8-1-2016

Simultaneous Behavior Onset Detection and Task Classification for Patients with Parkinson Disease Using Subthalamic Nucleus Local Field Potentials

Nazanin Zaker Habibabadi
University of Denver

Simultaneous Behavior Onset Detection and Task Classification for Patients with Parkinson Disease Using Subthalamic Nucleus Local Field Potentials

A Thesis

Presented to

the Faculty of the Daniel Felix Ritchie School of Engineering and Computer Science

University of Denver

in Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Nazanin Zaker

August 2016

Advisor: Jason Jun Zhang

Author: Nazanin Zaker
Title: Simultaneous Behavior Onset Detection and Task Classification for Patients with Parkinson Disease Using Subthalamic Nucleus Local Field Potentials
Advisor: Jason Jun Zhang
Degree Date: August 2016

Abstract

This thesis aims to develop of methods for behavior onset detection of patients with Parkinson's disease (PD), as well as to investigate the models for classification of different behavioral tasks performed by PD patient. The detection is based on recorded Local Field Potentials (LFP) of the Subthalamic nucleus (STN), captured through Deep Brain Stimulation (DBS) process.

One main part of this work is dedicated to the research of various properties and features of the STN LFP signals of several patients' behavior conditions. Features based on temporal and time-frequency analysis of the signals are developed and implemented. Evaluation and comparison of the features is conducted on several patients' data during a classification process, using onset windows of preprocessed signals.

Another part of this research is concentrated on automated onset detection of behavioral tasks for patients with PD using the LFP signals collected during DBS implantation surgeries. Using time-frequency signal processing methods, features are extracted and clustered in the feature space for onset detection. Then, a supervised model is employed which used Discrete Hidden Markov Models (DHMM) to specify the onset location of the behavior in the LFP signal.

Finally, a method for simultaneous onset detection and task classification for patients with PD is presented, which classifies the tasks into motor, language, and combination of motor and language behaviors, using LFP signals collected during DBS implantation surgeries. Again, time-frequency signal processing methods are applied, and features are extracted and clustered in the feature space. The features extracted from automated detected onset are used to classify the behavior task into predefined categories. DHMM is merged with SVM in a two-layer classifier to boost up the behavior classification rate into 84%, and the presented methodology is justified using the experimental results.

Acknowledgments

I would like to express my appreciation to my kind husband, Hessam, and my dearest parents, Lila and Mansour, who provided me their continuous support during the work on this research.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Parkinson Disease (PD) is a chronic progressive neurodegenerative movement disorder, with motor signs of tremor and rigidity [5]. As reported in [1], approximately 6.3 million people worldwide are affected by PD, for which no cure or medication has recognized yet. In 2013, PD caused 103,000 deaths and this number is increasing each year [4]. Medications are directed to lessen motor symptoms of PD such as tremor and rigidity [26]. PD is considered as an idiopathic disease with no specific cause [36]. The main area of the brain affected by PD is the basal ganglia [14]. The basal ganglia is a collection of neuronal nuclei that assists in the coordination of voluntary movement, and is strongly interconnected with the cerebral cortex, thalamus, and brainstem, as well as several other deep brain areas, [38, 11]. Subthalamic nucleus (STN) is one of the main components of basal ganglia. Increased abnormal activity of the STN occurs in movement disorders like Parkinsons disease [42].

One of the typical and apparent symptoms displayed by patients in PD is rest tremor. It is most intense when limb is at rest, and vanishing during the sleep or voluntary movements. Hypokinesia or slowness of movement is another distinct symptom of PD that causes difficulties during the whole process of movement, i.e. the commencement of an action through the completion. Also, performing sequential or simultaneous movements results in delay or obstruction. The other cardinal clinical motor symptoms in PD are rigidity, postural instability, gait and posture

disturbances such as festination. The non-motor symptoms (neuropsychiatric disturbances) can also be caused by PD such as speech, mood, memory, behavior problems and slowness in cognitive speed, thinking and planning.

There is a need to improve current open-loop DBS therapies and diminish side-effects [21, 44, 34]. DBS may lead to detriments in cognition, speech and balance [21, 32, 12, 35]. Also, the conventional constant DBS therapies are not adaptive to patients specific needs. Design and advancement of closed-loop IPG is considered as the next step for development of DBS therapy systems, where DBS will sense the physiological signal as well as respond to it. The bidirectional signals flow in both sensing and responding directions, and the sensed signals can modulate the stimulation output [21].

1.1 Motivation and Problem Information

Great advancement in technology during the past decades, caused important improvements in modern medicine. Recently, technology has become an essential part of modern medicine presented in approximately all divisions of this rapidly developing field. In addition to diagnosis, surgical and therapeutic tools and other extremely technical fields of health care, technical devices and their development is a great assist. Specially, in cases where traditional techniques are not successful.

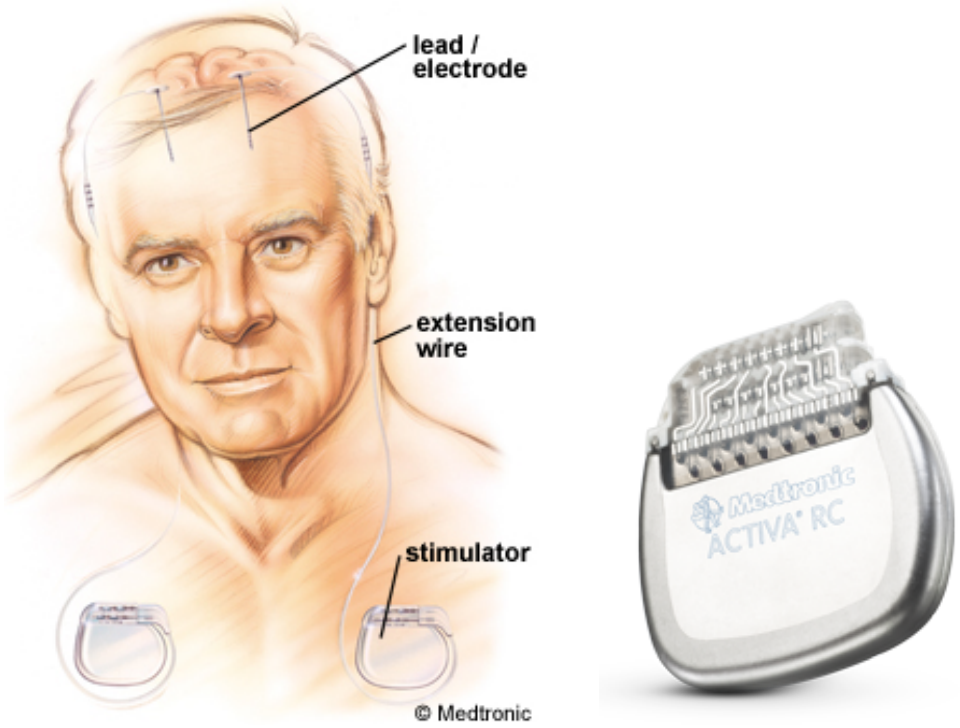
As an example, back in the 60's, the emergence of implantable heart pacemaker made the chromic electrical stimulation of human tissue as an assistive technology which aided the life of numerous patients. Passing several decades of this incident, the electrical stimulation of the heart became a usual procedure, and the similar stimulation for human brain was a new milestone. It was discovered that stimulation for the human brain can be helpful to diminish the problems related to motor,

psychiatric and other neurodegenerative disorders. The electrical stimuli to particular parts and structures in the brain could lessen to a good extend or remove symptoms of these kind of diseases, which will help the patient go back through his/her normal life.

In order to reach this goal, scientists have studied different implantable devices and developed special instruments that are able to deliver the electrical impulses directly into the brain, typically referred to as DBS devices or Deep Brain Stimulators (DBS). DBS is an established treatment of motor symptoms in PD [19]. DBS involves a surgical process during which electrodes are implanted in the brain. The electrical impulses are sent from an Implantable Pulse Generator (IPG) implanted in the chest to the electrodes to treat the motor symptoms of PD. DBS leads are implanted in the basal ganglia, typically in the STN or globus pallidus internus (GPi). These areas are stimulated with a constant pulse train of a specific amplitude, voltage, pulse width and frequency that are programmed by a physician (see Figure 1.1(a) and Figure 1.2).

The DBS devices (see Figure 1.1(b)) are used in various clinical application recently, but the details of their positive effects still remain mostly unclear and has not been studied precisely. The electrical impulses sent by embedded electrodes to parts of the brain such as STN help the treatment of movement and other disorders displayed by patients in PD. The STN itself is partitioned into sensorimotor, associative and limbic areas and it is not completely obvious which part of the STN is responsible for speech functions.

Local Field Potential (LFP) is electro-physiological signal and refers to electric potential around neurons. This signal is produced by summation of electric current



(a) (b)

Figure 1.1: (a) LFP Signal recording from the Subthalamic Nucleus [22] (b) Neurostimulator that contains Battery and Micro-Electronic Circuitry [29].

flowing from group of neighboring neurons ($\sim$300μm scale), and has been applied to distinguish cortical regions activity and sub-cortical nuclei [17]. Based on previous studies on time-frequency analysis of motor cortex electrocardiography (ECoG), suppression of β ($1330Hz$) frequency spectral power occurs while doing motor tasks [31, 30]. The same observation has been offered by analyzing the STN LFB signals [10, 25, 3, 24].

Another group of researches revealed the association of changes in β with speech production and audition in cortex [13, 15]. In [20], the authors provided a mechanism in which subjects performed different behavioral tasks while they were under DBS surgery, and STN LFP data were simultaneously collected. The authors re-

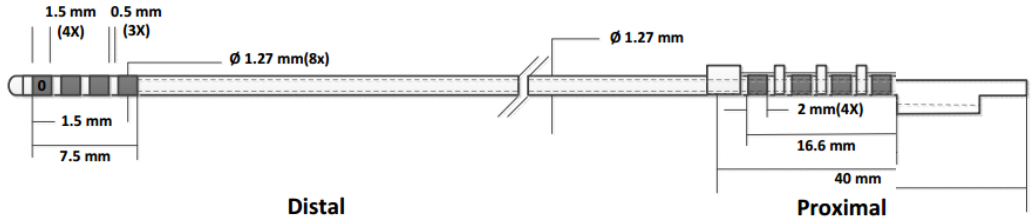


Figure 1.2: The recording electrodes' schematic representation for LFP recordings. Medtronic 3389 DBS lead [20].

.

marked the relation between bilateral switching of β power states and changing of the behavioral tasks performed by PD patients.

Research and experiments reveal the need of improvements in current DBS therapies and open-loop implantable pulse generator (IPG) to diminish their drawbacks and flaws [21] such as cognitive, speech and balance side-effects. In the current DBS therapy, a train of electrical pulses is sent with predefined and adjusted stimulation parameters. Although the designed open-loop therapy is practical and successful for movement disorders, there is a chance for an optimized therapy system being able to sense and deliver signals in both direction by applying and designing a closed loop model. The growing assistive technology to record bio-signals has improved the understanding of clinical state of the patients. They contribute in collecting essential information to customize neuro-modulation therapy.

Another need is to customize DBS therapy to specific patients tasks and control the side-effects to be non-damaging to the patients overall therapy process. Design and advancement in closed-loop IPG is considered as the next step for development of DBS therapy systems. The closed-loop means being able to sense the physi-

ological signal as well as responding to it. Here the bidirectional signals flow in both sensing and responding directions and the sensor signals can provide feedback modulation of stimulation. Reaching closed-loop DBS therapy require better understanding of LFP signals characteristics during the different behavior tasks. As a consequence, detection and learning the informative LFP features, classification of behavioral tasks, and detecting the onset of each behavior in PD patients play an important role in development and advancement in the closed-loop IPGs which are called the next generation DBS therapy systems.

