31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Feature Selection for Classification Based on Fine Motor Signs of Parkinson's Disease

B. R. Brewer, *Member, IEEE*, S. Pradhan, G. Carvell, and A. Delitto

***Abstract*—Effective evaluation of potential neuroprotective interventions for Parkinson's disease (PD) requires precise quantification of the motor signs associated with this disease. We have created a protocol that uses force tracking in a simultaneous task paradigm to quantify the fine motor control deficits in individuals with PD. We have used this protocol to collect data from 30 individuals with early to moderate PD and 30 age-matched controls. Based on this data, we computed 60 variables. We generated all possible combinations of three of these variables, and we then computed the classification accuracy of a support vector machine (SVM) trained on each variable combination. We were able to correctly classify 85% of subjects as with or without PD. We found that root-mean-square error variables were the most important features for classification and that utilizing a simultaneous task paradigm improves classification accuracy.**

I. Introduction

MANY research groups are trying to develop neuroprotective interventions for Parkinson's disease (PD), interventions that slow or halt the progress of the disease. However, testing the effectiveness of potential neuroprotective interventions requires sensitive, precise ways to diagnose the disease and measure its progression. Various types of imaging have been proposed as diagnostic and progression biomarkers for PD [1, 2], but the correlation between imaging and functional measures is low [3]. Robotic and sensing technology can enable quantitative assessment of the motor signs of PD in order to promote more complete evaluation of potential neuroprotective interventions [4-6].

We have developed an experimental protocol that uses high-precision force/torque sensors to measure an individual's performance during a force tracking tack. Our assessment is unique in that we utilize a simultaneous task paradigm in which the user performs a cognitive task (counting down from 100) while performing the force tracking task. Our previous work has shown that the simultaneous cognitive task causes greater deterioration in the force tracking performance for individuals with PD relative to age-matched control subjects [7]. In addition, we have shown that performance on our assessment can account for 76% of the variance of scores on the Unified Parkinson Disease Rating Scale [8].

This paper describes the classification of individuals as with or without PD based on our experimental protocol. We used the force information measured during the protocol to compute 60 variables. We generated all possible combinations of three of these variables, and we then computed the classification accuracy of a support vector machine (SVM) trained on each variable combination. Using this method, we estimated the classification accuracy we can obtain using our experimental protocol and identified the features important for classification of individuals as with or without PD.

II. Methods

A. Subjects

Thirty individuals with PD participated in this experiment. Each individual had a Hoehn-Yahr score between I and III (median H-Y score 2, mean H-Y score 2.03) [9]. All individuals remained off medication for PD for 12 hours before completing the testing protocol. Thirty age-matched control subjects also participated in this experiment. These individuals had no history of neurological disease or injury.

B. Experimental Protocol

The experimental environment for this experiment included two 6-axis force/torque NANO 17 sensors (ATI Industrial Automation). The sensors have a resolution of 0.003 N for force. The sensors were mounted to a portable platform using custom-made hardware (Fig. 1). The user exerted force on the sensors using the index finger and thumb. The mean force exerted by the user was shown on a computer screen, and the user modulated his or her force to track a target wave. Two target waveforms were used, a sine wave with a period of 7.5 seconds and pseudorandom waveform (Fig. 2). The subject tracked each waveform for three minutes. During the first minute, the subject tracked the target force. During the second minute, the subject tracked the target force while simultaneously counting down from 100 by 1. During the third minute, the subjects tracked

Manuscript received April 7, 2009. This work was supported in part by the Foundation for Physical Therapy Promotion of Doctoral Studies II Scholarship 2007 and by resources and the use of facilities of the Human Engineering Research Laboratories, VA Pittsburgh Healthcare System.

B. R. Brewer is with the Department of Rehabilitation Science and Technology, University of Pittsburgh, Pittsburgh, PA 15260 USA (phone: 412-383-6594; fax: 412-383-6597; e-mail: bbrewer@pitt.edu).

S. Pradhan was with the Department of Physical Therapy, University of Pittsburgh, Pittsburgh, PA 15260 USA. She is now with the Department of Rehabilitation Medicine, University of Washington, Seattle, WA 98195 USA (e-mail: sujatap@u.washington.edu).

