Feature Selection for Accelerometer-Based Posture Analysis in Parkinson's Disease

Luca Palmerini, Laura Rocchi, Sabato Mellone, Franco Valzania, and Lorenzo Chiari

***Abstract*—Posture analysis in quiet standing is a key component of the clinical evaluation of Parkinson's disease (PD), postural instability being one of PD's major symptoms. The aim of this study was to assess the feasibility of using accelerometers to characterize the postural behavior of early mild PD subjects. Twenty PD and 20 control subjects, wearing an accelerometer on the lower back, were tested in five conditions characterized by sensory and attentional perturbation. A total of 175 measures were computed from the signals to quantify tremor, acceleration, and displacement of body sway. Feature selection was implemented to identify the subsets of measures that better characterize the distinctive behavior of PD and control subjects. It was based on different classifiers and on a nested cross validation, to maximize robustness of selection with respect to changes in the training set. Several subsets of three features achieved misclassification rates as low as 5%. Many of them included a tremor-related measure, a postural measure in the frequency domain, and a postural displacement measure. Results suggest that quantitative posture analysis using a single accelerometer and a simple test protocol may provide useful information to characterize early PD subjects. This protocol is potentially usable to monitor the disease's progression.**

***Index Terms*—Accelerometer, Parkinson's disease (PD), feature selection, posture.**

I. Introduction

THE human postural control system has been the object of many investigations in the last few decades in various fields, including orthopaedics, neurology, and rehabilitation [1]–[4]. Instrumented analysis of posture is usually performed by means of force platforms, able to measure the ground reaction force necessary to maintain upright posture and counteract gravity. From the measurement of the ground reaction force components, it is possible to compute the position of the center of pressure (CoP), defined as the application point of the reaction vector on the force platform surface [5].

Postural analysis is often integrated with human movement analysis. The latter requires, besides the force platform for force measurement, optical motion analysis systems based on infrared video cameras and markers on the body or, more recently, markerless techniques, and standard video cameras [6], [7]. Although these movement analysis techniques have provided essential and accurate information for at least two decades [8], their use in clinical practice has been partly limited by cost and the need for qualified personnel.

For these reasons, inertial sensors are being increasingly investigated for movement analysis and to date the results are promising. Such studies have the potential to greatly increase the use of quantitative methods for movement analysis in clinical practice [9]. This approach overcomes the limits of traditional movement analysis laboratories, and has been shown to be a valid alternative both for dynamic motor tasks [10]–[12] and static trials [13], [14]. Inertial sensors are now well established as measurement devices [15], monitor systems, and activity monitors [16]. Additionally, they are much less cumbersome and expensive than motion analysis and force platform systems, even if these two still remain the gold standards in the analysis of human movement and balance.

Analysis of posture and balance in quiet standing is one of the components of the clinical evaluation of Parkinson's disease (PD). Postural instability is an important and disabling symptom of PD, often leading to falls [17] and injuries; therefore, assessing the standing balance of a PD patient is a promising method for identifying individuals with a high risk of falling due to balance impairments [17]. Several studies, mostly using force plates, have investigated posture in PD subjects both in static and dynamic conditions [3], [17]–[22]. Results from these studies suggest that quantitative posturography may be useful, in addition to clinical measures, for the objective assessment of postural impairments and a better understanding of the pathological deficits in postural control of PD subjects. Even though accelerometers are not yet routinely used for measuring quiet standing in PD, they have already been used to assess several motor complications in PD [10], [11], [23], [24]. This reflects the recent growing interest in quantitative tools for the assessment of PD motor symptoms, with an emphasis on a less invasive and more ubiquitous approach.

Since subjects in early stages of PD do not show severe clinical signs of postural instability [25], [26], quantitative evaluation of the postural performance could provide early markers for later developing problems [21]. Subjects in the present study are in an early mild stage of Parkinson's disease, evaluated OFF medication. We hypothesized that differences in postural behavior between controls and early mild PD subjects could be detected and described by quantitative features derived from accelerometer signals recorded during quiet stance (QS). Data mining techniques have recently been used in PD

Manuscript received July 26, 2010; revised December 15, 2010; accepted January 12, 2011. Date of publication February 24, 2011; date of current version May 4, 2011. This work was supported by the Regione Emilia-Romagna, High Technology Network initiative (AER-TECH—Automation, Electronics, and Bioengineering: Technologies for Manufacturing and People).

L. Palmerini, L. Rocchi, S. Mellone, and L. Chiari are with the Department of Electronics, Computer Science, and Systems, University of Bologna, Bologna 40136, Italy (e-mail: luca.palmerini@unibo.it; l.rocchi@unibo.it; sabato.mellone@unibo.it; lorenzo.chiari@unibo.it).

F. Valzania is with the Department of Neuroscience, University of Modena, and Reggio Emilia, Modena 41100, Italy (e-mail: f.valzania@ausl.mo.it).

Digital Object Identifier 10.1109/TITB.2011.2107916

studies on datasets recorded using quantitative tools (mostly wearable technology) [20], [24], [27]–[29]. These techniques were also applied in [30] to a semiquantitative dataset formed by epidemiological data, clinical scales, and questionnaires. In all these studies, data mining techniques served as suitable tools to extract information of clinical and scientific relevance from the available datasets. In our study, we use feature selection and classification to identify, from the high-dimensional dataset available, the features providing the best discrimination between controls and PD subjects.

In order to evaluate the feasibility and the impact of accelerometry-based posturography, we investigate measures computed from acceleration signals including tremor measures, postural acceleration measures, and postural displacement measures. We use a classification procedure based on *a priori* knowledge about the presence or absence of the pathology. At the same time, within the classification procedure, we implement a feature selection technique aimed to identify the subset of measures that best discriminate between early-mild PD subjects and control subjects. Different testing conditions are considered, to identify those sensory and attentional demands that, in QS, are more capable of disclosing postural differences between PD and control subjects. Conditions include manipulations of the visual and peripheral somatosensory input, and a dual task paradigm challenging attentional and cognitive resources of subjects.

The data analysis techniques applied to accelerometer data during QS are used to:

1) compute summary measures that characterize postural behavior of subjects with PD;
2) identify the most relevant summary measures (feature selection) able to discriminate between PD and control subjects.

II. Methods

A. Experimental Sessions

1) *Subjects:* The subjects were 20 early-mild PD OFF medication (Hoehn & Yahr ≤ 2.5, 62 ± 7 years old, 12 males) and 20 healthy age-matched control subjects (CTRL, 64 ± 6 years old, 7 males). The OFF condition was obtained by a levodopa washout of at least 18 h and a dopamine agonist washout of at least 36 h.

