

Automatic UPDRS Evaluation in the Sit-to-Stand Task of Parkinsonians: Kinematic Analysis and Comparative Outlook on the Leg Agility Task

Matteo Giuberti, *Member, IEEE*, Gianluigi Ferrari, *Senior Member, IEEE*, Laura Contin, Veronica Cimolin, Corrado Azzaro, Giovanni Albani, and Alessandro Mauro

I. INTRODUCTION

ACCORDING to the global declaration for Parkinson's disease (PD), 6.3 million people suffer from PD worldwide [1]. The prevalence of PD is about 0.3% of the whole population in industrialized countries, rising up to 1% over the age of 65 and to 4% over 80. The clinical picture of PD is characterized by a progressive deterioration of the motor performance. Although the symptoms can be improved by dopaminergic

drugs, such as L-dopa, over time its effectiveness worsens and motor fluctuations may occur as well as dyskinesias and involuntary movements. Furthermore, variations in the severity of these symptoms are observed during dosing intervals.

The clinical picture assessed during an outpatient checkup in the medical office poorly represents the real (actual) clinical status, especially in fluctuating patients. Indeed, repeated daily assessments of motor symptoms would be required (as suggested by the guidelines of the Movement Disorder Society, MDS [2] and this can be done by asking the patient or someone close to him/her (a relative or home nurse personnel) to annotate the numbers of hours of OFF (i.e., when drugs are not effective) and ON (i.e., when they are effective) conditions. However, patient-related annotation is not fully reliable because of perceptual bias and, in recent years, a number of studies on automatic systems to evaluate motor fluctuations of PD patients have been developed [3]. The most common approach is leveraging sensing technology to automatically evaluate the performance of specific motor tasks, such as sit-to-stand (S2S) [4], [5], gait [6], [7], tremor [8], and leg agility (LA) [9]–[14]. The basic idea is to develop a system able to get an evaluation of the motor status of a patient as close as possible to the evaluation of neurologists when they apply semiquantitative evaluation scales, such as the Unified Parkinson's Disease Rating Scale (UPDRS) [15]. The design and implementation of on-body sensor-based automatic systems for continuous assessment of PD has, overall, several benefits, ranging from medical (automaticity, objectivity) to socioeconomical (less burden for the healthcare system, less specialized personnel would be needed).

In [12]–[14], a novel approach for quantitative evaluation of relevant kinematic features, representative of the UPDRS score of the LA task performed by a PD patient, is presented. In the current study, we apply the same approach to kinematically characterize the S2S task. We identify relevant kinematic features, in both time and frequency domains, representative of the UPDRS level of the S2S task using a single body-worn inertial sensor placed on the chest of a PD patient. An experimental analysis is carried out considering 24 PD patients, identifying the most significant kinematic features associated with the S2S task characterization, by mapping them with the UPDRS scores attributed by expert neurologists. The use of principal component analysis (PCA) is also considered in order to reduce the dimensionality of the kinematic features' space. Finally, we propose an outlook on the LA task, analyzing the correlation between UPDRS scoring in the S2S and LA tasks and comparing the

Manuscript received October 31, 2014; revised February 28, 2015; accepted April 13, 2015. Date of publication April 22, 2015; date of current version May 7, 2015. This work was supported by the Italian Ministry of Health (RF-2009-1472190).

M. Giuberti was with the CNIT Research Unit of Parma and the Department of Information Engineering, University of Parma, Parma I-43124, Italy. He is now with Xsens Technologies B.V., 7500 AN Enschede, The Netherlands (e-mail: matteo.giuberti@xsens.com).

G. Ferrari is with the CNIT Research Unit of Parma and the Department of Information Engineering, University of Parma, Parma I-43124, Italy (e-mail: gianluigi.ferrari@unipr.it).

L. Contin is with the Research & Prototyping, Telecom Italia, Turin 10148, Italy (e-mail: laura.contin@telecomitalia.it).

V. Cimolin is with the Department of Electronics, Information, and Bioengineering, Polytechnic of Milan, Milan 20133, Italy (e-mail: veronica.cimolin@biomed.polimi.it).

C. Azzaro and G. Albani are with the Division of Neurology and Neurorehabilitation, Istituto Auxologico Italiano IRCCS, Piacavallo (VB) 28824, Italy (e-mail: c.azzaro@auxologico.it; g.albani@auxologico.it).

A. Mauro is with the Division of Neurology and Neurorehabilitation, Istituto Auxologico Italiano IRCCS, Piacavallo (VB) 28824, Italy, and also with the Department of Neurosciences, University of Turin, Turin 10126, Italy (e-mail: mauro@auxologico.it).

Color versions of one or more of the figures in this paper are available online.

distributions of the probability of decision error of our automatic detection system with respect to the neurologists' decisions.

We remark that our starting design point was to map to the S2S task the automatic UPDRS classification approach proposed in [14] for the LA task. We believe that its applicability, notwithstanding a radically different body sensor network (BSN) configuration (there is only one inertial node on the chest, instead of one node per thigh) and the considered kinematic features (as will be discussed in the remainder of the paper), is a strong indicator of the validity of our BSN-based automatic classification approach.

This paper is structured as follows. In Section II, we provide the reader with preliminaries and an overview on related works. Section III describes the experimental setup, detailing the used hardware and the considered PD subjects. In Section IV, the methods used in the paper for feature characterization and automatic classification are presented. In Section V, the obtained experimental results are presented, considering, first, an exploratory feature analysis and, then, evaluating the performance of the proposed classification algorithms. Section VI is dedicated to discussions related to i) a comparative outlook of the S2S task with the LA task and ii) interesting future research directions. Finally, Section VII concludes the paper.

II. PRELIMINARIES AND RELATED WORK

In [16] and [17], it is shown that different UPDRS tasks are representative of different aspects of the PD and allow to evaluate different motor/functional abilities. Concerning the evaluation of the S2S transition, in [4] an accurate kinematic study of the S2S movement pattern is conducted. In [18], a review about the determinants that can influence the movement is presented. Different approaches used to estimate some of the spatiotemporal parameters typical of the movement are described, through devices equipped with accelerometers and gyroscopes, in [5], [19] and [20] and, through devices based on accelerometers, in [21]–[23]. Moreover, the accuracy of an inertial approach to the measurement of the kinematics of arising from a chair with accelerometers and gyroscopes is discussed in [24].

The LA task aims at evaluating the severity of motion impairments of a PD patient, with specific focus on the lower limbs [9]–[11]. In this exercise, the patient is asked to sit on a chair provided with rigid backrest and armrests. The patient must place both his/her feet on the floor in a comfortable position and, then, alternately raise up and stomping the feet on the ground, as high and as fast as possible, ten times per leg (in order to test each leg separately). In [13] and [14], a detailed investigation, through the use of a BSN of inertial nodes (one sensor per thigh), of the LA task, with the final aim of characterizing it kinematically, is presented. The observed kinematic parameters are associated with the UPDRS evaluation according to the guidelines of the MDS.

In the S2S exercise, the patient is asked to sit on a straight-backed chair with armrests. The patient must keep both his/her feet on the floor and lean against the backrest of the chair (if the patient is not too short). The exercise consists in crossing the arms across the chest (in order to avoid their use in the

TABLE I
UPDRS MAPPING FOR THE S2S TASK

UPDRS	Failed attempts & use of armrests	Slowing	Move forward on chair
0	0 failed attempts, no need to use armrests	no	no
1	≥ 1 failed attempts, no need to use armrests	yes	yes
2	0 failed attempts, need to use armrests	–	–
3	≥ 1 failed attempts, need to use armrests	–	–
4	not able to stand up alone		

movement) and raising from the chair. In case of failure, the patient can retry to raise. After a maximum of three failed trials, the patient can move forward on the chair to facilitate the movement (still keeping the arms crossed on the chest). If he/she is still not able to stand up, the patient is allowed to push off using his/her hands on the armrests. After a maximum of three unsuccessful trials with the help of the arms, the examiner can eventually help the patient to stand up.

