Single Trial Behavioral Task Classification Using Subthalamic Nucleus Local Field Potential Signals*

Soroush Niketeghad, Adam O. Hebb, Joshua Nedrud,
Sara J. Hanrahan, Mohammad H. Mahoor, *Senior Member, IEEE*

***Abstract*— Deep Brain Stimulation (DBS) has been a successful technique for alleviating Parkinson's disease (PD) symptoms especially for whom drug therapy is no longer efficient. Existing DBS therapy is open-loop, providing a time invariant stimulation pulse train that is not customized to the patient's current behavioral task. By customizing this pulse train to the patient's current task the side effects may be suppressed. This paper introduces a method for single trial recognition of the patient's current task using the local field potential (LFP) signals. This method utilizes wavelet coefficients as features and support vector machine (SVM) as the classifier for recognition of a selection of behaviors: speech, motor, and random. The proposed method is 82.4% accurate for the binary classification and 73.2% for classifying three tasks. These algorithms will be applied in a closed loop feedback control system to optimize DBS parameters to the patient's real time behavioral goals.**

***Index Terms*— Deep Brain Stimulation, Parkinson's Disease, Support Vector Machine, Wavelet Transform**

I. Introduction

Parkinson's Disease (PD) is a movement disorder characterized by tremor, rigidity and bradykinesia (slow movements) and is caused by the premature death of dopaminergic neurons in the brain. PD increases with advancing age and raises among people in their 60s and 70s. There is no cure for PD and instead therapy is directed at treating the motor manifestations of PD such as tremor and rigidity.

The basal ganglia is a collection of deep brain nuclei that are involved in parallel processing loops involving the cerebral cortex, thalamus, and spinal cord for motor, associative, and limbic functions. The motor processing loop is also called the "extrapyramidal motor system", and modulates motor control of basic movements such as posture, reaching, walking, speech, and others. In neurological disorders such as PD, there is disrupted regulation of basal ganglia nuclei such that nuclei are either overactive or underactive.

*This research is partially supported by a grant from the Knoebel Institute for the Longevity and Health at University of Denver..

Soroush Niketeghad is a research assistant at electrical and computer engineering dept., University of Denver, Denver CO (email: soroush.niketeghad@du.edu)

Adam Hebb is a neurosurgeon at Colorado brain and spine institute, Englewood CO, and a research scholar at electrical and computer engineering dept., University of Denver, Denver CO (email: adam.hebb@aoh.md)

Joshua Nedrud is is an Investigator Scientist at Colorado Neurological Institute, Englewood, CO (email: jnedrud@thecni.org)

Sara J. Hanrahan is an Investigator Scientist at Colorado Neurological Institute, Englewood, CO shanrahan@thecni.org)

Mohammad H. Mahoor is an assistant professor at electrical and computer engineering dept., University of Denver, Denver CO (email: Mohammad.Mahoor@du.edu)

Deep Brain Stimulation (DBS) is an advanced FDA-approved therapeutic technique for alleviating the PD symptoms especially for whom drug therapy is no longer efficient. DBS of the subthalamic nucleus (STN-DBS) improves motor signs of PD and permits reduction of dopaminergic medication. DBS provides relief of PD's motor signs, but may have cognitive, speech, and balance side effects. Existing DBS therapy is open-loop, providing a time invariant stimulation pulse train that is not customized to a patient's current behavioral task goals. By customizing DBS therapy to a patient's task using signal processing methods, these side effects of stimulation may arise only when they are non-detrimental to the patient's current goals. This could therefore allow more aggressive DBS parameters and closed loop control to reduce the risk for therapy limiting side effects.

The electrical potential recorded from the implanted DBS electrodes may provide substantial information related to a subject's behavioral goals, and therefore is a candidate signal for an adaptive or closed loop DBS system. Further, electrical potentials recorded from the cerebral cortex have successfully been used in brain-computer interface tasks. This paper aims at exploring the subcortical electrical potentials in the human brain, obtained during surgical implantation of a DBS system and classification of three different tasks (speech, motor and random) as a part of designing a closed loop DBS system.

LFP recordings, which represent coherent activity of small cell assemblies, have been used in humans to characterize activity within cortical regions and subcortical nuclei [1]. Time-frequency analysis of motor cortex ECoG [2] and subthalamic nucleus (STN) LFPs [3], has revealed characteristic suppression of beta (13-30Hz) band and augmentation of gamma (30-70Hz) band power preceding and during motor behaviors. Thus, the human subthalamic nucleus in PD exhibits oscillatory behavior in a broad frequency band is modulated by motor activity.

