33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Longitudinal Monitoring of Patients with Parkinson's Disease via Wearable Sensor Technology in the Home Setting

Shyamal Patel, *Student Member, IEEE*, Bor-rong Chen, Chiara Mancinelli, Sabrina Paganoni, Ludy Shih, Matt Welsh, Jennifer Dy and Paolo Bonato, *Senior Member, IEEE*

***Abstract*—Objective longitudinal monitoring of symptoms related motor fluctuations can provide valuable information for the clinical management of patients with Parkinson's disease. Current methods for long-term monitoring of motor fluctuations, such as patient diaries, are ineffective due to their time consuming and subjective nature. Researchers have shown that wearable sensors such as accelerometers can be used to gather objective information about a patient's motor symptoms. In this paper, we present preliminary results from our analysis on wearable sensor data gathered during longitudinal monitoring of 5 patients with PD. Our results indicate that it is possible to track longitudinal changes in motor symptoms by training a regression model based on Random Forests.**

I. Introduction

PARKINSON'S disease (PD) is a common 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 are development of rest tremor, bradykinesia, rigidity, and impairment of postural balance. Current therapy is based on augmentation or replacement of dopamine, using the biosynthetic precursor levodopa or drugs that activate dopamine receptors. These therapies for PD are often successful for limited period of time but most patients eventually develop motor complications, including abrupt loss of efficacy at the end of each drug dosing interval, and involuntary and sometimes violent writhing movements [1, 2]. Monitoring these motor complications can assist adjustment of drug dosage or call for change of treatment.

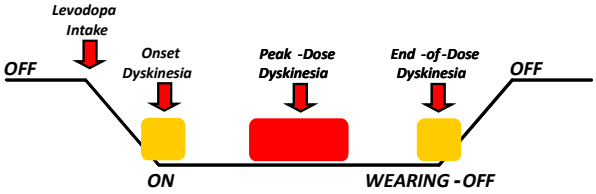

Fig. 1 Parkinson's disease motor fluctuation cycle.

Currently available tools for monitoring motor fluctuations (figure 1) are quite limited. In clinical practice, information about motor fluctuations is usually obtained by asking the patient to recall the number of hours of ON and OFF time they have experienced in the recent past. "ON time" is used to refer to periods when medications are effective in attenuating symptoms. "OFF time" is used to refer to periods when symptoms are present. This kind of self-report is subject to both perceptual bias and recall bias. A reliable quantitative tool for long term monitoring of motor complications in PD patients would be valuable both for routine clinical care of patients as well as for trials of novel therapies.

Advances in wearable technology [3] make it possible to develop monitoring systems to capture movement patterns associated with motor fluctuations. Recent work has shown initial success on both long term data collection with wearable sensors [4] and quantitative classification of PD symptom severity using data gathered from wearable sensors [5-7]. Building on these foundations, we are developing an integrated remote monitoring platform that extends sensor based monitoring into the home environment. Such a platform will eliminate the need for frequent hospital visits by PD patients for supervised data collection sessions and therefore reduce the cost and simplify logistics for PD monitoring.

We have designed a web-based remote motion monitoring system called MercuryLive [8], which allows clinicians to easily interact with patients who are at home, record annotated live motion data, and configure the motion monitoring parameters remotely. We are currently deploying this remote monitoring system in patients' home to gather wearable sensor data. The focus of our current study is on the longitudinal aspect of monitoring motor symptoms. Hence, we are working towards the development of techniques to extract clinically meaningful information from

Manuscript received on April 17, 2011. This work was supported by the Michael J. Fox Foundation for Parkinson's research under the project entitled "Wearable Sensors and a Web-Based Application to Monitor Patients with Parkinson's Disease in the Home Environment."

S. Patel is with the Department of Physical Medicine and Rehabilitation, Harvard Medical School, Spaulding Rehabilitation Hospital, Boston MA 02114, USA and Department of ECE, Northeastern University, Boston, MA 02115 USA (e-mail: spatel19@partners.org).

C. Mancinelli and S. Paganoni are with the Department of Physical Medicine and Rehabilitation, Harvard Medical School, Spaulding Rehabilitation Hospital, Boston MA 02114, USA (e-mail: cmancinelli@partners.org, s-paganoni@partners.org).

B. Chen and M. Welsh are with the Harvard School of Engineering and Applied Sciences, Cambridge, MA 02138 USA (e-mail: brchen@eecs.harvard.edu, mdw@eecs.harvard.edu).