According to the guidelines of the MDS, the S2S task must be evaluated observing the following parameters and features: number of failed attempts before succeeding, need to use of armrests, slowing, and need to move forward on chair [15], [25]. In particular, in Table I, these parameters/features are mapped with the corresponding UPDRS evaluation. Since the quantitative evaluation of most of the features typically relies on the experience of neurologists, interneurologist score variations cannot be *a priori* excluded.

III. EXPERIMENTAL SETUP

A. Experimental Testbed

The experiments were carried out at the San Giuseppe Hospital, Istituto Auxologico Italiano, in Piancavallo, Verbania, Italy, at a fully equipped last generation motion analysis laboratory. The kinematic analysis was carried out, in a comparative way, considering i) an optoelectronic system and ii) a wireless BSN-based system.

In particular, the BSN was formed by a single Shimmer node [26], [27]), which is a small (size: 53 mm \times 32 mm \times 25 mm; weight: 22 g) and low-power wireless sensing device equipped with a triaxial accelerometer, a triaxial gyroscope, and a triaxial magnetometer. The Shimmer device was placed on the patient's chest, attached to the body using a Velcro strap, as shown in Fig. 1. Specifically, for ease of the following analysis, it was placed trying to align the plane defined by the x - and y -axes of the device with the frontal plane of the user and trying to align one of these two axes with the direction of the spine. The Shimmer device was streaming data (via Bluetooth) at 102.4 Hz to a personal computer, where the signal processing analysis was performed.

In [14], for the purpose of accuracy validation, the considered BSN-based inertial system is directly compared with a



Fig. 1. Considered experimental S2S testbed.

reference optoelectronic system (Vicon, Oxford, U.K.). Specifically, the optoelectronic system has been used to estimate the three-dimensional (3-D) position of passive markers positioned on specific anatomical landmarks of the subject. Passive markers position data were collected for all body segments and the Davis marker-set was chosen as the protocol of choice to acquire the motion of lower limbs and trunk based on [28], [29]. The results in [14] highlight the accuracy of the proposed BSN-based inertial system. Therefore, in the current study we rely on these findings and assume that the measurements carried out through our inertial system are sufficiently accurate.

B. Subjects

The experimental results presented in the following refer to a group of 24 PD patients (17 males and seven females) with age ranging from 31 to 79 years (with mean equal to 65.9 years and standard deviation equal to 12.3 years). The patients have been asked to perform a S2S task, providing them instructions as described in Section II. A total of 34 S2S trials have been collected.¹ The patients' UPDRS scores, assigned by neurologists, ranged between 0 and 4. However, note that, unlike what the MDS document indicates (namely, that only discrete integer scores should be assigned), noninteger (0.5-type) scores have also been used in the case of indecision between two consecutive integer UPDRS values. In Fig. 2, we report the distribution of the 34 UPDRS scores assigned to the considered S2S trials—note that, in this experiment, not all possible values of UPDRS scores were observed in the considered patients.

IV. METHODS

A. Feature Extraction

As in [14], the 3-D orientation of the Shimmer device is estimated through an orientation estimation filter [30] and the inclination of the chest θ is then computed. Since the typical shape of θ , during the S2S task, is the one shown in Fig. 3,

¹Note that, even if only 24 patients have been considered, some patients have performed the S2S task multiple times, at different times, and/or for different PD conditions.

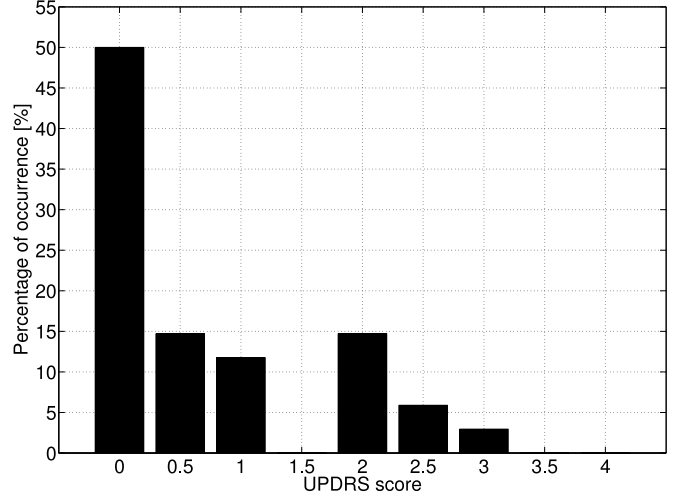


Fig. 2. Distribution of the UPDRS scores assigned to the S2S trials considered in the experimental analysis. A total of 34 UPDRS scores were given, performed by 24 PD patients.

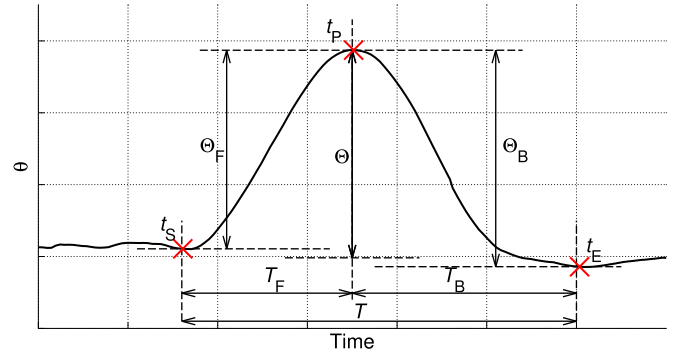


Fig. 3. Time evolution of the chest inclination θ during a typical S2S transition: the main segmentation events, namely the epochs t_S , t_P , and t_E , are shown as red crosses. Some of the kinematic features outlined in Table II are also shown.

the following relevant time labels are evaluated: i) the starting epoch t_S of the S2S (i.e., when the chest starts bending forward); ii) the epoch of maximal bending of the chest t_P (placed around the middle of the S2S exercise); and iii) the ending epoch t_E of the S2S (i.e., when the chest returns in the vertical position). Starting from the previous labels (t_S , t_P , and t_E) and the chest inclination θ , it is then possible to easily extract the 12 geometric features summarized in Table II and partly shown in Fig. 3. Note that the previous features are chosen in order to provide the most intuitive and complete (though possibly redundant) description of the S2S task from a kinematic perspective and, specifically, they contain information about the duration, the amplitude, and the speed of the considered movement.

B. Principal Component Analysis

Since the chosen features may likely be (at least partly) correlated, the following analysis has been performed by considering both the original feature space and a new one obtained by performing a PCA of (a subset of) the presented features, which

TABLE II
CONSIDERED FEATURES, PARTLY SHOWN IN FIG. 3 IN RELATION TO A TYPICAL
S2S TRANSITION, FOR THE S2S TASK

Name	Definition	Dimension
Forwards bending duration	$T_F \triangleq t_P - t_S$	[s]
Backwards bending duration	$T_B \triangleq t_E - t_P$	[s]
Total duration	$T \triangleq T_F + T_B =$ $t_E - t_S$	[s]
Forwards/backwards duration difference	$D_T \triangleq T_F - T_B$	[s]
Forwards bending amplitude	$\Theta_F \triangleq \theta(t_P) - \theta(t_S)$	[deg]
Backwards bending amplitude	$\Theta_B \triangleq \theta(t_P) - \theta(t_E)$	[deg]
Average bending amplitude	$\Theta \triangleq \frac{\Theta_F + \Theta_B}{2}$	[deg]
Forwards/backwards average bending amplitude difference	$D_\Theta \triangleq \Theta_F - \Theta_B$	[deg]
Forwards bending speed	$\Omega_F \triangleq \frac{\Theta_F}{T_F}$	[deg/s]
Backwards bending speed	$\Omega_B \triangleq \frac{\Theta_B}{T_B}$	[deg/s]
Average bending speed	$\Omega \triangleq \frac{\Theta_F + \Theta_B}{T} =$ $\Omega_F \frac{T_F}{T} + \Omega_B \frac{T_B}{T}$	[deg/s]
Forwards/backwards bending speed difference	$D_\Omega \triangleq \Omega_F - \Omega_B$	[deg/s]

aims at minimizing the redundancy and the dimensionality of the data while still preserving and, possibly, maximizing the variance in the original data.