G. Carvell is with the Department of Physical Therapy, University of Pittsburgh, Pittsburgh, PA 15260 USA (e-mail: gcarvell@pitt.edu).

A. Delitto is with the Department of Physical Therapy, University of Pittsburgh, Pittsburgh, PA 15260 USA (e-mail: delitto@pitt.edu).

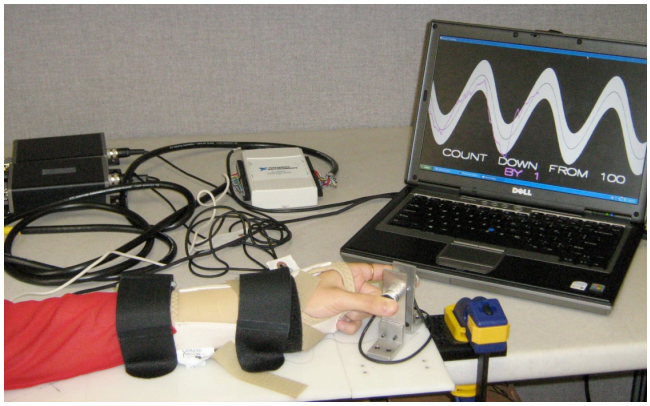

Fig. 1. The experimental environment. The subect grasped the force/torque sensors with the index finger and thumb and modulated the grip force to track a target wave shown on the computer screen.

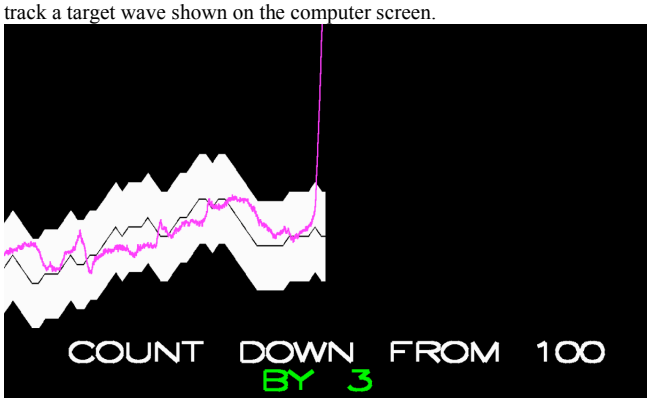

Fig. 2. An example of the visual feedback shown on the computer screen. The subject was shown the target wave, in this case a pseudorandom wave, as well as a 12.5 s history of his or her tracking performance. The current target force and subject response were shown in the horizontal center of the screen. This screen shot also shows the visual prompt to count down from 100 by 3 that was displayed during the third minute of tracking.

the target force while simultaneously counting down from 100 by 3. Each subject tracked each target waveform (sine or pseudorandom) with both the right and left hands. Thus, each subject completed a total of four trials.

C. *Data Analysis*

We considered a number of summary variables with the goal of determining which variables best separated individuals with and without PD. First, we computed the power spectral density for the subject response and calculated the integral of this function between 2 and 8 Hz. This variable quantified the tremor present in an individual's performance. The 2-8 Hz window contains the range of 4-5 Hz that is typical of Parkinsonian tremor. After computing the tremor integral, we filtered the subject response using a 2^{nd} order low-pass Butterworth dual-pass filter with a cut-off frequency of 2 Hz. All other summary variables were computed from the filtered data.

After filtering the data, we calculated the root-mean-square error (RMSE) between the target wave and the subject response. Because case and control individuals were expected to differ most in tracking performance at the extremes of the target waveforms, we also computed the RMSE error for a 1 s window centered on each peak and valley of the sine wave target. The extremes of the pseudorandom wave were defined as periods during which the target force was greater than 5 N or less that 3.2 N; the RMSE was computed for each of these periods. The final summary variable that we computed from the filtered data was the lag between the target waveform and the subject response. The lag was computed as the time period that maximized the cross-covariance between the target and the response.

The mean of each summary variable was computed for each cognitive load condition (no counting, counting down by 1, counting down by 3) for each target waveform (sine or pseudorandom) and each hand. This gave us a total of 60 variables (5 summary variables x 3 cognitive load conditions x 2 target waveforms x 2 hands). Because the motor signs of PD are typically asymmetrical, we divided the data collected for both hands based on the side of better/worse performance, rather than based on the right/left side. The side of better performance was the side with the lower mean RMSE error over both the sine and pseudorandom targets. The value of each summary variable in minute 1 was subtracted from the value of that summary variable in minute 2 and minute 3.

