr>
<td>S9</td>
<td>0.31</td>
<td>0.31</td>
<td>0.39</td>
<td>0.32 (0.35)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.20</td>
<td>0.19</td>
<td>0.30</td>
<td>0.27</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0005</td>
<td>0.0010</td>
<td>0.0011</td>
<td>0.0014</td>
</tr>
</tbody>
</table>

Table 2.4. Classification errors for same direction, different action. There were no above chance performances in classifications using only EOG and peripheral scalp channels.

<table>
<thead>
<tr>
<th>Subject</th>
<th>L walk/L point</th>
<th>R walk/R point</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.30</td>
<td>0.31</td>
</tr>
<tr>
<td>S2</td>
<td>0.41</td>
<td>0.43</td>
</tr>
<tr>
<td>S3</td>
<td>0.38</td>
<td>0.42</td>
</tr>
<tr>
<td>S4</td>
<td>0.24</td>
<td>0.26</td>
</tr>
<tr>
<td>S5</td>
<td>0.44</td>
<td>0.53</td>
</tr>
<tr>
<td>S6</td>
<td>0.41</td>
<td>0.39</td>
</tr>
<tr>
<td>S7</td>
<td>0.44</td>
<td>0.40</td>
</tr>
<tr>
<td>S8</td>
<td>0.39</td>
<td>0.36</td>
</tr>
<tr>
<td>S9</td>
<td>0.30</td>
<td>0.30</td>
</tr>
<tr>
<td>Mean</td>
<td>0.37</td>
<td>0.38</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0003</td>
<td>0.0010</td>
</tr>
</tbody>
</table>
Table 2.5. Classification errors for different direction, different action. There were no above chance performances in classifications using only EOG + peripheral channels.

<table>
<thead>
<tr>
<th>Subject</th>
<th>L walk/ R point</th>
<th>R walk/ L point</th>
<th>F walk/ R point</th>
<th>F walk/ L point</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.05</td>
<td>0.12</td>
<td>0.10</td>
<td>0.09</td>
</tr>
<tr>
<td>S2</td>
<td>0.39</td>
<td>0.35</td>
<td>0.39</td>
<td>0.33</td>
</tr>
<tr>
<td>S3</td>
<td>0.10</td>
<td>0.05</td>
<td>0.10</td>
<td>0.13</td>
</tr>
<tr>
<td>S4</td>
<td>0.03</td>
<td>0.05</td>
<td>0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>S5</td>
<td>0.13</td>
<td>0.04</td>
<td>0.09</td>
<td>0.06</td>
</tr>
<tr>
<td>S6</td>
<td>0.25</td>
<td>0.30</td>
<td>0.36</td>
<td>0.43</td>
</tr>
<tr>
<td>S7</td>
<td>0.45</td>
<td>0.34</td>
<td>0.41</td>
<td>0.47</td>
</tr>
<tr>
<td>S8</td>
<td>0.11</td>
<td>0.07</td>
<td>0.07</td>
<td>0.12</td>
</tr>
<tr>
<td>S9</td>
<td>0.24</td>
<td>0.32</td>
<td>0.29</td>
<td>0.34</td>
</tr>
<tr>
<td>Mean</td>
<td>0.20</td>
<td>0.17</td>
<td>0.20</td>
<td>0.23</td>
</tr>
<tr>
<td>ρ-value</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0016</td>
</tr>
</tbody>
</table>

Table 2.6. Classification errors for action vs. standing. Above chance performances in the classification using only EOG+peripheral scalp channels are shown in parentheses next to the reported performances that used only scalp channels. These values were not included in the calculation of the mean, and are crossed out.

<table>
<thead>
<tr>
<th>Subject</th>
<th>L walk/ stand</th>
<th>R walk/ stand</th>
<th>F walk/ stand</th>
<th>L point/ stand</th>
<th>R point/ stand</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.12</td>
<td>0.16</td>
<td>0.20</td>
<td>0.16</td>
<td>0.25</td>
</tr>
<tr>
<td>S2</td>
<td>0.38</td>
<td>0.33</td>
<td>0.42</td>
<td>0.46</td>
<td>0.48</td>
</tr>
<tr>
<td>S3</td>
<td>0.25</td>
<td>0.24</td>
<td>0.20</td>
<td>0.13</td>
<td>0.15</td>
</tr>
<tr>
<td>S4</td>
<td>0.19</td>
<td>0.36</td>
<td>0.33</td>
<td>0.10</td>
<td>0.07</td>
</tr>
<tr>
<td>S5</td>
<td>0.16</td>
<td>0.27</td>
<td>0.05</td>
<td>0.08</td>
<td>0.05</td>
</tr>
<tr>
<td>S6</td>
<td>0.27</td>
<td>0.49</td>
<td>0.37</td>
<td>0.36</td>
<td>0.35 (0.32)</td>
</tr>
<tr>
<td>S7</td>
<td>0.47</td>
<td>0.47</td>
<td>0.41</td>
<td>0.39</td>
<td>0.55</td>
</tr>
<tr>
<td>S8</td>
<td>0.10</td>
<td>0.32</td>
<td>0.17</td>
<td>0.10</td>
<td>0.16</td>
</tr>
<tr>
<td>S9</td>
<td>0.24</td>
<td>0.35</td>
<td>0.27</td>
<td>0.24</td>
<td>0.28</td>
</tr>
<tr>
<td>Mean</td>
<td>0.24</td>
<td>0.35</td>
<td>0.27</td>
<td>0.25</td>
<td>0.28</td>
</tr>
<tr>
<td>ρ-value</td>
<td>0.0001</td>
<td>0.0007</td>
<td>0.0003</td>
<td>0.0002</td>
<td>0.0030</td>
</tr>
</tbody>
</table>
Table 2.7. Classification errors for combined directions. Above chance performances in the classification using only EOG+peripheral scalp channels are shown in parentheses next to the reported performances that used only scalp channels. These values were not included in the calculation of the mean, and are crossed out.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Walk/stand</th>
<th>Point/stand</th>
<th>Walk/point</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.10 (0.38)</td>
<td>0.30 (0.39)</td>
<td>0.24</td>
</tr>
<tr>
<td>S2</td>
<td>0.34 (0.38)</td>
<td>0.41</td>
<td>0.40</td>
</tr>
<tr>
<td>S3</td>
<td>0.21</td>
<td>0.25</td>
<td>0.39</td>
</tr>
<tr>
<td>S4</td>
<td>0.19 (0.33)</td>
<td>0.22 (0.36)</td>
<td>0.24</td>
</tr>
<tr>
<td>S5</td>
<td>0.21</td>
<td>0.26</td>
<td>0.40</td>
</tr>
<tr>
<td>S6</td>
<td>0.35</td>
<td>0.41</td>
<td>0.40</td>
</tr>
<tr>
<td>S7</td>
<td>0.22 (0.34)</td>
<td>0.28</td>
<td>0.42</td>
</tr>
<tr>
<td>S8</td>
<td>0.18</td>
<td>0.17</td>
<td>0.28</td>
</tr>
<tr>
<td>S9</td>
<td>0.24</td>
<td>0.32</td>
<td>0.28</td>
</tr>
<tr>
<td>Mean</td>
<td>0.24</td>
<td>0.30</td>
<td>0.34</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0005</td>
<td>0.0005</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