L. Shih is with the Department of Neurology, BIDMC, Boston MA 02215 USA (e-mail: lshih@caregroup.harvard.edu)

J. Dy is with the Department of ECE, Northeastern University, Boston, MA 02115 USA (e-mail: jdy.ece@gmail.com).

P. Bonato is with the Department of Physical Medicine and Rehabilitation, Harvard Medical School, Spaulding Rehabilitation Hospital, Boston MA 02114, USA (corresponding author, phone: 617-573-2770; fax: 617-573-2769; e-mail: pbonato@partners.org).

sensor data and track changes in symptom severity as they our over a period of several months. In this paper, we present preliminary results on the estimation of clinical scores for two tasks from sensor data recorded longitudinally from 5 patients over a period of three days. Our main goal is to estimate clinical scores for the tests conducted on the last day of monitoring using data gathered during pervious monitoring sessions. We optimize the feature space by performing feature selection and implement a Random Forest regression based method for estimating the clinical scores. Preliminary analysis shows that it is feasible to monitor longitudinal changes in the severity of PD motor symptoms.

II. METHODS

A. Data Collection

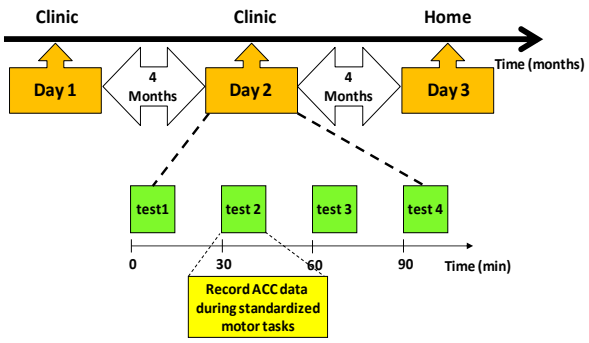

Fig. 2 Schematic of the data collection process. Data collection was performed for day 1 and day 2 in the clinic. MercuryLive remote monitoring system was used for day 3 data collection in the home.

Fourteen individuals with moderate to advanced PD have been recruited in the study so far. In this paper, we will present results on data gathered from 5 subjects who have completed all aspects of the study. A schematic of the data collection process is shown in figure 2. Subjects undergo motor assessments on three separate days. First two days of monitoring were performed in the clinical setting while the third day of monitoring was performed in the home setting using a web-based remote monitoring application called MercuryLive [8]. The clinical user-interface of MercuryLive is shown in figure 3. MercuryLive contains three tiers: a data collection engine that relies upon the wearable sensors, web services for live streaming and storage of sensor data, and a user interface for two-way communication between patient and clinician.

Four testing sessions are performed on each day of monitoring. During each of these tests, subjects perform a set of tasks from the Unified Parkinson's Disease Rating Scale (UPDRS) as well as activities of daily living. Each task is performed for approximately 30s. The UPDRS motor tasks included finger-to-nose (reaching and touching a target), index finger and thumb tapping, repeated hand movements (opening and closing of hand), heel tapping, quite sitting and alternating hand movements (repeated pronation/supination with an outstretched arm). Subjects are allowed to rest for approximately 20mins between two tests.

As shown in figure 4, subjects were instrumented with tri-axial accelerometer sensors based on the SHIMMER platform [9]. Accelerometers were placed bilaterally at the midpoint of forearm, the midpoint of upper arm, on the shank approximately 10cm about the ankle, and on the thigh approximately 10cm above the knee. In addition, video recordings were made during each testing session for later clinical assessments performed by an expert clinician. Using the UPDRS the clinicians provide us a score from 0-4 for each task performed by the subject during every monitoring session. A score of 0 indicates the absence of a symptom while a score of 4 indicates a high level of symptom severity.