Once we computed these variables, we examined which subset of variables could best classify individuals as with or without PD. The classification technique we used was a linear support vector machine (SVM) [10]. The SVM is a mathematical function that uses a vector of variables (the feature vector) to classify an individual as with or without PD. It can be thought of as an automated way of finding the surface that divides the feature vectors of PD and control subjects while maximizing the distance of each group from the dividing surface. In the simplest terms, the SVM finds the most reliable way to divide case from control individuals based on the elements of the feature vector. The SVM is created by "training" on a data set of individuals whose classification (PD or control) is known. Because the SVM function depends upon the data used for training, the classification accuracy of the SVM must be measured using an independent test set.

We generated all 34,220 possible combinations of three variables that could be chosen from our 60 variables. For each combination, we created an SVM using that subset of variables as the feature vector. We used leave-one-out cross-validation to determine the classification accuracy of the SVM corresponding to a particular combination of variables [10]. Each individual in term was considered to be the test set. The SVM was trained based on the data from every other individual (59 subjects) and was then used to classify the individual used as the test set. The process was repeated for all case and control individuals. The number of case and control individuals correctly classified was computed for every combination of variables, and we determined which combinations correctly classified the greatest number of individuals. All analyses were conducted in Matlab.

III. Results

Two combinations of variables were able to correctly classify 85% (51/60) of the individuals who participated in the experiment. The following subset of variables correctly classified 25/30 case subjects and 26/30 control subjects: RMSE at peaks for sine wave/minute 1/side of worse performance; RMSE for pseudorandom wave/minute 2/side of worse performance; and RMSE at peaks for sine wave/minute 3/side of better performance. These variables are shown in Fig. 3-5. The following combination of variables correctly classified 24/30 case subjects and 27/30 age-matched controls: RMSE at peaks for sine wave/minute 1/side of better performance; RMSE for pseudorandom wave/minute 2/side of worse performance; and RMSE for sine wave/minute 3/side of better performance.

Three combinations of variables were able to correctly classify 83.3% (50/60) of the individuals who participated in the experiment. All of these correctly classified 24/30 case subjects and 26/30 control subjects. The first combination of variables was RMSE at peaks for sine wave/minute 1/side of worse performance; RMSE for pseudorandom wave/minute 1/side of worse performance; and RMSE at peaks for sine wave/minute 3/ side of better performance. The second combination of variables was RMSE at peaks for sine wave/minute 1/side of worse performance; RMSE for sine wave/minute 1/side of better performance; and RMSE for sine wave/minute 2/side of better performance. The third combination was RMSE at peaks for sine wave/minute 1/side of better performance; RMSE; RMSE for sine wave/minute 3/side of better performance; and lag for pseudorandom wave/minute 2/side of better performance.

IV. Discussion

A. Performance of SVM Classification

We observed a wide range of fine motor abilities for both case and control subjects in this study. Some individuals were recently diagnosed with PD, while in others the disease was much more advanced. The range of motor abilities for age-matched control subjects was also quite large, as demonstrated by the plots in the results section. In fact, since more than 60% of the neurons in the basal ganglia are destroyed before clinical signs of PD are observed [2], we cannot be sure that all control subjects were without neurological impairment. As the above plots show, there was substantial overlap between case and control individuals for all variables. We choose to use the SVM method for classification because the boundary separating case from control individuals is chosen based only on those individuals who lie close to this dividing plane (the individuals determining the boundary are termed the support vectors). The classification boundary is not influenced by individuals who lay far it, individuals are relatively easy to classify as case or control. In choosing the SVM technique, we hoped that this characteristic would enable us to obtain good

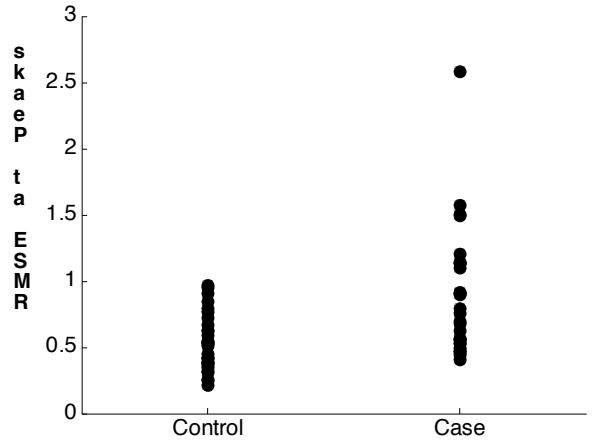

Fig. 3. The RMSE at the peaks of the sine wave for minute 1 (tracking with no counting task) for the side of worse performance. Each point represents the RMSE for a single individual.