1.2 Background

As presented in [37], the state of macaques behavior such as planning or cascade, as well as the direction of intended movement can be predicted by LFP signals, which reveals the prospective and potential appeal of LFP signals in future as valuable features for brain-machine interface. In our recent works [44], machine learning methods were applied in order to differentiate various behavioral tasks of PD patients. The behaviors were categorized as language, motor and combination of language-motor tasks. Time-frequency features were extracted from the LFP signals collected during the DBS surgery process. The classification frameworks such as hybrid Hidden Markov Models (HMM) and Support Vector Machines (SVMs) as well as adaptive learning with Dirichlet process Gaussian mixture models were designed and applied to categorize the behavioral tasks with high accuracy.

Achieving closed-loop DBS therapy requires better understanding of LFP signals characteristics during different behavior tasks. In [26], adaptive DBS (aDBS) of the subthalamic nucleus was tested on PD patients where the beta oscillations in the LFP recorded from the stimulation electrodes, was used as the feedback infor-

mation. Experimental results provided in [26], validated the effectiveness of aDBS comparing to constant stimulation and random intermittent stimulation.

Apart from classification of different behavioral tasks, detection of onset of signals and behaviors is another state of the art problem grabbed the attention of many researchers in different areas of signal processing [7]. Not many studies have been done on behavior onset detection using LFP signals for patients in PD. In [33], the authors applied an artificial neural network to predict the onset of PD tremors in one human subject. They accomplished to specify the pattern of the onset tremor. They detect and predict the onset of the PD tremors in human subjects with good accuracy. Radial Basis Function Neural Network (RBFNN) model based on Particle Swarm Optimization (PSO) was applied on the LFP signals in another work [43]. However, non of these works present an approach to detect the patients different behavior tasks onset.

The precise control of movement execution onset is absolutely important for patients in PD. According to a study from the Parietal Reach Region (PRR), the LFPs in cortical area might be one of the useful features in order to decode the execution time information. The striking difference in the LFP spectrum between the plan and execution states is the main reason for this hypothesis [37].

1.3 Thesis Contribution

In this work, we study the bipolar recording from DBS leads, implanted in STN of human subjects underwent implantation of DBS IPG. This data is collected from four patients and was previously used in [21] where patients underwent DBS implantation for treatment of idiopathic PD. In this research, we first proposed models to

only classify the PD patients behavior tasks for the best recognition rates. We improved the classification rate of our system to more than 90 percent in average for different behavioral tasks. The behavioral tasks were grouped as 1) motor tasks, 2) language tasks, 3) language with motor task and motor onset 4) language with motor task and speech onset [20, 44].

Further more, we proposed a new framework to detect the onset of the behavior from the collected LFP signals after the cue of the behavior [45]. We collected LFP values in small windows of 2000 samples and learned the essential features of onset windows for language with motor tasks of patients. Our proposed method was successful to detect onsets of behavior with average delay of 1500 ms through all the trials of the behavior task [45].

In our latest work, we have merged the onset detection and classification procedures for different behavior tasks. The proposed model is capable of detecting the onset of different behaviors, and after detection of the onset with an acceptable delay, it deals with classifying the behavior into predefined categories. This achievement, is a great step forward in order to being able to customize the DBS therapy in closed-loop models. Here, detecting the cue, the type of the behavior is mathematically recognized by applying machine learning classification models with an average of 80% recognition rate through all the subjects.

To locally monitor the signal for onset detection, the LFP signal is divided into small consecutive sliding windows with an overlap. We explore and learn the characteristics of windows of signal containing the onset of the behavior in addition to differentiating behavior tasks based on their specific features. Moreover, the suitability and relevance of the different approaches was compared. The classification

results were then used to assess the actual applicability of the proposed method. The results obtained in this study will assist us in advancing our knowledge of PD patients behaviors and help us develop the next generation of DBS systems.

To the best of our knowledge, this is the first study and framework designed to simultaneously detect the onset of different behavior tasks and specify the type of task (i.e. language or motor) as well. We applied Matching Pursuit Decomposition (MPD) for time-frequency analysis of the LFP. Considering the physical meaning of Gaussian atoms, the MPD with Gaussian atom dictionary (GD) can decompose a signal into a linear composition of Gaussian atoms with the features of amplitude, variance scale, time-shift and frequency-shift. In this thesis, the MPD with GD is employed to extract the useful features from LFP signals.

1.4 Thesis Organization

This thesis is organized as follows. Chapter 2 provides a background on DBS and explains the need for a closed loop DBS system for patients with PD. It also provides a background on LFP signals and how they are collected. Chapter 3 explains the data collection procedure and provides information regarding the subjects and behavioral study. The feature extraction method including Matching Pursuit Decomposition (MPD) algorithm is provided in Chapter 4. Here, the procedure of extracting informative features from LFP recordings using a Gaussian dictionary is explained in details. In addition, the clustering methods used in this work and the most suitable modified version is introduced.

Chapter 5 discusses the proposed onset detection algorithm. The implementation using Hidden Markov Model (HMM) and Support Vector Machines (SVM) is

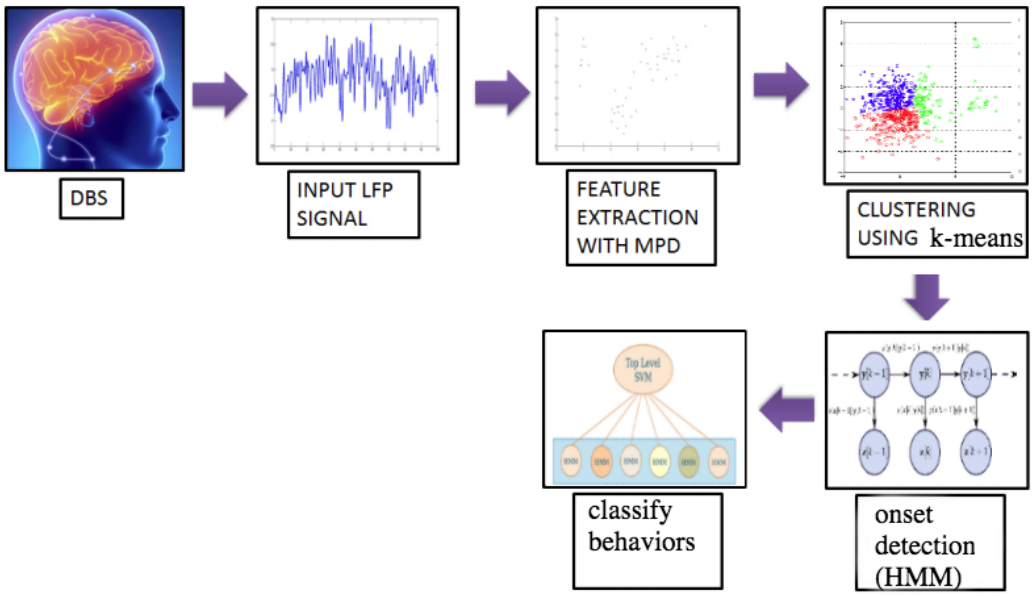


Figure 1.3: Block diagram summarizing the proposed method for simultaneous onset detection and classification.

elaborated, and explains our proposed integrated multi layer classifier. Chapter 6 provides the classification of behavioral tasks, and Chapter 7 introduces the proposed joint onset detection and classification model. In Chapter 8, the experiments and discussion of the results are provided and evaluated. Finally conclusion and future work are elaborated in Chapter 9. A block diagram representing the main contribution of this work is shown in Figure 1.3. Also, the acronyms used throughout this thesis are summarized in Table 1.1 and 1.2.

Table 1.1: Alphabetical list of acronyms used in this Thesis

Acronym	**Definition**	**Explanation**
BG	Basal Ganglia	Brain structure positioned mainly in the midbrain, involved also in modulation of movement. A dysfunction of this structure causes the Parkinsons disease.
DBS	Deep Brain Stimuation	Therapeutic method based on application of electrical impulses to the structures in the human brain.
DHMM	Discrete Hidden Markov Model	A discrete-time stochastic process is a stochastic process for which the index variable takes a discrete set of values.
ECoG	Electrocorticography	The practice of using electrodes placed directly on the exposed surface of the brain to record electrical activity from the cerebral cortex.
EEG	Electroencephalography	Method to record electrical activity of the brain along the scalp.
GMM	Gaussian Mixture Model	A probabilistic model that assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters.
HMM	Hidden Markov Model	Statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states.
IPG	Implantable Pulse Generator	A battery powered device designed to deliver electrical stimulation to the brain.
LFP	Local Field Potential	Summarized activity of the neuronal tissue in a specific region as captured by an electrode pair.

Table 1.2: Alphabetical list of acronyms used in this Thesis, continued.

Acronym	**Definition**	**Explanation**
MPD	Matching Pursuit Decomposition	A sparse approximation which involves finding the best matching projections of multidimensional data onto an over-complete dictionary D.
PD	Parkinson's Disease	Chronic neurological disorder, affecting mainly motor function of the diseased. Tremor, akinesia and rigidity are among the most common symptoms.
ROC	Receiver operating characteristic	Graphical plot of binary classifier system properties, used for system evaluation.
STN	Subthalamic Nucleus	Structure in the brain, functional unit of the Basal ganglia, involved in the modulation of motor functions. A target structure for DBS in Parkinson's disease.
SVM	Support Vector Machine	A supervised learning model with associated learning algorithms that analyze data and recognize patterns which used for classification and regression analysis.

Chapter 2

Deep Brain Stimulation and Local Field Potential

As mentioned in previous chapter, PD is considered as a progressive neurological condition. The studies showed that it causes from the degeneration of special neurons that produce dopamine in the substantia nigra which resides at the lower part of the brain [20]. PD has impact on functional activities like writing, typing, walking, speech as well as slowness in thinking and cognitive tasks. The very early treatments for managing the motor and cognitive symptoms of this disease is effective, however, by the progress of the disease, drug therapies may eventually considered as an ineffective option. In this regard, deep brain stimulation (DBS) treatment can be used as a therapy to relieve the motor symptoms [21].

2.1 Deep Brain Stimulation

Deep Brain Stimulation (DBS) is known as a surgery applied in order to cure or reduce several disabling neurological symptoms. It is usually used to devitalize the motor symptoms of PD, such as tremor, rigidity, stiffness, slowed movement, and walking problems. Accomplishments in clinical purposes by applying DBS therapies, has made it possible to extensive use of this devices for a great scope of neurological disorders [21].

Currently, open loop DBS is used vastly for treatment of PD and essential tremor. This kind of DBS, directs one-way signal through the brain of patient constantly.

The constant stimulation which is not adjusted to patients' different needs and tasks may cause some side effects such as impaired cognition, speech and balance.

In Figure 1.2, DBS device is shown. The thin coated leads is used to transmit electrical energy (LFPs) to the targeted portion of the brain. This area is considered mostly as subthalamic nucleus for PD therapies. Here, the LFP signals are recorded by the invasive micro-electrodes. These signals represent the oscillatory activity within the nuclei of the Basal Ganglia (BG). In Figure 1.1(b), a neurotransmitter is demonstrated. This type of neurotransmitter contains a computer chip controlling waveform and electric impulses transfered to the PD patient's brain, and is programmable to fine tune the system to the patient.

As mentioned in Chapter 1, the design and development of closed-loop IPG being able to transfer and sense physiological signals is the next border in brain stimulation research and therapy. It will absolutely extend the applications of DBS systems and introduce new ones in the fast pacing improving technology. In the closed-loop systems, the bidirectional signals move in both sensing and responding directions which let the sensing signals provide feedbacks based on the responding ones. This feedback loop can cause recovering the functionality of the targeted parts of the brain.

Having knowledge of the LFP features of patients operating behavioral tasks under conditions with no tremor and motor symptoms, DBS devices is adjusted to restore LFP signals in spacial parts of the brain while the patient experiencing sever tremor. To do so, correct understanding of the properties of LFP signals during different behaviors of the patient is of great value, and considered as an essential factor toward the success of closed-loop DBS systems.