C. Classification Algorithms

Since the aim of this study is to evaluate the feasibility of an automatic detection system able to associate a measured set of kinematic features to a specific UPDRS score, three well-known classification algorithms have been considered: nearest centroid classifier (NCC), k nearest neighbors (k NN), and support vector machine (SVM) [31]. In a few words: the NCC method classifies a new (unknown) point according to the same label of the nearest centroid (in terms of Euclidean distance); the k NN method classifies the new point according to the labels of the k nearest points (still in terms of Euclidean distance) through a majority rule; and the SVM method classifies the new point according to decision regions (associated with the UPDRS classes) that are constructed in order to maximize the separation between different classes.

D. Performance Analysis

As mentioned earlier, the three presented classifiers (namely, NCC, k NN, and SVM) have been run on both original data and “PCA-projected” data. In order not to bias the performance of the classifiers, a leave-one-out cross-validation method is considered. According to this method, each point of the original dataset is used, in turn, as the new (unknown) point and the remaining points are used to train the classifiers. The overall performance of the system is then evaluated by considering together the single observed performance results. The metric that we considered to quantify the performance of the aforementioned classifiers is the absolute UPDRS error e , which we define as follows:

$$e \triangleq |\hat{u} - u| \quad (1)$$

where u is the actual UPDRS score (assigned by neurologists), \hat{u} is the estimated one (using NCC, k NN, or SVM), and $u, \hat{u} \in \{0, 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4\}$. Note that the absolute value of the error is taken since, when evaluating the performance of the classifiers, we are mainly interested in quantifying the absolute deviation of the estimated UPDRS score with respect to the one given by neurologists.

In order to investigate the best performance achievable with the proposed system and with the considered features, an exhaustive performance analysis has been carried out,² by testing the system performance: for all possible combinations of features; for all possible values of k (when the k NN method, which will turn out to be the best, is used); and the number of considered principal components (when PCA data are used, instead of original data). In particular:

- 1) combination of the following 12 features, presented earlier, are evaluated: $T, T_B, T_F, D_T, \Theta, \Theta_B, \Theta_F, D_\Theta, \Omega, \Omega_B, \Omega_F$, and D_Ω ;
- 2) when using the k NN method, the following values of k are considered: $1, 2, \dots, 10$;
- 3) up to 12 principal components (as the number of features) are used when considering PCA data.

V. EXPERIMENTAL RESULTS

A. Explorative Feature Analysis

As anticipated in Section I, the aim of the following analysis is to devise an approach to automatically assign a UPDRS score to a specific S2S task. To this end, it is crucial to determine if there exists a relationship between the UPDRS score assigned by neurologists to a task and the values of the kinematic features introduced in Section IV-A. In [14], the presence of smooth UPDRS “trajectories,” in proper kinematic (multidimensional) feature spaces, is clearly shown in the case of the LA task: the existence of these smooth trajectories indicates that it may be feasible to define unambiguous decision regions in the feature space, based on which the UPDRS scores can be automatically and correctly estimated. We now extend this investigation to the S2S task, trying to identify possible UPDRS characteristic trajectories in proper (multidimensional) feature spaces, which would suggest that a classification approach similar to the one presented in [14] is still feasible.

For the sake of thoroughness, in Fig. 4 the distributions of the values of all the considered features, with reference to the UPDRS scores of the corresponding S2S tasks, are shown using a “boxplot” representation. In particular, given a feature and a UPDRS score, the box shows where the 50% of the feature values around the median (namely, the red horizontal segment) lie and, therefore, its vertical width gives a direct visual indication of the value of the interquartile range (which quantifies the data dispersion). Furthermore, the black “whiskers” and red crosses outside the box give information about the data skewness and highlight the presence of outliers (from a statistical

²Note, in particular, that an exhaustive search was feasible due to the relatively small parameter space that we evaluated. That allowed us to derive an exact performance analysis, rather than an approximated one that would have derived, for instance, by the use of heuristic methods for feature selection.

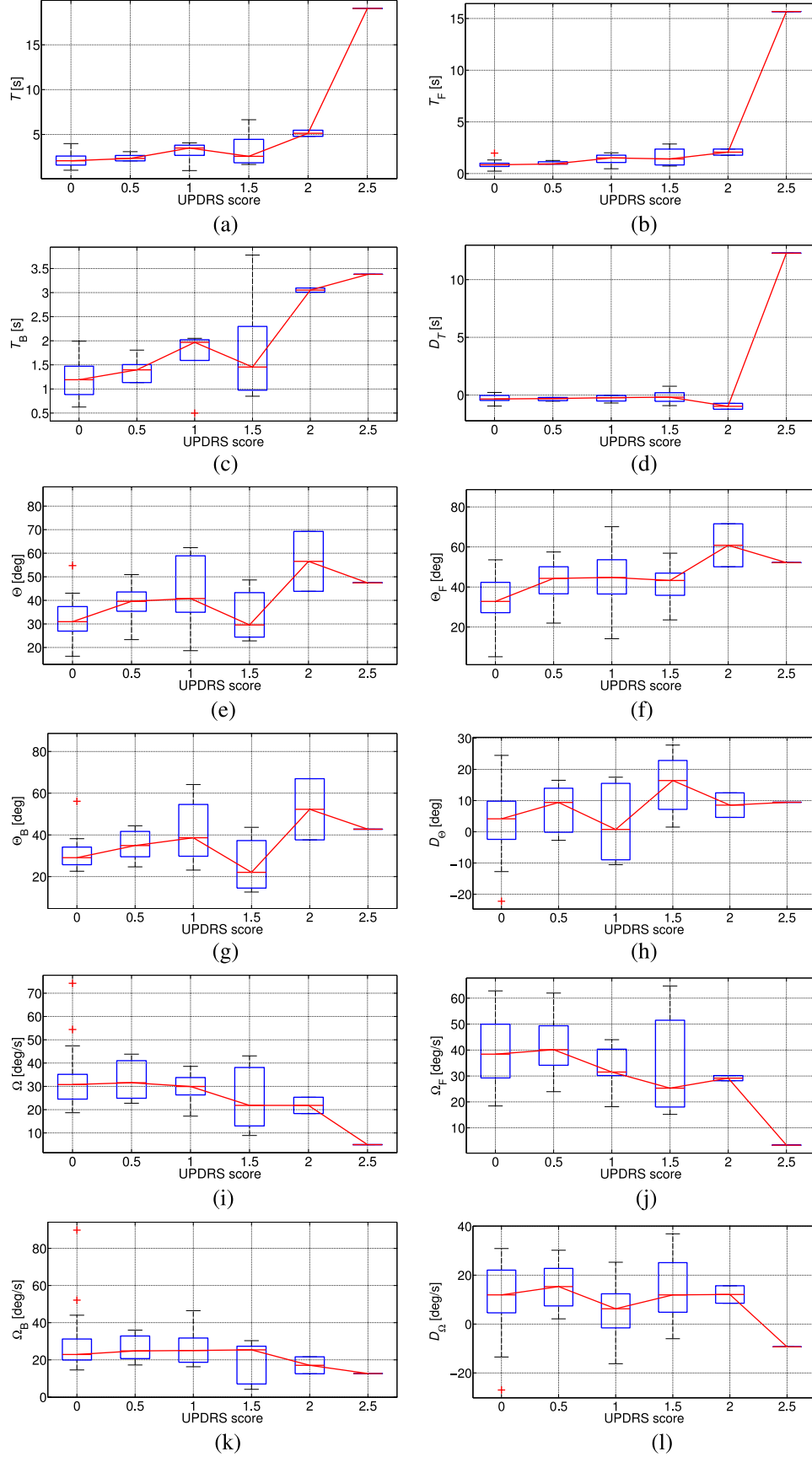


Fig. 4. “Boxplot” diagrams for all the considered kinematic features (defined as in Table II): (a) T , (b) T_F , (c) T_B , (d) D_T , (e) Θ , (f) Θ_F , (g) Θ_B , (h) D_Θ , (i) Ω , (j) Ω_F , (k) Ω_B , and (l) D_Ω . Feature values are treated separately according to the corresponding UPDRS score. All 24 patients are considered.