The unified PD rating scale (UPDRS) was assessed for each subject by an expert neurologist the same day of the experimental sessions. In this study, Sections II and III of the UPDRS were considered. The average value of motor UPDRS (Section III) in PD subjects was $26.6 \pm 7.1/108$.

2) *Tests and Conditions:* Subjects were asked to stand upright, barefoot, with arms crossed on the chest, looking at a visual marker (a black circle, 5 cm in diameter) placed on a wall 2.5 m in front of them. Foot placement was kept consistent over trials using an averaged preferred position traced on the floor [31]. The subjects were tested in five different QS conditions. Descriptions of conditions, acronyms, and perturbed subsystems are reported in Table I.

The dual task administered in the eyes open dual task (EODT) condition consisted of a concurrent cognitive task: counting audibly backward from 100 by 3's. In each condition, a different aspect of postural control was perturbed in order to detect a possible deterioration of a particular postural control mechanism in PD subjects. The measurement session was organized in three sequential blocks. Each block was made up of five consecutive trials, corresponding to the different conditions presented in the following order: eyes open (EO), eyes closed (EC), EODT, eyes open on a foam-rubber support (EOF), and eyes closed on a foam-rubber support (ECF). Each trial lasted 30 s.

TABLE I
QS Conditions

Acronym	Description	Perturbed postural sub-system
EO	Eyes Open on a rigid surface	None
EC	Eyes Closed on a rigid surface	Visual
EODT	Eyes Open with Dual Task on a rigid surface	Attention
EOF	Eyes Open on a Foam-rubber support	Somato-sensory
ECF	Eyes Closed on a Foam-rubber support	Somato-sensory + Visual

3) *Measurement System*: We used a triaxial accelerometer, McRoberts Dynaport Micromod, with a sample rate of 100 Hz, a range of ± 2 g, and a resolution of 1 mg. The accelerometer was worn on the lower back by means of an elastic waist belt, at the level of the fifth lumbar vertebra.

B. Preprocessing and Measure Computation

The acceleration signals along the two orthogonal axes of the accelerometer were considered for the analysis: the first aligned with the back/forward direction of sway and coincident with the biomechanical anteroposterior (AP) axis of the body; the second in the left/right direction and coincident with the biomechanical mediolateral (ML) axis of the body.

Three different versions of each signal were considered: raw acceleration, tremor-free acceleration, and tremor-free displacement (approximate center of mass, CoM, displacement). From each of these versions specific measures were computed (see Fig. 1).

The preprocessing of the signals was accomplished following the steps listed as follows (see Fig. 1).

1) Ad hoc measures were computed from the raw signals to evaluate the presence and amplitude of tremor (compare Fig. 1(a) and Table II).
2) A filtering procedure based on Hilbert–Huang transform(HHT) [32] was applied to the raw signals [see Fig. 1(b)] in order to consider only properties related to postural control (i.e., tremor free). HHT can deal with nonstationary processes (such as tremor) and nonlinear systems (such as the postural control systems) and can be used for signal artifact reduction [33]; therefore, this technique was used to isolate and later suppress the effect of tremor. Since tremor frequency in PD is usually in the band 4–7 Hz [34], the 0– 3.5 Hz interval was the

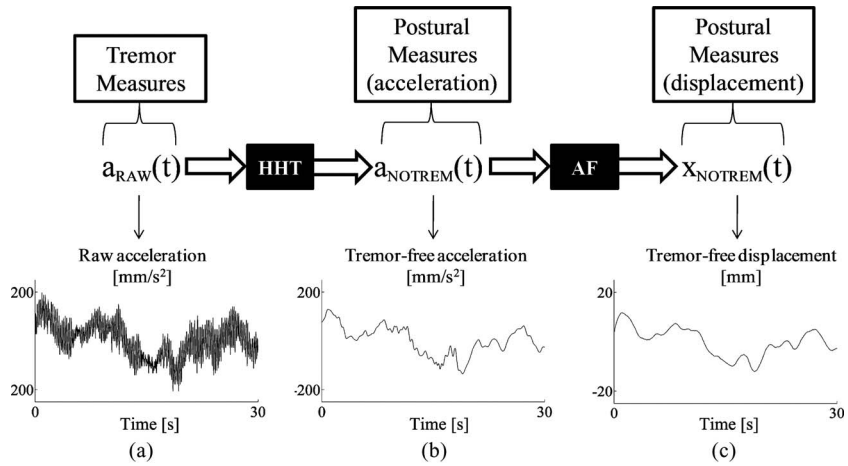

Fig. 1. Schematic representation of the different processing stages of the acceleration signals. (a) Raw acceleration is filtered by the HHT, obtaining (b) tremor-free acceleration; then the anthropometric filter (AF) is applied to get (c) tremor-free displacement.

TABLE II
TREMOR AND POSTURAL MEASURES: ACRONYMS AND BRIEF DESCRIPTIONS. THE LAST COLUMN REPORTS THE DIRECTIONS ALONG WHICH THEY WERE COMPUTED

Tremor Measures	**Domain**	**Description**	**Directions**
Power HF	Frequency	Fraction of power of the signal for high frequencies (between 4 and 7 Hz) [%]	AP, ML
Peak HF	Frequency	Frequency of the maximum of the PSD for high frequencies (> 4 Hz) [Hz]	AP, ML
RHL	Frequency	Power ratio of the high (3.5 – 15 Hz) to low (0.15 – 3.5 Hz) frequency components (unitless)	AP, ML
Postural Measures (acceleration)	**Domain**	**Description**	
F50	Frequency	50% power frequency: frequency containing 50% of the total power [Hz]	AP, ML
F95	Frequency	95% power frequency: frequency containing 95% of the total power [Hz]	AP, ML
CF	Frequency	Centroidal Frequency: the frequency at which spectral mass is concentrated [Hz]	AP, ML
FD	Frequency	Frequency Dispersion: a unitless measure of the variability of the PSD frequency content (zero for pure sinusoid, increases with spectral bandwidth to one) (unitless)	AP, ML
Entropy	Frequency	Power spectrum entropy of acceleration (unitless)	AP, ML
JI	Time	Jerk Index: a function of the time derivative of the acceleration; it is commonly viewed as an index of smoothness [mm^2/s^5]	AP, ML
NJI[a]	Time	Normalized Jerk Index: JI is normalized by dividing it by SP^2 [$1/s^5$]	AP, ML
Postural Measures (displacement)	**Domain**	**Description**	
MD	Time	Mean Distance from center of CoM trajectory [mm]	AP, ML
RMS	Time	Root Mean Square distance from center of CoM trajectory [mm]	AP, ML
Range	Time	Range of CoM displacement [mm]	AP, ML
SP	Time	Sway Path: total CoM trajectory length [mm]	AP, ML
MV	Time	Mean Velocity of the CoM, computed as the median value of the absolute value of the time series obtained through the derivative of the displacement [mm/s]	AP, ML
SA	Time	Sway Area: area included in CoM displacement per unit of time [mm^2/s]	planar
CEA	Time	Confidence Ellipse Area: area of 95% confidence ellipse [mm^2]	planar
mSCEA	Time	Minor Semiaxis of CEA [mm]	planar
MSCEA	Time	Major Semiaxis of CEA [mm]	planar
|90-Mdir|	Time	Angular deviation from AP sway of the Max Variance Direction [deg]	planar