Contributions from channels and frequency ranges

To pinpoint the origin and nature of the signal that the classifier was using, the contributions of different frequency ranges and channels were visualized by plotting the weightings by different PCs and the LDA classifier. The results for subject 4 (S4) and subject 8 (S8) in the two most difficult classifications, L walk/L point and R walk/R point, and in the three classifications collapsed across directions best provided insight into the classification process. Both subjects had significantly low errors for all classifications, but S4 probably had contributions from EOG or peripheral channels in the walk/stand and point/stand classifications while S8 had no such contributions. S8 had much higher error in the L walk/L point and R walk/R point classifications compared to S4.
The frequency bands with the highest weighting were visualized by plotting the final weight for each of the 358 coefficients (Figures 2, 3). The final weight was computed by multiplying the PC weights from all 10 PCs by their respective LDA weights. This was done using training data from all 10-folds and 10 runs. Additionally, scalp topographies of channel weights were plotted using only coefficients within a given frequency band (smaller scalp plots) and using all coefficients (large scalp plot). The former indicated which brain regions had informative activity within a given frequency band while the latter gave a sense of which channels were most crucial for the LDA classification. The plots show that coefficients corresponding to lower frequency bands were more highly weighted. The scalp map averaged over all coefficients indicated that channels Cpz, C4, and FC5 were most highly weighted in S4 whereas S8 only had one channel, C3, highly weighted. The scalp maps within different frequency bands showed very variable topographies.
Figure 2.2. L walk/ L point feature visualization for S4 and S8. The smaller scalp maps show topographies of final weights of channels within different frequency bands. The larger scalp map shows the weighted average of the final weights from all frequencies, and represents what is used by the classifier. Each line represents the absolute value of the final weights for all coefficients for an individual channel.
Figure 2.3. Similar to Figure 2.2, but with walk/stand, point/stand, and walk/point feature visualization for S4 and S8.

Using EOG and peripheral channels alone resulted in significantly low errors in the walk/stand and point/stand classifications in S4. Scalp topographies suggested weighting of the frontal electrodes in a pattern that suggested vertical eye movements, though this reliance on frontal electrodes was diminished in the walk/point classification. An example of possible neck muscle artifact can be seen in the scalp map for the 16–32 Hz frequency band of the R walk/R point classification. There were no scalp topographies with exclusively frontal or peripheral electrode weighting, which would suggest that eye or muscle movements were driving the classification.

The most weighted PCs did not necessarily have the highest eigenvalues (Figure 4). There were three distinct topographies with slight
variations represented by the 10 PCs, and the first five PCs summed to 75% of the total contribution to classification.

**Figure 2.4.** Scalp topographies of the weighted average of the contributions from all frequencies for all 10 PCs used in the L walk/L point classification in S4 are ordered from the PC with the highest LDA weight to that with the lowest LDA weight. Percentage contribution to LDA was calculated by normalizing LDA weights to one. PC1 has the highest eigenvalue and PC10 the lowest.

**Error patterns from MDA classification**

In the confusion matrix for the three-class problem of left walk/front walk/right walk (Figure 5a) most left and right trials were correctly classified, less likely to be classified one location away, and least likely to be classified two locations away. Most front trials were correctly classified, with similar misclassification rates as left or right. This same structure more or less exists in the confusion matrix for the three-class problem of left point/front walk/right point (Figure 5b). In the four-class problem of left walk/right walk/left point/right point (Figure 5c) once again trials were most likely to be correctly classified. Direction was a stronger factor than type of action for correct classification. For example, left walk trials were more likely to be misclassified as left point trials but less likely and at similar rates to be misclassified as right walk or right point trials.
Figure 2.5. (A) Confusion matrix for three-class MDA of left walk/front walk/right walk. Row labels indicate true classes and column labels indicate classifier labels. Color of cells reflect fraction of trials classified as trial type indicated by the column label, i.e., the second cell in the top row indicates the fraction of left trials classified as front trials. Larger fractions (lighter colors) represent higher number of correct classifications. (B) Confusion matrix for three-class MDA of left point/front walk/right point. (C) Confusion matrix for four-class MDA of left walk/right walk/left point/right point.
Classification performance of different features

**Figure 2.6.** Comparison of wavelet classification to classification using autoregressive coefficients and band power. Actions that were different but in the same direction were classified.
Figure 2.7. Comparison of wavelet classification to classification using autoregressive coefficients and band power. The same actions were classified for different directions.
Figure 2.8. Comparison of wavelet classification to classification using autoregressive coefficients and band power. Each walk and point action was classified against the action stand.
Figure 2.9. Comparison of wavelet classification to classification using autoregressive coefficients and band power. Different actions in different directions were classified from each other.
Figure 2.10. Comparison on wavelet classification to classification using autoregressive coefficients and band power. The directions were collapsed so that the walk class included walk left, walk front, and walk right and the class point included point left and point right.
Figure 2.11. Power over the 7-30 Hz frequency band of all channels for each trial is shown. Red indicates L point, black indicates R point. Circles are training trials, and squares are test trials. Test trials are colored to label their true class.
Figure 2.12. Sixteen AR coefficient values are plotted for each trial. Red indicates L point, black indicates R point. Circles are training trials, and squares are test trials. Test trials are colored to label their true class.
Figure 2.13. All wavelet coefficient values of Channel Cz are plotted for each trial. Red indicates L point, black indicates R point. Circles are training trials, and squares are test trials. Test trials are colored to label their true class.
Figure 2.14. Final LDA weights are plotted for each trial. Red indicates L point, black indicates R point. Circles are training trials, and squares are test trials. Test trials are colored to label their true class.
Figure 2.15. Final LDA weights are plotted for each trial. Red indicates L point, black indicates R point. Circles are training trials, and squares are test trials. Test trials are colored to label their true class.
Figure 2.16. Final LDA weights are plotted for each trial. Red indicates L point, black indicates R point. Circles are training trials, and squares are test trials. Test trials are colored to label their true class.
Figure 2.17. Final LDA weights are plotted for each bandpower value for all channels and trials. Red indicates L point, black indicates R point. Circles are training trials, and squares are test trials. Test trials are colored to label their true class.
Figure 2.18. Final LDA weights are plotted for each AR coefficient for all channels and trials. Red indicates L point, black indicates R point. Circles are training trials, and squares are test trials. Test trials are colored to label their true class.
Figure 2.19. Final LDA weights of each wavelet coefficient for channel Cz in all trials. Red indicates L point, black indicates R point. Circles are training trials, and squares are test trials. Test trials are colored to label their true class.
Figure 2.20. Final LDA weights vs. weights of each principal component for all trials in the classification using bandpower over 7-30 Hz as features.
Figure 2.21. Final LDA weights vs. weights of each principal component for all trials in the classification using AR coefficients as features.
Figure 2.22. Final LDA weights vs. weights of each principal component for all trials in the classification using wavelet coefficients as features.
Classification performance of different features

In addition to wavelets, autoregressive coefficients and bandpower were used as features for classification. Autoregressive coefficients were calculated using the Burg method, and the model order was set to 16 so that the result was a feature space of approximately 16 coefficients x 64 channels x 100 training trials where the number of channels and training trials varied between subjects depending on artifact rejection. Bandpower was calculated over the 7-30 Hz frequency range, resulting in a feature space of approximately 1 power value x 64 channels x 100 training trials.

Classification performance for actions specified by direction was best when wavelets were used as features (Figure 2.6A-D); in fact, using AR and bandpower features yielded chance or near chance classification performance. However, when actions were collapsed over direction, AR features and bandpower features resulted in below chance performance (Figure 2.6E). One explanation for this is that AR and bandpower feature spaces require more trials to sufficiently train the classifier; another is that AR and bandpower feature spaces are not very good at discriminating between movements of different directions.

