A Body Sensor Network to Monitor Parkinsonian Symptoms: Extracting Features on the Nodes

Shyamal Patel[1,2], Konrad Lorincz[3], Richard Hughes[1], Nancy Huggins[4], John Growdon[4],

David Standaert[5], Jennifer Dy[2], Matt Welsh[3], Paolo Bonato[1]

[1] *Department of Physical Medicine and Rehabilitation, Harvard Medical School, 125 Nashua Street, Boston, MA, USA*

spatel19@partners.org

rhughes1@partners.org

[2] *Department of Electrical and Computer Engineering, Northeastern University, 360 Huntington Avenue, Boston, MA, USA*

jdy@ece.neu.edu

[3] *Computer Science, Harvard University, 33 Oxford St., Cambridge, MA, USA*

konrad@eecs.harvard.edu

mdw@eecs.harvard.edu

[4] *Department of Neurology, Harvard Medical School, 55 Fruit Street, Boston, MA, USA*

nhuggins@partners.org

growdon@helix.mgh.harvard.edu

[5] *Department of Neurology, University of Alabama at Birmingham, 1720 7th Ave S, Birmingham, AL, USA*

dstand@uab.edu

***Abstract*— The work presented in this paper concerns the development of a body sensor network to monitor changes in the severity of symptoms in patients with Parkinson's disease. We analyzed the impact of different features derived from wearable sensor (i.e. accelerometer) data on the reliability of the prediction of clinical scores that capture the severity of Parkinsonian symptoms. We implemented an off-line feature estimation procedure on the nodes to assess the computational complexity associated with estimating each feature. Our analyses revealed that good prediction performance can be achieved by extracting a subset of features that are not computationally demanding, thus making it possible to envision estimating features and saving and/or transmitting them to a base station (e.g. PDA) rather than raw accelerometer data. This strategy would avoid numerous problems with transmission of data recorded at relatively high sampling rate including power consumption and the need for dealing with transmission errors. These findings suggest that a wireless wearable system could be a viable tool for long term monitoring of patients with Parkinson's disease.**

I. Introduction

Parkinson's disease (PD) is the most common cause of movement disorder, affecting about 3% of the population over the age of 65 years and more than 500,000 US residents. The characteristic motor features of PD include tremor, bradykinesia (i.e. slowness of movement), rigidity (i.e. resistance to externally imposed movements), and impaired postural balance. Current therapy for Parkinson's disease is based primarily on augmentation or replacement of dopamine, using the biosynthetic precursor levodopa or other drugs that activate dopamine receptors. These therapies are often successful for some time in alleviating the abnormal movements, but most patients eventually develop motor complications as a result of these treatments [1][2]. These complications include wearing-off, the abrupt loss of efficacy at the end of each dosing interval, and dyskinesia, i.e. involuntary and sometimes violent writhing movements [3][4].

We are working towards the application of wearable sensors to monitor motor fluctuation cycles in patients with PD. Our goal is to develop a wireless wearable sensor system that PD patients can wear in the home environment over a period of several days to weeks. Two key points toward developing the tools necessary to achieve continuous monitoring of motor function are (1) development of a robust and deployable wearable wireless network of sensors and (2) the development of analysis techniques to derive clinically relevant information from miniature sensor data.

We are working to develop a system that can efficiently process data on the wireless nodes without the need for transmitting raw data for off-line processing. An advantage of doing so is the reduction in the amount of storage space required on the nodes and power saving, which would enable longer recordings. Also when monitoring patients for up to one week, it is desirable to assess the quality of the recordings beyond a few spot checks. This objective can be achieved by estimating data features on the nodes with available computational resources, and by wirelessly transmitting only features, as opposed to raw data. Building this capability into a body sensor network would allow clinical personnel to check that data captured during the monitoring interval are satisfactory, and carry the information they need. One challenge to implementing this strategy is the limited availability of computational resources on the nodes of the body sensor network. This necessitates a trade-off between the relevance of the clinical information captured by a given data feature and the computational cost associated with its estimation. In this paper, we present results obtained from simulations to determine the amount of computational time

required to extract features from raw sensor data. We also estimate the importance of different features in predicting clinical information. We aim to determine a set of features that are both computationally inexpensive and lead to accurate prediction of clinical information.

II. Wireless Wearable Sensor System

The sensor platform we are using is the Intel Digital Health Group's Sensing Health with Intelligence, Modularity, Mobility, and Experimental Reusability (SHIMMER). SHIMMER consists of a TI MSP430 microprocessor; a Chipcon CC2420 IEEE 802.15.4 2.4 GHz radio; a MicroSD card slot; a triaxial MEMS accelerometer, the Freescale MMA7260Q; and optionally, a Bluetooth radio which allows streaming of sensor data at high rates, as a radio-agnostic solution.

