

Characterization of Gait Abnormalities in Parkinson's Disease Using a Wireless Inertial Sensor System

Iris Tien, Steven D. Glaser, and Michael J. Aminoff

Abstract— Gait analysis is important in diagnosing and evaluating certain neurological diseases such as Parkinson's disease (PD). In this paper, we show the ability of our wireless inertial sensor system to characterize gait abnormalities in PD. We obtain physical features of pitch, roll, and yaw rotations of the foot during walking, use principal component analysis (PCA) to select features, and use the support vector machine (SVM) method to create a classification model. In the binary classification task of detecting the presence of PD by distinguishing between PD and control subjects, the model performs with over 93% sensitivity and specificity, and 97.7% precision. Using a cost-sensitive learner to reflect the different costs associated with misclassifying PD and control subjects, performance of 100% specificity and precision is achieved, while maintaining sensitivity of close to 89%. In the multi-class classification task of characterizing parkinsonian gait by distinguishing among PD with significant gait disturbance, PD with no significant gait disturbance, and control subjects, 91.7% class recall for control subjects is achieved and the model performs with 84.6% precision for PD subjects with significant gait disturbance. The features selected for this classification task indicate the features of gait that are principal in discriminating gait abnormalities due to PD compared to a normal gait. These results demonstrate the ability of our wireless inertial sensor system to successfully detect the presence of PD based on physical features of gait and to identify the specific features that characterize parkinsonian gait.

I. INTRODUCTION

PARKINSON'S disease (PD) is a neurodegenerative disease in which patients often exhibit a disturbance of gait. Gait is usually evaluated subjectively by clinicians, and it may be difficult to distinguish between different disorders of gait. We are developing a system that utilizes wireless inertial sensors to perform quantitative gait analysis and assist clinicians in making an accurate diagnosis of PD based on physical features of the gait.

II. BACKGROUND AND RELATED WORK

The diagnosis of disease can be viewed as a classification task that sorts subjects into two classes: those with disease and those without it. Machine learning methods and, in

particular, support vector machines (SVMs) are a powerful tool for use in classification tasks, especially in high dimensions. In Parkinson's disease (PD), SVM has been used to estimate the severity of tremor, bradykinesia, and dyskinesia [1]. Studies of gait in PD using SVMs have been limited to plantar pressure data [2] and ground reaction forces [3]. The equipment used for these studies only provides information related directly to the foot hitting the ground. For example, stride time can be obtained, but no information can be obtained regarding how the foot moves during a stride. This limits the understanding of the physical characteristics of gait that distinguish PD patients from healthy subjects.

III. METHOD

A. Experiments

In this study, we use wireless inertial sensors attached to the foot to continuously monitor the motion of the foot during walking. The sensors used are assembled from the Micro-AHRS model of sensors manufactured by MicroStrain, Inc [4]. They are small and lightweight, measuring 41 mm x 63 mm x 32 mm and weighing 39 grams. The sensors include a 50-g triaxial accelerometer and 1200 deg/s triaxial rate gyroscope, outputting raw 3D acceleration and angular rate data. This raw data is transformed into 3D displacements as previously described using the principal of "zero velocity updating" [5]. This technique assesses the sensor drift at the end of every step, i.e. when foot velocity is zero, and allows the removal of cumulative effects of drift over one step to obtain accurate integration results [6].

Data were collected from 23 subjects with a clinical diagnosis of PD attending the UCSF Parkinson's Disease Clinic and Research Center in San Francisco, and from 16 age-matched control subjects without history of neurological disease. Of the subjects diagnosed with PD, 11 had a clinically significant disturbance of gait, and 12 had no such disturbance. The wireless inertial sensors were attached to subjects' feet using a foot mount as shown in Figure 1(a). Subjects then walked along a predetermined path along a hallway in the UCSF clinic. Data were transmitted over Bluetooth to a nearby handheld PDA and stored on the device for later processing. In total, collecting data from sensors on both feet in some cases, 21 PD with significant gait disturbance, 24 PD with no significant gait disturbance, and 24 control data points were obtained.

Manuscript received April 23, 2010. This work was supported in part by the National Science Foundation Graduate Research Fellowship Program.