A system for recognition of patient's activities by analyzing the LFP brain signals is introduced. We use the LFP signals collected from nine patients with PD. Wavelet coefficients are the features that describe the LFP signal.and support vector machine (SVM) is the classifier that predicts the patient's task.

Section II illustrates the recording procedure and details of STN LFP signals. Section III describes the methods used for classification including wavelet transform, analysis of the LFP data, and SVM classifier structure. Section VI includes classification results for SVM and k-nearest neighbor (KNN)

Fig. 1. The electrode tip is placed precisely in the subthalamic nucleus (STN). A coil of lead wire is left under the skin for later attachment to the stimulator. Adapted from Figure 6 in [4] with permission.

Fig. 2. Schematic representation of recording electrodes used for LFP recordings. (above) Medtronic 3389 DBS lead (reprinted with permission of Medtronic Inc. © 2008). (below) Alpha omega neuroprobe microelectrode (reprinted with permission of Alpha Omega, USA © 2011).

classifier. Section V discusses the results provided by different classifiers and concludes the presented method.

II. Data Recordings

Nine subjects undergoing DBS were enrolled in a prior study of LFP recording during behavioral tasks (see Figure 1). Fourteen recordings were performed which there were seven left, one right and four bilateral recordings. All subjects were in the off medication state [5]. Subjects provided informed consent for participation in this research study, in a manner approved by the institutional review board at University of Washington.

Paired microelectrodes (pME) or the DBS lead were used for LFP recording(as seen in Table I). For pME, macro collars on the microelectrodes were used (Alpha Omega, Israel). After optical isolation and amplification, the signals were digitized (4 kHz), band pass filtered (5-300 Hz), and combined with event markers and subject response signals. The microelectrode recording (MER) guide tube was used as common reference. For the DBS lead recording, four contacts were used. Signals were amplified, digitized (5 kHz), band pass filtered (1-1000 Hz), and combined with event markers and subject response signals. A linked mastoid common reference was used for recording.

Behaviors included motor and speech tasks. The motor task block consisted of cued repetitions of a button press using either the ipsilateral or contralateral thumb. For speech initiation task subjects were asked to name the months of the year. Speech tasks were also completed in a block of several repetitions. For task initiation and completion, subjects received an audio cue from a laptop computer.

III. Classification Method

The LFP data was analyzed in time-frequency domain using continuous wavelet transform. Proper wavelet coefficients were selected based on the analysis and fed to an SVM based classifier.

A. Wavelet Transform

Wavelet analysis methods have been widely used in biomedical applications [6]. Wavelet transform allows us to see changes in frequency over time and has been used for non-stationary signals such as event related synchronization/desynchronization (ERS/ERD) analysis [7]. The wavelet transform can be defined as [8]:

$$g(a,b) = a^p \int_{-\infty}^{\infty} g(t)\psi(\frac{t-b}{a})dt \tag{1}$$

where $\psi(t)$ is the analyzing wavelet, b is a time-like translation variable, a is dimensionless frequency scale variable, and p is a real normalization parameter. One of the popular

TABLE I

Recording information for each subject

Recording	Subject	Side	Design	Speech trials	Motor trials
1	1	Bi	DBS	54	59
2	2	Bi	DBS	75	60
3	3	Lt	pME	107	59
4	3	Lt	DBS	58	29
5	4	Bi	DBS	117	53
6	5	Bi	DBS	57	59
7	5	Lt	pME	30	29
8	6	Lt	DBS	117	53
9	6	Lt	pME	53	22
10	7	Rt	pME	81	56
11	7	Lt	pME	111	60
12	8	Lt	DBS	118	58
13	9	Lt	pME	111	58
14	9	Rt	pME	107	55

Fig. 3. Time-frequency representation of averaged z-scored power spectral density (PSD) for motor task (above) and speech (bellow). Time 0 corresponds to the onset of button press for the motor task and onset of speaking for the speech task. Dotted lines display the segments extracted for classification.

wavelets in biomedical signal processing is complex Morlet wavelet which is the product of a Gaussian and a sinusoidal function:

$$\psi(t) = e^{(-t/c)^2} e^{i2\pi f_0 t} \tag{2}$$

B. Analysis

For this work, time-dependent power spectral density estimates were calculated for LFP data and averaged for each task. For the speech task, voice was recorded and used to detect the onset and offset of each speech trial. For the motor task a digital channel input from a button press is utilized to detect the onset of movement. As illustrated in Figure 3, for both motor and speech tasks there is beta power suppression proceeding and during the tasks. For the motor task, beta suppression is followed by a considerable increase right after movement, while for the speech task, this increase is not considerable.