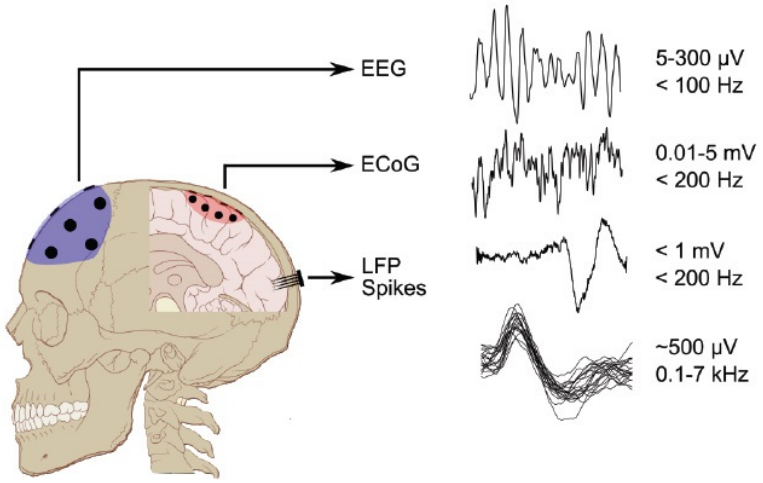


Figure 2.1: LFP signals recording from deeper areas of the brain comparing to EEG signals [39].

2.2 Local Field Potential

Local Field Potentials (LFPs) are electrical events at deeper locations in the brain which can be recorded by invasive metal or glass electrodes, or silicon probes into the brain. They are also known as micro-EEG (see Figure 1.2), and considered as the most informative brain signal representing action potentials and other membrane potentials-derived fluctuations in a small neuron volume [9]. LFP is different from normal EEG or ECoG signals. The range of LFP is less than 1 mV with frequency less than 200 Hz as shown in Figure 1.2.

The LFP signals used to assess the performance of our proposed methods were obtained from a study involving twelve patients undergoing DBS implantation for treatment of idiopathic PD at University of Washington and Colorado Neurological Institute (CNI) [44]. The LFP signals recorded during behavioral tasks . The tasks described four types of behaviors as mentioned in Chapter 1.

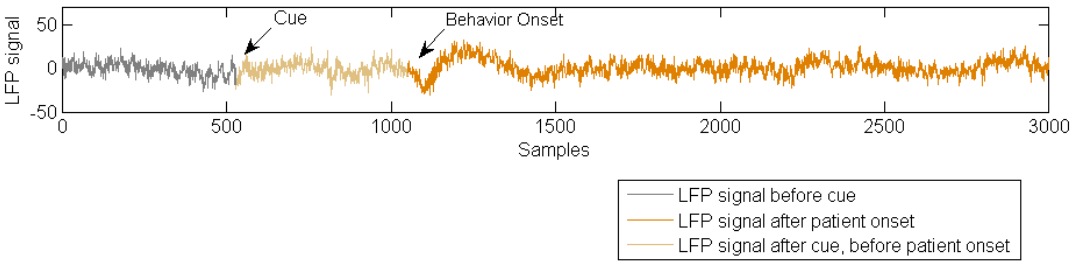


Figure 2.2: In the LFP signal, the cue and patients' behavior onset are shown. The sampling frequency of LFP is 5k Hz in this example.

The recordings were obtained from each of the four contacts of the DBS lead (Medtronic 3389, see Figure 1.2). Although primarily designed for stimulation, these electrodes have been used for LFP recording in humans, as they do not require modification of standard surgical practice. The DBS lead contact is platinum/iridium, has a surface area of 6.0 mm^2 and impedance of 1.7$k\Omega$. Signals were amplified, sampled using a sampling frequency of 4.8kHz or 5kHz (depending on the subject and place data collected), and combined with event markers and subject response signals. A typical LFP signal taken from one of the subjects with Idiopathic PD is shown in Figure 2.2. Here, the cue and onset of the behavior are also marked on the signal.

Chapter 3

Data Collection

For the data collection procedure, LFP signals were collected from several patients which is discussed in more details in the following.

3.1 Subjects

In this research, the data we used was collected from four subjects who underwent DBS surgery as approved standard treatment for idiopathic PD. All subjects were in the off-medication state, and experienced DBS surgery per clinical routine. They provided informed consent for their participation in the manner approved by the Institutional Review Board (IRB) of the institutes that data were collected, University of Washington and Colorado Neurological Institute (CNI). Four independent recordings were measured from the four participants. One of the subjects provided sequential recordings from each side. The remaining three patients provided bilateral recordings. Overall, there were four left, and four right hemisphere recordings for our experiments and analysis.

3.2 Data Acquisition Design and DBS Surgery

Data recording were performed at University of Washington and Colorado Neurological Institute (CNI) using Medtronic 3389 DBS leads implanted in the right and left STN. The DBS lead is shown in Figure 1.2 [1]. Each DBS lead has 4 contacts, and the DBS lead contact is made of platinum/iridium [20]. The LFP signals were

amplified, digitized and collected simultaneously with event markers and subject response signals using SynAMPS2 (Neuroscan, Victoria, Australia) or g.USBamp (g.tec, Graz, Austria). We used linked mastoid reference and one ground plate.

Table 3.1: Behavior condition code and specification for different PD patients.

Condition code	**Description**	**Number of trials**	**Patient ID**	**Center**
I	button press	90	P1	University of Washington
II	language	45	P1	University of Washington
III	language and button press, with button press onset	45	P1	University of Washington
IV	language and button press with speech onset	45	P1	University of Washington
I	button press	55	P2	University of Washington
II	language	53	P2	University of Washington
III	language and button press with button press onset	54	P2	University of Washington
IV	language and button press with speech onset	54	P2	University of Washington
V	left hand button press	46	P3	Colorado Neurological Institute
VI	right hand button press	47	P3	Colorado Neurological Institute
II	language	43	P3	Colorado Neurological Institute
V	left hand button press	45	P4	Colorado Neurological Institute
VI	right hand button press	46	P4	Colorado Neurological Institute
II	language	45	P4	Colorado Neurological Institute

3.3 Behavioral Study

The behavior tasks studied in this thesis included motor, speech and combinational motor-speech tasks. All subjects performed left and right hand button-press, and for some of the subjects, we further classified the button-press between left and right hands. Different speech initiation tasks were also investigated, i.e, naming the months of the year, repeating names of objects, counting upwards from one. Combinational speech-motor tasks include naming the months of the year with a simultaneous button press marking the first month, and counting with a simultaneous button press marking the first number. Six specific behavior tasks were selected for our analysis, which are 1) button press, 2) language, 3) language and button press with button press onset, and 4) language and button press with speech onset. For subjects whose left and right hand data is available, we also considered 5) left hand button press, and 6) right hand button press.

Behaviors were performed in 3-6 blocks of 15 trials for subject "P1", 2-4 blocks of 11-15 trials for subject "P2", 2-4 blocks of 15-20 trials for subjects "P3", and "P4". The total number of trials for each subject's behavioral task is shown in Table 3.1. For task initiation and completion, subjects received an audio cue from a presentation laptop computer running EPrime 2.0 (Psychology Software Tools, Sharpsburg, PA), or a custom kivy script. In order to remove the effect of anticipation of the cue, random time factor was programmed into task length in each trial. If the patient did not respond within three seconds, the trial was marked as invalid.

Chapter 4

Feature Extraction

Let m denote the mth behavior task of PD patients, where $m = I$ represents motor task, $m = II$ represents language task, $m = III$ represents language plus motor task with motor onset, $m = IV$ represents language with motor task and language onset, $m = V$ represents left hand motor task, and $m = VI$ represents right hand motor task. We denote the set of LFP signals collected from different behavior tasks as $\mathcal{L} = \{\mathbf{L}^m | m = I, II, III, IV, V, VI\}$.

We utilized the recorded LFP signals after the cue of a behavior for 10 seconds to detect the patients' behavior onset. In Figure 4.1, a sample signal from a trial of the experiment is displayed and marked with the cue and patient onset. Let the LFP signal from the ith ($i = 1, \ldots, I$) experiment of mth behavior task be given by $\mathbf{l}_i^m$, where $\mathbf{l}_i^m = [l_i^m[1]\ l_i^m[2] \cdots\ l_i^m[n] \cdots l_i^m[N]]$, $l_i^m[n]$ denotes the discrete LFP signal values collected at time n, and N stands for the number of collected samples. We divided the signal into small sliding windows as $\mathbf{w}_{i,d}^m = [l_i^m[(d-1)U + 1] \cdots l_i^m[(d-1)U + V]]$ where $d = 1, \cdots, D$. The window length is indicated by V that is equal to 0.4 second, and $V - U$ is the window overlap that is equal to 0.2 second. As a consequence, $\mathbf{l}_i^m$ is decomposed into D consecutive sliding windows (Figure 4.1).

This work is considered as a hard problem since any flaw or drawback in detecting the onset position will also lead to misclassification of the behavioral tasks. Also, there is a high similarity between the features extracted from combinational

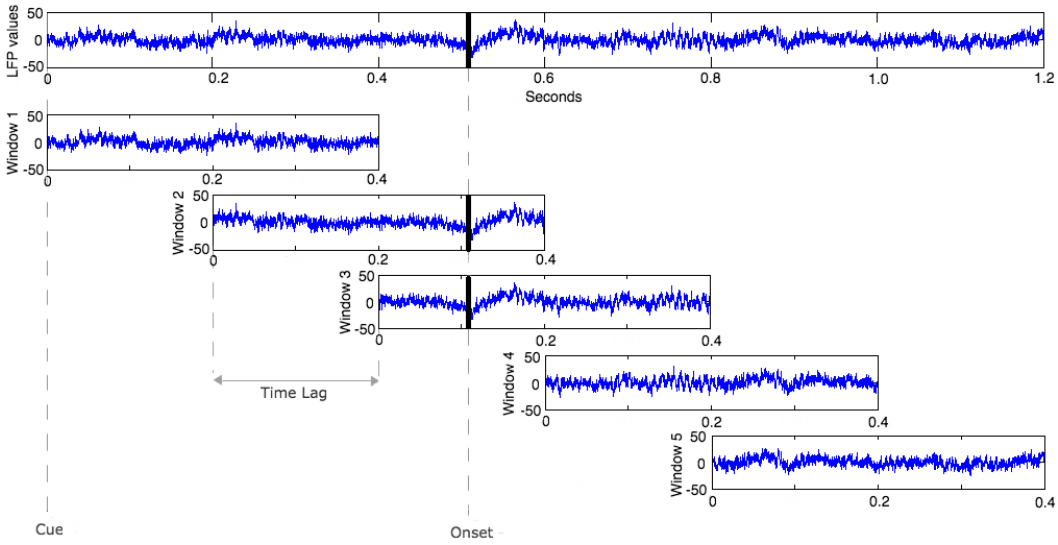


Figure 4.1: The LFP signal extracted from each patient was monitored for 10 seconds after the cue and divided into sliding windows with length of 0.4 second, and an overlap of 0.2 second with their adjacent windows (49 windows overall). This figure shows the LFP signal extracted from patient "P1" for 1.2 seconds (the top figure), and divided into sliding windows as demonstrated. The *Cue, Onset* and *Time Lag* are shown in the figure. The sampling frequency of LFPs is 5000 Hz for this patient.

tasks. As an example, "language with button press and button press onset" is very similar to "language with button press and speech onset" in terms of collected features which makes the classification task very difficult.

As the aim of the advanced closed-loop DBS system is to customized the therapy for each patient, we focused on subject specific onset detection and classification for each patient. Several recordings (i.e. trials) of performing a specific task are available for each of the patients, which are considered as separate samples of each behavior task. To train onset detection classifier, $K-$fold cross validation is applied to separate the training and testing sets. The same procedure is used in behavior task classification problems.