To examine how the feature space is transformed before classification, plots of feature value versus trial (Figure 2.7), LDA weight versus trial (Figure 2.8), LDA weight versus feature value (Figure 2.9), and LDA weight versus PC weight (Figure 2.9) were plotted for L point/R point and walk/stand
classifications in Subject 4. The classifier performance for this subject in L point/ R point was 48% using AR coefficients, 50% using bandpower, and 1% using wavelets. For walk/stand, the classifier performance was 22% using AR coefficients, 26% using bandpower, and 28% using wavelets.

All of the original feature spaces, regardless of feature type, show no clear separation between classes (Figure 2.7). Once the dimensionality of the features is reduced by PCA and LDA, bandpower and AR coefficients result in poor separation of classes in the L point/R point classification, but better separation in the walk/stand classification (Figure 2.8A-B). Wavelet coefficients result in near perfect separation in the L point/R point classification, and similar separation to the other features in the walk/stand classification (Figure 2.8C). This same pattern can also be seen when the final LDA weight for each trial is assigned to the features originally associated with that trial (Figure 2.9A-C). Finally, plots of LDA weights versus PC weights separated by PC with PC1 having the highest eigenvalue (most informative) and PC10 with the lowest eigenvalue (least informative) show good separation of classes using wavelet features (Figure 2.10C), but not with bandpower or AR features (Figure 2.10A-B). There is also not much variability in feature space and separation between the different PCs when using wavelet features, but considerable variability when using bandpower and AR features.
2.5 Discussion

This study demonstrated that single trial EEG data is (1) classifiable for walk intent before the onset of natural movement, (2) classifiable between two motor plans (walking and pointing) that activate overlapping muscles, and (3) classifiable for an action at different target spatial locations. The largest contributors to successful classification were low frequencies (0–4 Hz) and channels located over areas involved in motor planning or motor production.

It is important to note that “movement intent” as used in this paper refers to the preparation of a movement by the subject in response to an external experimental cue. Further, movement was elicited by a warning stimulus (S1) that was followed by an imperative stimulus (S2). The resultant MRP from this paradigm is the CNV, which contains contributions from both motor preparation and attention for the upcoming stimulus (Luck and Kappenman, 2011). The motor aspect of the CNV is often equated to the late phase of the BP (Bereitschaftspotential), which is another slow cortical DC potential that is an indicator of human voluntary activity from as early as 2 s (the “early BP”) to 400 ms (the “late BP”) before the onset of movement (Shibasaki and Hallet, 2006). While the CNV results from external regularly timed cued movement, the BP results from internally self-paced movement. The beginning of the early BP starts in the pre-supplementary motor area (pre-SMA) and lacks somatotopic organization, but the rest of the early BP and the
entire late BP have generators with discrete spatial locations corresponding to body parts. Application of this classifier paradigm to subject driven gait intent may show better classification as a result of the greater specificity of the BP compared to the CNV (Jankelowtiz and Colebatch, 2002; Lew E, Chavarríaga R, Silvoni S, Millán J, 2012).

A large challenge of using MRPs is finding informative activity within single trials since the low signal to noise ratio of a single trial presents a considerable obstacle in feature detection. Though the majority of MRPs in EEG literature are presented as ERPs averaged over hundreds of trials in the time-voltage domain, efforts are being made to capture the MRP within a single trial. One group succeeded in single trial detection of the MRP using features from wavelet analysis for predicting foot torque movement while subjects were seated (Farina D, do Nascimento OF, Lucas MF, Doncarli C, 2007).

Another feature commonly used in BCIs is the event related desynchronization (ERD)/event related synchronization (ERS), or the decrease/increase in power of mu (8–13 Hz) or beta (14–25 Hz) rhythms at somatotopically distinct regions before and during motor planning, action or imagery (McFarland DJ, Miner LA, Vaughan TM, Wolpaw JR, 2000; Pineda JA, Allison BZ, Vankov A, 2000). In select subjects, it is visually observable during single trials and has been used to detect hand, foot, or tongue motor

In this study we used all frequency information available in the signal by including coefficients from the entire wavelet decomposition. Previously mentioned fMRI, PET, and NIRS studies on gait found task-relevant activity in the supplementary motor cortex (SMA), medial primary motor cortex, and medial sensorimotor cortices. Cortical folding is different between individuals, which results in highly variable topographic distributions of useful EEG signal within a group of subjects. We created unsupervised custom spatial filters for each subject by including all noise-free channels available from the 64-channel montage covering the scalp and then applying PCA to transform and reduce this large feature space to the 10 dimensions with the most variance in the signal.
The variable ranking of PCs by LDA weights was in line with a study of single trial EEG classification using PCA-reduced feature space and LDA for an imagined left-and-right hand movement task (Lugger K, Flotzinger D, Schloegl A, Pregenzer M, Pfurtscheller G, 1998). Though the first few PCA components had the greatest variance, they were poor in discriminating between left vs. right hand motor imagery. These components were thought to represent background cortical EEG activity that was unrelated to the task.

Visualizations of the features used by the classifier revealed that frequency components within the range of the MRP were most heavily weighted. The channels most weighted in S4 may have reflected activity from the PMd (FC5), leg motor area (CPz), and arm motor area (C4), which could explain the lower errors for the L walk/L point and R walk/ R point classifications compared with S8, who only had high weighting of channel C3. Notably, the classification error for walk/point was similar for both S4 and S8, possibly because more trials were available for training as the classification used trials from both left and right conditions.

The dorsal premotor cortex (PMd) is postulated to be activated in externally cued movement preparation vs. movement that is internally driven by the subject, and vice versa in the SMA. Evidence for this has been circumstantial in humans, though one group demonstrated double-dissociation of PMd and SMA activity on MRPs by repetitive transcranial magnetic
stimulation (rTMS) in humans during a right digit task in externally cued vs. internally driven conditions (Lu MK, Arai N, Tsai CH, Ziemann U, 2011). Most relevant to this EEG study is an experiment that demonstrated that EEG surface potential configurations had the same order and same strength but longer duration over the PMA compared to the SMA in externally cued vs. internally driven right digit movement (Thut G, Hauert CA, Viviani P, Morand S, Spinelli L, Blanke O, Landis T, Michel C, 2000).

BCIs also must minimize or eliminate contributions from non-cortical electrical activity, such as from EOG and EMG. These signals not only obscure the cortical signal but can misdirect classification as in a situation when a subject is looking at a target that is not located in the intended direction of gait. Topographies suggestive of eye and muscle movements did appear in some scalp maps but were not the only feature present in the maps or had smaller weights. Classification using EOG and peripheral scalp channels resulted in significantly low errors in only 6% of all classifications, and these error values were poorer or comparable to those from classifications using only scalp channels. The classifiers appear to have minimized the contribution of eye or muscle movement on the scalp channel cortical signals for correct classification of movement intent.

Another challenge was in differentiating between two different motor plans to ensure that discrimination was based on a signal unique to gait
movement intent and not generalized to any movement intent. Further, this was not a straightforward classification of an upper limb activity from a lower limb activity. Gait requires not only the participation of lower limb muscles, but trunk and upper limb muscles as well. In EMG studies, the posterior deltoid muscle was consistently activated in all 35 subjects in a walking task (Barthelemy and Nielsen, 2010), and contributed to horizontal abduction, external rotation, and depression at the shoulder as well as flexion and supination at the elbow during arm reaching (Vandenberghne A, Bosmans L, De Schutter J, Swinnen S, Jonkers I, 2012). Though regions of the cortex involved in planning movements of these muscles have specific somatopically distribution, they are contiguously located on the sensorimotor cortex (Bakker M, Verstappen CCP, Bloem BR, Toni I, 2007) and may not be easily separable by the poor spatial resolution of EEG. This was probably the biggest factor in the higher errors for the walk vs. point classification problems compared to point vs. stand and walk vs. stand classification problems, though it is notable that the errors were still significantly below chance. With more training trials and the use of a non-linear classifier such as kernel support vector machines, these errors may be further lowered.