SHIMMER (Figure 1) is smaller than other wireless systems, with conventional board technology and a lithium-polymer battery for easy maintenance. Internal and external connectors allow new sensor boards to be interfaced to the device, expanding its capabilities. A triaxial gyroscope board using two InvenSense IDG-300 dual-axis gyroscope chips was designed for internal expansion. One of these gyroscopes is mounted perpendicularly to the main sensor board.

The SHIMMER device combines computation, radio communication, high-fidelity triaxial sensors, and a large flash memory into a tiny, wearable rugged plastic enclosure. It measures 1.75" x 0.8" x 0.5" and weighs just 10 g. SHIMMER utilizes a MicroSD card slot that allows up to 2 GBytes of flash memory. This permits continuous data recording of uncompressed, 50 Hz sampled data from 3 channels for more than 80 days. This amount of on-board storage is unprecedented in wearable sensor design.

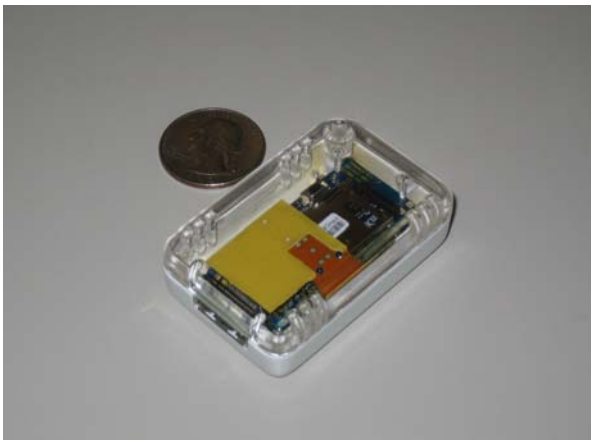

Figure 1. SHIMMER wearable sensor platform.

III. Methods

A. Data Collection

Twelve individuals were recruited in the study, ranging in age from 46 to 75 years, with a diagnosis of idiopathic Parkinson's disease (Hoehn & Yahr stage 2.5 to 3, i.e. mild to moderate bilateral disease with ability to recover from sudden postural disturbance or with some postural instability). Subjects were assessed on the day of the experiment by a clinician using the Unified Parkinson's Disease Rating Scale (UPDRS) [5].

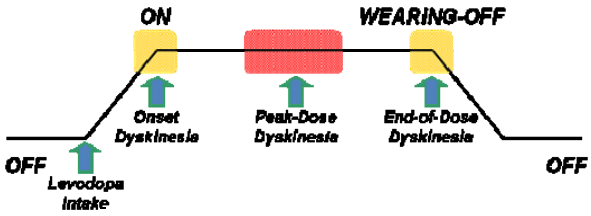

Figure 2. PD Motor fluctuation cycle.

Figure 2 schematically represents a motor fluctuation cycle. Subjects were asked to delay their first medication intake in the morning so that they could be tested in a "practically-defined OFF" state. This approach is clinically used to observe patients in the period of most severe Parkinsonian symptoms. Subsequently, patients took their medications and were tested at 30-minute intervals thereafter until the following medication intake, in order to gather sensor data during an entire motor fluctuation cycle.

Accelerometer sensors were used to gather biomechanical signals during standardized motor tasks utilized for clinical assessment including sitting, finger-to-nose movements, finger tapping, alternating hand movements, leg agility, sit-to-stand, walking, and stand-to-sit. For each task, 30 s of sensor data were recorded. Accelerometers were placed on the right and left upper arm, right and left forearm, right and left thigh, right shin, and left shin. The sensors were connected to an ambulatory system (Vitaport 3, Temec BV, The Netherlands) equipped with data acquisition hardware and software to collect and store the signals. Subjects were videotaped throughout the experiment, and motor UPDRS and dyskinesia scores were assigned for each task by review of the videotapes.

B. Feature Extraction

Raw data were high-pass filtered with a cutoff frequency of 1 Hz to remove gross orientation changes, and low-pass filtered with a cutoff frequency of 15 Hz to remove high frequency noise. Sets of thirty 5-second epochs were randomly selected from sensor data corresponding to each task, for each testing interval. Six different types of features were extracted from different body segments from each epoch. The features were chosen to represent characteristics such as intensity, modulation, rate, periodicity, and coordination of movement. Intensity was measured as the root-mean-square (RMS) value of the detrended accelerometer signal. The modulation of the output of each sensor was used to represent

dynamic characteristics of the tasks, and was calculated as the range of the auto-covariance of each channel. Rate of movement was represented by the dominant frequency component below 10 Hz. Periodicity was measured by computing the ratio of energy in the dominant frequency component to the total energy below 10 Hz. Coordination between body segments on the left and right side and proximal and distal segments was captured in three aspects: magnitude (obtained by calculating the correlation coefficient), delay (estimated as the time lag corresponding to the peak of the cross-correlation function) and similarity (measured by the value of the peak of the cross-correlation function). Approximate entropy was used as a measure of signal complexity.