C. Support Vector Machines (SVM)

SVM is a state-of-the-art method for classification and regression which was introduced by Cortez and Vapnik [9]. It has been used widely in pattern recognition and brain computer interfaces [10]. SVM finds an optimal hyperplane that separates the two training classes. As seen in Figure 4 the samples of a class located in the closest distance of the other class are called support vectors and the margin is the distance between the hyperplane and the support vectors. SVM orients the margin in a way that it is maximized.

Given (x_i, y_i), $i = 1, 2, \ldots, N$ as N training set of samples where $x_i \in R^d, y_i \in \{1, +1\}$, SVM solves the following optimization problem described in [11]:

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}\mathbf{w}^t\mathbf{w} + C\sum_{i=1}^{N}\xi_i \\ &\text{Subject to} \quad y_i(\mathbf{w}^t\phi(x_i) + w_0) \geq 1 - \xi_i, \quad \xi_i \geq 0 \end{aligned} \tag{3}$$

The function $\phi(.)$ maps the vectors x_i in another space so that $\phi(x_i)$'s are linearly separable. $C \geq 0$ is the penalty parameter of the error term. Lagrangian method is used to

Fig. 4. SVM uses risk optimization to compare various separating hyperplanes and chooses the model with the largest margin of separation [12].

solve the optimization problem. One maximizes the dual variable lagrangian:

$$\begin{aligned} &\text{Maximize} \quad \sum_{i=1}^{N} \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j \mathbf{x_i}^t \mathbf{x_j} \\ &\text{Subject to} \quad 0 \leq \lambda_i \leq C, \quad \sum_{i=1}^{N} \lambda_i y_i = 0 \end{aligned} \tag{4}$$

A kernel function is defined as $K(\mathbf{x_i}, \mathbf{x_j}) = \phi(\mathbf{x_i})^t \phi(\mathbf{x_i})$. The Radial Basis Function (RBF) is given as $K(\mathbf{x_i}, \mathbf{x_j}) = \exp \gamma ||\mathbf{x_i} - \mathbf{x_j}||^2$.

A proper parameter setting improves SVM classification accuracy. There are two parameters to be set in the SVM model with RBF kernel: C and γ. Instinctively the γ parameter defines the distance a single training example can reach, which low values correspond to far distances and vice versa. The C parameter trades off training examples misclassification against decision surface simplicity. A low C ensures a smooth decision surface while a high C attempts to classify training examples correctly. Experiments are undertaken to evaluate SVM performance through variations of the C and γ parameters.

IV. Experimental Results

A. Preprocessing

Both training and testing data were low-pass filtered using a 80 order butterworth anti-aliasing filter and downsampled to 100Hz (from 5 kHz or 4kHz) to avoid high computational load. In the case of DBS lead recordings, LFP channels were subsequently bipolar re-referenced (0-1, 1-2, 2-3) .

B. Feature Extraction

Wavelet coefficients corresponding to the frequencies between 8 Hz and 30 Hz with the frequency interval of 1 Hz and a time window from 500 ms before onset to 3500 ms after onset were used as features for motor and speech task samples. For the random segments, a time window with the same size is applied to the random parts of signal. Note that the time frequency representation of a random segment can be any randomly shifted rectangle in Figure 3 along the time axis.

C. Classification

LibSVM toolbox is used for classification [13]. Empirically (using the evaluation data) C and γ parameters are assigned to be 500 and 1 respectively. Three binary classifiers are used for different permutations of tasks (i.e. motor Vs. speech, motor Vs. random, and speech Vs. random). Also a three class classification was done for all the tasks. The results of the SVM classifier is compared with k-nearest neighbor (KNN) classifier [14] with three different k values($k = 1, 3, 5$) and euclidean distance. In both methods, principal component analysis (PCA) [15] was performed before classification for reducing the dimensionality of data from 9200 to 47 features. For the evaluation of the classifiers, 10-fold cross validation method is used and the results are presented in Table II.

TABLE II

Averaged percentage of classification accuracies for SVM and KNN classifiers

	SVM		KNN		
	linear	RBF	k=1	k=3	k=5
speech Vs. motor	**81.44**	81.36	66.20	66.42	66.90
speech Vs. random	**81.69**	76.56	76.72	69.46	67.35
motor Vs. random	**84.08**	82.03	76.93	74.89	73.26
all three classes	**73.24**	66.50	62.69	58.75	57.23

Two different kernel functions has been used for SVM classification: RBF and linear. The RBF function is discussed in Section III. The linear function is claculated by the inner product of two vectors: $k(\mathbf{x}, \mathbf{y}) = \mathbf{x}^t\mathbf{y} + c$ where c is an optional constant. Figure 5 presents the comparison results.