Finally, we tested the ability to predict the direction of movement toward a target either to the left, in front, or right of the subject. For all three targets, subjects initiated walking with the same leg and same arm, so classification
was not based on contralateral cortical activation of the limb used. Previously, reaching to three different (left, right, center) locations using the same arm revealed an underlying structure in the pre-movement EEG signal that corresponded with target space (Hammon PS, Makeig S, Poizner H, Todorov E, de Sa VR, 2008). Trials were most likely to be correctly classified, less likely to be classified one target away, and least likely to be classified two targets away. We found a similar underlying structure that corresponded to the target space in our data.

We should note that in an online real-time BCI, the classifier would be trained on initial data and then be tested on future data. This means it must be robust to non-stationarity in the signals that can cause drifts in the features over time. In this experiment there were not enough trials to create a large enough training set for adequate training and a separate (later) test set with an adequate number of trials to reliably assess the classifier accuracy, so we were not able to test this aspect.

In conclusion, this study has shown that the EEG signal can be used for predictive classification between walk, point and stand actions as well between different target directions for these actions. Spatial and spectral contributions were from areas involved in motor planning or production and mostly from low frequency cortical activity, with smaller contributions from mu and beta
frequency bands. It remains to be tested whether this signal can be detected in real-time non-stationary data.

2.6 Acknowledgements

Chapter 2 includes material as it appears in Velu PD, Mullen T, Noh E, Valdivia MC, Baram Y, Poizner H, de Sa VR, “Effect of visual feedback on the occipital-parietal-motor network in Parkinson’s disease patients with freezing of gait,” *Front Neurol.* 4: 209, 2013. The dissertation author was the primary author of this publication. The copyright for this publication is owned by Frontiers and used in accordance with their terms and conditions.
2.7 References


do Nascimento OF, Farina D, 2006. Classification of movement-related cortical potentials according to variations of rate of torque development during imaginary isometric plantar-flexion. Proceedings of the 3rd International Brain-Computer Interface Workshop and Training Course Graz,
Austria 70-71.


Friedman JH, 1989. **Regularized Discriminant Analysis.** *JASA* 84(405), 165-175.

Graybiel AM, 2005. **The basal ganglia: learning new tricks and loving it.** *Curr. Opin. in Neurobio.* 15:638-44.


Hammon PS, Makeig S, Poizner H, Todorov E, de Sa VR, 2008. **Predicting reach targets from human EEG.** *IEEE Sig Processing Mag* 70:69-77.


Chapter 3

Effect of visual feedback on the occipital-parietal-motor network in Parkinson’s disease patients with freezing of gait

3.1 Abstract

Freezing of gait (FOG) is an elusive phenomenon that debilitates a large number of Parkinson’s disease (PD) patients regardless of stage of disease, medication status, or deep brain stimulation implantation. Sensory feedback cues, especially visual feedback cues, have been shown to alleviate FOG episodes or even prevent episodes from occurring. Here, we examine cortical information flow between occipital, parietal, and motor areas during the pre-movement stage of gait in a PD-with-FOG patient that had a strong positive behavioral response to visual cues, one PD-with-FOG patient without any behavioral response to visual cues, and age-matched healthy controls, before and after training with visual feedback. Results for this case study show differences in cortical information flow between the responding PD-with-FOG patient and the other two subject types, notably, an increased information flow in the beta range. Tentatively suggesting the formation of an alternative
cortical sensory-motor pathway during training with visual feedback, these results are proposed as subject for further verification employing larger cohorts of patients.

3.2 Introduction

Freezing of gait (FOG) is a debilitating phenomenon in a subset of patients with Parkinson’s disease (PD). FOG occurs in 53% of PD patients who are in advanced stages of disease but can occur even in early stages. It consists of episodic periods of motor inability in which patients feel like they are frozen or that their feet are glued to the floor, rendering them unable to move their legs. FOG can happen during initiation of walking or while reaching a destination, upon turning or walking through narrow passages such as doorways or when entering an elevator. Patients also report FOG when feeling stressed, anxious or pressed for time, or when presented with multiple sensory inputs such as when crossing a busy intersection. The freezing episodes usually last a few seconds to a minute, though longer durations are not uncommon. (Nutt JG, Bloem BR, Giladi N, Hallet M, Horak FB, Nieuwboer A, 2011; Browner and Giladi, 2010; Okuma and Yanagisawa, 2008; Okuma 2006)

Behavioral studies have shown that cadence of gait increases and stride length decreases before a freezing episode (Nieuwboer A, Dom R, De
Weerdt W, Desloover K, Fieuws S, Broens-Kaucsik E, 2001). The timing and activation of the tibialis anterior (TA) and gastrocnemius muscles of the lower leg involved in the starting and swinging phases of gait are abnormally timed and activated (Nieuwboer A, Dom R, De Weerdt W, Desloover K, Janssens L, Stijn V, 2004). Research since 1966 on FOG has been generalized to five different hypotheses, of which two focus on more downstream dysfunction that can be measured through behavioral studies, such as abnormal gait pattern generation and abnormal coupling of posture with gait. The other three suggest that dysfunction is primarily due to more upstream causes such as conflict between central drive and automaticity of movement, perceptual malfunction, and frontal executive dysfunction, and benefit from insight from measured brain activity (Nutt JG, Bloem BR, Giladi N, Hallet M, Horak FB, Nieuwboer A, 2011; Heremans E, Nieuwboer A, Vercruysse S, 2013).

Earth-stationary visual feedback has been shown to improve various gait parameters, help reduce freezing episodes, and eliminate or reduce the eventuality of freezing in patients with PD (Baram Y, Aharon-Peretz J, Simionovici Y, Ron L, 2002). Even simple visual cues have been shown to stop FOG episodes. By using lines drawn on the floor in the form of railroad tracks, the light from a laser pointer, or the bottom of a walking cane as visual targets, patients resumed their intended movement (Jiang Y, Norman KE, 2006). A more dynamic approach to providing visual cues was demonstrated
by connecting visual feedback to the user’s movements by an accelerometer attached to the hip. The visual cue was in the form of a black and white checkerboard pattern that moved in concert with the user’s walking speed, forming a closed loop between the visual stimulus and user’s musculoskeletal system (Baram Y, Aharon-Peretz J, Simionovici Y, Ron L, 2002). Studies also showed that mitigation of FOG symptoms persisted after visual targets were used (Espay AJ, Baram Y, Dwivedi AK, Shukla R, Gartner M, Gaines L, Duker AP, Revilla FJ, 2010; Griffin HJ, Greenlaw R, Limousin P, Bhatia K, Quinn NP, Jahanshad M., 2011).