C. Feature Benchmarking

To estimate the actual computational cost of extracting each feature on the SHIMMER platform, the time needed to compute each feature for a 5-second epoch of simulated accelerometer data was measured. Data were synthesized as sampled at 100 Hz. Estimates were derived for 100 epochs and average values were derived to compare the computational cost associated with each feature. The overhead associated with measuring the start and end times was negligible because start and end times were measured by reading the value of a time register incremented by a 32 kHz clock, converted later into milliseconds. The processor currently utilized by the SHIMMER nodes (MSP430) does not have a floating-point unit. Floating-point operations were therefore simulated via software, thus significantly increasing the computation time required for the operation compared to integer operations.

D. Support Vector Machines

The SVM [6] method was selected due to its success in many classification problems, including gene classification, speaker identification; face image detection, and text categorization. SVM success can be attributed to several properties. SVM optimize an objective function that is convex, hence guaranteed to find an optimal solution; many other classification algorithms only guarantee that local optima be reached. SVM have the ability to generate nonlinear decision boundaries, by mapping the feature space into a higher dimensional space (using kernels) where classes are linearly separable. The kernel "trick" allows SVM to project data to a high dimensional space without added computational cost by replacing inner products in the higher dimensional space with kernels in the lower dimensional input space. SVM demonstrate good generalization performance because they implement the structural risk minimization principle [6]. SVM construct linear hyperplanes, in high dimensional space, that maximize the margin (separation) between classes. The PRTools4 toolbox was used to implement SVM [7]. The specific SVM implementation we used relies on the one-vs.-rest approach to tackle the multi-class classification problem.

IV. Results

The results of our feature benchmarking tests showed that frequency features appear to be the most demanding from a computational standpoint. Cross-correlation based features appear to require the largest power consumption, given the need for transmitting data among nodes in order to combine accelerometer time series recorded by different nodes. The approximate entropy is ranked next, followed by the root mean square value; the data range is the least computational demanding feature to be estimated. Estimating data range feature values required approximately 2 ms, root mean square feature values required about 20 ms, and approximate entropy feature values required about 1 s. Estimating Fast Fourier Transform outputs required on average longer than 4.5 s. These simulation results suggest that the data range and root mean square features can be estimated on the nodes without major interference with other operations that occur on the node (e.g. data sampling and transferring to the SD card, synchronization of multiple nodes), that estimating the approximate entropy feature is more challenging but compatible with available technology, and that frequency features should be only utilized for off-line analysis of the data. Furthermore, our simulations indicated that cross-correlation based features can be estimated at a relatively low cost from a computational standpoint. Estimation of these features required about 80 ms for each 5 s segment of accelerometer data. However, estimation of these features required transmitting at least one of the accelerometer time series between nodes, leading to significant power consumption.

TABLE 1
PREDICTION ERROR (%) FOR FEATURE COMBINATIONS
(A) Tremor, (B) Bradykinesia and (C) Dyskinesia

	1 feature	2 features	3 features	4 features	5 features
A	6.6 (±8.0)	3.1 (±2.9)	2.5 (±3.3)	2.7 (±3.3)	2.8 (±3.6)
B	2.2 (±3.3)	1.6 (±2.7)	1.4 (±2.2)	1.4 (±3.0)	1.7 (±3.7)
C	3.7 (±3.0)	1.9 (±2.7)	1.8 (±1.8)	1.6 (±2.1)	1.2 (±0.9)

Table 1 summarizes the results of the simulations performed to assess the effect of individual data feature types and combinations of feature types on the reliability of the estimates of clinical scores. The values shown in this table were obtained by applying SVM with a third order polynomial kernel to the feature sets. Results are shown for simulations performed using a single feature type, two feature types, etc. up to all five feature types. Numerical values shown in Table 1 are the minimum average prediction error values (standard deviation values are shown in parenthesis) for each symptom and motor complication when the estimates of clinical scores were derived using a different number of feature types.

Results derived when estimating clinical scores of tremor show that the use of a single feature type can lead to an average prediction error value as low as 6.6 %. This result was obtained using the approximate entropy feature. A further decrease in prediction error value is shown when two and three feature types were utilized. With two feature types, the lowest average error (3.1 %) was achieved by utilizing root mean square and data range features. A prediction error value slightly higher (i.e. 3.5 %) was achieved by using root mean square and approximate entropy features. When three feature types were utilized, the lowest average prediction error (2.5 %) was achieved by using the root mean square, the data range, and the approximate entropy feature types. Interestingly, none of these combinations of feature types includes the frequency feature type and the cross-correlation based feature type. These feature types are not particularly desirable in the context of using the SHIMMER platform because they are either very demanding from a computational standpoint (the frequency features) or they require wireless transmission of data (the cross-correlation based features). It should also be noted that approximate entropy appears to be a key feature in estimating the severity of tremor. Our results appear to confirm previous observations by Vaillancourt et al. [8] that approximate entropy is relevant to studying Parkinsonian and other types of tremor.