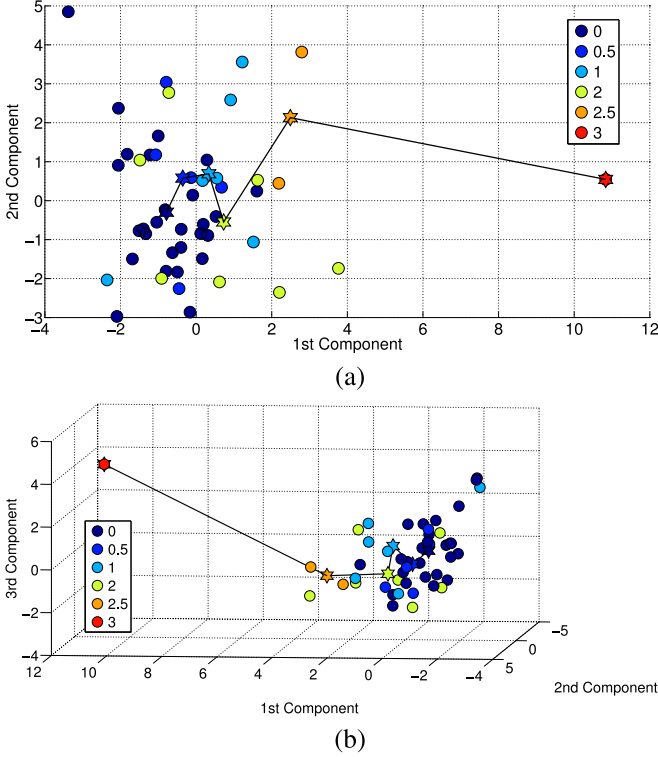


Fig. 5. Representation of original 12-tuples of features (obtained from the 34 trials) projected onto a (a) 2-D plane and a (b) 3-D space using PCA. Points are colored according to the corresponding UPDRS scores. Centroids of each cluster of points (drawn as filled stars), corresponding to the same UPDRS score, are shown and are linked in UPDRS-wise order from 0 to 3.

point of view). In particular, whiskers extend to the last values which is distant less than 1.5 times the interquartile range from the box. Beyond this limits, the values are considered as outliers and depicted as red crosses. Finally, the red line links all the median values. In order to investigate the existence of monotonic behaviors of the considered features, as functions of the UPDRS value, the medians (for the various UPDRS values) are connected with solid (red) lines. It is apparent that only in some cases [e.g., in cases (a) and (i)] monotonic trends can be partially observed.

Motivated by the fact that the results in Fig. 4 do not highlight any evident monotonic behavior of the considered features, as anticipated in Section IV-B, we apply a PCA to the original data (namely, the values of the original 12 features) in order to exploit correlation (if any) between different features [31]. In particular, in Fig. 5, the behaviors of the first two principal components [see Fig. 5(a)] and of the first three principal components [see Fig. 5(b)], respectively, are investigated. In order to visually highlight, in the reduced dimensionality spaces, the presence of parametric (in terms of UPDRS score) “trajectories,” the centroids (drawn as filled stars) of each cluster of points associated with the same UPDRS score are linked in UPDRS-wise order from 0 to 3. Although it can be observed that, moving from UPDRS 0 to UPDRS 3, the corresponding centroids are positioned in well-separated portions of the reduced dimensionality space, the centroids relative to intermediate UPDRS values are

not uniformly distributed along the trajectories. In particular, clusters of points, corresponding to different consecutive UPDRS scores, are not disjoint. This fact suggests that the use of simple centroid-based classification algorithms will likely fail, as we will see in the case of the NCC in the following analysis.

B. Automatic UPDRS Evaluation

In Section V-A, we investigated the possibility of intuitively defining decision regions in the feature space, with which it would have been possible to automatically detect the UPDRS scores of specific S2S tasks. Even though it was not possible to derive insightful considerations by looking at the values of each feature separately and by applying a PCA on the complete set of features, it is well known that machine learning algorithms are particularly effective at automatically identifying significant data patterns, especially when the data are characterized by a high dimensionality. Therefore, as explained in Section IV-D, an exhaustive analysis has been performed where the performances of the classifiers have been evaluated for all possible combinations of the 12 features presented in Section IV-A. The classifiers performances have been studied for all the selected subsets of features as well as for the same subsets projected into new feature spaces using the PCA. Specifically, in Fig. 6, a direct (exhaustive) comparison of the cumulative distribution functions (CDFs) of the error e defined in (1) with NCC, k NN, and SVM (using the PCA and not using it) is carried out. In particular, in Fig. 6(a), the CDFs for all possible number and combinations of features are shown. Different colors have been used to highlight when different classifiers are used and when PCA is applied to the features. Furthermore, in order to provide a more concise idea of the classifiers performance, in Fig. 6(b), the average CDFs (averaged over all possible combinations of features) for each classifier are also shown. Due to the definition of CDF, the depicted curves are monotonically increasing and the overall performance associated with a specific curve may be mainly evaluated in two ways: i) by looking at its value in $e = 0$, which represents the *accuracy* of the classifier (i.e., how often the classifier precisely estimates the UPDRS score), and/or ii) by computing the area under the curve (AuC), which ideally needs to be as large as possible and which provides a more general indication of the classifier performance. In particular, in the following, we will use the AuC to determine the best combination of features and the best classifier.

Looking at Fig. 6(a), it can be immediately noticed that the choice of the classifier has a more pronounced impact on the system performance rather than the choice of the features (whose proper tuning is, however, crucial to the achievement of the best performance, when considering a specific classifier). Indeed, the groups of CDFs for each classifier tend to lie in the same portion of the plane even for different combinations of features and even if using or not the PCA. This is particularly evident, for instance, in the case of NCC, which performs considerably worse than k NN and SVM. Note that similar considerations have been also made in [14], for the case of the LA task.