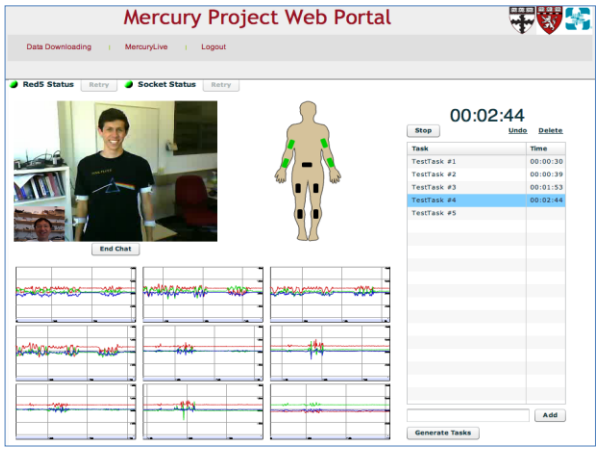

Fig 3. Screen shot of the MercuryLive web-application

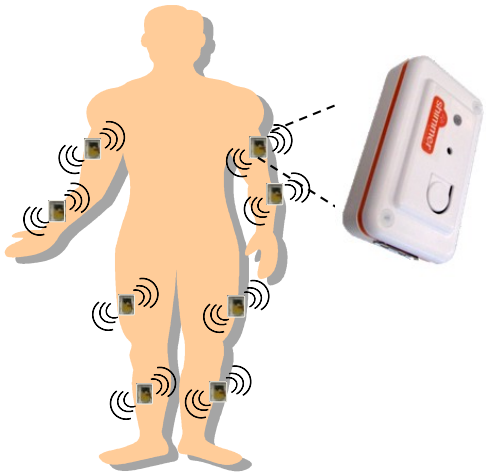

Fig 4. Schematic representation of the sensor setup

B. Feature Extraction

Raw tri-axial accelerometer data for each task were high-pass filtered with a cutoff frequency of 0.5Hz to remove gross changes in the orientation of body segments. The time series were further low-pass filtered with a cutoff frequency of 5Hz. All filters were implemented as IIR filters based on an elliptic design. Features were extracted from the each accelerometer time series (X, Y & Z axis) using a 5s rectangular windows randomly positioned throughout the

recordings performed during performance of each motor task. Features were extracted from 30 such window segments (i.e. epochs) for each motor tasks. In this paper we focus on the heal tapping and alternating hand movement tasks. We extract features that capture qualitative characteristics of movement such as intensity, modulation, frequency, periodicity, and smoothness of movement. Intensity was measured as the root mean- square value of linearly detrended accelerometer signal. The dynamic characteristics of the tasks were represented by the modulation of each accelerometer sensor output and measured as the range of amplitude of each channel. Two frequency-based features were estimated. The dominant frequency component was used to capture the rate of movement; the algorithm identified such value in the range between 0.5 and 5Hz as the peak of the Fast Fourier Transform output. Range of auto-covariance was derived as a measure of signal modulation. The ratio of energy associated with the dominant frequency component to the total energy in the range 0.5 to 5Hz was utilized to measure periodicity. Signal entropy [10] estimates were derived as a measure of signal complexity. A total of 6 features are extracted per accelerometer axis (X, Y & Z) making a total of 18 features per sensor node and 144 features in total.

C. *Feature Selection*

During the feature extraction process we extract a large number of features from each tri-axial accelerometer sensor. Hence, to escape the curse of dimensionality we perform feature selection. Before performing feature selection, we filter the feature space by discarding all the features that were derived from sensors not on the body segments involved in the task. For example, for alternating hand movement with right hand (AHR) we discard all the features except the ones derived from the sensors located on the right forearm and right upper arm.

Feature selection was implemented as a two step process. First, using the ReliefF [11] algorithm in conjunction with a ranker search method, we derive the order of importance of the features. The ReliefF algorithm iterates through every instance updating the weights assigned to a feature based on its ability to correctly classify the instance. To calculate the weights, for each instance, it searches for K nearest neighbors from the same class, and K nearest neighbors from each of the other classes. The number of nearest neighbors K was set to 10 as suggested by Robnik-Sikonja and Kononenko [11]. The ReliefF algorithm is computationally simple and robust. We used the WEKA [12] implementation of the algorithm.

The second step is to select a subset of the ranked features. To do this we use the Davies-Bouldin (DB) cluster validity index [13] as an objective measure of the clustering quality. The DB index measures how well-separated clusters belonging to different classes are as well as the spread within each cluster. It is calculated for each pair of clusters as a ratio of within-class scatter and between-class separation. A low value of the DB index indicates tight well-separated clusters and vice versa. Using the ranked set of features from the ReliefF step, we calculate the DB index by incrementally adding one feature at a time in the order of ranking. A feature subset is selected when we see no significant improvement or an increase in the DB index.

D. *Clinical Score Estimation*

To estimate the UPDRS clinical scores we implemented a regression Random Forest (RF) based on the R implementation by Liaw et al [14]. RFs, introduced by Breiman [15], are ensembles of weakly correlated decision trees that generate an output as an aggregate result of predictions by individual trees. RFs introduce an additional level of randomness to bagging by training individual trees using a randomly selected subset of features. They have been shown to outperform several other techniques while being robust against overfitting [15].