The structural and functional neuroanatomical properties of FOG have been studied by using voxel-based morphometry (VBM), single-photon emission computed tomography (SPECT) and functional magnetic resonance imaging (fMRI). VBM showed grey matter atrophy in the frontal and parietal cortices, specifically in the left cuneus, precuneus, lingual gyrus, and posterior cingulate cortex in PD FOG compared to PD without FOG and controls, with clinical severity of FOG correlated significantly with gray matter loss in posterior cortical regions (Tessitore et al. 2012a). SPECT showed decreased brain perfusion in prefrontal, orbitofrontal, and anterior cingulate regions in PD FOG compared with PD without FOG (Imamura K, Okayasu N, Nagatsu T, 2012). In a virtual-reality walking task in PD FOGs, fMRI showed decreased blood oxygen level-dependent response in sensorimotor regions and an
increased response in frontoparietal cortical regions (Shine JM, Matar E, Ward PB, Bolitho SJ, Gllat M, Pearson M, Naismith SL, Lewis SJ, 2013). Another fMRI study used motor imagery of gait for the task and found reduced activity in the superior parietal lobule and the anterior cingulate cortex in both patient groups compared to controls, and a statistical trend toward increased activity in the left supplementary motor cortex and R superior parietal lobe in PD without FOG compared to PD with FOG with ROI analysis (Snijders A, Leunissen I, Bakker M, Overeem S, Hemlich RC, Bloem BR, Toni I, 2011). Whole brain analysis in the same study revealed increased task related activity in the posterior mid-mesencephalon of PD FOG compared to PD without FOG. However, it is important to note that these studies required a stationary head during imaging and so provided an indirect look at neural activity of FOG by using motor imagery and stepping in virtual reality environments as proxies for actual gait tasks.

Patients with FOG also exhibited differences in cortical activity during rest when compared to controls. An fMRI study showed decreased functional connectivity in the resting state within a network consisting of the right middle frontal gyrus and angular gyrus and a network consisting of the right-occipito-temporal gyrus in PD FOG compared to PD without FOG (Tessitore A, Amboni M, Esposito F, Russo A, Picollo M, Marcuccio L, Pellecchia MT, Vitale C, Cirillo M, Tedeschi G, Barone P, 2012). These areas are regarded as
executive-attention and visual networks, respectively, with the executive-attention network recognized as one of the cognitive resting state networks.

A gap in the literature remains regarding how visual cues change neural responses to overcome or prevent FOG. By using EEG, we were able to examine time varying directed connectivity as opposed to relative levels of activation of different anatomical regions, which allowed for a dynamic picture of effectors and their targets with millisecond resolution. We hypothesized that visual cues result in an increase in information flow from visual and parietal areas to the motor cortex in the pre-movement time period, and that this effect has residual staying power. We used time-series data from EEG recordings of controls and PD FOGs in a task that required them to turn and walk through a doorway. Analysis was performed on the pre-movement period with the idea that the dysfunction in FOG that occurs in the preparation period before any movement is crucial towards understanding the phenomenon. This allowed us to avoid any confounding effects of cerebral activity due to motor activity, performance, or sensory feedback. The portable EEG also permitted ambulation so that the preparation period reflected planning for actual gait, and was not restricted to motor imagery as in fMRI studies since a concern in interpreting motor imagery data is that PD patients are not as capable as their age matched healthy peers in estimating walking during motor imagery tasks (Cohen RG, Chao A, Nutt JG, Horak FB, 2011).
3.3 Materials and Methods

Subjects

Two subjects with PD-with-FOG, one responding (PDr) and one non-responding (PDnr) to visual feedback, and six age-matched healthy individuals (Control, ages 57–75, 1 female) were analyzed. Both PD patients took medications as scheduled so as to remain in the “ON” state throughout the experiment. All subjects read and signed informed consent forms that were approved by the UCSD Human Research Protections Office. Clinical characteristics of both PD patients are specified in Table 1.
Table 3.1. Characteristics of PD subjects.

<table>
<thead>
<tr>
<th>Subject</th>
<th>UPDRS III</th>
<th>H&amp;Y</th>
<th>FOGQ</th>
<th>Duration</th>
<th>DBS</th>
<th>Age/Sex</th>
<th>Medications</th>
</tr>
</thead>
<tbody>
<tr>
<td>TB (PDv)</td>
<td>50</td>
<td>3</td>
<td>8</td>
<td>16</td>
<td>-</td>
<td>69/F</td>
<td>Lev, LevR, Pr, Am, Ras</td>
</tr>
<tr>
<td>DO (PDn)</td>
<td>40</td>
<td>3</td>
<td>10</td>
<td>8</td>
<td>-</td>
<td>48/M</td>
<td>Lev, RopXL, Ras</td>
</tr>
<tr>
<td>JF (PDn)</td>
<td>23</td>
<td>3</td>
<td>11</td>
<td>5</td>
<td>-</td>
<td>64/M</td>
<td></td>
</tr>
<tr>
<td>CR* (PDn)</td>
<td>38</td>
<td>3</td>
<td>8</td>
<td>12</td>
<td>1y</td>
<td>56/F</td>
<td></td>
</tr>
<tr>
<td>JA^* (PDn)</td>
<td>n/a</td>
<td>n/a</td>
<td>20</td>
<td>19</td>
<td>2y</td>
<td>81/M</td>
<td></td>
</tr>
<tr>
<td>DV^ (PDn)</td>
<td>48</td>
<td>4</td>
<td>21</td>
<td>9</td>
<td>-</td>
<td>88/M</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>Average without JA, DV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>59</td>
<td></td>
</tr>
</tbody>
</table>

*DBS patient with harmonics at 30 Hz, removed from EEG analysis
^Unable to complete task, removed from behavioral analysis

Scores on the MMSE and BDI were available for PDnr 5 months prior to testing, and for PDr, 2 years prior to testing.


FOGQ: Freezing of Gait Questionnaire (Giladi et al., 2000). Raw score range 0-24. Higher scores indicate greater impairment.

Duration is years since first remembered parkinsonian symptom.

DBS indicates amount of time since implantation of deep brain stimulation device.

Medication codes:
LevR – Carbidopa/levodopa sustained release
Lev – Carbidopa/levodopa (regular formulation)
Pr – Pramipexole
Rop – Ropinirole
RopXL – Ropinirole extended release
St – Stalevo (Carbidopa/levodopa/entacapone)
Ras - Rasagiline
Am - Amantadine

Table 3.2. Characteristics of control subjects.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age/Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>TM</td>
<td>60/M</td>
</tr>
<tr>
<td>JD</td>
<td>71/F</td>
</tr>
<tr>
<td>MM</td>
<td>57/M</td>
</tr>
<tr>
<td>DG</td>
<td>69/M</td>
</tr>
<tr>
<td>TG</td>
<td>75/M</td>
</tr>
<tr>
<td>PR</td>
<td>70/M</td>
</tr>
<tr>
<td>Average</td>
<td>67</td>
</tr>
</tbody>
</table>
**Visual feedback**

A cellphone-size belt-mounted box containing inertial sensors and a microprocessor generated an earth-stationary visual cue in the form of checkerboard tiles that were displayed by VR glasses and moved in accordance with the patient's own motion. A lens, centrally embedded between two non-transparent stereo micro-displays, provided a see-through capability for patient safety.

**Task**

The task consisted of five blocks. In Blocks A, C, and E, the subject walked on a set path that consisted of three maneuvers: 1) start at a designated start spot and walk forward; 2) turn either left or right to approach a doorway; and 3) enter and pass through doorway to a designated end spot. In Blocks B and D, the subject was seated and asked to remain still with eyes open for five minutes in order to record resting state EEG. Subjects wore the VR glasses throughout the experiment, but visual stimuli were only shown in Block C. The visual stimulus consisted of a moving black and white checkerboard pattern presented with a refresh rate of 60 Hz in the upper half portion of a display built into glasses. The glasses were synced with an accelerometer that was attached to the subject’s hip on a belt or the waistband on clothing so that the pattern moved in response to the subject’s speed.