[a]NJI's computation requires information both from acceleration (JI) and displacement (SP); it was considered as an acceleration-based measure because it mainly describes properties of the acceleration signal.

bandwidth isolated by the HHT procedure and then considered for computation of postural acceleration measures.

3) In order to get measures related to displacement of CoM on the horizontal plane, HHT-filtered acceleration signals underwent a low-pass filtering, with a cutoff frequency of 0.5 Hz and a static gain of $-1/g$, where g is the gravitational acceleration. The transfer function was obtained from a simple biomechanical model, in the sagittal plane, based on inverted pendulum modeling of the human body during QS (see Fig. 2) [5]. This assumption and the small

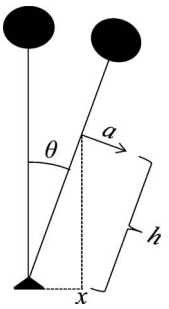

Fig. 2. Inverted pendulum model of standing posture.

angles of sway in quiet standing trials lead to the following equation, which relates acceleration to sway angle and CoM height

$$a(t) = h\,\ddot{\theta}(t) - g\,\sin\,\theta(t) \approx h\,\ddot{\theta}(t) - g\,\theta(t) \quad (1)$$

where a is the accelerometer output in the AP direction, θ is the sway angle with respect to vertical, and h is the height of the inertial sensor (which is assumed to be the same height as the CoM). The corresponding transfer function is

$$H(s) = \frac{\theta(s)}{a(s)} = -\frac{1}{g - h\,s^2}. \quad (2)$$

Equation (2) in the frequency domain can be written as

$$H(j\omega) = -\frac{1}{g\ +\ h\,\omega^2} = -\frac{1}{g}\frac{1}{1 + (\omega/\omega_n)^2}, \qquad \omega_n = \sqrt{\frac{g}{h}}. \quad (3)$$

With the reasonable assumption of h being, on average, approximately 1 m, displacement of the CoM projection on the ground ($x = h\,\sin\theta$) can be approximated as $x = \sin\theta \approx \theta$. Thus, (3) represents the relation between the acceleration and the AP projection of CoM in the frequency domain (i.e., the frequency response of the filter obtained from the biomechanical model). Typical postural displacement measures, traditionally obtained from CoP [35], were computed from x (see Fig. 1(c) and Table II). The same processing was performed on the ML signals, to obtain a displacement signal in this direction. It may be considered that the aim of the procedure was not a precise estimation of the CoM, but rather the achievement of a signal reasonably approximating the characteristics of a displacement, in order to compute corresponding displacement-related measures (such as area and main direction of oscillation during QS, as described in detail in Table II).

Table II summarizes the measures computed from the accelerometers in each test condition. In particular, tremor measures describe the characteristics of the acceleration signal in the frequency domain, assuming that PD tremor is localized between 4 and 7 Hz [34]. Then, several postural measures were computed from tremor-free accelerations to characterize postural sway in the time and frequency domains; the mathematical definitions of these measures are parallel to those traditionally computed from the CoP in posturographic studies and used in clinical practice [36], with the exception of some measures that strictly rely on the acceleration.

Regarding postural measures computed from acceleration, five of them quantify the properties of the acceleration in the frequency domain (F50, F95, CF, FD, and Entropy), and two in the time domain (JI, NJI). JI is computed as a function of the time derivative of the acceleration:

$$\mathrm{JI} = \frac{1}{2}\int \dot{a}^2(t)\,dt. \quad (4)$$

Measures derived from displacement are computed in the time domain and describe the amount and direction of sway. Among these, mean velocity is obtained through the derivative of the displacement.

Both tremor and postural measures were computed for the AP and ML directions, with the exception of five postural measures that describe planar (bidimensional) characteristics of the displacement (SA, CEA, mSCEA, MSCEA, and 90-Mdir, see Table II).

For each subject, the mean values of the measures from the three trials in each QS condition were considered in the following feature selection procedure.

C. Feature Selection Procedure

As shown in Table II, 35 measures were computed from accelerometer signals for each QS condition (EO, EC, EODT, EOF, ECF); considering the five conditions jointly, the total number of measures (features) is very high (175) with respect to the available sample of subjects (40).

The feature selection process is characterized by an objective function measuring the discriminative ability of a subset of features, and by a search strategy to select candidate subsets. We implemented a wrapper feature selection approach [37] whose objective function is the predictive accuracy of a given classifier. The classifiers that we considered in this study are some of the most commonly used classifiers which indeed could be easily integrated into the specific feature selection procedure that we designed. These are the linear and quadratic discriminant analysis (LDA and QDA, respectively), mahalanobis classifier (MC), logistic regression (LR), K-nearest neighbors (KNN), and support vector machines (SVM). A brief description of each of them is reported in Table III.