The system configuration that achieves the best performance, chosen as the one that maximize the AuC of the CDFs, turns

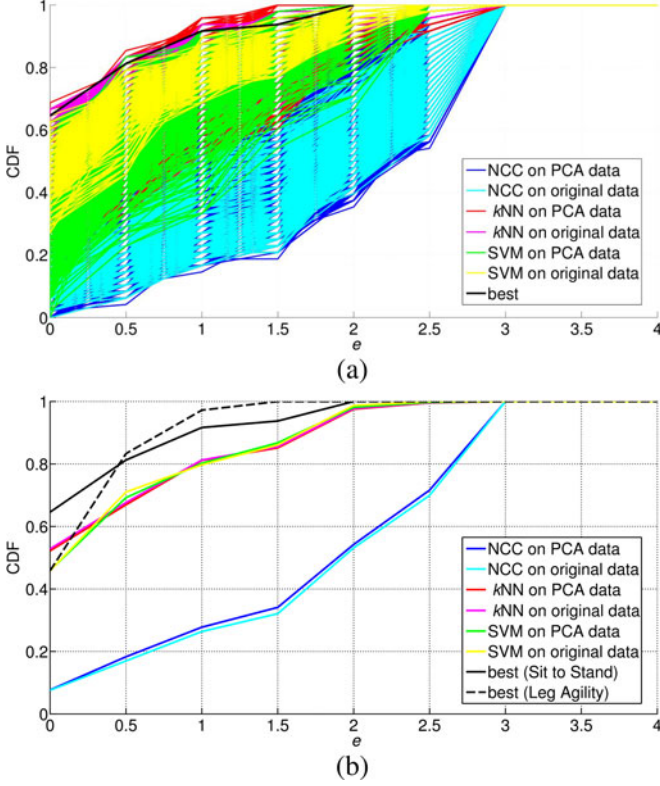


Fig. 6. CDFs of the absolute UPDRS error e for the S2S task using NCC, k NN, and SVM. In (a), the CDFs for all possible combinations of parameters and features are shown, whereas, in (b), the average CDFs for every classifier are shown. The black solid line is the CDF of the best case (i.e., k NN with $k = 4$ applied to the first principal component of PCA-projected data, where PCA is only applied to the following subset of features: T , T_F , T_B , D_T , Θ , and Ω). For comparison purposes, in (b) we show the best average CDF for the LA task, obtained in [14] (black dashed line).

out to be the one that uses k NN with $k = 4$ applied to the first principal component of PCA-projected data, with PCA applied to the following subset of features: T , T_F , T_B , D_T , Θ , and Ω . In Fig. 6(b), the black solid line represents the CDF associated with this optimized system configuration. Note that, for comparison purposes, the CDF of the best case for the LA task, obtained in [14], is also shown: this CDF refers to the use of the k NN with $k = 3$ and $(\Theta, R, P_{X_\Theta})$ as features (as defined in [14]). In Section VI, further considerations will be made on the system performance by examining, in a more comparative way, the results obtained for the S2S and the LA tasks.

Finally, since PCA is used on the six features selected in the optimized configuration (namely, T , T_F , T_B , D_T , Θ , and Ω), in Fig. 7 more details about the actual weight given to them in the transformation process are provided. In particular, in Fig. 7(a), the PCA loadings of each feature (namely, the coefficients given to each feature in the linear combination performed by the PCA) are shown for all considered trials. Furthermore, in Fig. 7(b), the average PCA loadings over all trials are also shown. Since the features, before applying the PCA, have been standardized (namely, they have been centered by their mean and scaled by their variance), larger values of the loadings correspond to a larger importance of the features in the linear combination

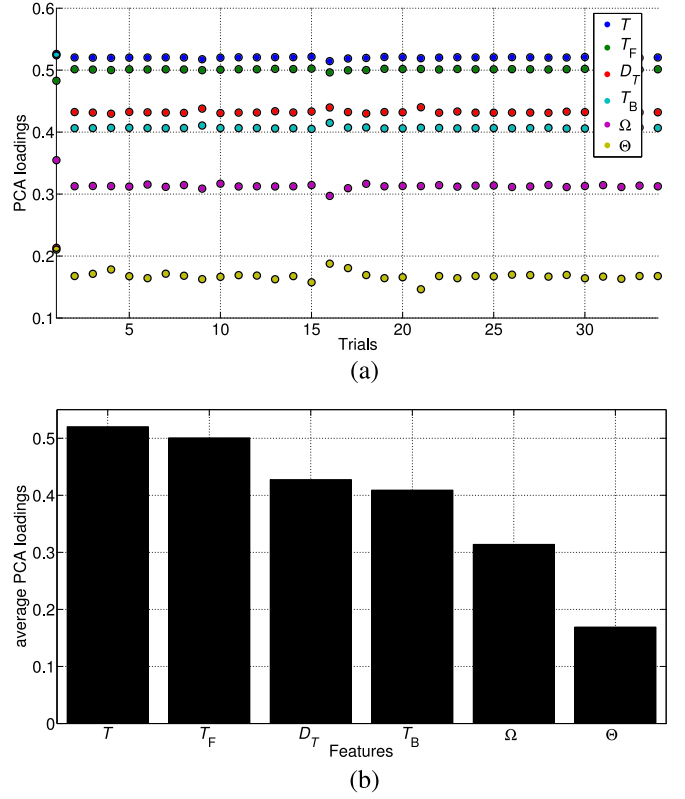


Fig. 7. PCA loadings of the features considered in the optimized system configuration (namely, T , T_F , T_B , D_T , Θ , and Ω). In (a), the PCA loadings are shown for all trials. In (b), their averages over all trials are shown. Features are sorted in descending order of importance (according to their weights in the PCA linear combination process).

process. It is easy to observe that the largest weight in the PCA is given to the time-related features (namely, T , T_F , T_B , D_T), whose aggregate weight represents almost 80% of the total. Note also, by looking at Fig. 7(a) that the variance of the feature loadings over all trials is very small. Therefore, the previous consideration is almost always (and not just on average) valid.

VI. DISCUSSION

A. Comparative Outlook of S2S Task on the LA Task

As anticipated in Section I, the main motivation of this study, as well as [14], is to investigate the feasibility of a unique portable system able to automatically detect UPDRS scores of functional tasks performed by PD patients, possibly leveraging on a common algorithmic approach. The integrated evaluation of multiple functional tasks (such as the LA and S2S tasks) is even more significant when such tasks are not correlated and provide diverse information of the clinical status of the PD patients. In order to investigate the (un)correlation of the LA task and the S2S task, the first interesting consideration can be made by looking at the distribution of the UPDRS scores assigned by neurologist to the considered 24 PD patients. To this end, for comparison purposes, in Fig. 8 we show the distribution of the UPDRS scores for the LA task (presented in [14]), where the same set of 24 PD patients is considered. By comparing

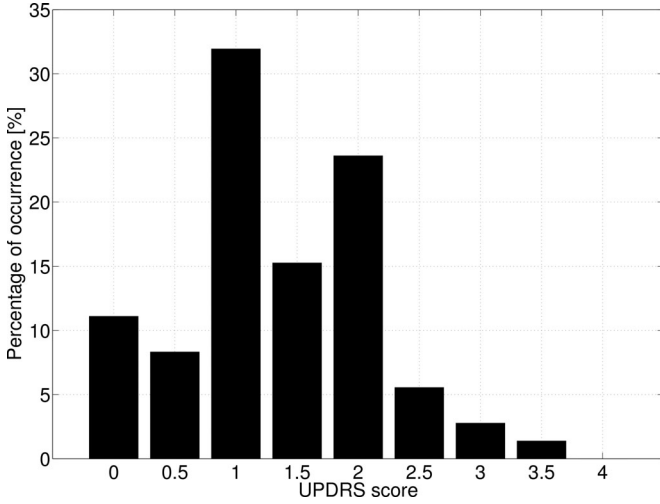


Fig. 8. Distribution of the UPDRS scores assigned to the LA trials (76 UPDRS scores were given), already presented in [13] and [14]. The same set of 24 PD patients is considered.

Fig. 8 with Fig. 2, it can readily be observed that the two tasks, carried out by the same group of patients, have led to different UPDRS evaluations. In particular, the S2S task appears to be less “challenging,” for the considered PD patients, than the LA task: in fact, in the S2S task the majority of the patients have been given UPDRS scores equal to 0.