A single trial started with an auditory command, “stand still”, to prepare
subjects for a “beep” noise that indicated that the subject could begin walking (go cue). The interval between the preparation cue and the go cue was randomly chosen from three possible values: 1250ms, 1000ms, or 750ms. Trials in which subjects started walking before the go cue sounded or that involved any freezing were eliminated from analysis. Each block lasted until 30 properly performed trials were collected. Subjects could request breaks during a block in which they could sit down, though only patients TB, DV, and JA and no controls took advantage of this option.

Subjects wore the EEG cap and electrodes during the entire task. A spotter followed all subjects with a backpack containing the EEG amplifier and battery pack, and from this, data were transmitted through a 40 ft long fiber optical cable to the recording computer. An additional spotter assisted in experiments with patients for fall prevention.
Figure 3.1. Subjects had to walk from a starting point, turn, then go through a constructed doorway to finish at a designated endpoint.

Electrophysiological recording

Data collection

Continuous EEG and EOG/mastoid/EMG (EXG) was recorded from 64 Ag/AgCl scalp electrodes positioned on a BioSemi nylon head cap according to the 10-20 International System and 8 EXG electrodes placed on the surface of cleaned skin. The signal was amplified with fixed gain BioSemi ActiveTwo amplifiers, band-passed from 0.1-100 Hz, and digitized at 512 Hz with 24-bit resolution. The independent software package DataRiver was used to read and record EEG signals as well as to integrate EEG signals with events from the Stim2007 stimulus presentation software (Vankov A, Bigdely-Shamlo N,
Two EOG electrodes were placed to record eye movements (one on the right outer canthus and one above the right eye). Right and left mastoid electrodes were averaged off-line to serve as reference. To minimize movement artifacts, subjects were encouraged to remain still and look forward until the cue to move was heard. Two electrodes were placed on the anterior tibialis muscle (the first muscle to activate in gait) on each leg to detect premature muscle contraction during trials (Mann RA, Hagy JL, White V, Liddell D, 1979).

**Pre-processing and artifact rejection**

Pre-processing utilized various functions from the EEGLAB software package (Delorme A, Mullen T, Kothe C, Zeynep A, Bigdely-Shamlo N, Vankov A, Makeig S, 2011). Data were referenced to L and R mastoid electrodes and band pass filtered from 1 Hz to 50 Hz. Data were split into epochs starting four seconds before and ending four seconds after the go cue with -2.75 s to -2.25 s before the go cue used as a baseline. Trials in which leg EMG indicated movement before the go cue sounded and trials that had excessive noise by visual inspection were eliminated. Channels were also visually inspected for noise and removed. These data were then further cleaned using EEGLAB automatic artifact rejection functions that removed channels and epochs that had kurtosis measures 5 standard deviations from
the mean kurtosis value (Delorme and Makeig, 2004). Kurtosis is the fourth moment measure of a probability distribution, and large positive kurtosis values indicate increased peaky shape whereas large negative kurtosis values indicate abnormally flat shape in the distribution. In EEG, these may represent undesirable artifacts in the data.

To remove eye and electronic artifacts, Independent Component Analysis (ICA) was performed for each subject on the pre-processed and cleaned data from all three blocks concatenated to form one dataset. The resulting independent components (ICs) were then analyzed by an automatic algorithm from the ADJUST plug-in for EEGLAB. This algorithm identified ICs for elimination by looking for stereotyped spatial and temporal features present in eye blinks, eye movements, and generic discontinuities such as impedance fluctuations or electronic device interference.

**Behavioral recording**

**Data collection**

The amount of time the subject took to reach the doorway from the starting position and that the subject took to pass through the doorway in order to stop at the designated end spot were recorded by an observer with a stopwatch. The corresponding number of steps taken by the subject at these
timed portions were counted and reported by the spotter holding the backpack.

Data analysis

Analysis of variance (ANOVA) was performed for average steps to door, steps to finish, time to door, and time to finish on the different blocks and subject groups.

Pre-movement EEG analysis using SIFT

EEG analysis was focused on spectral power and connectivity in a network composed of occipital (Oz), parietal (P4), and motor (Cz) channels. These channels were chosen based on anatomical findings from prior fMRI and PET experiments on paradoxical gait in PD (13, 21). Power spectra and connectivity measures were obtained using a multi-trial sliding window adaptive multivariate vector autoregressive (AMVAR) modeling approach applied to the non-stationary channel time series.

Pre-processing

For optimal model fitting, data went through several pre-processing steps. They were first down-sampled to 128 Hz from 512 Hz in order to minimize the number of coefficients required to adequately fit a model, as increased model order leads to increased variability of spectral and causal
estimates (22). Then, piecewise linear de-trending was applied using a least squares fit to eliminate remaining drift in the data. Finally, data were normalized by point-wise subtraction of the ensemble mean and division by ensemble standard deviation over all trials and the subtraction of the temporal mean and division by the temporal standard deviation for each trial.

Model fitting

The Vieira-Morf algorithm performs better than Arfit and Levinson algorithms for small sample sizes (Schlögel, 2006) and was used for all AMVAR fitting in this analysis. The window length was set to 1s and the step size 0.01s. Model order was selected by minimizing Hannan-Quinn criterion, which optimizes a tradeoff between the prediction error of the model and the number of freely estimated parameters in the model (Lütkepohl, 2006).

Model validation

The AMVAR model was validated using whiteness, consistency, and stability and stationarity measures. Whiteness measures the amount of correlation left in the residuals of the model; an ideal fit has no correlation structure left in the residuals. A portmanteau test for whiteness, specifically the Li-Mcleod which is considered the most conservative, was used to determine if residuals were white (Lütkepohl, 2006). Consistency of the model was
determined by generating an ensemble of simulated data of equal dimensionality as the original data using the AMVAR model and calculating the auto and cross correlations between all variables to determine if the generated data has a similar correlation structure to the real data to at least 85% (Ding MZ, Bressler SL, Yang WM, Liang HL, 2000) similar correlation structure to the real data. Stability implies stationarity, and stability in the M-dimensional AMVAR models with order $p$ was checked by ensuring that the eigenvalues of the $(M_p \times M_p)$ augmented coefficient matrix had moduli less than 1 (Lütkepohl, 2006). This was done by using a stability index based on the log of the largest eigenvalue of the coefficient matrix.

*Information flow analysis*

Connectivity between nodes in networks can be structural, functional, and effective in nature (Bullmore and Sporns, 2009). Here we attempted to map effective connections, or causal interactions, between different areas of the brain by calculating directed coherence. There are several causal estimators that can be derived from the AMVAR model coefficients. Here we use renormalized partial directed coherence (rPDC) as it provides a scale-free estimator, avoids arbitrary normalization by inflow or outflow, and provides a constant (frequency-independent) point-wise significance threshold (Schelter B, Timmer J, Micahel E, 2009).
Statistics

We made no assumptions about the probability distributions of the spectra and connectivity measures and applied rigorous non-parametric surrogate statistical methods to calculate the significance of these estimators. Bootstrap statistics with 1000 resamples was performed on all datasets, and confidence intervals were calculated at 95%. Multiple comparisons were accounted for using the false discovery rate control.