I. Tien and S. D. Glaser are with the Center for Information Technology Research in the Interest of Society (CITRIS), University of California, Berkeley, CA 94720 USA (emails: itien@berkeley.edu and glaser@berkeley.edu).

M. J. Aminoff is with the Neurology Department, University of California, San Francisco, CA 94143 USA (email: aminoffm@neurology.ucsf.edu).

Figure 1(b) shows an example of the 3D displacement signal for a 10-second time segment during the course of an experiment for one subject. The pitch, roll, and yaw directions are as given in Figure 1(a).

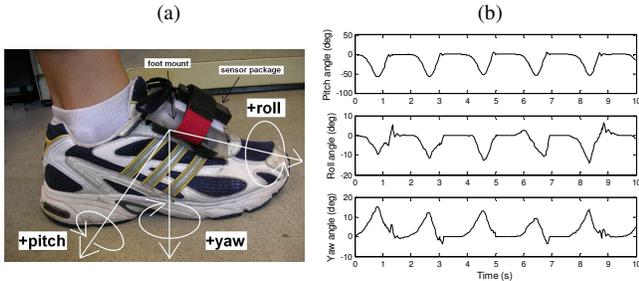


Figure 1: Orientation of axes on foot with sensor package attached using foot mount (a) and sample 3D displacement signal in pitch, roll, and yaw directions over a 10-second segment (b).

B. Data Analysis

The data analysis is performed for two separate classification tasks. The first is to detect the presence of PD, i.e. a binary classification task of distinguishing between PD and control. The second is to characterize parkinsonian gait, i.e. a multi-class problem of distinguishing among PD with significant gait disturbance, PD with no significant gait disturbance, and control. A similar data analysis algorithm is employed for both tasks. This algorithm involves multiple modules (Figure 2). Each analysis module is discussed in detail in their following subsections.

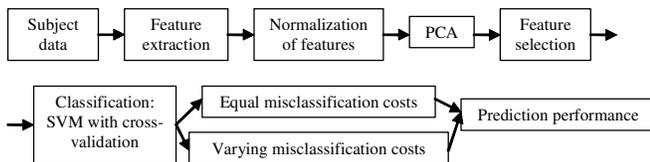


Figure 2: Flow chart of data analysis algorithm.

1) *Feature extraction*: From the 3D displacement data signals for each subject, 67 features of physical relevance are extracted. These are time-domain features of 3D displacements, so values of features clearly correspond with physical behaviors. For example, range of motion in the pitch direction corresponds with maximum motion from plantar flexion to dorsiflexion of the foot during walking, while a maximum positive rotation of the foot in the pitch direction corresponds with the foot's maximum angle of dorsiflexion reached during a step.

2) *Normalization of features*: The values of each feature are in varying ranges. For example, the range of motion in the pitch direction spans from 18 degrees to 102 degrees whereas the maximum angle of dorsiflexion spans from 1 degree to 33 degrees. In order to accurately compare the different features, the features are normalized. Figure 3 is a plot of two features extracted from the data: maximum range of motion in the pitch direction, and maximum angle of dorsiflexion reached during a step as averaged over all steps of the walk. Figures 3(a) and 3(b) show the normalized feature values for the binary and multi-class tasks, respectively.

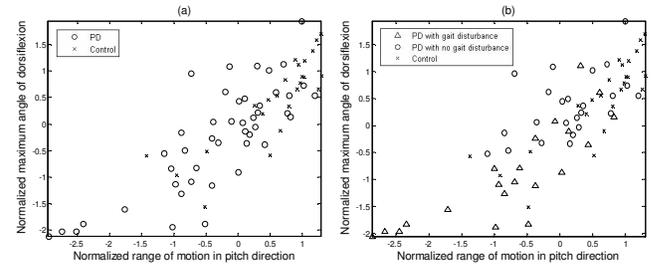


Figure 3: Normalized values for two features extracted for the binary (a) and multi-class (b) tasks.

3) *Principal Component Analysis (PCA)*: The 67 features extracted from the data produce a 67-dimensional space. However, not all the features contribute equally to overall data variance. To reduce dimensionality while maximizing the data variance in the lower-dimensional space, principal component analysis (PCA) is used. The eigenvalue decomposition of the covariance matrix allows for each feature to be ranked in order of its contribution to overall data variance.