4.1 Matching Pursuit Decomposition

A time-frequency analysis algorithm, namely Matching Pursuit Decomposition (MPD), is used in our proposed approach in order to extract informative features from each LFP signal windows.

The MPD algorithm is considered as a sparse approximation method which tries to find the best matching projections of multi-dimensional data onto an over-complete dictionary. The idea behind this algorithm is to represent a time-domain signal, $f(t)$, as a weighted sum of Gaussian atom functions $g_p(t)$. Compared with many time-frequency representations which may result in cross-terms and cause information distraction, MPD feature extraction algorithm with GD can decompose the original signal into highly dense Gaussian atoms in time-frequency domain and generate minimum residual signal energy. In (4.1.1), α_p denotes the coefficient for the Gaussian atom, $g_p(t)$ stands for the Gaussian atom selected from a given GD, and p represents the iteration index.

$$f(t) = \sum_{p=1}^{\infty} \alpha_p g_p(t). \tag{4.1.1}$$

The convergence of the above representation is discussed in [28]. Although orthogonality is not required for the GD $\mathcal{D}$, the completeness is required for the MPD. Using this method, we can reconstruct the original signal $f(t)$ by applying decomposition using finite iterations, and a remainder $r_P(t)$ with small energy residual. The formulation is illustrated in the following equation as

$$f(t) = \sum_{p=1}^{P-1} \alpha_p g_p(t) + r_P(t). \tag{4.1.2}$$

Considering (4.1.2), the MPD algorithm is described as follows. Let $r_1(t) = f(t)$ which denotes the energy residual at first step. The atom $g_p(t)$ is searched in the GD $\mathcal{D}$ for the one which has the maximum magnitude of the projection in $r_p(t)$, $p = 1, 2, 3, \ldots, P$. (4.1.3) explains the selection of $g_p(t)$ in details.

$$g_p(t) = \underset{g^{(e)}(t) \in D}{\arg\max} \left| \int_{-\infty}^{+\infty} r_p(t) g^{(e)}(t) dt \right|, \tag{4.1.3}$$

where, $e = \{\tau, \nu, \sigma\}$ stands for the time-shifting, frequency-shifting and related normalization coefficient for the Gaussian atoms. Consequently, the corresponding coefficient α_p is calculated after $g_p(t)$ is obtained as

$$\alpha_p = \int_{-\infty}^{+\infty} r_p(t) g_p(t) dt. \tag{4.1.4}$$

The remainder $r_p(t)$ and $r_{p+1}(t)$ are related according to the equation below

$$r_{p+1}(t) = r_p(t) - \alpha_p g_p(t). \tag{4.1.5}$$

The Pth remainder after $(P-1)$th iteration is calculated as

$$r_P(t) = r_1(t) - \sum_{p=1}^{P-1} \alpha_p g_p(t). \tag{4.1.6}$$

Using MPD, the signal $f(t)$ is decomposed into a number of Gaussian atoms, and can be represented by the parameters of the selected atoms, i.e., $[\alpha_{i,d}^{m,p}\ \tau_{i,d}^{m,p}\ \nu_{i,d}^{m,p}\ \sigma_{i,d}^{m,p}]$, $p = 1, \cdots, P$.

Specifically for this study, after extracting the MPD features from the LFPs of each window $\mathbf{w}_{i,d}^m$, P 4-dimensional feature vectors are extracted from each LFP windows. P denotes the number of MPD algorithm iterations. The MPD features extracted from the LFPs in $\mathbf{w}_{i,d}^m$ after P iterations are denoted as $\mathbf{M}_{i,d}^m = [M_{i,d}^{m,1}\ \ \ldots\ \ M_{i,d}^{m,P}]^T$ with T denoting matrix transpose. The pth feature vector is given by $M_{i,d}^{m,p} = [\alpha_{i,d}^{m,p}\ \ \tau_{i,d}^{m,p}\ \ \nu_{i,d}^{m,p}\ \ \sigma_{i,d}^{m,p}]$, corresponds to amplitude, time-shift, frequency-shift and variance parameters of the MPD Gaussian atom.

4.2 Clustering

After signal feature vectors are extracted by MPD, the resulted P feature vectors are quantified into P feature symbols by K-means clustering algorithm, i.e., each feature vector is mapped to one feature symbol [27]. The resulting feature symbols are used for HMM based classification. The MPD atoms are composed of time-frequency parameter vectors, and the K-means clustering method is applied to quantify these vectors into K symbols [27]. The feature vectors $\mathbf{M}_{i,d}^m = [M_{i,d}^{m,1}\ \ \ldots\ \ M_{i,d}^{m,P}]^T$ are mapped into $\mathbf{K}_{i,d}^m = [k_{i,d}^{m,1}\ \ \ldots\ \ k_{i,d}^{m,P}]^T$, where the element $k_{i,d}^{m,p}$ is one of the K symbols, i.e., $k_{i,d}^{m,p} \in \{S_1, S_2, \cdots, S_K\}$. The quantification results, $\mathbf{K}_{i,d}^m$, are used to train and test HMM classifier.

4.2.1 K-means Clustering Method

K-means is one of the most commonly used unsupervised learning algorithms that can easily be used for clustering problems. The procedure includes a straight forward method to cluster a given set of data without considering or having any knowledge for their labels. Here, a certain priori number of clusters can be specified say k. The main idea is to select k centers or centroids each related to each of the clusters. Selecting different centers will cause different clustering results. As a consequence, the best guess is to select centroids which has the most distance from each other.

Here, the definition of distance may vary from euclidean, hamming, cosine and other distances which best suited the data. Later, each point will be assigned to the nearest centroid and the process continues till no data point remains. Since, this clustering may not be the best regarding the data, new centroids will be selected in the next step based on the calculated center of each cluster. The process of assigning data point to the new centroids will be done, and again new centroids will be selected based on newly built clusters. As a result of this loop we may notice that the k centroids change their locations step by step, and this procedure will continue till no more changes are done. In other words, centroids do not move any more.

Finally, this algorithm aims at minimizing an objective function, in this case a squared error function. The objective function is provided in Equation 4.2.1.

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \|x_i^{(j)} - c_j\|^2 \tag{4.2.1}$$

where $\|x_i^{(j)} - c_j\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster center c_j. It is an indicator of the distance of the n data points from their respective cluster centers or centroids.

4.2.2 Semi-Supervised K-means Clustering

K-Means algorithm is known as one of the most used clustering algorithm for Knowledge Discovery in Data Mining which usually lead to acceptable results. Seed based K-Means is a semi-supervised learning algorithms such that integrates a small set of labeled data which are called seeds to the K-Means algorithm to boost the results and decrease its sensitivity to the initial centers. The centers in the first step of the algorithm are usually generated at random or assumed to be available for each of the clusters [6].

Another efficient algorithm is active seeds selection [18]. This method employs a Min-Max approach in order to help the coverage of the whole data points. By applying active seeds selection method, the seeds are collected so that each of the cluster has at least one seed. Also, the number of convergence iteration of K-Means clustering will be reduced, which is one of the essential factors for many applications [41].

Chapter 5

Onset Detection

In order to detect the behaviors onset, we conducted several experiments. We collected the recorded LFP values after the cue for 10 seconds. As it is demonstrated in Figure 2.2, a sample signal from a trial of the experiment is displayed and marked with the cue and patient onset. We divided the signal into small sliding windows of 2000 samples, which have overlaps of 1000 samples. There is a time lag of approximately 200 ms between two consecutive windows. Figure 4.1 better demonstrates this process.

We collected MPD features from the signal in each window. Since MPD atoms are composed of four-dimensional parameter vectors (i.e. amplitude, variance scale, time-shift and frequency-shift), the K-means clustering method is applied to cluster these atoms to 64 nodes [27]. In other words, we quantify the feature vectors and map them into clusters using K-means method. Then, the deep brain signal obtained during PD patient's behavior is represented by a one-dimensional feature vector. The quantified vectors which are mapped into K-means nodes used to train HMM classifier as explained previously.

In the onset detection procedure, we employed supervised probabilistic model based on DHMM classifiers. We also use a two-layer detection model by applying SVMs as a discriminative model in the top layer, and DHMMs as generative models in the lower layer of the classifier. In the following, each method is elaborated.

Hidden Markov Model (HMM) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved or hidden state. We utilize the standard type of hidden Markov model, in which the state space of the hidden variables is discrete, i.e. Discrete Hidden Markov Model (DHMM) [8].

The feature symbol set is denoted as β which contains the observation symbols to train different HMMs for detecting and identifying the behavior onset. In our HMM say Λ, we specify the initial state distribution vector, the hidden state transition matrix, and the state-dependent observation density matrix as $\lambda = [\pi, A, B]$, respectively. The maximum-likelihood estimate for λ is given by the following equation using the Baum-Welch algorithm [8].

$$\lambda_{ML} = \arg\max_{\lambda} \log P(\beta|\lambda, \Lambda), \tag{5.0.1}$$

Where β is the observation data and λ is the parameter set of the HMM Λ. After passing i iterations, the λ is calculated as:

$$\lambda^{(i+1)} = \arg\max_{\lambda} \sum_{\mathbf{H}} P(\mathbf{H}|\beta, \lambda, \Lambda) \log P(\mathbf{H}, \beta|\lambda^{(i)}, \Lambda) \tag{5.0.2}$$

Here, $\mathbf{H}$ indicates the hidden states. The summation over $\mathbf{H}$ signifies the overall possible state sequences in the HMM. Finally, the probability of the test is calculated as

$$P(\beta|\lambda, \Lambda) = \sum_{\mathbf{H}} \pi_{H_1} \prod_{n=1}^{N_{obs}} a_{H_n, H_{n+1}} \prod_{n=1}^{N_{obs}} b_{H_n}(\beta_n), \tag{5.0.3}$$

Where π_{H_1} is defined as initial state probability of state H_1. Also, the state transition probability from state H_n to H_{n+1} is defined as $a_{H_n, H_{n+1}}$. $b_{H_n}(\beta_n)$ spec-

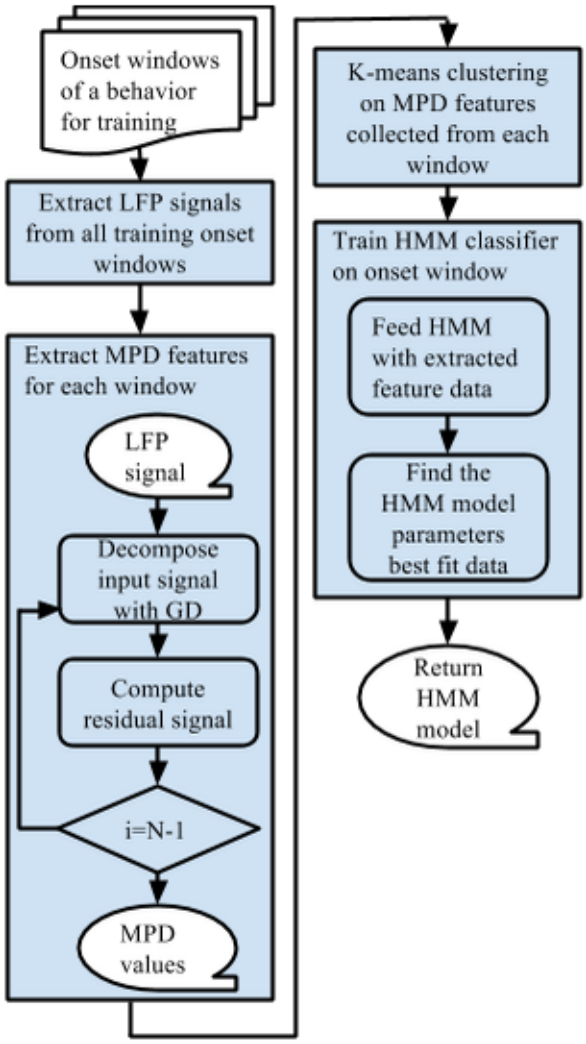


Figure 5.1: Flowchart demonstrating the train procedure for onset detection.

ifies the probability of observing β_n in state b_{H_n} [8], and N_{obs} is the number of observations. The training procedure is shown in a flowchart in Figure 5.