Fig. 5. Three class classification results for linear and RBF kernel SVM.

V. Discussions and Conclusions

The main finding of this study is the characteristic time-frequency patterns of STN-LFP signals during particular tasks which lead to designing a single trial classification system for decoding patient's behavioral goals. This can further be a part of a closed-loop DBS system that generates proper stimulation pulse train optimized for patient's current behavioral goal. The proposed method which uses continuous wavelet coefficients as features and SVM as classifier, is able to perform binary classification with an average accuracy of 82.40% and three class classification 73.24% accurate.

Table II shows that SVM outperforms KNN in accuracy. For KNN, as dimensionality increases, the distant to the nearest data point approaches to the distance to the furthest data point. Therefore high number of features decreases the performance of the KNN method [16]. Also due to the high dimensionality, linear kernel function provides a more proper mapping rather than RBF [17].

The variety of behavioral tasks in this research was limited by the restricted time in the operating room to perform the tasks. The reason behind using random segments for classification is to train the classifier to distinguish "other tasks" rather than speech and motor. The classification results for different pairs of tasks (rows 1 to 3 in Figure II) show that the classifiers performances for speech Vs. motor is nearly the same when one is random. This verifies that random segments are proper representations for "other tasks".

Figure 5 demonstrates that some recordings lead to a highly accurate classification (recording 5) while for some others the results are relatively poor (recordings 3, 9, and 10). Even for different recordings of the same subject (recordings 8 and 9 from subject 6) classification accuracy is considerably different (see Table I for information about recordings). This may be the result of different recording electrode positioning.

The classification results show that SVM classier is more accurate in distinguishing all the tasks rather than KNN. Also linear kernel function provides a better mapping for the SVM rather than the RBF function and can be considered a proper tool for task recognition for a closed loop DBS system.

The proposed classification technique can be utilized as the initial step of designing a high level DBS system. Recognizing patient's current task can be led to an optimal DBS parameter adjustment to decrease the side effects.

References

[1] A. K. Engel, C. K. Moll, I. Fried, and G. A. Ojemann, "Invasive recordings from the human brain: clinical insights and beyond," *Nature Reviews Neuroscience*, vol. 6, no. 1, pp. 35–47, 2005.

[2] K. J. Miller, G. Schalk, E. E. Fetz, M. den Nijs, J. G. Ojemann, and R. P. Rao, "Cortical activity during motor execution, motor imagery, and imagery-based online feedback," *Proceedings of the National Academy of Sciences*, vol. 107, no. 9, pp. 4430–4435, 2010.

[3] F. Kempf, A. A. Kühn, A. Kupsch, C. Brücke, L. Weise, G.-H. 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*, vol. 25, no. 10, pp. 3137–3145, 2007.

[4] G. Mandybur, "deep brain stimulation (dbs)," "http://www.mayfieldclinic.com/PE-DBS.htm#.UyX3boWmXoE", Feb. 2013.

[5] 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*, vol. 202, pp. 218–233, 2012.

[6] M. Unser and A. Aldroubi, "A review of wavelets in biomedical applications," *Proceedings of the IEEE*, vol. 84, no. 4, pp. 626–638, 1996.

[7] J. Yordanova, V. Kolev, O. A. Rosso, M. Schürmann, O. W. Sakowitz, M. Özgören, and E. Basar, "Wavelet entropy analysis of event-related potentials indicates modality-independent theta dominance," *Journal of Neuroscience Methods*, vol. 117, no. 1, pp. 99–109, 2002.

[8] C. Liner, "An overview of wavelet transform concepts and applications," *University of Houston*, pp. 1–17, 2010.

[9] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.

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

[11] N. Jrad, M. Congedo, R. Phlypo, S. Rousseau, R. Flamary, F. Yger, and A. Rakotomamonjy, "sw-svm: sensor weighting support vector machines for eeg-based brain–computer interfaces," *Journal of neural engineering*, vol. 8, no. 5, p. 056004, 2011.

[12] A. Chandrasekhar and K. Raghuveer, "Intrusion detection technique by using k-means, fuzzy neural network and svm classifiers," in *Computer Communication and Informatics (ICCCI), 2013 International Conference on*. IEEE, 2013, pp. 1–7.

[13] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.

[14] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.

[15] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2005.

[16] K. Beyer, J. Goldstein, R. Ramakrishnan, and U. Shaft, "When is nearest neighbor meaningful?" in *Database TheoryICDT99*. Springer, 1999, pp. 217–235.

[17] C.-W. Hsu, C.-C. Chang, C.-J. Lin *et al.*, "A practical guide to support vector classification," 2003.