III. RESULTS

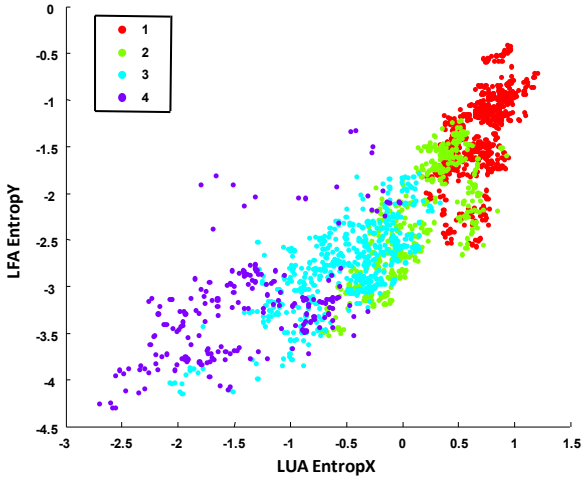

Fig. 5 Scatter plot using two top ranked features for the Alternation Hand Movement task with left hand (AHL). Signal entropy derived from the X-axis on the left upper arm (LUA EntropyX) is on the abscissa and signal entropy derived from Y-axis on the left forearm (LFA EntropyY) is on the ordinate. The scatter plot is color labeled by clinical score.

In this paper, we focus our analysis on two specific UPDRS tasks 1) heel tapping with left (LAH) and right (LAR) leg and 2) alternating hand movement (pronation/supination) with left (AHL) and right (AHR) hand. Our goal is to derive reliable estimates of the clinical scores for tests performed during day 3 of the monitoring. Figure 5 shows a scatter plot of the combined data from all the subjects and all days of monitoring for two of the top ranked features provided by the ReliefF algorithm for the AHL task. We can observe a clear trend from a score of 1 to a score of 4. The overlaps between clusters are expected as the level of severity, in reality, is continuous unlike the UPDRS scoring system, which can only provide discrete scores from 0 to 4. Also, there can be slight variations in the performance of motor tasks from subject to subject. An attractive aspect of using wearable sensors is the possibility of deriving more continuous estimates for changes in severity and thus tracking clinical scores with a much higher resolution that currently possible.

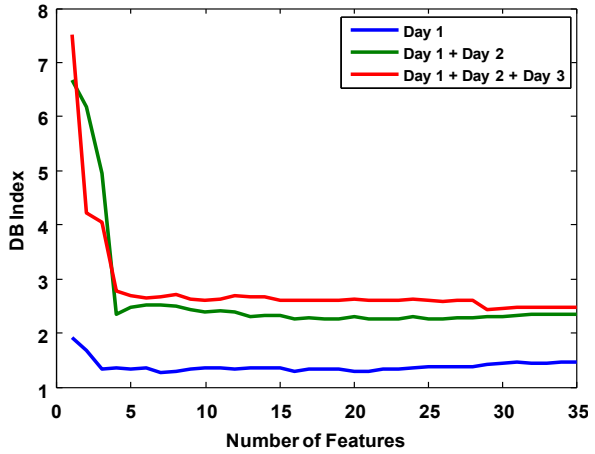

Fig. 6 DB Index as a function of the number of top ranked features (provided by ReliefF) included in the dataset. Features are incrementally added, one at a time, in the decreasing order of ranking.

Figure 6 shows the aggregate DB index for all the subjects and all the tasks. As we combine data from day 1 with data from day 2 and day 3 we see an increase in DB index which indicates that the quality of clustering has deteriorated. This increase in DB index can be attributed to the differences in sensor placement, more levels of observed severity (i.e. more clusters) and change in the way the task was performed by the subject. As we increase the number of top ranked features, the DB index decreases significantly for the combined datasets. However, we observe no significant increase beyond 5 top ranked features.

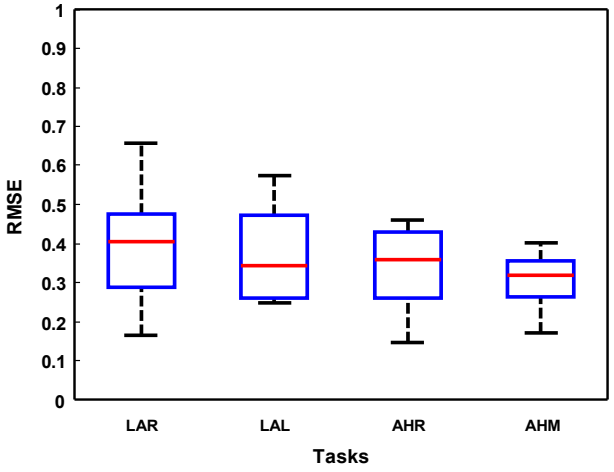

Fig. 7 Root mean square (RMS) error for estimation of UPDRS clinical scores.