Chapter 6

Classification of Behaviors

In order to classify the behaviors, we applied several experiments. As a preprocessing step, we quantify the feature vectors and map them into new feature values using the K-means classifier. The feature vectors are clustered into 64 cluster nodes. Each consecutive 30 feature vectors represent deep brain signal attained during a PD patient's behavior. Then, for learning the best model for classification, we employed a combination of SVM classifiers as a discriminative model and HMM as a generative model in the first two experiments. In the later one, we applied Dirichlet Process Gaussian Mixture Models (DP GMMs) to depict arbitrarily complex statistical data distributions. In the following, each method is elaborated with more details separately.

6.1 Support Vector Machine

To improve the performance and robustness of the classifier, several DHMMs were trained. The log-likelihood returned by all DHMMs constructs a new vector of features as the input of the SVM. According to [40], given data points $D = \{(\varphi_i, y_i)|\varphi_i \in R^p, y_i \in \{-1, 1\}\}_{i=1}^n$, where φ_i is a $p-$dimensional real vector. The objective function of the SVM classifier is provided below, which tries to find the maximum-margin hyperplane that divides the data points based on their related label value.

$$\min_{w,b} \frac{\|w\|^2}{2}, \quad \text{subject to: } \forall i,\ y_i(w \cdot \varphi_i - b) \geq 1 \tag{6.1.1}$$

The offset of the hyperplane from the origin along the normal vector w is calculated by the parameter $\frac{b}{\|w\|}$. The top-level SVM fuses all the low-level DHMMs and returns one value. This approach improved the recognition results [44].

6.2 Hidden Markov Model

An HMM is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved or hidden states which can be presented as the simplest dynamic Bayesian network. Here the standard type of hidden Markov model is considered, in which the state space of the hidden variables is discrete and called Discrete Hidden Markov Model (DHMM). In our experiments, we applied the same equations as described in Chapter 5 for training and testing HMMs.

6.3 Fusing HMMs by Top Level SVM Classifier

Since the features have temporal patterns, we used HMM models as a good candidate to model and classify cognitive tasks patterns. Based on the probability nature of HMMs and the random assignment of prior probabilities, the HMMs may lead to different classification results in different runs of the model. In our first approach, to avoid the randomness and improve the performance of the classifier, we trained several DHMM classifiers for each behavior and put the log-likelihood returned by all DHMMs in a new vector. The resulted vector is given as an input of a SVM to decide between the outputs of the DHMMs. According to [40], the objective function of the SVM classifier can be found in Equation 6.1.1.

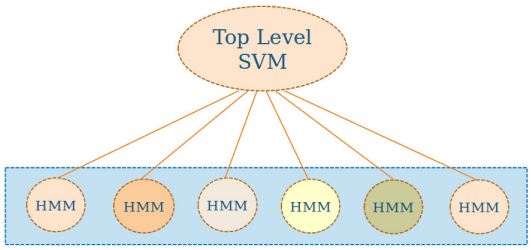


Figure 6.1: Independent DHMMs are fused by the top level SVM classifier.

The top-level SVM fuses all the low-level DHMMs and returns one value as demonstrated in Figure 6.3. This approach improved the recognition result for seven percent comparing to the previous work [23]. The training procedure is shown in Figure 6.4.

6.4 Hybrid SVM-HMM Model

In the second alternative, we trained a hybrid SVM-HMM model which is an implementation of structural SVMs for sequence tagging based on a combination of SVM and HMM [2]. It handles dependencies between neighboring labels by Viterbi decoding. In contrary to basic HMM, the learning procedure is based on a maximum soft margin criterion but it also shares the major advantages with other discriminative methods, specially the potential to deal with overlapping features. In the SVM-HMM, models that are isomorphic to an k^{th}-order HMM are discriminatively trained utilizing the Structural Support Vector Machine (SVM) concept. Consider an input sequence $x = (x_1, ..., x_l)$ of feature vectors $x_1, ..., x_l$, the SVM-HMM tries to train a model that predicts a tag sequence $y = (y_1, ..., y_l)$. The following linear discriminant function is applied.

$$y = \arg\max_{y} \Sigma_{i=1}^{l}[\Sigma_{j=1}^{k}(X_i.W_{y_{i-j,...,y_i}}) + \varphi_{trans}(y_{i-j},...,y_i).W_{trans}] \tag{6.4.1}$$

Where $W_{y_{i-k},...,y_i}$ is the emission weight vector learned for each different k^{th}-order tag sequence $y_{i-k},...,y_i$, and W_{trans} is the transition weight vector for the transition weights between adjacent tags. $\varphi trans(y_{i-j},...,y_i)$ stands for an indicator vector which has an entry set to 1 for the sequence $y_{i-j},...,y_i$. An optimization problem in which the training entries are $(x^1,y^1),...,(x^n,y^n)$. The feature vectors are $x^j = (x_1^j,...,x_l^j)$, and the training sequence tags are $y^j = (y_1^j,...,y_l^j)$. For a model with first-order transitions and zero-order emissions the optimization problem is defined as follows:

$$\min \frac{1}{2}ww + \frac{C}{n}\Sigma_{i=1}^{n}\xi_i \tag{6.4.2}$$

$$\begin{aligned} s.t.\forall y : [\Sigma_{i=1}^{l}(x_i^1.W_{y_i^1}) + \varphi_t rans(y_{i-1}^1, y_i^1).W_{trans}] \\ \geq [\Sigma_{i=1}^{l}(x_i^n.W_{y_i}) + \varphi trans(y_{i-1}, y_i).W_{trans}] + \Delta(y^n, y) - \xi_n \end{aligned} \tag{6.4.3}$$

Here, C is the trades off parameter for the margin size and training error.

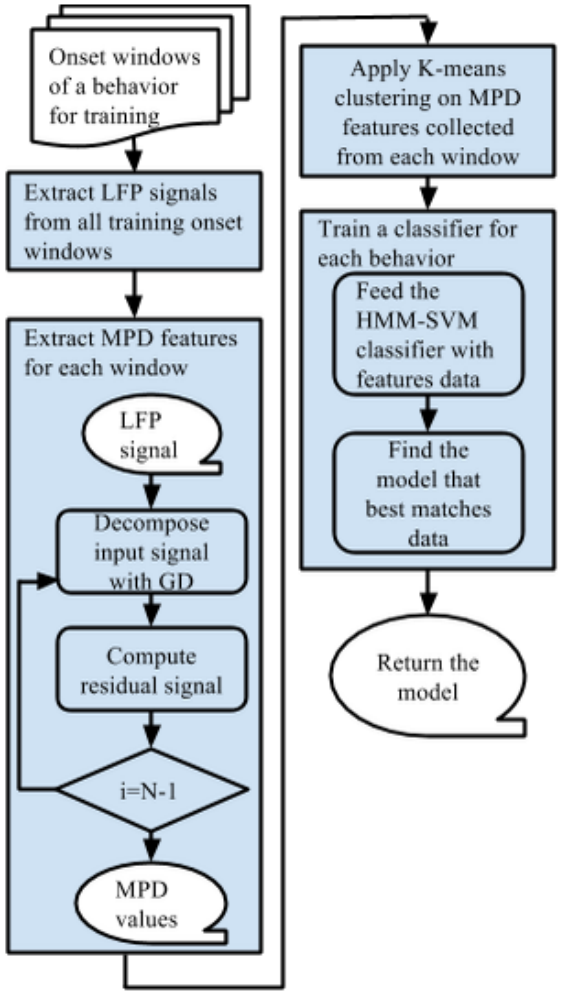


Figure 6.2: Flowchart demonstrating the train procedure for classification.

Chapter 7

Jointly Onset Detection and Classification

The proposed simultaneous onset detection and behavior classification method is a two-phase approach. In the first phase, the behavior onset detection is conducted. As the onset detection outcome, one of the sliding LFP signal windows is recognized as the onset of patient behavior. This window is further used in the second phase to classify the behavior type. The details of the behavior classification approach are elaborated as follows.

Similar to onset detection, we applied HMM to model behavioral patterns since it has inherent temporal transitional models. Due to the probability nature of HMM and the random assignment of prior probabilities, the HMMs may lead to slightly different classification results in multiple trials. To improve the performance and robustness of the classifier, c HMMs are trained for each behavior, i.e., for the mth behavior, c HMMs, say $\Lambda_1^m, \ldots, \Lambda_c^m$, are trained.

Consider a binary classification between the m_1th behavior and the m_2th behavior. Given the feature symbols $\mathbf{K}_{i,\hat{d}}^m$ from the detected onset window indexed by $\hat{d}$, the log-likelihood, $\eta_{i,c}^{m_1}$ returned by all HMMs corresponding to behavior m_1, construct the log-likelihood feature vector $\boldsymbol{\eta}_i^{m_1} = \{\eta_{i,1}^{m_1}\ \eta_{i,2}^{m_1}\ \cdots\ \eta_{i,c}^{m_1}\}$. We note that in $\eta_{i,c}^{m_1}$, the window index $\hat{d}$ is removed for simplicity. A similar procedure is conducted for HMMs trained on behavior m_2 to calculate the log-likelihood vector

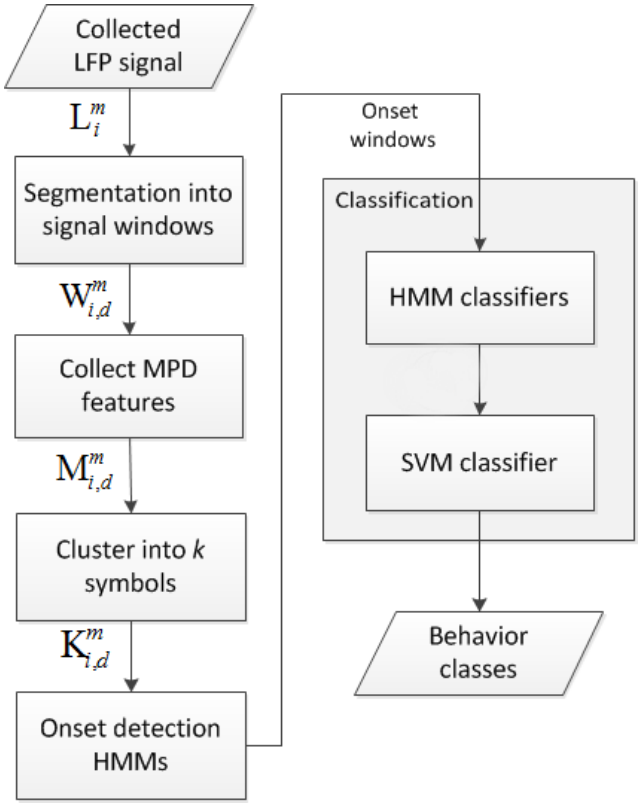


Figure 7.1: The flowchart on simultaneous detection of patient's onset and classification of behaviors.

$\boldsymbol{\eta}_i^{m_2}$. The log-likelihood feature vector of $\{\boldsymbol{\eta}_i^{m_1}|\forall i\}$ and $\{\boldsymbol{\eta}_i^{m_2}|\forall i\}$ were used for train and test an SVM model for classifying behavior m_1 and m_2. Similarly, we used 10-fold cross validation to separated training and testing sets, and binary classification was performed between different behaviors. The classification model will not only rely on one HMM, but also considers the log-likelihood value returned by a pool of HMMs. As a consequence, the recognition results are improved [44].