Nonparametric significance thresholds on between-condition differences in power and rPDC were obtained using an Efron bootstrap approach. In brief, for each T-trial dataset, a surrogate dataset was constructed containing T trials randomly sampled with replacement from all trials. The surrogate dataset was then subjected to the aforementioned pre-processing and modeling procedure. This procedure was repeated 750 times yielding empirical distributions of power and rPDC for each time window and frequency bin. For each time-frequency “pixel,” a pointwise two-sided empirical p-value for rejecting the null hypothesis of equal power or rPDC between any two conditions was then obtained by computing the quantile at which zero occurs in the between-condition distribution of surrogate differences. Finally, pointwise significance estimates were corrected for multiple comparisons across time, frequency, and channel (pair) using the Benjamini-Hochberg False Discovery Rate procedure and thresholded at $p = 0.05$. 
3.4 Results

Behavioral measures

Behavioral measures showed marked decreases in the time and number of steps taken to reach and exit the doorway in stage C in the patient that responded to visual feedback (PDvr) compared to control subjects and patients that did not respond to visual feedback (PDn) (Figure 1b,c). In agreement with previous studies (Espay AJ, Baram Y, Dwivedi AK, Shukla R, Gartner M, Gaines L, Duker AP, Revilla FJ, 2010; Griffin HJ, Greenlaw R, Limousin P, Bhatia K, Quinn NP, Jahanshahi M., 2011), PDvr retained these behavioral effects in stage E suggesting that there are residual effects from prior feedback. Notably, PDvr took double the time to reach the endpoint compared to the time to reach the door in block A. This normalized to control and PDn levels by block C. Only PDvr subject TB reported a subjective feeling of improvement in gait in response to the visual cues.
Figure 3.2. Steps were counted from the designated start point until the subject entered the doorway (Number of steps to door) and from the doorway until the designated end point was reached (Number of steps to finish). PDn includes subjects DO, JF, and CR.
Figure 3.3. Time in seconds was counted from the designated start point until the subject entered the doorway (Time to door) and from the doorway until the designated end point was reached (Time to finish). PDn includes subjects DO, JF, and CR.
Figure 3.4. Histogram showing number of steps taken to reach the doorway (Steps to door) and to reach the endpoint from the doorway (Steps to finish) for PDvr subject TB.
Neural activity

EEG spectra and RPDC connectivity as a function of frequency are shown for control subjects in blocks A, C, and E (Figure 3.6) and PD subjects in blocks A, C, and E (Figure 3.7). Similar plots are also shown depicting the difference between blocks A and C in controls (Figure 3.8A) and PD patients (Figure 3.9A) and the difference between blocks A and E in controls (Figure 3.8B) and PD patients (Figure 3.9B). Spectra for each channel are on the diagonal of the three by three matrix, and the RPDC between channels are on the upper and lower triangles of the matrix. The originating channels are shown on the first row and the receiving channels are shown on the first column so that the cell in row 2, column 1 shows the RPDC from channel Oz to channel Pz.

In block A, controls showed a general pattern of peaks in alpha (8-12Hz), beta (12-30 Hz), and/or gamma (30-50 Hz) frequency ranges in channel spectra. RPDC measures ranged from little or no activity to peaks in alpha and beta. Controls JD, DG, and PR showed no significant change in spectra or RPDC between blocks A and C. The other three subjects showed an increase in theta (4-7 Hz) and alpha, increase or decrease in beta, and/or a decrease in gamma in channel spectra. Two controls showed decreases in the theta/alpha range in RPDC between channels Oz to Cz and/or channels Oz to Pz, and one control showed decreased connectivity in the beta range from channel Pz to Cz.
PD subjects PDvr TB, PDnr DO, PDnr JF, and PD patient DV who could only complete block A had prominent peaks at theta in the RDPC from Oz to Pz and Oz to Cz in block A. This peak went away in block C for TB, but not for DO or JF. There were no other significant changes in connectivity between blocks A and C for the PD subjects. The spectra showed increased delta and/or theta power and decreased mu and/or beta power in block C compared to block A in channels Oz, Pz, and Cz for subjects TB and DO.

PDvr subject TB maintained essentially the same difference in spectral and connectivity patterns in block A-E compared to block A-C, while PDn DO showed no change and PDn JF showed minimal change in the gamma band in block A-E. Controls JD and TM show no significant differences between blocks A and E, while the rest of the controls variably showed increases in delta, alpha, and theta and decreases in gamma activity in spectra and RPDC.
Figure 3.5. Block A, control subjects. Spectra of channels are shown on the diagonal. The RPDC measures between channels are shown off the diagonal.
Figure 3.6. Block C, control subjects. Spectra of channels are shown on the diagonal. The RPDC measures between channels are shown off the diagonal.
Figure 3.7. Block E, control subjects. Spectra of channels are shown on the diagonal. The RPDC measures between channels are shown off the diagonal.
Figure 3.8. Block A, PD subjects. Spectra of channels are shown on the diagonal. The RPDC measures between channels are shown off the diagonal.
Figure 3.9. Block C, PD subjects. Spectra of channels are shown on the diagonal. The RPDC measures between channels are shown off the diagonal.
Figure 3.10. Block E, PD subjects. Spectra of channels are shown on the diagonal. The RPDC measures between channels are shown off the diagonal.
Figure 3.11. Block A-C, control subjects. The difference between block A and block C of the spectra of channels are shown on the diagonal. The difference between block A and block C of the RPDC measures between channels are shown off the diagonal.
Figure 3.12. Block A-E, control subjects. The difference between block A and block E of the spectra of channels are shown on the diagonal. The difference between block A and block E of the RPDC measures between channels are shown off the diagonal.
Figure 3.13. Block A-C, PD subjects. The difference between block A and block C of the spectra of channels are shown on the diagonal. The difference between block A and block C of the RPDC measures between channels are shown off the diagonal.
Figure 3.14. Block A-E, PD subjects. The difference between block A and block C of the spectra of channels are shown on the diagonal. The difference between block A and block C of the RPDC measures between channels are shown off the diagonal.
3.5 Discussion

In this study we compared EEG signals prior to and after presentation of closed loop visual feedback in a walking task in a responding patient (PDvr) and non-responding patients with PD (PDn) with histories of FOG in addition to those of age-matched healthy control subjects. EEG was recorded during a real, natural, ambulation task that required subjects to take multiple steps to and past a door. The purpose of this study was to identify possible modes of dysfunction and sensory compensation within nodes of a network involved in sensory-modulated gait in responders and non-responders with PD.

Of note, only one of the six PD FOG subjects tested reported a subjective feeling of improved gait in response to the visual cues. This paralleled the behavioral results, though PDn subjects DO, CR, and JF performed similarly to controls and thus were likely performing at ceiling on these measures. This may have impaired our ability to determine any unconscious effects on behavior in these subjects.

The clearest effect was a decrease in prominent theta band connectivity from Oz to Pz and Oz to Cz in PDvr. in blocks C and E compared to A. This theta band peak was also present in PDn DO and JF in block A, but did not change in blocks C and E. An analysis of the only block that PD patient DV could complete, block A, also showed this theta peak. None of the controls had this prominent theta band peak in RPDC measures. In a recent
experiment, PD patients showed increased spectral power in the theta band compared to controls during rest periods in a virtual reach and grasp task (Lainscsek C, Hernandez ME, Weyhenmeyer J, Sejnowski T, Poizner H, 2013). This difference in theta band power was larger in more lateral channels than in central channels. Increased theta spectral power was also observed in electrodes over anterior cingulate and sensorimotor cortex during motor errors (Anguerra JA, Seidler RD, Gehring WJ, 2009; Ferdinand NK, Mecklinger A, Kray J, Gehring WJ, 2012). Theta band power increase was also seen in healthy subjects walking on a balance beam compared to a treadmill, and upon losing balance while on the balance beam. The specific cortical regions involved included the superior dorsolateral prefrontal, sensorimotor, anterior cingulate, and anterior parietal areas as identified by ICA of EEG data (Sipp AR, Gwin JT, Makeig S, Ferris DP, 2013). The group concluded that there is baseline theta band activity in a multi-focal network involving motor control, error detection, and sensory information integration that exists during walking and increases with loss of balance.