4) *Feature selection*: Using the results of the PCA, features are selected for use in classification. Of the 67 features, the first 11 principal components are selected for the first classification task of distinguishing between PD and control. These 11 features cumulatively account for 80.2% of the variance in the data. The subsequent contributions of additional features to the data variance are less than 2.3% per feature, decreasing to 0.0%. An additional four features are selected for the second classification task of characterizing parkinsonian gait for a total of 15 features. All selected features for both classification tasks are listed in Table I. These features include statistical variability as measured using standard deviation (SD). The (+) or (-) next to the feature names indicate whether PD patients exhibit higher or lower values for the features, respectively, relative to controls. These features allow us to see which components of gait are most significant in characterizing parkinsonian gait.

TABLE I
PHYSICAL FEATURES SELECTED BASED ON PCA FOR USE IN CLASSIFICATION MODEL, WITH PD SUBJECTS HAVING HIGHER (+) OR LOWER (-) VALUES FOR THE FEATURES RELATIVE TO CONTROLS

Pitch features (deg)	Roll features (deg)	Yaw features (deg)
Range of motion (-)	Range of motion (-)	Range of motion (-)
Maximum angle of dorsiflexion (-)	Maximum positive roll angle (-)	Maximum positive yaw angle (-)
Maximum angle of plantar flexion (-)	Maximum negative roll angle (+)	Maximum negative yaw angle (+)
Plantar flexion SD (+)		
Single-step maximum of maximum angle of plantar flexion (-)		
Additional features utilized in characterizing parkinsonian gait		
Overall 3D SD (-)	Single-step maximum of maximum negative roll angle (-)	
Maximum cadence (+)	Single-step minimum of maximum negative roll angle (-)	

5) *Classification using support vector machine (SVM)*: With the selected features, the support vector machine (SVM) method is then used for classification. SVM optimally separates classes by maximizing the margin between classes and minimizing the classification error. In Figure 3, there is clear overlap between the two classes of PD and control, and among the three classes of PD with significant gait disturbance, PD with no significant gait

disturbance, and control. The data are therefore linearly inseparable and a nonlinear kernel must be used in the SVM process. The nonlinear radial basis function (RBF) kernel is chosen, and this kernel function maps the data into a new space, where a k -dimensional hyperplane is used to separate the classes, and k is the number of features used in the model. Parameters for the kernel are also selected to optimize the tradeoff between maximizing the margin between classes and minimizing the misclassification costs. Figure 4 is an illustration of SVM in a 2D space.

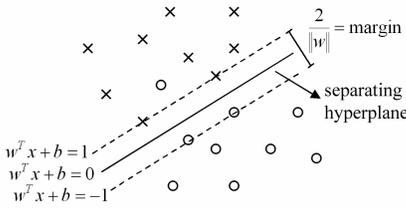


Figure 4: Illustration of SVM in a 2D space.

Utilizing the LIBSVM package [7], the RBF kernel is used and a cost-sensitive SVM classification model is built for the binary classification task by solving the optimization problem of (1).

$$\min_{w, b, \xi} \frac{1}{2} \|w\|^2 + C_{PD} \sum_{i=1}^{n_{PD}} \xi_i + C_{Control} \sum_{j=1}^{n_{Control}} \xi_j \quad (1)$$

$$\text{subject to } y_i (w^T x + b) \geq 1 - \xi_i, \quad \xi_i \geq 0$$

where w, b, ξ are optimization parameters; C_{PD} and $C_{Control}$ are the costs for misclassifying PD and control subjects, respectively; $y \in \{1, -1\}^n$ is the vector of labels, i.e. 1 for PD and -1 for control, for n_{PD} and $n_{Control}$ PD and control data points, respectively; and x is the feature value for the data point to be classified. Points are classified based on the sign of $w^T x + b$, i.e. on which side of the hyperplane the data point falls on. Thus, data points are classified according to (2).

$$\text{Class} = \begin{cases} \text{PD if } w^T x + b \geq 1 - \xi_i, y_i = 1 \\ \text{control if } w^T x + b \leq -1 + \xi_i, y_i = -1 \end{cases} \quad (2)$$

In the case of equal misclassification costs, $C_{PD} = C_{Control} = 1$. In predicting the presence of PD, however, the cost of misclassifying a subject as PD or control may differ. A cost-sensitive learner is used to accommodate these varying misclassification costs. In a more conservative approach, one errs on the side of caution and classifies a subject as normal rather than PD in the face of uncertainty. This translates into a higher cost for misclassifying a control subject as PD compared to misclassifying a PD subject as control and $C_{PD} < C_{Control}$.