Our proposed approach aims to detect the onset and classify the behavior tasks simultaneously. The general flowchart of proposed approach is demonstrated in Figure 7.1. As an example, consider LFP signals collected for $m = I$, which is button press, and $m = II$ which stands for language behavior. The collected LFP sig-

nals, L_i^m, where $m = 1$ are segmented into consecutive windows $W_{i,d}^m$ for all trials $i = 1, .., I$ as shown in Figure 7.1. Then the MPD features, $M_{i,d}^m$, are collected from each window, and these features are quantized into cluster nodes, $K_{i,d}^m$. As a consequence, each window is represented as one-dimensional vector $K_{i,d}^m$.

For training the onset detection HMMs, 10 fold cross validation is used to select $\frac{9}{10}$ of onset windows to learn parameters of behavior $m = I$. The same procedure is performed for behavior $m = II$. In the next section, the onset windows are also used to train one-against-one SVM-HMM classifier to categorize different behavioral tasks as described previously. After learning the parameters of onset detection and behavior classification sections, the remaining $\frac{1}{10}$ fraction of trials from $m = I$ and $m = II$ are used for testing. The procedure for testing can be followed according to Figure 7.2.

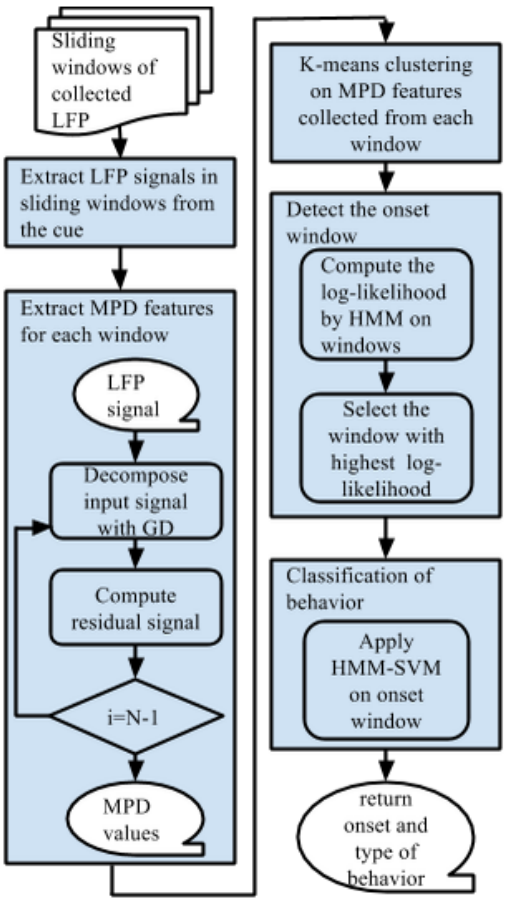


Figure 7.2: The flowchart displays the test procedure to jointly detection of patient's onset and classification of behaviors.

Chapter 8

Experimental Results

For demonstrating the effectiveness of our approach, we employed LFP signals collected from patients with PD in our experiments. All subjects provided informed consent for participation in this research study, in a manner approved by the internal review board of University of Washington (UW) and Colorado Neurological Institute (CNI) where the signals were collected. We used signal segments associated with *language*, *motor*, and combinational tasks of *language* and *motor* performed by PD patients during the data collection procedure. The sampling rate of our system is either $4.8kHz$, or $5kHz$, depending on the subjects and the place data was collected. The number of trials each patient performed the tasks in addition to other specific details of data collection is provided in Table 3.1.

8.1 Behavior Classification

We use the LFP signals collected as explained in Chapter 3, from patients with PD for demonstrating our approach. The signal segments associated with different behavioral tasks were labeled by physicians during data collection. The behavioral tasks are: motor task with condition code $m = I$, language and motor task with motor task onset $m = III$, and language task $m = IV$. The language tasks $m = II$ combines tasks III and IV.

The sampling rate of our system is $4kHz$, and for different behavioral tasks, the number of data segments varied from 80 to 109. We used K-fold cross validation

Table 8.1: Confusion matrix of PD patients behavior tasks using DHMMs by top-level SVM.

(a) language and motor task with motor onset $m = III$ vs language and motor task with speech onset$m = IV$

mth task	III	IV
III	**0.93**	0.07
IV	0.09	**0.91**

(b) motor task $m = I$ vs language $m = II$

mth task	I	II
I	**0.94**	0.06
II	0.10	**0.90**

(c) motor task $m = I$ vs language and motor task with speech onset $m = IV$

mth task	I	IV
I	**0.94**	0.06
IV	0.09	**0.91**

(d) motor task $m = I$ vs language and motor task with motor onset $m = III$

mth task	I	III
I	**0.89**	0.11
III	0.10	**0.90**

technique to separate the training and test data for the DHMMs, and four hidden states for the DHMMs. Linear SVM provided the best performance for both the top-level and hybrid approaches. Tables 8.1 and 8.2 provide the confusion matrices that summarize the classification results. On average, the classification rates are 92.02% and 92.1%, respectively.

Table 8.2: Confusion matrix of PD patients behavior tasks using hybrid HMM-SVM

(a) language and motor task with motor onset $m = III$ vs language and motor task with speech onset $m = IV$

mth task	III	IV
III	**0.92**	0.08
IV	0.08	**0.92**

(b) motor task $m = I$ vs language $m = II$

mth task	I	II
I	**0.93**	0.07
II	0.08	**0.92**

(c) motor task $m = I$ vs language and motor task with speech onset $m = IV$

mth task	I	IV
I	**0.94**	0.06
IV	0.08	**0.92**

(d) motor task $m = I$ vs language and motor task with motor onset $m = III$

mth task	I	III
I	**0.91**	0.09
III	0.10	**0.90**

8.2 Onset Detection

We collected LFP signal from the *cue* moment for 10 seconds by consideration of the success of all of the subjects to display the response during this time span. As explained in Chapter 3, the sample signal from a trial of behavior task is marked with the cue and patient's response onset in Figure 2.2. As the first step of the proposed model, we aim to detect the behavior onset position. Consequently, the signal is divided into small sliding windows of 2k samples, which have overlaps of 1k samples. The time lag between two consecutive sliding windows varies between 200 ms to 208 ms considering the frequency of data collection. This process is illustrated in Figure 4.1.

Later, we collected MPD features from the signal in each window. As MPD atoms are composed of four-dimensional parameter vectors (i.e. amplitude, variance scale, time-shift and frequency-shift), the K-means method is applied for clustering [27]. The quantified feature vectors resulted from K-means are used as features for HMM classifier. In Figure 8.1, the histogram of feature values or cluster nodes, is shown for all trials of two different behavioral tasks. As it is illustrated, the collected features are good representatives of each behavior condition, and specific feature values are explicit indication of a condition. For example, the first and second feature vector values with frequency of less than 50, only occurred in condition II trials. As a consequence they can be regarded as indication of condition II. The same conclusion is made for third and forth feature values that represent condition III.

In order to train the onset detection module, we used the windows which contain the onset to learn the HMM parameters for detection. We applied K-fold cross val-

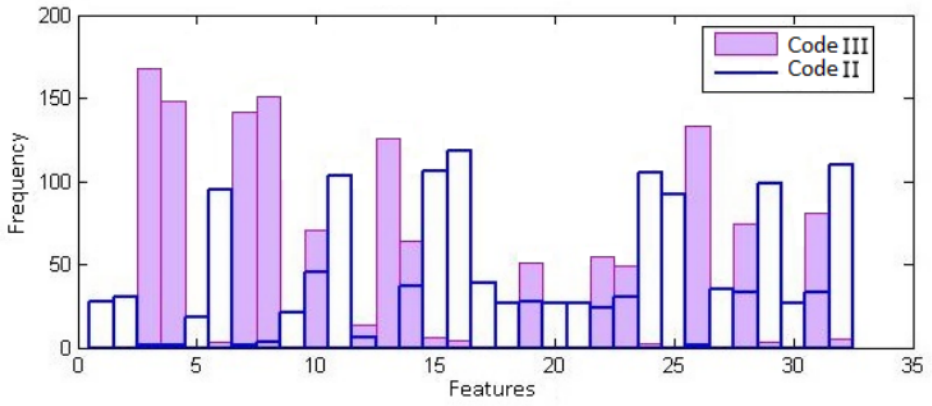


Figure 8.1: The histogram of feature symbols after applying $K-$means clustering method on MPD features extracted from patient "P1", for two different behaviors coded as II and III. As illustrated, different behaviors depict distinct features.

idation technique to separate the training and test set for each behavior condition [16]. In each fold, one of the trials was used for testing and the others to train.

Afterwards, we train the SVM-HMM classifier to classify different behavioral tasks. The probability nature of HMMs, causes the results of experiment vary for a small proportion in different trials. To avoid these variations and improve the robustness, several HMMs were trained and combined by a top layer SVM classifier. The two layer classifier has shown the best performance to classify and detect the onset [44, 45]. Here, we also applied k-fold cross validation technique to partition training and testing sets. For all the classification experiments, k is set to 10.

8.3 Jointly Detect the Onset and Classify Behavioral Tasks

For demonstrating the effectiveness of our approach, we present the following detection and classification results using LFP signals collected from PD patients. We used signal segments associated with language, motor, and combinational tasks performed by PD patients during the data collection procedure. The sampling rate

Table 8.3: Onset detection results using the Hidden Markov Models for 45 trials of behavioral task.

Code	**Patient ID**	**Trials**	**Average delay (second)**	**Standard deviation (second)**
I	P1	90	0.9	0.3
II	P1	45	0.8	0.4
III	P1	45	0.7	0.5
IV	P1	45	0.7	0.4
I	P2	55	0.8	0.4
II	P2	53	0.8	0.4
III	P2	54	0.8	0.4
IV	P2	54	0.9	0.3
V	P3	46	0.8	0.5
VI	P3	47	0.9	0.4
II	P3	43	0.8	0.5
V	P4	45	0.7	0.5
VI	P4	46	0.8	0.4
II	P4	45	0.8	0.4

of our system were either $4.8kHz$ or $5kHz$, depending on the subjects. The number of trials each patient performed during each task and other specific details of data collection are provided in Table 3.1.

The LFP signals were extracted after the cue moment for 10 seconds in all trials of behavioral tasks. LFP signals from trials of behavior task are marked with the cue and patient's response onset as shown in Figure 4.1. Consequently, the signal is divided into small sliding windows of 0.4 second, and each two consecutive windows have an overlap of approximately 0.2 second (49 windows overall). The time lag between two consecutive sliding windows is similar to the windows overlap, and is 0.2 second. This process is illustrated in Figure 4.1.

In the next step, we collected MPD features from the signal in each window. As MPD atoms are defined by time-frequency parameters, i.e., variance scale, time-shift and frequency-shift, the K-means method is applied for feature vector quantization. The quantified feature vectors resulted from K-means are used as feature symbols

for HMM classifiers. In order to train the onset detection HMMs, we used the windows which contain the onset to learn the HMM parameters for detection. We applied 10-fold cross validation technique to separate the training and test set for each behavior condition. In each fold, $\frac{1}{10}$ trials was used for testing and the others for training. This process is illustrated in the block diagram in Figure 8.2.

Afterwards, we train the SVM-HMM classifier to classify different behavioral tasks. Again, a number of HMMs were trained for each behavior and were combined by a top layer SVM classifier to avoid small proportion of differences of the results in different trials [44, 45]. Here, we also applied 10-fold cross validation technique to partition training and testing sets (See Figure 8.2).

Our experimental results for onset detection of behavior *language and button press with button press onset* for subject "P1" is demonstrated in Figure 8.3. The vertical axes of the figure shows the time delay and the horizontal axes represents the trial number. The time lag between two consecutive windows is approximately 200 ms, which is the time step for calculating delay. According to Figure 8.3, in nine trials the onsets were detected with no delay in the same window where they occurred. Also, in 38 trials, the onsets were detected with less or equal to 1.4 second delay. Table 8.3 summarized these results.