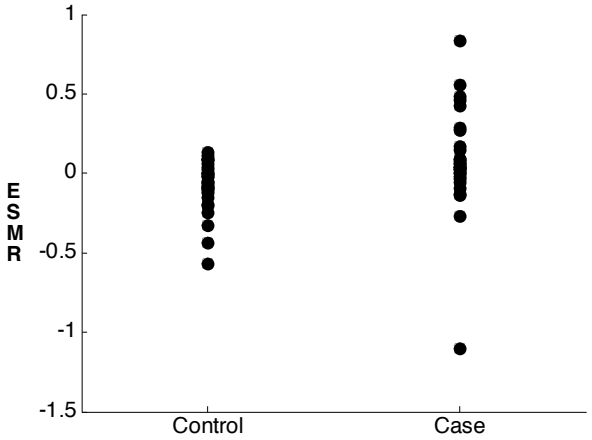

Fig. 4. The RMSE for the pseudorandom wave for minute 2 (tracking while counting down by 1) for the side of worse performance. Each point represents the RMSE for a single individual.

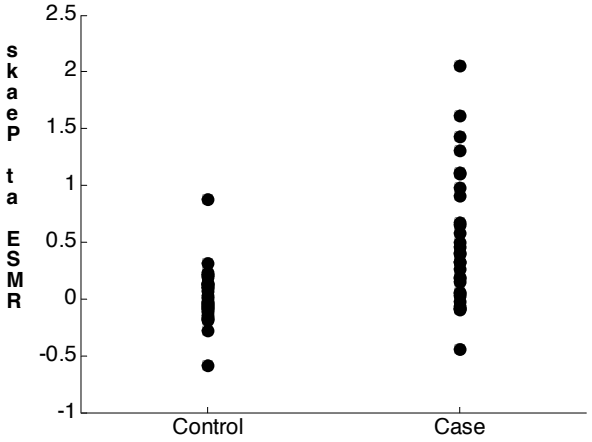

Fig. 5. The RMSE at the peaks of the sine wave for minute 3 (tracking while counting down by 3) for the side of better performance. Each point represents the RMSE for a single individual.

classification performance despite the overlap in motor abilities between case and control subjects. In addition, this method has yielded good empirical results for previous

applications. For example, Jiang et al. [11] used SVM to classify the degree of liver fibrosis in individuals with Hepatitis C with 87% accuracy. Van Calster et al. [12] used SVM to successfully classify ovarian tumors as benign or malignant.

The best SVM had a classification accuracy of 85% overall, 83.3% for case subjects, and 86.7% for control subjects. While not perfect, we feel this is a respectable performance, given the small size of our data set and the substantial overlap in motor abilities between case and control subjects.

B. Selected Features

While our relatively small number of subjects limits us from a data mining perspective, this small sample size made it computationally feasible to exhaustively search all possible combinations of three variables in order to select the features that best predict whether or not an individual has Parkinson's disease. We identified five subsets of variables that led to the high classification accuracy; examination of these variables thus identified as important can inform the future work of our group and others interested in the use of technology to quantify neurological motor deficits. Variables that were included in the five best subsets were almost exclusively RMSE variables, including the overall RMSE for both sine and pseudorandom targets and the RMSE at the peaks for the sine target. Four of the five best subsets included variables from both hands (side of better performance and side of worse performance). All five of the best subsets include variables from different minutes of tracking (different cognitive load conditions). Three of the five best subsets contained variables from all three cognitive load conditions. Therefore, we conclude that the RMSE is a useful variable for distinguishing individuals with and without PD, and that utilizing a simultaneous task paradigm improves classification accuracy.