A further evidence of the loose correlation between the two tasks can be observed in Fig. 9, where each PD patient is associated with the pair of UPDRS scores he/she has been given in the two tasks. In particular, we distinguish between: 1) UPDRS scores assigned by our automatic detection system (for S2S, using the classification approach outlined in Section V-B and, for the LA task, using the similar classification approach proposed in [14]) and 2) UPDRS scores assigned by neurologists. As expected, a generally increasing trend can be observed, i.e., for increasing values of the UPDRS score in S2S task, the possible UPDRS scores in the LA task tend to increase. However, this phenomenon is not very pronounced and, for each UPDRS value assigned in a task, several UPDRS values are possible for the other task—this is slightly more pronounced in case 2, i.e., with UPDRS pairs assigned by neurologists. In fact, this apparently noncorrelated behavior of UPDRS scoring is rooted in the very definition, in the medical viewpoint, of the various UPDRS tasks, which aim at characterizing different aspects of the PD. In [16], a factor analysis of the various UPDRS tasks is presented and it is shown that different tasks belong to different characterizing subclasses of the PD: for instance, the S2S task belongs to the subclass “posture,” whereas the LA task belong to the “rigidity” subclass. Furthermore, in [17], following a physiatric perspective, it is shown that the S2S task allows to measure functional abilities, whereas the LA task is representative of the motor activity level. A comprehensive investigation of the correlation between scoring in all UPDRS tasks is an interesting research direction.

After highlighting the importance of an automatic system able to automatically detect UPDRS scores of multiple functional

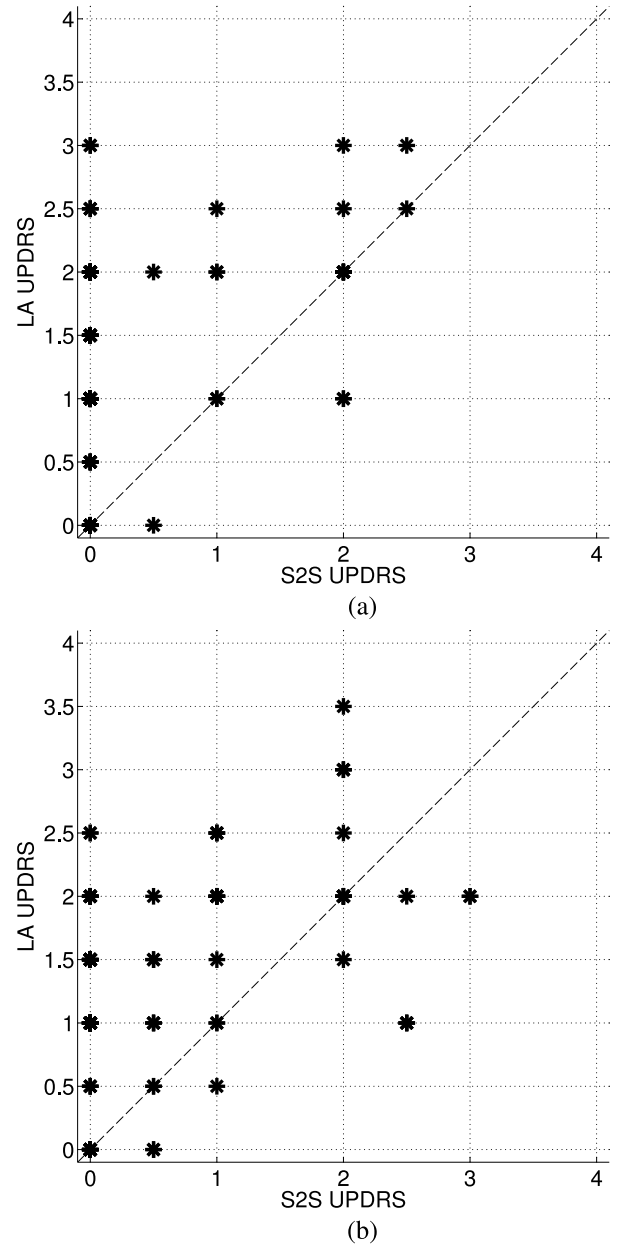


Fig. 9. Per-patient UPDRS value assigned for the S2S task as a function of the UPDRS value assigned to the same patient for the LA task: (a) through the proposed automatic detection system and (b) on the basis of the neurologists’ scoring.

tasks, we proceed by further comparatively characterizing the performance achieved for the LA task and the S2S task. In order to directly compare the kinematic characterization of the S2S task with the kinematic characterization of the LA task, investigated in [14], the overall BSN on each PD patient is composed by three Shimmer nodes, i.e., one per thigh and one on the chest. For ease of clarity, we show the typical placement of Shimmer nodes on a patient in Fig. 10. Note that, concerning the evaluation of the UPDRS in the LA task, placing the nodes on the thighs and, thus, implicitly focusing on the thighs’ inclination, rather than directly focusing on the heel elevation (as the definition of the task, presented in Section II, would suggest)



Fig. 10. Considered experimental testbed for a direct comparison between LA and S2S tasks.

has been motivated in [14], by verifying the strong correlation between the heel elevation and the thigh inclination. In particular, the heel elevation was measured through an optoelectronic system (by positioning a marker on the heel), whereas the inclination was measured with the Shimmer node positioned on the thigh. The obtained results showed indeed that the correlation between the two signals is over 0.98.

On the basis of thigh inclination/heel elevation equivalence, various relevant kinematic features for the LA task (associated with the inertial nodes on the thighs) are identified in the time and frequency domains. Eventually, the classifier that guarantees the best performance for the LA task is the k NN (as for the S2S task) and the best features are the following:

- 1) in the time domain: the arithmetic average of the angular amplitudes of ascent and descent, concisely denoted as angular amplitude Θ (dimension: [deg]), and the regularity of execution R (dimension: [s]), which intuitively quantifies the regularity among the ten repetitions per leg foreseen in the LA task;
- 2) in the frequency domain: the power of the spectrum of inclination of the thighs (denoted as P_{X_θ}).

Note that in [14], as in Section V-B, the best performance is assumed to correspond to the system configuration that maximizes the AuC of the CDF of the error e given by (1). For comparison purposes, the CDF corresponding to this optimized configuration is also shown in Fig. 6(b).

In Fig. 11, we further investigate the error e 1) in the S2S task (using the best configuration of the proposed classification method identified at the end of Section V-B) and 2) in the LA task (using the optimized classification method proposed in [14]). It can be observed that the probability mass functions have different behaviors in the two tasks. In particular: in the S2S case, most of the mass concentrates in 0 (i.e., the automatic classification system makes no error), but there are some cases where the error is 2; in the LA task, the masses in 0 and 0.5 are

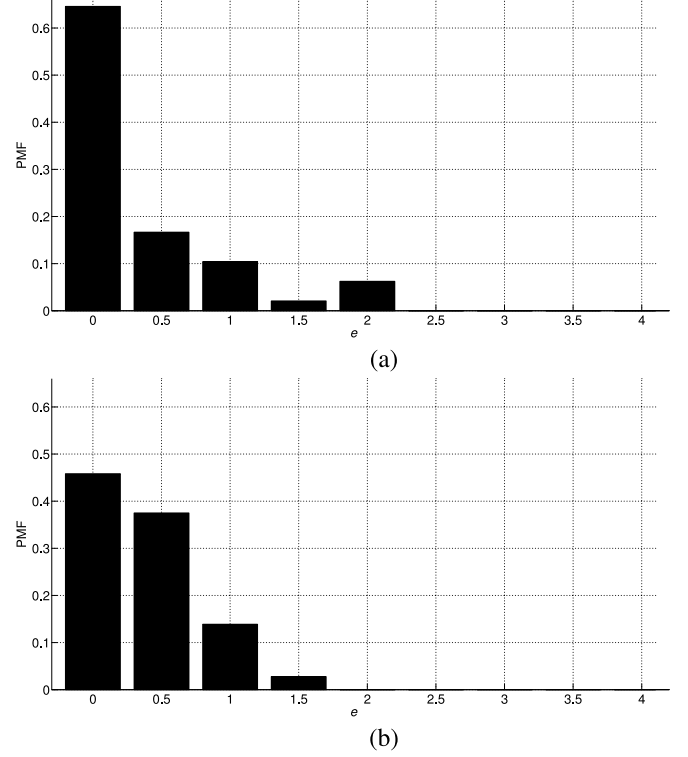


Fig. 11. Distribution of the error e [defined in (1)] between the proposed automatic detection systems and the decisions of neurologists: (a) S2S task (using the best classification method proposed at the end of Section V-B) and (b) LA task (using the best classification method proposed in [14]).