Beta band power decreased in block C in TB and DO, but not in JF. Recent studies in PD patients with deep brain stimulation (DBS) have demonstrated the existence of cortico-subthalamic networks that differ in dominant frequency and spatial location (Fogelson N, Williams D, Tjissen M, van Bruggen G, Speelman H, Brown P, 2006; Lalo E, Thobois S, Sharott A,
Polo G, Mertens P, Pogosyan A, Brown P, 2008). One such network exhibits a decrease in dominant beta frequency power between the supplementary motor area (SMA) and the subthalamic nucleus (STN) during voluntary movement compared to rest (Lalo E, Thobois S, Sharott A, Polo G, Mertens P, Pogosyan A, Brown P, 2008). Synchronization in cortical beta frequency in healthy subjects has been postulated to favor existing motor state over novel movement (Gilbertson T, Lalo E, Doyle L, Lazzaro VD, Cioni B, Brown P, 2005) and in PD patients off therapy vs. on therapy, it is excessive with the degree of synchrony correlated to the level of motor impairment (Silberstein P, Pogosyan A, Kuhn A, Hotton G, Tisch S, Kupsch A, Dowsey-Limousin P, Hariz M, Brown P, 2005). The decreased power in the beta band in PDvr in stages C and E compared to stage A suggest a correlation between the presentation of visual cues and entrainment of beta band oscillations to allow for movement to occur. PDn DO may not have shown a similar effect either due to ceiling performance or because a decrease in theta connectivity did not occur concurrently. Another possibility is that DO may have altered his movement in more subtle ways that were not captured by the behavioral measurements used here. The lack of a decrease in theta connectivity and decrease in beta band power along with no behavioral response in PDn JF may also be linked, and requires further study.
These results, which suggest that visual cues can affect activity and information flow in nodes of an occipital-parietal-motor network in PD patients with FOG, provide insights into cortical neural processes underlying gait improvement with visual feedback in FOG and encourage further studies involving larger cohorts of patients. One caveat is that the alleviation of FOG with visual cues does not guarantee that the perceptual-visual hypothesis is correct. Visual cues may simply activate an alternate cortico-cerebellar pathway that compensates for the impairment produced by FOG regardless of its true neural origination. Recently a group demonstrated the feasibility of reproducing FOG in MPTP-treated macaque monkeys (Revuelta GJ, Uthayathas S, Wahlquist AE, Factor SA, Papa SM, 2012), giving hope that lesion studies may one day definitively define the anatomical and pathophysiological correlates of FOG.

### 3.6 Acknowledgements

Chapter 3 includes material as it appears in Velu PD, de Sa VR, “Single trial classification of gait and point movement preparation from human EEG,” *Front Neurosci.* 7:84, 2014. The dissertation author was the primary author of this publication. The copyright for this publication is owned by Frontiers and used in accordance with their terms and conditions.
3.7 References


Chapter 4

Conclusion

4.1 Summary of findings

Wavelet derived features from human EEG before movement onset and within single-trials can be distinguished between motor movements with overlapping muscle activation such as walking, standing, and pointing through classification with linear discriminant analysis. The highest contributions to classification within this cohort of subjects came from slow cortical potentials in the sensorimotor cortex, with smaller contributions from mu and beta frequency ranges. Given that gait action was predicted before movement from EEG signals, cortical involvement in gait production is not merely corrective or adaptive, but also initiative.

In this study with a limited numbers of subjects, there is evidence of prominent theta oscillations within the directed connectivity from occipital to parietal and motor regions in PD FOG before movement that sharply decrease with the presentation of visual feedback only in PD FOG that respond to cueing as measured by improved behavioral parameters and positive subjective experience. This change is retained even after feedback is removed. None of the healthy controls showed theta oscillations in directed connectivity measures. The PD FOG who responded to visual feedback was
also the only subject in the study to show decreased beta spectral power in occipital, parietal, and sensorimotor areas when visual cues were provided and to retain these effects when they were removed. These preliminary results are encouraging, and pave the way for studies with a much larger cohorts of patients as well as lesion studies to elucidate if the origin of these theta oscillations is from the cortex, subcortical structures, or brainstem.

4.2 Connection to broader literature

The past couple of decades have seen a resurgence of interest in the cortical contributions to locomotive control, including intracortical electrode recordings from monkeys performing gait on treadmills, non-invasive NIRS and mobile EEG characterization of cortical activity in humans during gait on treadmills, and EEG, fMRI and PET analysis of humans preparing for or performing gait-related tasks such as step initiation or motor imagery of gait. Our findings continue in this tradition by being the first to characterize the preparatory stage for gait movement from recordings of direct neural activity at millisecond resolution during a natural gait task, and to isolate this activity within single-trials. We are also the first to see if these features can be distinguished from those of actions that activate overlapping muscle groups by using a classification scheme.
The past five years have seen a convergence of evidence implicating theta oscillations in motor control, especially in detecting or correcting for error. Multi-focal increases in theta band spectral power have been seen in healthy humans during imbalance and motor errors, and PD patients appear to have spatially widespread increases in theta oscillations at rest compared to healthy humans. Our findings are in concert with these, and further, show that this effect is also seen in directed connectivity measures and can be modulated with sensory feedback.

4.3 Applications and future work

Not only do the findings from this work contribute to current literature on the neural processes underlying gait preparation in healthy subjects and Parkinson’s disease patients with FOG, but they also provide results that can be directly applied toward brain-computer interfaces (BCIs), brain-monitoring devices, and clinical diagnostics.

Teams in Belgium have developed BCIs consisting of lower limb prostheses with central pattern generators that are driven to start by neural commands such as the steady state visual evoked potential or P300 potential from real-time EEG in humans. They have also studied how the movement of the artificial limbs modulates the EEG signal. Our results on the EEG features of preparation of gait provide a more natural cortical signal command to initiate
movement of these prosthetic limbs and also allow for differentiation of gait intent from similar actions such as pointing and standing.

These results may also benefit stroke patients with loss of lower limb motor control by continuously monitoring brain activity through the EEG signal and providing tactile and proprioceptive feedback to limbs and muscles when gait intent is detected. In this way, cortical pathways for gait preparation can be maintained and rehabilitation enhanced and extended to twenty-four hours instead of daily sessions lasting a few hours or less.

Constant brain-monitoring of the directed connectivity from occipital to parietal and motor regions in PD patients with freezing of gait can be used to assess if feedback intervention is necessary or as a neural measure of feedback success. Detection of increased theta oscillations in directed connectivity may also be used as a clinical diagnostic tool for PD.

For these applications, accuracies must approach 100% to prevent user frustration and ensure maximal functionality of these devices. Future work is needed to 1) collect larger amounts of training data within subjects on different days to account for natural drift and variability of the EEG signal; 2) for the first study, see if these findings can be generalized to whole populations by testing many more subjects of all ages and many different demographics; 3) for monitoring of FOG state, test a much larger cohort of Parkinson’s disease patients with FOG, test other modalities that improve FOG parameters such as
auditory cueing and traditional motor rehabilitation, and analyze the temporal changes in connectivity; and 4) for diagnosis of PD with FOG, test normal PD patients and patient populations with other movement disorders such as stroke to ensure that the theta band connectivity feature is exclusive to PD with FOG.

This is an exciting time in history for patients with neurological deficits and psychological disorders thanks to advances in neural decoding and interfacing with BCIs as well as neural activation using transmagnetic cranial stimulation. We are currently in a nascent stage, but with continuing improvements in our understanding of basic neuroscience and continuing development of more advanced signal processing methods and user-friendly neuroimaging devices, the outlook is bright.