This work is limited by the fact that we searched only combinations of three variables. It is possible that a combination of four or more variables might have better classification accuracy. However, preliminary tests showed that using combinations of more variables led to overfitting due to our limited sample size. When a larger data set is available, we will investigate other methods of variable selection. In addition, the results of SVM should be compared to the results obtained with other methods of classification.

C. Future Work

We have expanded our experimental environment to include two haptic robots in addition to the force/torque sensors. This new environment enables us to precisely measure the force and position of the index finger and thumb as the user interacts with a virtual object. In particular, we will examine the simulated functional task of moving a virtual object while maintaining a grip force within certain limits (simultaneous motor tasks). We are currently recruiting individuals with early PD to measure the test-retest reliability of the new experimental protocol. In addition, we wish to compare the current results with our ability to classify individuals as with or without PD based on the data collected in our new protocol. Finally, we plan to follow individuals longitudinally in order to validate our assessment as a measure of progression of the motor signs of PD.

ACKNOWLEDGMENT

The authors would like to thank Dr. Robert Moore of the Department of Neurology of the University of Pittsburgh for his help in recruiting individuals with PD to participate in our experiment.

REFERENCES

[1] D.J. Brooks, K.A. Frey, K.L. Marek, D. Oakes, D. Paty, R. Prentice, C.W. Shults, and A.J. Stoessl, "Assessment of neuroimaging techniques as biomarkers of the progression of Parkinson's disease," *Exp Neurol*, vol. 184 Suppl 1, pp. S68-79, Nov 2003.

[2] G. Becker, A. Muller, S. Braune, T. Buttner, R. Benecke, W. Greulich, W. Klein, G. Mark, J. Rieke, and R. Thumler, "Early diagnosis of Parkinson's disease," *J Neurol*, vol. 249 Suppl 3, pp. III/40-8, Oct 2002.

[3] W. Pirker, "Correlation of dopamine transporter imaging with parkinsonian motor handicap: how close is it?," *Mov Disord*, vol. 18 Suppl 7, pp. S43-51, Oct 2003.

[4] R. Cordell, H.C. Lee, A. Granger, B. Vieira, and A.H. Lee, "Driving assessment in Parkinson's disease-A novel predictor of performance?," *Mov Disord*, Jun 4 2008.

[5] J.P. Giuffrida, D.E. Riley, B.N. Maddux, and D.A. Heldman, "Clinically deployable KinesiaTM technology for automated tremor assessment," *Mov Disord*, Jan 9 2009.

[6] S. Papapetropoulos, J.R. Jagid, C. Sengun, C. Singer, and B.V. Gallo, "Objective monitoring of tremor and bradykinesia during DBS surgery for Parkinson disease," *Neurology*, vol. 70, (no. 15), pp. 1244-9, Apr 8 2008.

[7] S. Pradhan, B.R. Brewer, G.E. Carvell, P.J. Sparto, A. Delitto, and Y. Matsuoka, "Effects of a secondary cognitive task on assessment of fine motor control using force tracking in individuals with Parkinson's disease," accepted for publication.

[8] B.R. Brewer, S. Pradhan, G. Carvell, and A. Delitto, "Application of modified regression techniques to validate a quantitative assessment for the motor signs of Parkinson's disease," accepted for publication.

[9] M.M. Hoehn and M.D. Yahr, "Parkinsonism: onset, progression and mortality," *Neurology*, vol. 17, (no. 5), pp. 427-42, May 1967.

[10] V. Vapnik, *Statistical learning theory*: John Wiley & Sons, Inc., 1998.

[11] Z. Jiang, K. Yamauchi, K. Yoshioka, K. Aoki, S. Kuroyanagi, A. Iwata, J. Yang, and K. Wang, "Support vector machine-based feature selection for classification of liver fibrosis grade in chronic hepatitis C," *J Med Syst*, vol. 30, (no. 5), pp. 389-94, Oct 2006.

[12] B. Van Calster, D. Timmerman, C. Lu, J.A. Suykens, L. Valentin, C. Van Holsbeke, F. Amant, I. Vergote, and S. Van Huffel, "Preoperative diagnosis of ovarian tumors using Bayesian kernel-based methods," *Ultrasound Obstet Gynecol*, vol. 29, (no. 5), pp. 496-504, May 2007.