TABLE III
CONFUSION MATRIX FOR THE S2S TASK USING THE CLASSIFICATION METHOD DEVELOPED IN SECTION V-B

[%]	0	0.5	1	1.5	2	2.5	3	3.5	4
0	96.3	0	3.7	0	0	0	0	0	0
0.5	100	0	0	0	0	0	0	0	0
1	50	0	50	0	0	0	0	0	0
1.5	—	—	—	—	—	—	—	—	—
2	42.9	14.3	0	0	28.6	14.3	0	0	0
2.5	0	0	0	0	100	0	0	0	0
3	0	0	0	0	100	0	0	0	0
3.5	—	—	—	—	—	—	—	—	—
4	—	—	—	—	—	—	—	—	—

dominant and e never exceeds 1.5. This behavior is consistent with the UPDRS distributions based on neurologists' evaluation, shown in Figs. 2 and 8, where it can be observed that: in the S2S task, most of the patients have a UPDRS equal to 0; whereas, in the LA task, the UPDRS distribution is less concentrated. In order to better characterize the distributions of the error e , in Tables III and IV, the confusion matrices associated with the S2S and LA tasks, respectively, are also shown. It can be observed that the confusion matrix in the S2S case is "less sparse" than in the LA case. This is consistent with UPDRS scores distributions of the available PD patients. Obviously, a larger set of patients, with a uniform distribution of the UPDRS values, would allow to obtain more relevant (from a statistical

TABLE IV
CONFUSION MATRIX FOR THE LA TASK, USING THE BEST CLASSIFICATION
METHOD PROPOSED IN [14]

[%]	0	0.5	1	1.5	2	2.5	3	3.5	4
0	75	0	25	0	0	0	0	0	0
0.5	16.7	0	66.7	0	16.7	0	0	0	0
1	4.3	8.7	65.2	13	8.7	0	0	0	0
1.5	9.1	0	18.2	18.2	45.5	9.1	0	0	0
2	0	0	23.5	5.9	58.8	11.8	0	0	0
2.5	0	0	0	0	75	0	25	0	0
3	0	0	0	0	0	100	0	0	0
3.5	0	0	0	0	0	0	100	0	0
4	—	—	—	—	—	—	—	—	—

point of view) perspective. However, this goes beyond the scope of the paper, which focuses on proposing a novel approach, rather than an exhaustive medical investigation.

B. Research Extensions

While the proposed classification approach is designed to output “discrete” UPDRS scores, since aimed at replicating the neurologists’ decisions, a challenging research extension consists in extending the proposed approach to develop an automatic UPDRS detection system able to output continuous values. In other words, rather than providing a UPDRS score in a discrete set of values, an automatic detection system could return a real UPDRS value between 0 and 4 and quantify all the subtle differences that could be observed in the PD symptoms. In order to reach this goal, the key point is to understand how to rely on a set of discretized ground truth decisions (by the neurologists) and generate a continuous output. Considering Fig. 5, one could think of parametrizing, in a continuous way, the obtained trajectories identifying “virtual centroids” corresponding to any possible value of UPDRS between 0 and 4. At this point, for each observed patient, one could simply “project” the multidimensional point corresponding to the patient onto the trajectory and select the corresponding continuous UPDRS score. However, as already observed in Section V-A, clusters of points associated with the same UPDRS values are not disjoint and, thus, the use of this projection-based approach needs to be carefully investigated. Another very interesting approach consists of the use of regression techniques, which are particularly effective at dealing with continuous outputs [32].

Finally, the proposed automatic UPDRS classification system for the S2S task relies on the use of a personal computer to process the inertial data received from the body-worn inertial node on the chest—from the two nodes on the thighs for the LA task. Owing to the ever increasing processing capabilities of embedded systems, it is very likely that the proposed classification algorithm could be run directly inside wearable devices. As of now, the use of a smartphone could be a viable option. Although the best classification algorithm turns out to be the k NN, from the results in Fig. 6(b) it can be concluded that the SVM algorithm has a performance comparable to (actually, slightly worse than) that of the k NN algorithm, but with a lower computational complexity. In fact, while the SVM al-

gorithm builds a compact classification model on the basis of the training data, the k NN algorithm, as most of the so-called “lazy learning” algorithms, relies on the entire training dataset for each classification act. Therefore, the SVM algorithm seems an attractive processing strategy to make the implementation of the proposed automatic detection algorithm on board of future wearable devices feasible. This represents an interesting experimental research extension.

VII. CONCLUSION

In this paper, we have investigated how kinematic variables, collected through a single inertial node placed on the chest, can be representative of the UPDRS value assigned by neurologists for the S2S task. The experimental investigation has been carried out considering 24 PD patients. Various kinematic features of the chest inclination signal, in the time domain, have been investigated, relying also on the use of PCA to investigate the presence of correlation between the considered kinematic variables. After an exploratory investigation of the extracted kinematic features, the performance of automatic UPDRS evaluation systems, considering various classification methods (NCC, k NN, and SVM), has been carried out. The best system configuration, chosen as the one that maximize the AuC of the CDFs of the classification error, turns out to be that which uses k NN with $k = 4$ applied to the first principal component of PCA-projected data, with PCA applied to the following subset of features: T , T_F , T_B , D_T , Θ , and Ω . A comparative outlook of the S2S task on the LA task has then been carried out, highlighting the rather uncorrelated relation between the two tasks and the impact of the UPDRS distributions of the patients. The obtained results make the design and implementation of an automatic UPDRS detection system, based on a simple BSN of inertial sensors applied to various tasks, feasible. Our future research activities include the application of the proposed approach to the gait analysis task.

REFERENCES

- [1] European Parkinson’s Disease Association [Online]. Available: <http://www.epda.eu.com/en/parkinsons/in-depth/parkinsonsdisease/>
- [2] Movement Disorder Society [Online]. Available: <http://www.movementdisorders.org>
- [3] B. Chen, S. Patel, T. Buckley, R. Rednic, D. McClure, L. Shih, D. Tarsy, M. Welsh, and P. Bonato, “A web-based system for home monitoring of patients with Parkinsons disease using wearable sensors,” *IEEE Trans. Biomed. Eng.*, vol. 58, no. 3, pp. 831–836, Mar. 2011.
- [4] S. Nuzik, R. Lamb, A. Vansant, and S. Hirt, “Sit-to-stand movement pattern. A kinematic study,” *Phys. Therapy*, vol. 66, no. 11, pp. 1708–1713, 1986.
- [5] D. Giansanti, G. Maccioni, F. Benvenuti, and V. Macellari, “Inertial measurement units furnish accurate trunk trajectory reconstruction of the sit-to-stand manoeuvre in healthy subjects,” *Med. Biological Eng. Comput.*, vol. 45, no. 10, pp. 969–976, Oct. 2007.
- [6] A. Salarian, H. Russmann, F. J. G. Vingerhoets, C. Dehollain, Y. Blanc, P. R. Burkhard, and K. Aminian, “Gait assessment in Parkinson’s Disease: Toward an ambulatory system for long-term monitoring,” *IEEE Trans. Biomed. Eng.*, vol. 51, no. 8, pp. 1434–1443, Aug. 2004.
- [7] A. Salarian, C. Zampieri, F. B. Horak, P. Carlson-Kuhta, J. G. Nutt, and K. Aminian, “Analyzing 180° turns using an inertial system reveals early signs of progression of Parkinsons Disease,” in *Proc. IEEE 31st Annu. Int. Conf. Eng. Med. Biol. Soc.*, Minneapolis, MN, USA, Sep. 2009, pp. 224–227.