The search strategy used in this study consisted of an exhaustive search for subsets containing one to three features; the limit of three was chosen to permit a clinical interpretation of the results (it would be difficult to associate too many measures with different aspects of the disease). In addition, this limit keeps the search strategy computationally acceptable. Since feature selection is part of the design of the classifier, it should be performed only on the training set, in order to avoid the so-called feature selection bias in the final evaluation of the accuracy of the classifier [43]. This bias may occur when the accuracy of a classifier is estimated using all the available data (instead of the training set only). The most common solution to this problem, when the sample size is not large enough to split the data into a training set for feature selection and a testing set for accuracy evaluation, is to use a nested cross-validation procedure [37]. In this case, the internal cross validation for feature selection is repeated for every training set resulting from the external cross validation,

TABLE III
CLASSIFIERS: ACRONYMS AND BRIEF DESCRIPTION

CLASSIFIER	ACRONYM	DESCRIPTION	Notes
Linear Discriminant Analysis [38], [39]	LDA	LDA assumes normal distribution of the data, with equal covariance matrix for both classes. The separating hyperplane is obtained by seeking the projection that maximizes the distance between the two classes' means and minimizes the interclass variance.	
Quadratic Discriminant Analysis [38], [39]	QDA	Similar to LDA, but it does not assume the equal covariance matrix for both classes.	
Mahalanobis Classifier [38], [39]	MC	It assumes a normal distribution for each prototype of a certain class and the feature vector is assigned to the nearest prototype according to Mahalanobis distance.	
Logistic Regression [40]	LR	It aims to translate the information from continuous independent features to the [0, 1] interval which contains the probability associated with the dichotomous choice outcomes (0, 1).	In this study the dichotomous choice outcomes (0,1) are CTRL and PD.
K-Nearest Neighbors [41]	KNN	It assigns an unseen point to the dominant class among its k nearest neighbors within the training set. These neighbors are obtained using a distance metrics.	Euclidean distance and k=1 were chosen in this study.
Support Vector Machines [42]	SVM	To identify classes, it uses a discriminant hyperplane which maximizes the margins, i.e. the distance from the nearest training points.	The version with linear boundaries was used in this study.

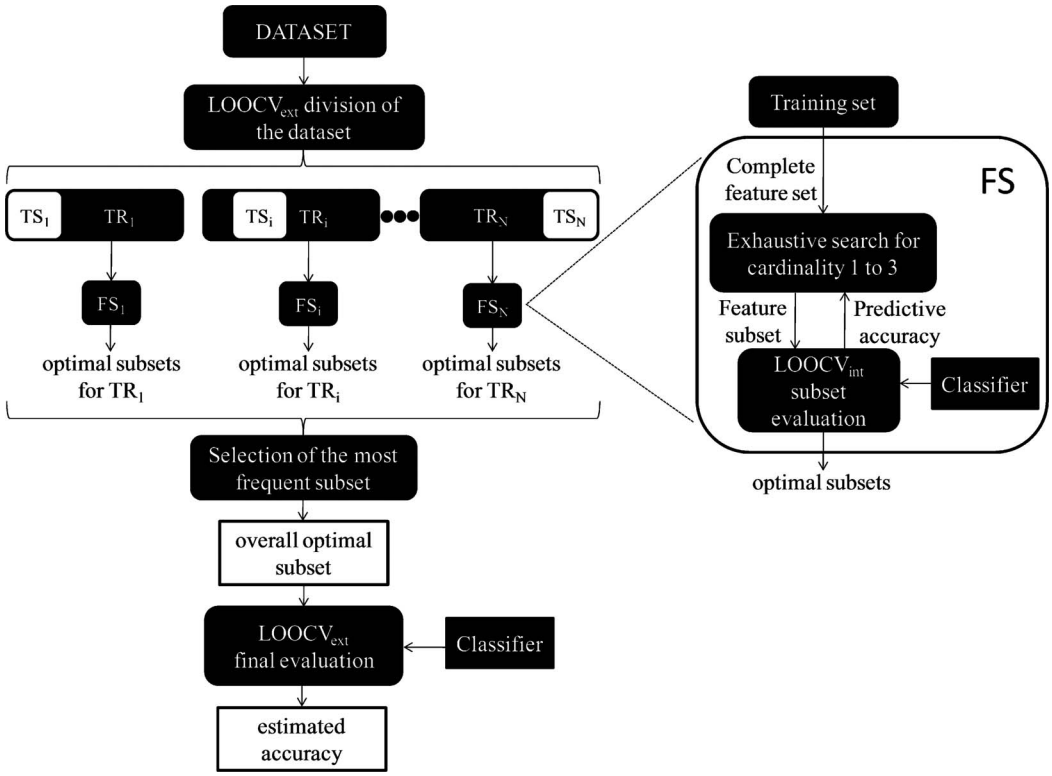

Fig. 3. Feature selection procedure.

which is used to estimate the accuracy of the classifier. In this study, because of the small sample size, a leave-one-out cross validation (LOOCV) was implemented both for the internal feature selection ($\mathrm{LOOCV_{int}}$) and for the external evaluation of the classifier ($\mathrm{LOOCV_{ext}}$).

As summarized in Fig. 3, $\mathrm{LOOCV_{ext}}$ splits the dataset into 40 different training and testing sets (TR_i, TS_i with $1 \leq i \leq 40$); for each TR_i, a different feature selection was performed (FS_i, $1 \leq i \leq 40$). The objective function (predictive accuracy) of each feature selection was evaluated by $\mathrm{LOOCV_{int}}$.

After each FS_i, a list of optimal subsets of features was generated. There was generally more than one optimal subset with the highest $\mathrm{LOOCV_{int}}$ accuracy. Following the typical nested procedure, TS_i should be classified from the classifier built with a single subset chosen by FS_i; however, since more than one optimal subset was found, it was not possible to make a unique

TABLE IV
CLASSIFIERS PERFORMANCE. FOR EACH CLASSIFIER THE OVERALL OPTIMAL SUBSETS ARE REPORTED WITH THE RELATIVE MR. ONLY SUBSETS WHICH WERE SELECTED AS OPTIMAL MORE THAN 20 TIMES (OUT OF 40) ARE REPORTED

CLASS	OVERALL OPTIMAL SUBSETS Measure (direction, condition)			SELECTION TIMES (out of 40)	MR % [CI][a]
LDA	Power HF(ML,EODT)	Entropy(AP,EODT)	F50(AP,EOF)	26	5 [1.4-16.5]
QDA	Power HF(ML,EODT)	CF(AP,EODT)	Range(AP,ECF)	25	5 [1.4-16.5]
	Power HF(ML,EODT)	FD(AP,EOF)	Range(AP,ECF)	24	5 [1.4-16.5]
MC	Power HF(ML,EODT)	FD(AP,EOF)	Range(AP,ECF)	36	5 [1.4-16.5]
	Power HF(ML,EODT)	CF(AP,EODT)	Range(AP, ECF)	21	5 [1.4-16.5]
LR	JI(AP,EO)	F95(AP,EODT)	RHL(ML,EODT)	30	5 [1.4-16.5]
KNN	Entropy(AP,EODT)	RHL(ML,EODT)	Entropy(ML,EOF)	27	7.5 [2.6-19.9]
	Power HF(ML,EO)	Power HF(ML,EC)	Power HF(ML,EODT)	22	7.5 [2.6-19.9]
	Power HF(ML,EO)	Power HF(ML,EODT)	mSCEA(EOF)	22	7.5 [2.6-19.9]
SVM	F95(AP, EO)	Power HF(ML,EODT)	|90-Mdir|(EOF)	31	5 [1.4-16.5]

[a] 95% Confidence Intervals (CI) were computed based on [44].

choice. Moreover, different FS_i's led to different lists of optimal subsets, so we decided to extract those subsets, which were selected more frequently as optimal (overall optimal subsets, see Fig. 3). The number of times a certain subset was selected as optimal (selection times) serves as an index of how robust that subset is to changes in the training set, and therefore to selection bias. Finally, the accuracy of the classifier was computed by $\mathrm{LOOCV}_{\mathrm{ext}}$ for the overall optimal subsets (see Fig. 3) and quantified by the misclassification rate (MR), which is the proportion of incorrectly classified subjects. Subsets of different cardinality were considered separately.