In Tables 8.4, 8.6, 8.5, and 8.7, the classification results for different behavioral tasks and subjects are presented. The detected onset windows by the HMM detector are used for classification. We tested classification performance of motor vs. language tasks. We set button press as $m = I$ representing motor behaviors, and merged behaviors with speech onset as $m = II + IV$, representing speech onset behaviors. These results are provided in Tables 8.4(a), 8.4(b), 8.5(a), and 8.5(b).

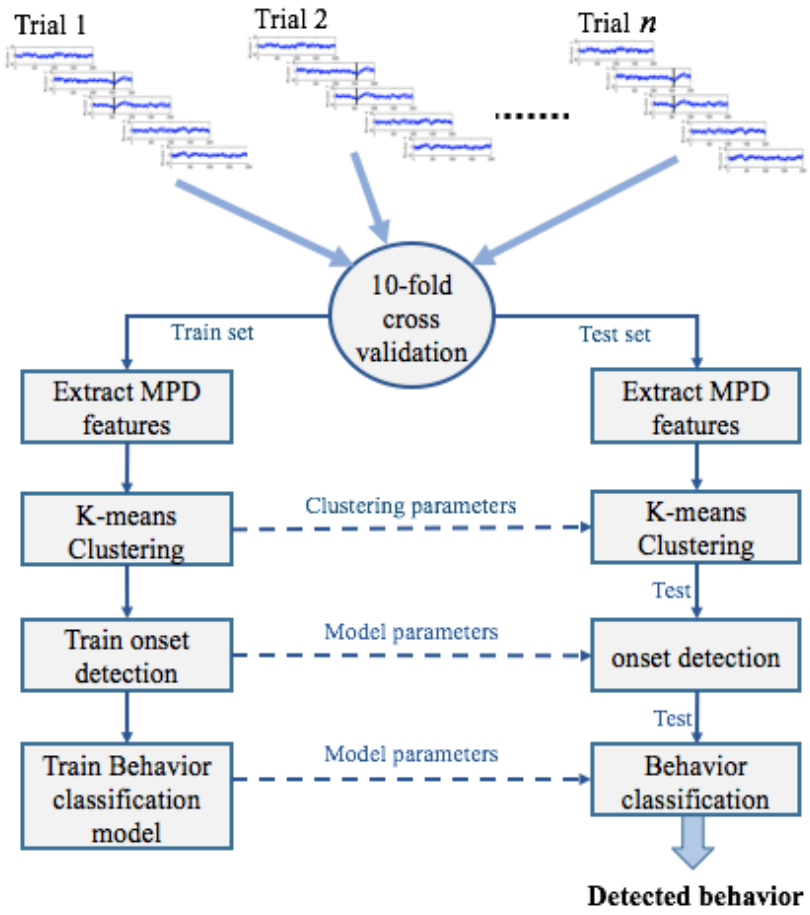


Figure 8.2: Block diagram of the implemented method. For each subject, the trials are divided into train and test sets. After extracting features, and clustering, the onset of the behavior is detected. Then, the recognized onset is classified into behavior classes.

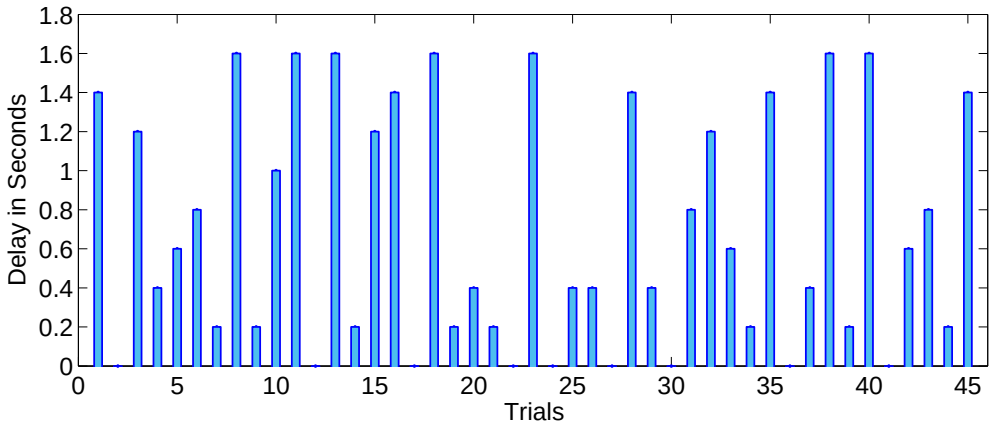


Figure 8.3: Time delay of onset detection for 45 trials. In 29 out of 45, the onset is detected within one second after the true onset.

We also assessed the proposed approach in classifying combinational tasks such as *language and button press with button press onset*, vs. non-combinational tasks such as *language* only. As shown in confusion matrices provided in Tables 8.6(a), and 8.6(b), we are able to classify tasks with an accuracy of at least 79%, and 86% for subjects "P2", and "P1", respectively. According to Table 8.4, and 8.6, subject "P1"displays higher classification results comparing to all other subjects.

For subjects "P3"and "P4", we merged the *left* and *right* hand recordings for classifying motor tasks vs. language tasks. We performed another assessment of the approach by classifying *left* vs. *right* hand *button press* for subjects "P3"and "P4", as provided in Tables 8.7(a) and 8.7(b).

Table 8.4: Confusion matrix of PD patient (a) "P2", and (b) "P1" behavior tasks using combined onset detection and classification method

(a) button press onset $m = I$, vs. speech onset $m = II + IV$

mth task code	I	II+IV
I	**0.88**	0.12
II+IV	0.14	**0.86**

(b) button press onset $m = I$, vs. speech onset $m = II + IV$

mth task code	I	II+IV
I	**0.93**	0.07
II+IV	0.09	**0.91**

Table 8.5: Confusion matrix of PD patients (a) "P3", and (b) "P4" behavior tasks using combined onset detection and classification method

(a) button press onset $m = V{+}VI$ vs. speech onset $m = II$

mth task code	V+VI	II
V+VI	**0.78**	0.22
II	0.26	**0.74**

(b) button press onset $m = V{+}VI$ vs. speech onset $m = II$

mth task code	V+VI	II
V+VI	**0.75**	0.25
II	0.27	**0.73**

Table 8.6: Confusion matrix of PD patient (a) "P2", and (b) "P1" behavior tasks using combined onset detection and classification method

(a) speech onset $m = II$, vs. langauge and button press, with button press onset $m = III$

mth task code	II	III
II	**0.81**	0.19
III	0.21	**0.79**

(b) speech onset $m = II$, vs. langauge and button press, with button press onset $m = III$

mth task code	II	III
II	**0.90**	0.10
III	0.14	**0.86**

Table 8.7: Confusion matrix of PD patients (a) "P3", and (b) "P4" behavior tasks using combined onset detection and classification method

(a) left hand button press onset $m = V$ vs. right hand button press onset $m = VI$

mth task code	V	VI
V	**0.75**	0.25
VI	0.21	**0.79**

(b) left hand button press onset $m = V$ vs. right hand button press onset $m = VI$

mth task code	V	VI
V	**0.78**	0.22
VI	0.24	**0.76**

8.4 Discussion of the Results

In this thesis, we investigated the detection of behavioral tasks onset, as well as classification of behaviors using the stimulation electrodes applied for DBS surgery.

We suggested that LFP signals are sufficient for behavior onset detection and classification, and tested this hypothesis on four PD patients. This work is considered as a difficult problem, since any flow or drawback in detecting the onset position will also lead to misclassification of the behavioral tasks. Also, there is a high similarity between the features extracted from combinational tasks, for example, *language and button press with button press onset* is similar to *language* in terms of collected features, and this fact deteriorates the classification performance. In this section, we aim to evaluate and provide a discussion on experimental results and possible errors causing misclassification.

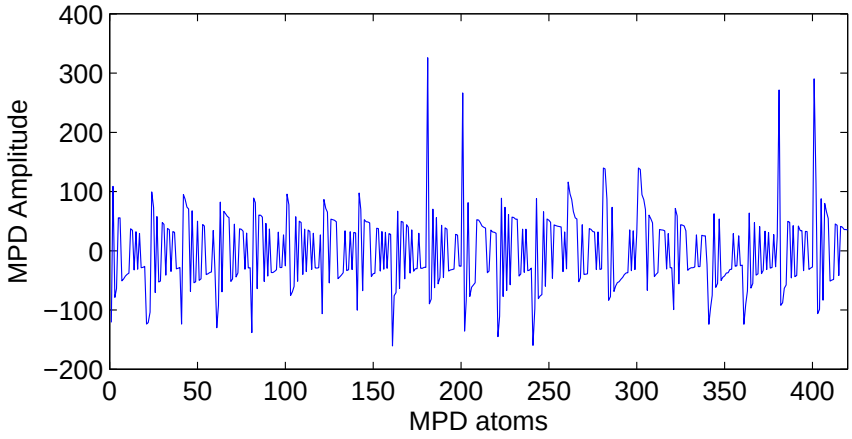
In the results, we demonstrated the classification performance of proposed method on different patients' data. Aside from the strength of the proposed classification approach, one of the main reasons for functionality of the suggested method is the informative collected features. In Figure 8.1, the histogram of the feature symbols, is shown for all trials of two different behavioral tasks. As illustrated, the collected features are able to well represent each behavior condition, and certain feature values are explicit indication of a behavior. In Figure 8.1, comparing the two behaviors, the first and second feature vector values with frequency of less than 50, only occur in *language* behavior trials. The same observation is made for third and forth feature values that represent *language and button press with button press onset*.

In order to analyze the onset detection procedure more precisely, we explored the MPD atoms' parameter vectors in the onset windows. In agreement with our findings, the onset windows were characterized to have greater peaks of amplitude comparing to non-onset ones. Another observation indicates that the onset windows also have the higher frequency shifts in the same atoms. These features can be regarded as justifications for the experimental results. In Figure 8.4(a) and 8.4(b),

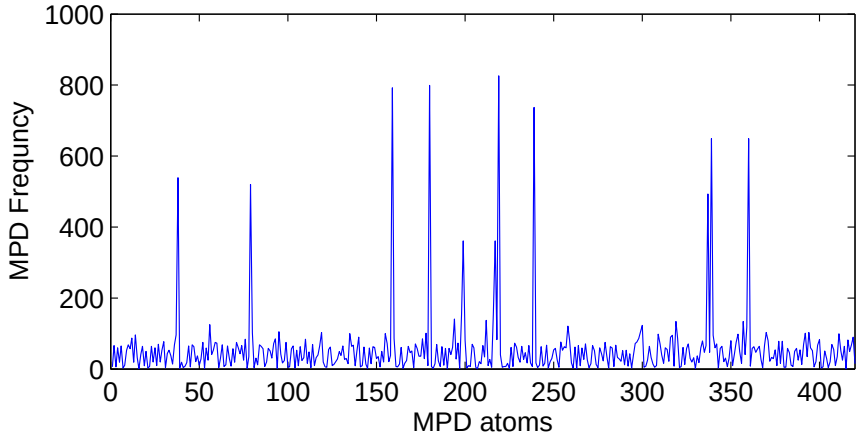
these characteristics are shown from a sample trial. In these figures, 21 sliding windows with 20 MPD iterations are shown. The onset window has higher peaks in MPD atoms' amplitudes and frequency shifts comparing to windows without onset.

We also noticed that there are possible differences in data collection procedure between the subjects. A subject's limited attention in an operating room during the recording, may influence the quality of our recordings. For example, during a *language* task, a subject may have moved his hand. Furthermore, DBS lead locations were driven by clinical benefits and may vary in location between subjects. This unavoidable differences may affect the onset detection results and cause the onset detector to recognize undesirable windows as the onset. As a consequence, the behavioral task classification will be affected. The mentioned causes as well as differences in behavior task and subject variability in responding to cue may cause slight difference in the classification results of Tables 8.4 comparing to Tables 8.6.