6) *K-fold cross-validation with stratified sampling*: Finally, K-fold cross-validation with stratified sampling is used to test the SVM classification model and estimate its predictive performance. K=10 was chosen for optimal model estimation from real-world data [8]. The data set is

partitioned into 10 subsets using stratified sampling where the class distribution in each subset, i.e. proportion of PD to control data points, is the same as the distribution in the entire data set. In each of 10 iterations, one subset is retained as testing data for the model, and the remaining nine subsets are used to train the model. Performing multiple repetitions of the cross-validation process checks for overfitting of the model, where the model fits the training data very well, but performs poorly in predicting class labels for new data. Overall performance of the model is obtained by averaging the performance over the 10 rounds.

IV. RESULTS

A. Detecting presence of PD: binary classification task

For the two cases of equal misclassification costs for PD and control data points, and varying misclassification costs where the cost of misclassifying a control subject as PD is assigned a 50% higher cost compared to misclassifying a PD subject as control, the prediction performance is given in Table II. Sensitivity measures how well the prediction model recognizes PD cases. Specificity measures how well the model identifies control cases. The false positive rate measures how often controls are misclassified as PD. Precision, or positive predictive value, measures how well the prediction of PD reflects the underlying presence of disease.

TABLE II
PREDICTION PERFORMANCE METRICS FOR CASES OF EQUAL OR VARYING MISCLASSIFICATION COSTS

	Equal costs	Varying costs
Sensitivity	93.3%	88.9%
Specificity	95.8%	100.0%
False positive rate	4.2%	0.0%
Precision	97.7%	100.0%

B. Characterizing parkinsonian gait: multi-class classification task

Table III gives the prediction performance in the multi-class problem of distinguishing among PD with significant gait disturbance, PD with no significant gait disturbance, and control. Because this classification task is no longer binary, the performance metrics of sensitivity and specificity do not apply. Precision is still applicable, and in this case, indicates the rate at which data points classified as a member of a certain class actually belong to that class. Recall is used to measure prediction performance by indicating the rate at which data points of a certain class can be correctly identified as members of that class. In this classification task, varying misclassification costs does not improve the performance of the classifier in any metric without a significant decrease in performance in other metrics.

TABLE III
PREDICTION PERFORMANCE FOR MULTI-CLASS CLASSIFICATION TASK

	Class recall	Class precision
PD with gait disturbance	52.4%	84.6%
PD with no gait disturbance	66.7%	64.0%
Control	91.7%	71.0%

V. DISCUSSION

From Figure 3, we can expect that it is difficult to separate the classes using only two features. Therefore, the SVM approach, using the RBF kernel, is useful for working in a higher-dimensional space where the values for the principal features are utilized to achieve the best nonlinear separation between the classes.

The features selected for use in classification as shown in Table I all together describe the 3D motion of the foot through the air during walking. These principal features are not obtainable from alternative gait analysis methods or equipment such as looking at plantar pressure [2] or ground reaction forces [3].

An examination of the features utilized in the multi-class classification task to distinguish PD subjects with gait disturbance from the other two classes allows us to characterize parkinsonian gait. The gait abnormalities of PD are characterized by decreased range of motion in all three angular directions. PD gait is characterized by decreased dorsiflexion and plantar flexion, a tendency to roll the foot towards the inside during the swing phase of the step, and a decreased swing of the foot in the positive yaw direction. PD subjects with significant gait disturbances exhibit greater variability in plantar flexion, a result consistent with existing findings about increased gait variability in PD [9]. Lower variability overall found here may be due to decreased magnitude of all variables, and therefore decreased magnitude of overall SD. Finally, increased maximum cadence is consistent with previous studies that have found parkinsonian patients to exhibit increased cadence compared to controls [10].