Two-sample T-tests were performed to detect differences in the values of certain measures between PD and CTRL subjects. Kendall's tau correlation analysis was performed in the PD group between tremor-related measures selected by the feature selection procedure and the clinical scores related to tremor. Kendall's tau correlation analysis was also performed in the PD group between measures selected by the feature selection procedure and UPDRS motor section subscores.

MATLAB R2009b was used for signal processing, postural measure computation, the entire classification procedure, and statistical analysis.

III. RESULTS

A. Feature Selection and Classifiers Performance

In Table IV, a summary of the results of the feature selection procedure is reported, with the subsets of three measures that were selected more frequently as optimal subsets (overall optimal subsets). For each classifier, subsets of three features selected more than 20 times are shown. The number of times a specific subset was selected (SELECTION TIMES) out of the 40 different feature selection procedures is reported in Table IV, as well as the corresponding $\mathrm{LOOCV}_{\mathrm{ext}}$ final accuracy estimates MR. Subsets of three features, with an MR of at most 7.5%, were more accurate than subsets of fewer features; in fact for all classifiers, the best MR for pairs of measures was 10% and the best MR for a single measure was 17.5% (not shown).

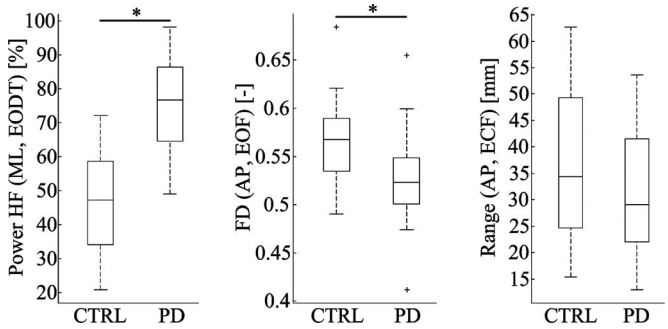

Fig. 4. Boxplots of the three measures selected for the MC with the overall highest SELECTION TIMES: power HF in ML direction during EODT condition, FD in AP direction during EOF, and range AP during ECF. Significant differences between the CTRL and PD groups are indicated with *($p < 0.05$).

It can be observed that subsets with the same MR may have different SELECTION TIMES: when this happens subsets with higher SELECTION TIMES should be preferred because they show a higher robustness with respect to feature selection bias, as discussed in Section II-C.

All classifiers performed at the same MR (5%), even on different subsets, except KNN, which has a MR of 7.5% in all its best subsets.SELECTION TIMES vary among subsets within the same classifier, and between different classifiers. The highest SELECTION TIMES value (36) is found for a subset chosen by the feature selection procedure based on MC, consisting of:

1) power of the signal for high frequencies (power HF) in ML direction during EODT condition;
2) frequency dispersion (FD) in AP direction during EOF condition;
3) range in AP direction during ECF condition.

In Fig. 4, boxplots of the values of these three measures for CTRL and PD groups are shown; a statistical difference was detected between the two groups for FD and Power HF. Although no statistical difference was found for range, its interaction with the other two measures acts to improve the classifier discriminative ability.

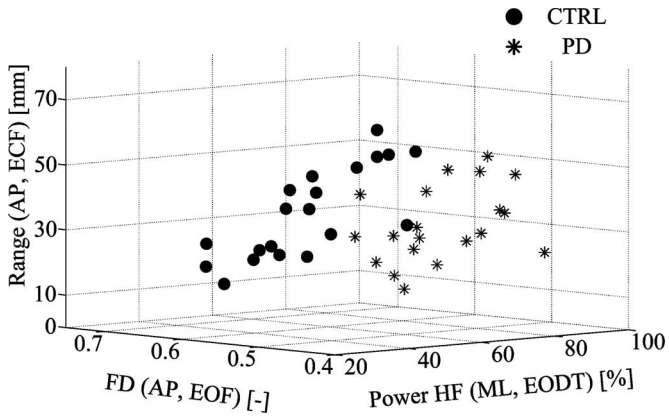

Fig. 5. 3-D view of the values of the three measures selected for the MC with the overall highest SELECTION TIMES.

In Fig. 5, the values of these three measures are shown in a 3-D view, a good separation between the two groups is graphically detectable.

It can further be noted in Table IV that for three classifiers (QDA, MC, SVM) the best (in terms of highest SELECTION TIMES) subsets consist of: 1) a tremor-related measure, computed from the signal along the ML direction during EODT condition; 2) a postural measure in the frequency domain from the AP signal in EOF , EO or EODT; and 3) a postural measure related to CoM displacement in EOF or ECF.

B. Relation Between Postural and Tremor Features and Clinical Measures

As described in Section II-C, a correlation analysis was performed in the PD group between the tremor-related measures in Table IV and the clinical scores related to tremor. A tremor score was calculated according to [45] as the average of the following UPDRS items: item 16 (tremor by history, from UPDRS section of "activities of daily living"), item 20 (tremor at rest, from motor UPDRS), and item 21 (action tremor, from motor UPDRS). A correlation ($R = 0.51$, $p = 0.003$) was found between Power HF in ML direction in EODT condition and this tremor score. A similar correlation ($R = 0.54$, $p = 0.001$) was found with ratio of the high to low frequency components (RHL) in ML direction in EODT condition.

None of the selected features showed a significant correlation with the total motor UPDRS score, with the exception of the angular deviation from AP sway, 90-Mdir, in EOF condition ($R = 0.33$, $p = 0.047$).