Results for bradykinesia show that, in general, clinical scores for this symptom can be predicted with higher reliability (i.e. lower prediction error) than scores for tremor. When using only one feature type, prediction error values as low as 2.2 % can be achieved by utilizing cross-correlation based feature types. The lowest value of prediction error that can be achieved without using cross-correlation based or frequency feature types is 7.1 %. This prediction error value can be achieved by using the approximate entropy feature. When two feature types were considered, the lowest prediction error achievable without using cross-correlation based or frequency feature types was 5.1 %, which was obtained combining root mean square and approximate entropy feature types. When three feature types were used, the best result was achieved by combining root mean square, cross-correlation based, and approximate entropy feature types. The only combination not using cross-correlation based or frequency feature types (i.e. the one utilizing root mean square, data range, and approximate entropy feature types) led to a prediction error equal to 3.4 %. Combining four and all five feature types led to a further reduction in prediction error values for clinical scores of bradykinesia. Overall, the results for bradykinesia showed that the use of cross-correlation based or frequency feature types was key to minimize prediction error values for a given number of feature types. However, not utilizing these feature types resulted in only a moderate increase in prediction error values of a few percentage points (i.e. 2 % to 5 % according to the number of utilized feature types). The use of the approximate entropy feature appeared to be key to achieve these results.

Results for dyskinesia showed prediction error values similar to those obtained for bradykinesia. When a single feature type was utilized, a minimum prediction error value of 3.7 % was achieved using the approximate entropy feature. When using two feature types, an average prediction error value of 1.9 % was obtained by using cross-correlation based features and the approximate entropy as input features to the SVM. Similar prediction error values were achieved by using root mean square value and approximate entropy (1.9 %) and data range and approximate entropy (2.0 %), thus suggesting that the approximate entropy feature might play a key role in predicting the severity of dyskinesia. The best result achieved when combining three feature types was a prediction error of 1.8 % (achieved when using frequency feature types, cross-correlation based types, and approximate entropy). A slight increase in prediction error values (i.e. a value of 5.6 %) was observed when using a combination of feature types, not including frequency feature types and cross-correlation based types. Combinations of four feature types led to a prediction error as low as 1.6 %, whereas using all five feature types led to a prediction error value of 1.2 %.

Overall, these results demonstrate that the severity of tremor, bradykinesia, and dyskinesia can be reliably estimated using the procedures proposed in this paper with as little as two features. Furthermore, the outcomes of our analysis suggest that such results can be achieved without including frequency and cross-correlation based feature types. Finally, the results demonstrate that an average prediction error not exceeding a few percentage points can be achieved for the symptoms and motor complications discussed above by solely utilizing root mean square value and approximate entropy features.

ACKNOWLEDGMENT

This work was supported by the National Institute of Neurological Disorders and Stroke, National Institutes of Health under the grant #R21NS045401-02.

REFERENCES

[1] TN Chase, “Levodopa therapy: consequences of the non physiologic replacement of dopamine”, Neurology, 50(Suppl5): S17-S25, 1998

[2] JA Obeso, CW Olanow, JG Nutt, “Levodopa motor complications in Parkinson's disease”, Trends Neurosci, 23: 2-7, 2000

[3] AE Lang, AM Lozano, “Parkinson's disease. First of two parts”, N Engl J Med, 339(16): 1044-1053, 1998

[4] AE Lang, AM Lozano, “Parkinson's disease. Second of two parts”, N Engl J Med, 339(16): 1130-1143, 1998

[5] S Fahn, RL Elton, “Unified Parkinson’s Disease Rating Scale”. In Fahn S (Ed), Recent Developments in Parkinson’s Disease, MacMillan Healthcare Information, 153-163, 1987

[6] V. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY: Springer-Verlag, 1995.

[7] R. P. W. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. de Ridder, and D. M. J. Tax, "PRTools4, A Matlab Toolbox for Pattern Recognition," in *Delft University of Technology*, 2004.

[8] D. E. Vaillancourt, M. M. Sturman, L. Verhagen Metman, R. A. Bakay, and D. M. Corcos,"Deep brain stimulation of the VIM thalamic nucleus modifies several features ofessential tremor," *Neurology*, vol. 61, pp. 919-25, 2003.