Figure 7 shows the root mean square (RMS) error in the estimation of the UPDRS score for data gathered during testing sessions performed on day 3. The estimates were derived using a regression Random Forest with 20 trees. Overall, the average RMS error is around 0.4. We observed a larger variability in the estimation performance for the LAR (0.39 ± 0.19) and LAL (0.37 ± 0.17) tasks. This can be attributed to the fact that the leg agility task is not as well defined as the alternating hand movement tasks and hence is not performed consistently between and within subjects.

IV. Discussion

Preliminary results indicate that longitudinal tracking of severity of motor symptom in PD using wearable sensors is feasible. In this paper, we showed that we were able to track UPDRS scores for two tasks by using a regression Random Forest to within 0.5 points on a scale of 0-4. We observed that the estimation of scores for alternating hand movement task (AHL/AHR) was more accurate than leg agility task (LAL/LAR). This can be attributed to the fact that AHL/AHR task is more systematic and well defined. In the future, we aim to expand our analysis to include all the UPDRS tasks and a larger set of subjects from our ongoing study. We believe that longitudinal tracking of motor symptoms in the home setting a challenging task and will require creative data processing techniques to account for the uncontrolled nature of the home environment.

Acknowledgment

The authors would like to thank Mr. Ben Kuris for providing technical guidance.

References

[1] W. J. Weiner, "Motor fluctuations in Parkinson's disease," *Rev Neurol Dis,* vol. 3, pp. 101-8, Fall 2006.
[2] J. A. Obeso, *et al.*, "Levodopa motor complications in Parkinson's disease," *Trends Neurosci,* vol. 23, pp. S2-7, Oct 2000.
[3] P. Bonato, "Advances in wearable technology for rehabilitation," *Stud Health Technol Inform,* vol. 145, pp. 145-59, 2009.
[4] S. Patel, *et al.*, "Monitoring motor fluctuations in patients with Parkinson's disease using wearable sensors," *IEEE Trans Inf Technol Biomed,* vol. 13, pp. 864-73, Nov 2009.
[5] K. Lorincz, *et al.*, "Mercury: a wearable sensor network platform for high-fidelity motion analysis," presented at the Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems, Berkeley, California, 2009.
[6] N. L. Keijsers, *et al.*, "Ambulatory motor assessment in Parkinson's disease," *Mov Disord,* vol. 21, pp. 34-44, Jan 2006.
[7] A. Salarian, *et al.*, "Quantification of tremor and bradykinesia in Parkinson's disease using a novel ambulatory monitoring system," *IEEE Trans Biomed Eng,* vol. 54, pp. 313-22, Feb 2007.
[8] C. Bor-Rong, *et al.*, "A Web-Based System for Home Monitoring of Patients With Parkinson's Disease Using Wearable Sensors," *Biomedical Engineering, IEEE Transactions on,* vol. 58, pp. 831-836, 2011.
[9] A. Burns, *et al.*, "SHIMMER: A Wireless Sensor Platform for Noninvasive Biomedical Research," *Sensors Journal, IEEE,* vol. 10, pp. 1527-1534, 2010.
[10] R. Moddemeijer, "On Estimation of Entropy and Mutual Information of Continuous Distributions," *Signal Processing,* vol. 16, pp. 233 - 246, 1989.
[11] M. Robnik-Sikonja and I. Kononenko, "Theoretical and empirical analysis of ReliefF and RReliefF," *Machine Learning Journal,* vol. 53, pp. 23 - 69, 2003.
[12] M. Hall, *et al.*, "The WEKA Data Mining Software: An Update," *SIGKDD Explorations,* vol. 11, 2009.
[13] J. C. Bezdek and N. R. Pal, "Some new indexes of cluster validity," *IEEE Transactions on System, Man, and Cybernetics,* vol. 28, pp. 301 - 315, 1998.
[14] A. Liaw and M. Wiener, "Classification and regression by Random Forest," *R News,* vol. 2, pp. 18-22, 2002.
[15] L. Breiman, "Random Forests," *Machine Learning,* vol. 45, pp. 5-32, 2001.