IV. Discussion

A. Summary

In this study, a quantitative approach for accelerometer-based analysis of posture in subjects with PD was presented. We first computed various measures from the acceleration signals during five QS conditions; these measures were analyzed in order to create a relevant metrics of postural sway in PD. Using a feature selection procedure, we found the appropriate combination of measures that could best discriminate between early-mild PD subjects and CTRL subjects. Our results revealed that feature selection procedures based on different classifiers were comparable in terms of accuracy in discriminating between PD and CTRL subjects (5% or 7.5% of MR). A nested cross validation was applied to limit the effect of possible feature selection bias on the estimated accuracy. Since it was not possible to follow the typical procedure (because several different combinations of features were selected as optimal), a value was derived (SELECTION TIMES) that can be seen as an index of the reliability of the estimation of the MR (with respect to feature selection bias). With equal MR, the subset selected with the MC was the one that had the highest SELECTION TIMES. We might then expect that it will be the most robust when applied to new subjects.

Obtaining good results with only three measures shows that several redundant and irrelevant features were in the dataset. The proposed feature selection hence is useful to: 1) optimize the experimental protocol and reduce the cost of data acquisition (see Section IV-C); 2) reduce the computational costs; 3) improve clinical understanding of the results (see Section IV-B); 4) help in data visualization (see Fig. 5); and 5) improve classification accuracy. For this last purpose, we compared different common classifiers: their similar performances suggest that in the present dataset the choice of the classifier is not critical. Therefore in similar circumstances, attention should be focused on the feature selection procedure rather than on building new or complex classifiers.

B. Selected Features

In our study, only subsets of one to three features were analyzed in the feature selection procedure. The limit of three measures was fixed *a priori* to select a set of measures easily acceptable from a clinical point of view (too much information may be misleading in clinical practice); in addition, this limit kept the exhaustive search for optimal features computationally acceptable. Choosing a maximum of three features was also a way to avoid the so-called curse of dimensionality: the difficulty for classifiers to learn effective models in spaces of high dimensionality (many features) when the number of samples is limited. High dimensionality may lead to overparameterization (the complexity is too high to identify all the coefficients of the model) or to poor performance of the classifiers. Subsets of three features were found to be the best for classification (compared to subsets of fewer features).

It is noteworthy that the best features (in subsets with the highest SELECTION TIMES) selected by half of the classifiers have common characteristics representing the same aspects of postural control. The similar subsets share: a tremor-related measure, a postural measure in the frequency domain, and a displacement-related postural measure. Our results confirm that all the different kinds of measures we proposed (tremor measures, postural measures computed on acceleration, and displacement) are essential, and can detect different characteristics of postural behavior of subjects with PD in the early stage.

The best combination of classifiers and selected measures highlighted by this study is the MC and the set composed of:

1) power HF in ML direction during dual task (EODT); 2) FD in AP direction with the subjects standing on foam rubber (EOF); 3) range in AP direction with the subjects on foam rubber, but with eyes closed (ECF). Power HF correlates with the UPDRS tremor score, and its values in PD subjects are significantly different from its values in CTRL subjects. Even when clinical signs of tremor are not severe (tremor score: $0.74 \pm 0.45/4$, range: 0–1.375), this measure can identify and objectively quantify tremor characteristics in PD subjects. One possible reason that Power HF was selected in EODT condition is that in stressful conditions (such as dual tasking) PD subjects usually show higher tremor amplitude [34]. FD values are significantly lower in PD than in CTRL subjects, being a sign of a bit more regular and less variable sway. No statistical difference was shown for the remaining measure, range, which exemplifies a feature that is not able to discriminate by itself but is important for classification because of its interaction with other measures.

From a clinical point of view, PD subjects in the present group did not show pronounced signs of postural impairments. In fact items 28 (posture) and 30 (postural stability) of motor UPDRS had average values of $(1.4 \pm 0.9/4)$ and $(0.9 \pm 0.3/4)$, respectively. Nonetheless quantitative postural measures were able to add discriminative power to tremor-related measures.

We found a significant positive correlation between the motor UPDRS and the angular deviation from AP sway (absolute value) in EOF condition: the higher the motor UPDRS, the more the sway deviated from the AP direction, toward ML. A possible explanation is that more pronounced motor impairments in PD subjects are reflected by an increased postural sway in the ML direction. This is in agreement with the results by Rocchi *et al.* [18] who found an increased ML sway for PD subjects, both ON and OFF levodopa, compared to CTRL subjects. Further supporting this interpretation, increased ML sway has also been associated with high fall risk in the elderly [46].

In summary, quantitative measures computed from accelerometers seem to be able to identify postural and tremor characteristics of PD subjects even when in an early-mild stage of the disease, these may not be evident from a clinical evaluation.

C. Testing Conditions

Another important outcome is identifying, which quiet standing conditions are capable of disclosing postural differences between PD and CTRL subjects. Interestingly, the condition with EC does not qualify, since it was not chosen by any classifier as part of the subsets with the highest SELECTION TIMES.

This study found that three conditions are enough to obtain good accuracy in PD/CTRL classification, even if patients are in the early stage of the disease. From a clinical point of view, it would, thus, be possible to shorten the postural test protocol accordingly. In addition, it is interesting to note that, for early stage subjects, when clinical signs are not so evident, just a small perturbation of the sensory channels or an attentional overload may create the conditions that differentiate between PD and CTRL subjects.

The EODT condition was shown to be particularly sensitive to the disease, since it was selected in all reported subsets. It is worth noting that, when subsets of three features from the same condition are considered, EODT provides the best performance. We found, in fact, a subset exclusively composed of EODT features that has a MR of 7.5% and SELECTION TIMES of 9 (it is not shown in Table IV since only subsets with SELECTION TIMES ≥ 20 are reported). This EODT subset, selected by the LR classifier, is made of JI AP, F95 AP, and RHL ML. Thus, a good discrimination can be achieved considering the EODT condition alone, even if this subset is less robust to changes in the training set than the ones presented in Table IV. This result further confirms the importance of dual tasks in differentiating between the postural behavior of PD and CTRL subjects [19].

D. Impact on PD

Identification of testing conditions that are specific to PD with respect to CTRL subjects may be considered important from a clinical point of view. In fact, this result may lead one to think that the method plausibly characterizes PD and, hence, may be used for evaluation, follow-up, and possibly remote monitoring of PD subjects after therapy. In addition, further research on PD may benefit from these results, since they introduce new insight on the influence of sensory inputs on PD subjects and how the central nervous system integrates them. In particular, the dual task condition is shown to be particularly discriminating (see Section IV-C).