- [8] A. Salarian, H. Russmann, C. Wider, P. R. Burkhard, F. J. G. Vingerhoets, and K. Aminian, "Quantification of tremor and bradykinesia in Parkinson's Disease using a novel ambulatory monitoring system," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 2, pp. 313–322, Feb. 2007.
- [9] 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.
- [10] S. Das, L. Trutoiu, A. Murai, D. Alcindor, M. Oh, F. D. la Torre, and J. Hodgins, "Quantitative measurement of motor symptoms in Parkinson's disease: A study with full-body motion capture data," in *Proc. Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2011, pp. 6789–6792.
- [11] D. A. Heldman, D. E. Filipkowski, D. E. Riley, C. M. Whitney, B. L. Walter, S. A. Gunzler, J. P. Giuffrida, and T. O. Mera, "Automated motion sensor quantification of gait and lower extremity bradykinesia," in *Proc. IEEE 34th Annu. Int. Conf. Eng. Med. Biol. Soc.*, San Diego, CA, USA, Aug. 2012, pp. 1956–1959.
- [12] M. Giuberti, G. Ferrari, L. Contin, V. Cimolin, N. Cau, M. Galli, C. Azzaro, G. Albani, and A. Mauro, "On the characterization of leg agility in patients with Parkinson's Disease," in *Proc. 10th Int. Conf. Wearable Implantable Body Sensor Netw.*, Cambridge, MA, USA, May 2013, pp. 1–6.
- [13] M. Giuberti, G. Ferrari, L. Contin, V. Cimolin, C. Azzaro, G. Albani, and A. Mauro, "Linking UPDRS scores and kinematic variables in the leg agility task of Parkinsonians," in *Proc. 11th Int. Conf. Wearable Implantable Body Sensor Netw.*, Zurich, Switzerland, Jun. 2014, pp. 115–120.
- [14] M. Giuberti, G. Ferrari, L. Contin, V. Cimolin, C. Azzaro, G. Albani, and A. Mauro, "Assigning UPDRS scores in the leg agility task of Parkinsonians: Can it be done through BSN-based kinematic variables?" *IEEE Internet Things J.*, vol. 2, no. 1, pp. 41–51, Feb. 2015.
- [15] S. Fahn and R. L. Elton, "Unified Parkinson's disease rating scale," in *Recent Developments in Parkinson's Disease*, vol. 2. Florham Park, NJ, USA: Macmillan Health Care Information, 1987, pp. 153–163.
- [16] S. D. Vassar, Y. M. Bordelon, R. D. Hays, N. Diaz, R. Rausch, C. Mao, and B. G. Vickrey, "Confirmatory factor analysis of the motor unified Parkinson's disease rating scale," *Parkinson's Disease*, vol. 2012, p. 10, 2012.
- [17] K. J. Brusse, S. Zimdars, K. R. Zalewski, and T. M. Steffen, "Testing functional performance in people with Parkinson disease," *Phys. Therapy*, vol. 85, p. 134, Feb. 2015.
- [18] W. G. M. Janssen, J. B. J. Bussmann, H. L. D. Horemans, and H. J. Stam, "Determinants of the sit-to-stand movement: A review," *Phys. Therapy*, vol. 82, no. 9, pp. 866–879, Sep. 2002.
- [19] R. Williamson and B. J. Andrews, "Detecting absolute human knee angle and angular velocity using accelerometers and rate gyroscopes," *Med. Biological Eng. Comput.*, vol. 39, no. 3, pp. 1–9, May 2001.
- [20] D. Giansanti and G. Maccioni, "Physiological motion monitoring: wearable device and adaptive algorithm for sit-to-stand timing detection," *Physiol. Meas.*, vol. 27, no. 8, pp. 713–723, Aug. 2006.
- [21] W. G. M. Janssen, J. B. J. Bussmann, H. L. D. Horemans, and H. J. Stam, "Analysis and decomposition of accelerometric signals of trunk and thigh obtained during the sit-to-stand movement," *Med. Biol. Eng. Comput.*, vol. 43, no. 2, pp. 265–272, Mar. 2005.
- [22] W. G. M. Janssen, J. B. J. Bussmann, H. L. D. Horemans, and H. J. Stam, "Validity of accelerometry in assessing the duration of the sit-to-stand movement," *Med. Biol. Eng. Comput.*, vol. 46, no. 9, pp. 879–887, Sep. 2008.
- [23] W. G. M. Janssen, D. G. Külcü, H. L. D. Horemans, H. J. Stam, and J. B. J. Bussmann, "Sensitivity of accelerometry to assess balance control during sit-to-stand movement," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 16, no. 5, pp. 479–484, Oct. 2008.
- [24] M. C. Boonstra, R. M. A. van der Slikke, N. L. W. Keijsers, R. C. van Lummel, M. C. de W. Malefijt, and N. Verdonschot, "The accuracy of measuring the kinematics of rising from a chair with accelerometers and gyroscopes," *J. Biomechanics*, vol. 39, no. 2, pp. 354–358, 2006.
- [25] C. G. Goetz, S. Fahn, P. Martinez-Martin, W. Poewe, C. Sampaio, G. T. Stebbins, M. B. Stern, B. C. Tilley, R. Dodel, B. Dubois, R. Holloway, J. Jankovic, J. Kulisevsky, A. E. Lang, A. Lees, S. Leurgans, P. A. LeWitt, D. Nyenhuis, C. W. Olanow, O. Rascol, A. Schrag, J. A. Teresi, J. J. V. Hilten, and N. LaPelle, "Movement disorder society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Process, format, and clinimetric testing plan," *Movement Disorders*, vol. 22, no. 1, pp. 41–47, Jan. 2007.
- [26] Shimmer [Online]. Available: <http://www.shimmersensing.com>
- [27] A. Burns, B. R. Greene, M. J. McGrath, T. J. O'Shea, B. Kuris, S. M. Ayer, F. Stroiescu, and V. Cionca, "SHIMMER—a wireless sensor platform for noninvasive biomedical research," *IEEE Sensors J.*, vol. 10, no. 9, pp. 1527–1534, Sep. 2010.
- [28] R. B. Davis, S. Ounpuu, D. J. Tyburski, and J. R. Gage, "A gait analysis data collection and reduction technique," *Human Movement Sci.*, vol. 10, pp. 575–587, 1991.
- [29] A. Ferrari, M. G. Benedetti, E. Pavan, C. Frigo, D. Bettinelli, P. Rabuffetti, P. Crenna, and A. Leardini, "Quantitative comparison of five current protocols in gait analysis," *Gait Posture*, vol. 28, pp. 207–216, 2008.
- [30] S. O. H. Madgwick, A. J. L. Harrison, and R. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," in *Proc. IEEE Int. Conf. Rehabil. Robot.*, Zurich, Switzerland, Jun. 2011, pp. 1–7.
- [31] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification and Scene Analysis*. 2nd ed. New York, NY, USA: Wiley-Interscience, 2000.
- [32] D. A. Freedman, *Statistical Models: Theory and Practice*. Cambridge, UK: Cambridge Univ. Press, 2005.