Due to the time complexity of feature extraction algorithm, the proposed method cannot be run in real time. However, it is considered as a tool that can be used to monitor patients activity off line. Moreover, it depicts the effectiveness and strength of collected LFP signals to observe different activities performed by the patients. According to the experimental results, it can successfully detect the onset of behaviors and classify them into predefined groups of behavioral tasks.



(a)



(b)

Figure 8.4: The amplitude and frequency of MPD coefficients for 21 consecutive sliding windows is shown for subject "P1". The length of each window was set as 0.4 second . The MPD parameter P was set to 20 for all 21 windows, which is the number of iterations. (a) shows the amplitude of MPD atoms. Amplitude peaks occur around atom number 200 (window number 10) which is the patient's onset. (b) shows the frequency shift of MPD atoms.

Chapter 9

Conclusion and Future Works

In this work, we proposed and investigated new approaches for classification of behavior tasks performed by PD patients, automatically detecting the onset of the behaviors, as well as simultaneous onset detection and classification of behaviors. We collected STN LFP signals from PD patients who underwent DBS surgeries. We applied MPD feature extraction method, and proposed learning models in order to detect the onset of the behavioral tasks. In addition we suggested classification models to improve the classification rate of behavioral task into 90%, comparing to the previous works.

Further, we employed the detected onsets to feed a two-layer classification model to recognize the type of PD patients behavior tasks in addition to the onset of behavior. Our result depict a reliable approach which is able to detect patient's onset with average delay of 1.43 seconds. It is also capable of recognizing behavior tasks with detection rate of 84% on average.

In the current data, there are six streams of LFP signal recording available in addition to the EEG signals, which can provide more informative features for the onset detection, and classification procedure. As the future work to improve the system detection rate and accuracy, we suggest to fuse and combine multiple recorded signals, and design a multi feature learning model which are able to detect the most informative features automatically and improve the recognition rates.

Bibliography

[1] Deep brain stimulation (dbs). http://www.medtronic.eu/your-health/parkinsons-disease/, urldate = 2010-09-22.

[2] Y. Altun, I. Tsochantaridis, and T. Hofmann. Hidden markov support vector machines. 2003.

[3] R. Amirnovin, Z.M. Williams, G.R. Cosgrove, and E.N. Eskandar. Visually guided movements suppress subthalamic oscillations in parkinson's disease patients. *The Journal of neuroscience*, 24(50):11302–11306, 2004.

[4] J. Ärnlöv, A. Larsson, et al. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the global burden of disease study 2013. *The Lancet*, 2014.

[5] M.G. Baker and L. Graham. The journey: Parkinson's disease. *BMJ: British Medical Journal*, 329(7466):611, 2004.

[6] S. Basu, A. Banerjee, and R. Mooney. Semi-supervised clustering by seeding. In *In Proceedings of 19th International Conference on Machine Learning (ICML-2002*. Citeseer, 2002.

[7] J.P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M.B. Sandler. A tutorial on onset detection in music signals. *Speech and Audio Processing, IEEE Transactions on*, 13(5):1035–1047, 2005.

[8] C.M. Bishop et al. *Pattern recognition and machine learning*, volume 4. springer New York, 2006.

[9] G. Buzsáki, C.A. Anastassiou, and C. Koch. The origin of extracellular fields and currentseeg, ecog, lfp and spikes. *Nature reviews neuroscience*, 13(6):407–420, 2012.

[10] M. Cassidy, P. Mazzone, A. Oliviero, A. Insola, P. Tonali, V. Di Lazzaro, and P. Brown. Movement-related changes in synchronization in the human basal ganglia. *Brain*, 125(6):1235–1246, 2002.

[11] V.S. Chakravarthy, D. Joseph, and R.S. Bapi. What do the basal ganglia do? a modeling perspective. *Biological cybernetics*, 103(3):237–253, 2010.

[12] R. Cilia, C. Siri, G. Marotta, D. De Gaspari, A. Landi, C.B. Mariani, R. Benti, I.U. Isaias, F. Vergani, G. Pezzoli, et al. Brain networks underlining verbal fluency decline during stn-dbs in parkinson's disease: an ecd-spect study. *Parkinsonism & related disorders*, 13(5):290–294, 2007.

[13] N. Crone, L. Hao, J. Hart, D. Boatman, R.P. Lesser, R. Irizarry, and B. Gordon. Electrocorticographic gamma activity during word production in spoken and sign language. *Neurology*, 57(11):2045–2053, 2001.

[14] C.A. Davie. A review of parkinson's disease. *British medical bulletin*, 86(1):109–127, 2008.

[15] E. Edwards, M. Soltani, W. Kim, S.S. Dalal, S.S. Nagarajan, M.S. Berger, and R.T. Knight. Comparison of time–frequency responses and the event-related potential to auditory speech stimuli in human cortex. *Journal of neurophysiology*, 102(1):377–386, 2009.

[16] B. Efron and R. Tibshirani. Improvements on cross-validation: the 632+ bootstrap method. *Journal of the American Statistical Association*, 92(438):548–560, 1997.

[17] T. Engel, J.J. Lucas, P. Gómez-Ramos, M.A. Moran, J. Ávila, and F. Hernández. Cooexpression of ftdp-17 tau and gsk-3β in transgenic mice induce tau polymerization and neurodegeneration. *Neurobiology of aging*, 27(9):1258–1268, 2006.

[18] N. Grira, M. Crucianu, and N. Boujemaa. Active semi-supervised fuzzy clustering. *Pattern Recognition*, 41(5):1834–1844, 2008.

[19] M. Hariz. Twenty-five years of deep brain stimulation: Celebrations and apprehensions. *Movement Disorders*, 27(7):930–933, 2012.

[20] A.O. Hebb, F. Darvas, and K.J. Miller. Transient and state modulation of beta power in human subthalamic nucleus during speech production and finger movement. *Neuroscience*, 202:218–233, 2012.

[21] A.O. Hebb, J.J. Zhang, M.H. Mahoor, C. Tsiokos, C. Matlack, H.J. Chizeck, and N. Pouratian. Creating the feedback loop: Closed-loop neurostimulation. *Neurosurgery clinics of North America*, 25(1):187–204, 2014.

[22] Deep brain stimulation (dbs), http://www.mayfieldclinic.com/PE DBS.htm.

[23] H. Jiang, J.J. Zhang, A. Hebb, and M.H. Mahoor. Time-frequency analysis of brain electrical signals for behavior recognition in patients with parkinsons disease. In *Asilomar Conference on Signals, Systems and Computers*, November 2013.

[24] F. Kempf, A.A. Kühn, A. Kupsch, C. Brücke, L. Weise, G. Schneider, and P. Brown. Premovement activities in the subthalamic area of patients with

parkinson's disease and their dependence on task. *European Journal of Neuroscience*, 25(10):3137–3145, 2007.

[25] R. Levy, P. Ashby, W.D. Hutchison, A.E. Lang, A.M. Lozano, and J.O. Dostrovsky. Dependence of subthalamic nucleus oscillations on movement and dopamine in parkinsons disease. *Brain*, 125(6):1196–1209, 2002.

[26] S. Little, A. Pogosyan, S. Neal, B. Zavala, L. Zrinzo, M. Hariz, T. Foltynie, P. Limousin, K. Ashkan, J. FitzGerald, et al. Adaptive deep brain stimulation in advanced parkinson disease. *Annals of neurology*, 74(3):449–457, 2013.

[27] D. MacKay. *Information theory, inference, and learning algorithms*, volume 7. Citeseer, 2003.

[28] S.G. Mallat and Z. Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, 41:3397–3415, 1993.

[29] Medtronic. Deep brain stimulation for movement disorders, http://professional.medtronic.com/pt/neuro/dbs-md/prod/.

[30] K.J. Miller, G. Schalk, E.E Fetz, M. den Nijs, J.G. Ojemann, and R. Rao. Cortical activity during motor execution, motor imagery, and imagery-based online feedback. *Proceedings of the National Academy of Sciences*, 107(9):4430–4435, 2010.

[31] K.J. Miller, S. Zanos, E.E. Fetz, M. Den Nijs, and J.G. Ojemann. Decoupling the cortical power spectrum reveals real-time representation of individual finger movements in humans. *The Journal of neuroscience*, 29(10):3132–3137, 2009.

[32] C. Moreau, L. Defebvre, A. Destee, S. Bleuse, F. Clement, J.L. Blatt, P. Krystkowiak, and D. Devos. Stn-dbs frequency effects on freezing of gait in advanced parkinson disease. *Neurology*, 71(2):80–84, 2008.

[33] S. Pan, K. Warwick, J. Stein, M.N. Gasson, S.Y. Wang, T.Z. Aziz, J. Burgess, et al. Prediction of parkinsons disease tremor onset using artificial neural networks. In *Proceedings of the fifth IASTED International Conference: biomedical engineering*, pages 341–345. ACTA Press, 2007.

[34] M. Rosa, M. Arlotti, G. Ardolino, F. Cogiamanian, S. Marceglia, A. Di Fonzo, F. Cortese, P.M. Rampini, and A. Priori. Adaptive deep brain stimulation in a freely moving parkinsonian patient. *Movement Disorders*, 30(7):1003–1005, 2015.

[35] M. Rousseaux, P. Krystkowiak, O. Kozlowski, C. Özsancak, S. Blond, and A. Destée. Effects of subthalamic nucleus stimulation on parkinsonian dysarthria and speech intelligibility. *Journal of neurology*, 251(3):327–334, 2004.

[36] A. Samii, J.G. Nutt, and B.R. Ransom. Parkinson's disease. *The Lancet*, 363(9423):1783–1793, 2004.

[37] H. Scherberger, M.R. Jarvis, and R.A. Andersen. Cortical local field potential encodes movement intentions in the posterior parietal cortex. *Neuron*, 46(2):347–354, 2005.

[38] A. Stocco, C. Lebiere, and J.R. Anderson. Conditional routing of information to the cortex: A model of the basal ganglias role in cognitive coordination. *Psychological review*, 117(2):541, 2010.

[39] N. Thakor. Building brain machine interfaces–neuroprosthetic control with electrocorticographic signals. *Newsletter*, 2015, 2015.

[40] V. Vapnik, S.E. Golowich, and A. Smola. Support vector method for function approximation, regression estimation, and signal processing. In *Advances in Neural Information Processing Systems 9*, pages 281–287. MIT Press, 1996.

[41] V. Vu, N. Labroche, and B. Bouchon-Meunier. Active learning for semi-supervised k-means clustering. In *Tools with Artificial Intelligence (ICTAI), 2010 22nd IEEE International Conference on*, volume 1, pages 12–15. IEEE, 2010.

[42] P. Watson and E.B. Montgomery. The relationship of neuronal activity within the sensori-motor region of the subthalamic nucleus to speech. *Brain and language*, 97(2):233–240, 2006.

[43] D. Wu, K. Warwick, Z. Ma, M. N. Gasson, J. G. Burgess, S. Pan, and T. Z. Aziz. Prediction of parkinson's disease tremor onset using a radial basis function neural network based on particle swarm optimization. *International journal of neural systems*, 20(02):109–116, 2010.

[44] N. Zaker, A. Dutta, A. Maurer, J.J. Zhang, S. Hanrahan, A.O. Hebb, N. Kovvali, and A. Papandreou-Suppappola. Adaptive learning of behavioral tasks for patients with parkinson's disease using signals from deep brain stimulation. In *Signals, Systems and Computers, 2014 48th Asilomar Conference on*, pages 208–212, Nov 2014.

[45] N. Zaker, J.J. Zhang, S. Hanrahan, J. Nedrud, and A.O. Hebb. A new approach for automated detection of behavioral task onset for patients with parkinson's disease using subthalamic nucleus local field potentials. In *2015 49th Asilomar Conference on Signals, Systems and Computers*, pages 780–784. IEEE, 2015.