Hence, the potential impact of this study will be in both clinical practice, since it introduces a low- cost and accessible method for data acquisition, and clinical research, since it allows easy monitoring of balance maintenance in subjects with PD. The proposed postural test protocol is potentially usable in home and clinical environments, to evaluate the follow-up of a rehabilitative procedure or to analyze the disease's progression and fluctuations. Since PD symptoms may vary throughout the day and from day to day, the timetable for administering the medication is very important: the proposed protocol could be used to define and control this timetable in order to optimize the effect of the medication on postural function. Moreover, this protocol could also be used to test the effect of physical exercise and physiotherapy on balance impairments in PD subjects. From this perspective the reliability and sensitivity of the selected features with respect to the dynamics of the disease need to be addressed next.

This accelerometer-based approach makes it easier to quantify postural impairments than the conventional protocol with force plates, which are more expensive and nonportable; however, the goal of this study was not to make a comparison between force-plates and accelerometers but to present a valid option, with some guidelines, for quantitative posture analysis in PD. The focus of this study was on the early-mild stage of the disease, where accelerometers have already proved their usefulness in detecting impaired anticipatory postural adjustments [10] and deficits in gait and turning [12].

E. Limitations and Future Developments

The lack of comparison with a gold standard is a limit of the present study, which can be overcome by future experiments. A further improvement could be gained by integrating this postural evaluation with movement information (such as gait evaluation from an instrumented timed-up-and-go test), permitting a more complete evaluation of the clinical symptoms of a PD subject.

The relatively small sample size is a limitation, reflected by the large confidence intervals in Table IV. This problem could be overcome by widening the presented dataset of PD subjects and by testing selected measures on other datasets. These approaches should make the present findings statistically more robust; in particular, to check if the selected optimal measures are reliable early biomarkers of PD, they could be tested on PD subjects in an earlier stage at their first diagnosis. It remains to be seen how specific the selected optimal measures are for PD, since other pathologies might be characterized by similar postural impairments.

On the other hand, it is worth mentioning that even if the presented subsets are optimal for classifying early-mild PD, there is no guarantee that they would be optimal for monitoring the disease (i.e., sensitive to postural changes during the progression of the disease) or for detecting changes after a medical treatment. In this context, a follow-up of the study is desirable to evaluate the performance of selected measures over time. Testing subjects at regular intervals would permit us to assess whether the selected measures are sensitive to the severity and/or the progression of the disease.

Future work could also deal with the feature selection procedure: different options for the search strategy of best subsets could be used; combinations of more than three features could be explored since it is possible that this would result in improved accuracy of the classification.

ACKNOWLEDGMENT

The authors would like to thank Luca Codeluppi and Valentina Fioravanti from the Department of Neuroscience, University of Modena and Reggio Emilia, Modena, Italy, for clinical supervision and assistance in data acquisition. They would also like to thank Kristina Mayberry for the English revision of the manuscript.

REFERENCES

[1] J. Massion, "Postural control system," *Curr. Opin. Neurobiol.*, vol. 4, no. 6, pp. 877–887, Dec. 1994.
[2] F. B. Horak, "Postural orientation and equilibrium: What do we need to know about neural control of balance to prevent falls?" *Age Ageing*, vol. 35, no. Suppl. 2, pp. ii7–ii11, Sep. 2006.
[3] A. Nardone and M. Schieppati, "Balance in Parkinson's disease under static and dynamic conditions," *Mov. Disord.*, vol. 21, no. 9, pp. 1515–1520, Sep. 2006.
[4] B. Missaoui, P. Portero, S. Bendaya, O. Hanktie, and P. Thoumie, "Posture and equilibrium in orthopedic and rheumatologic diseases," *Neurophysiol. Clin.*, vol. 38, no. 6, pp. 447–457, Dec. 2008.
[5] D. A. Winter, A. E. Patla, and J. S. Frank, "Assessment of balance control in humans," *Med. Prog. Technol.*, vol. 16, no. 1–2, pp. 31–51, May 1990.
[6] S. Corazza, E. Gambaretto, L. Mundermann, and T. Andriacchi, "Automatic generation of a subject specific model for accurate markerless motion capture and biomechanical applications," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 4, pp. 806–812, Apr. 2010.
[7] A. Cappozzo, U. Della Croce, A. Leardini, and L. Chiari, "Human movement analysis using stereophotogrammetry: Part 1. Theoretical background," *Gait Posture*, vol. 21, no. 2, pp. 186–196, Feb. 2005.
[8] T. P. Andriacchi and E. J. Alexander, "Studies of human locomotion: Past, present and future," *J. Biomech.*, vol. 33, no. 10, pp. 1217–1224, Oct. 2000.
[9] A. Paraschiv-Ionescu, E. E. Buchser, B. Rutschmann, B. Najafi, and K. Aminian, "Ambulatory system for the quantitative and qualitative analysis of gait and posture in chronic pain patients treated with spinal cord stimulation," *Gait Posture*, vol. 20, no. 2, pp. 113–125, Oct. 2004.
[10] M. Mancini, C. Zampieri, P. Carlson-Kuhta, L. Chiari, and F. B. Horak, "Anticipatory postural adjustments prior to step initiation are hypometric in untreated Parkinson's disease: An accelerometer-based approach," *Eur. J. Neurol.*, vol. 16, no. 9, pp. 1028–1034, Sep. 2009.
[11] A. Weiss, T. Herman, M. Plotnik, M. Brozgol, I. Maidan, N. Giladi, T. Gurevich, and J. M. Hausdorff, "Can an accelerometer enhance the utility of the Timed Up & Go Test when evaluating patients with Parkinson's disease?," *Med. Eng. Phys.*, vol. 32, no. 2, pp. 119–25, Nov. 2009.
[12] C. Zampieri, A. Salarian, P. Carlson-Kuhta, K. Aminian, J. G. Nutt, and F. B. Horak, "The instrumented timed up and go test: Potential outcome measure for disease modifying therapies in Parkinson's disease," *J. Neurol. Neurosurg. Psychiatry*, vol. 81, no. 2, pp. 171–176, Feb. 2010.
[13] L. Chiari, M. Dozza, A. Cappello, F. B. Horak, V. Macellari, and D. Giansanti, "Audio-biofeedback for balance improvement: An accelerometry-based system," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 12, pp. 2108–2111, Dec. 2005.
[14] R. E. Mayagoitia, J. C. Lotters, P. H. Veltink, and H. Hermens, "Standing balance evaluation using a triaxial accelerometer," *Gait Posture*, vol. 16, no. 1, pp. 55–59, Aug. 2002.
[15] D. Roetenberg, H. J. Luinge, C. T. Baten, and P. H. Veltink, "Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, no. 3, pp. 395–405, Sep. 2005.
[16] Y. L. Kuo, K. M. Culhane, P. Thomason, O. Tirosh, and R. Baker, "Measuring distance walked and step count in children with cerebral palsy: An evaluation of two portable activity monitors," *Gait Posture*, vol. 29, no. 2, pp. 304–310, Feb. 2009.
[17] J. W. Blaszczyk, R. Orawiec, D. Duda-Klodowska, and G. Opala, "Assessment of postural instability in patients with Parkinson's disease," *Exp. Brain Res.*, vol. 183, no. 1, pp. 107–114, Oct. 2007.
[18] L. Rocchi, L. Chiari, and F. B. Horak, "Effects of deep brain stimulation and levodopa on postural sway in Parkinson's disease," *J. Neurol. Neurosurg. Psychiatry*, vol. 73, no. 3, pp. 267–274, Sep. 2002.
[19] R. Marchese, M. Bove, and G. Abbruzzese, "Effect of cognitive and motor tasks on postural stability in Parkinson's disease: A posturographic study," *Mov. Disord.*, vol. 18, no. 6, pp. 652–658, Jun. 2003.
[20] L. Rocchi, L. Chiari, A. Cappello, and F. B. Horak, "Identification of distinct characteristics of postural sway in Parkinson's disease: A feature selection procedure based on principal component analysis," *Neurosci. Lett.*, vol. 394, no. 2, pp. 140–145, Feb. 2006.
[21] M. A. McVey, A. P. Stylianou, C. W. Luchies, K. E. Lyons, R. Pahwa, S. Jernigan, and J. D. Mahnken, "Early biomechanical markers of postural instability in Parkinson's disease," *Gait Posture*, vol. 30, no. 4, pp. 538–542, Nov. 2009.
[22] F. B. Horak, D. Dimitrova, and J. G. Nutt, "Direction-specific postural instability in subjects with Parkinson's disease," *Exp. Neurol.*, vol. 193, no. 2, pp. 504–521, Jun. 2005.
[23] R. J. Dunnewold, C. E. Jacobi, and J. J. V. Hilten, "Quantitative assessment of bradykinesia in patients with Parkinson's disease," *J. Neurosci. Methods*, vol. 74, no. 1, pp. 107–112, Jun. 1997.
[24] P. Bonato, "Clinical applications of wearable technology," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 1, pp. 6580–6583, Sep. 2009.
[25] D. J. Gelb, E. Oliver, and S. Gilman, "Diagnostic criteria for Parkinson disease," *Arch. Neurol.*, vol. 56, no. 1, pp. 33–39, Jan. 1999.
[26] 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, no. 3, pp. III/40–III/48, Oct. 2002.
[27] S. Patel, K. Lorincz, 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 Trans. Inf. Technol. Biomed.*, vol. 13, no. 6, pp. 864–873, Nov. 2009.
[28] N. L. Keijsers, M. W. Horstink, and S. C. Gielen, "Online monitoring of dyskinesia in patients with Parkinson's disease," *IEEE Eng. Med. Biol. Mag.*, vol. 22, no. 3, pp. 96–103, May 2003.