Using the physical features selected in Table I, the model performs both binary and multi-class classification tasks. Table II summarizes the performance of our model in the binary classification task of separating PD and control. Looking at the equal costs column shows that the data analysis algorithm produces a prediction model of high sensitivity and specificity of over 93%, with a low false positive rate of under 5%, and high precision of 97.7%. This performance exceeds the results of previous studies by around 10% in all metrics [3].

Comparing the results for an equal misclassification costs model with a cost-sensitive model with varying costs shows that the method successfully accommodates differing costs for misclassification. The difference in costs is a parameter that can be adjusted by clinicians. A 50% higher cost for misclassifying a control subject as PD compared to misclassifying a PD as control performs at 100% specificity and precision, while maintaining sensitivity of 88.9%. This performance exceeds the results of previous studies by close to 15% [2]. The chosen relative cost difference for misclassification is representative of a more conservative approach to diagnosis, where a clinician wants to be certain of the PD diagnoses that are made and avoid misdiagnoses of the presence of PD.

From the prediction performance in the multi-class classification task as shown in Table III, the model performs best with respect to recall for control subjects. A class recall for the control group of 91.7% indicates that 91.7% of control subjects are correctly diagnosed not to have PD. Having a high class recall for controls is significant in a conservative diagnostic approach as described previously. The model performs best with respect to precision for PD subjects with gait disturbance. A precision of 84.6% indicates that for 84.6% of subjects identified as PD with gait disturbance, the prediction from the model accurately reflects the underlying condition of the subject. It should be noted that the remaining 15.4% of subjects incorrectly identified as PD with gait disturbance are not controls, but PD subjects that have not been clinically observed to have a gait disturbance. The current clinical determination of gait disturbance is subjective, so some variability with respect to gait disturbance or not in PD subjects is expected.

VI. CONCLUSION

Our wireless inertial sensor system allows us to quantitatively measure differences between PD and healthy gait. This is an important step in creating a system to assist clinicians in diagnosing and evaluating patients using gait analysis, and in more fully understanding manifestations of neurological disease in disturbances of gait.

REFERENCES

- [1] S. Patel, K. Loriincz, R. Hughes, N. Huggins, J. Growdon, D. Standaert, M. Akay, J. Dy, M. Welsh, and P. Bonato, "Monitoring motor fluctuations in patients with Parkinson's disease using wearable sensors," *IEEE Transactions on Information Technology in Biomedicine*, Vol. 13, No. 6, pp. 864-873, 2009.
- [2] H. S. Jeon, J. Han., W. J. Yi, B. S. Jeon, and K. S. Park, "Classification of Parkinson gait and normal gait using spatial-temporal image of plantar pressure," *Proceedings of the 30th Annual International IEEE EMBS Conference*, pp. 4672-4675, 2008.
- [3] M. Yang, H. Zheng, H. Wang, and S. McClean, "Feature selection and construction for the discrimination of neurodegenerative diseases based on gait analysis," *Pervasive Health*, 2009.
- [4] MicroStrain, "3DM-GX2 Product Datasheet," accessed April 1, 2010, from <http://www.microstrain.com/pdf/3dm-gx2_datasheet_v1.pdf>.
- [5] I. Tien, S. D. Glaser, R. Bajcsy, D. S. Goodin, and M. J. Aminoff, "Results of Using a Wireless Inertial Measuring System to Quantify Gait Motions in Control Subjects (accepted for publication)," *IEEE Transactions on Information Technology in Biomedicine*, to be published.
- [6] L. Hutchings, *Rotational Sensor System*, US patent 6,305,221, to Acceleron Technologies, Patent and Trademark Office, 2001.
- [7] C. C. Chang and C. J. Lin, LIBSVM: a library for support vector machines, 2001. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [8] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," *International Joint Conference on Artificial Intelligence*, pp. 1137-1143, 1995.
- [9] J. M. Hausdorff, M. E. Cudkowicz, R. Firtion, J. Y. Wei, and A. L. Goldberger, "Gait variability and basal ganglia disorders: stride-to-stride variations of gait cycle timing in Parkinson's disease and Huntington's disease," *Movement Disorders*, Vol. 13, No. 3, pp. 428-437, 1998.
- [10] M. E. Morris, R. Ianssek, T. A. Matyas, and J. J. Summers, "The pathogenesis of gait hypokinesia in Parkinson's disease," *Brain*, Vol. 117, No. 5, pp. 1169-1181, 1994.