[29] B. R. Brewer, S. Pradhan, G. Carvell, and A. Delitto, "Feature selection for classification based on fine motor signs of Parkinson's disease," *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2009, pp. 214–217, Sep. 2009.

[30] M. Navío, J. J. Aguilera, M. J. del Jesus, R. González, F. Herrera, and C. Iribar, "Feature selection algorithms applied to Parkinson's disease," in *Proc. 2nd Int. Symp. Data Analysis (ISMDA)*, Madrid, 2001, pp. 195–200.

[31] W. E. McIlroy and B. E. Maki, "Preferred placement of the feet during quiet stance: Development of a standardized foot placement for balance testing," *Clin. Biomech. (Bristol, Avon)*, vol. 12, no. 1, pp. 66–70, Jan. 1997.

[32] N. E. Huang, Z. Shen, S. R. Long, M. L. C. Wu, H. H. Shih, Q. N. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. R. Soc. Lond. A*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[33] S. Liu, Q. He, R. X. Gao, and P. Freedson, "Empirical mode decomposition applied to tissue artifact removal from respiratory signal," in *Proc. 30th Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2008, pp. 3624–3627.

[34] K. E. Lyons and R. Pahwa, *Handbook of Essential Tremor and Other Tremor Disorders*. London: Taylor & Francis, 2005.

[35] T. E. Prieto, J. B. Myklebust, R. G. Hoffmann, E. G. Lovett, and B. M. Myklebust, "Measures of postural steadiness: Differences between healthy young and elderly adults," *IEEE Trans. Biomed. Eng.*, vol. 43, no. 9, pp. 956–966, Sep. 1996.

[36] L. Chiari, L. Rocchi, and A. Cappello, "Stabilometric parameters are affected by anthropometry and foot placement," *Clin. Biomech. (Bristol, Avon)*, vol. 17, no. 9–10, pp. 666–677, Nov. 2002.

[37] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artif. Intell.*, vol. 97, no. 1–2, pp. 273–324, Dec. 1997.

[38] G. A. F. Seber, *Multivariate Observations*, 1st ed. New York: Wiley, 1984.

[39] W. J. Krzanowski, *Principles of Multivariate Analysis a User's Perspective*. Oxford: Clarendon, 1988.

[40] D. W. Hosmer, S. Lemeshow, and E. D. Cook, *Applied Logistic Regression*, 2nd ed. New York: Wiley, 2000.

[41] T. M. Mitchell, *Machine learning*, Int. ed. New York: McGraw-Hill, 1997.

[42] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*. Cambridge: Cambridge Univ. Press, 2000.

[43] R. Simon, M. D. Radmacher, K. Dobbin, and L. M. McShane, "Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification," *J. Natl. Cancer Inst.*, vol. 95, no. 1, pp. 14–18, Jan. 2003.

[44] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed. San Francisco, CA: Morgan Kaufmann, pp. 146–149.

[45] J. Jankovic, M. McDermott, J. Carter, S. Gauthier, C. Goetz, L. Golbe, S. Huber, W. Koller, C. Olanow, I. Shoulson, M. Stern, C. Tanner, W. Weiner, and Parkinson Study Group, "Variable expression of Parkinson's disease: A base-line analysis of the DATATOP cohort. The Parkinson Study Group," *Neurology*, vol. 40, no. 10, pp. 1529–1534, Oct. 1990.

[46] B. E. Maki, P. J. Holliday, and A. K. Topper, "A prospective study of postural balance and risk of falling in an ambulatory and independent elderly population," *J. Gerontol.*, vol. 49, no. 2, pp. M72–M84, Mar. 1994.

Authors' photographs and biographies not available at the time of publication.