A Probabilistic Approach Based on Random Forests to Estimating Similarity of Human Motion in the Context of Parkinson's Disease

Andreas Kuhner[1] Tobias Schubert[1] Massimo Cenciarini[2] Christoph Maurer[2] Wolfram Burgard[1]

***Abstract*— The objective characterization of human motion is required in a variety of fields including competitive sports, rehabilitation and the detection of motor deficits. Nowadays, typically human experts evaluate the motor behavior. These evaluations are based on their individual experience which leads to a low inter- and intra-expert reliability. Standardized tests improve on the reliability but are still prone to subjective ratings and require human expert knowledge. This paper presents a novel method to characterize the motor state of Parkinson patients using full body motion capturing data based on a combination of multiple metrics. Our approach merges various metrics with a Random Forest and uses a probabilistic formulation to compute a one-dimensional measure for the performed motion. We present an application of our approach to the problem of relating subject motion to different classes like healthy subjects and Parkinson disease patients with deep brain stimulation switched on or off. In the experimental session we show that our measure leads to high classification rates and high entropy values for real-world data. Besides, we show that our method discriminates between Parkinson's subjects (with and without stimulation) and healthy persons as good as the Unified Parkinson's Disease Rating Scale (UPDRS).**

I. INTRODUCTION

Human motion analysis is highly relevant in several application domains including humanoid robotics, competitive sports, rehabilitation and the detection of movement disabilities. The motion of a human is heavily influenced by its neurological and physiological condition. Especially certain neurological disorders, such as Parkinson's disease (PD) cause movement impairments which need to be properly assessed in order to achieve a proper diagnosis. A popular way to analyze human motion is to rely on the visual inspection and lifelong experience of experts. In the context of PD, a therapy typically involves multiple estimates of the current state of a patient. However, these evaluations depend on previously collected knowledge and differ from one expert to the other. To arrive at an objective measure, experts nowadays evaluate a patient with the standardized Unified Parkinson's Disease Rating Scale (UPDRS) [11]. This rating scale collects scores from different short exercises, which typically cannot be executed correctly by PD patients. The UPDRS includes tests of the muscle tone (rigor) of arms, legs, and neck. Moreover, more complex movements (standing up) and the ability to perform fast repetitive movements is incorporated. However, the UPDRS suffers from a low inter- and intra-expert reliability (Richards *et al.* [14]) and, thus, is subjective. In addition, the UPDRS covers only a fraction of motion abnormalities.

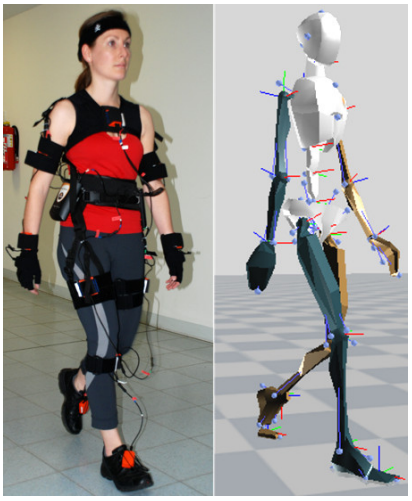


Fig. 1: The XSens motion capture suit and the visualized skeleton.

In this paper, we deal with the problem of measuring the performance of movements from PD patients using an XSens Motion Capture suit such as the one depicted in Fig. 1. We capture the motion of our subjects performing different tasks including walking, getting up or pouring water from one glass into another. All subjects belong either to the group of Parkinson's patients with deep brain stimulation, without deep brain stimulation or to the healthy control group.

Deep brain stimulation (DBS) has become a useful tool to reduce the motor impairments of PD patients. One can think of it as a pacemaker for the brain. A few weeks after the implantation of the electrodes, an expert activates the stimulator for the first time and adjusts the parameters like stimulus amplitude, frequency, or pulse width. However, up to now, adjustments are done on a trial and error base while getting feedback from a patient. The performance measure presented in this paper compares the executed motion with the one of healthy subjects and returns a score. Furthermore, our method detects the performance increase with an activated stimulation. Our long-term goal is a closed-loop system in which sensors give feedback to a control unit that autonomously computes a new set of improved parameters. The first step towards such a system is a performance metric for the motion quality, which we present in this paper.

[1] These authors are with the Department of Computer Science, Albert-Ludwigs University of Freiburg, Germany. [2] These authors are with the University Hospital in Freiburg, Germany. This research was supported by the German Research Foundation (DFG, grant number EXC 1086). kuhnera@informatik.uni-freiburg.de

II. RELATED WORK

A performance measure has an important value during a therapy. However, a lot of research focuses on classification in a healthy and PD group or between various symptoms. Stawarz *et al.* [16] and Boczarska *et al.* [2] recorded data with a Motion Capture (MoCap) System and analyzed resting tremor in limbs with different stimulation settings. They compared patients with stimulation on/off and medication on/off and found statistically significant differences between various parameter settings. Ruzicka *et al.* [8] deploy a small MoCap system that tracks the tips of two fingers during a tapping test which is usually part of the UPDRS and Lewek *et al.* [9] recorded the arm swing where they found differences in the magnitude of the left and the right arm swing among PD patients. There are several approaches which use marker based MoCap data of PD patients to evaluate the trajectory with several metrics and classify the subjects in different groups ([4, 12, 13]).

In contrast to the previous approaches we make use of a MoCap suit from XSens which tracks the motion through inertial measurement units. Similar approaches with a MoCap suit are those of Giuberti *et al.* [6, 7] who investigate the relation between leg motion and the UPDRS.

Overall, assigning scores to the performance quality of a motion is the topic of this paper. Rincon *et al.* [15] calculated four metrics on MoCap data and showed differences between patients with cerebral palsy and healthy subjects. We deploy various metrics from Balasubramanian *et al.* [1] where the authors researched the movement smoothness of stroke patients during their rehabilitation time. In combination with our own metric which calculates the probability of a trajectory to match a certain motion pattern (Fig. 2 shows three patterns) we combine these single measures with an altered Random Forest ([3]) into a global performance measure.

Mera *et al.* [10] showed that the UPDRS and quantitative variables characterizing tremor relate to stimulation parameters. They argued that various settings might be necessary to improve different motor tasks. Hence, we need a more generalized method to improve agility more globally. With our mixture of metrics we are able to compare PD patients with and without DBS to healthy subjects and rank them with a score similar to the UPDRS.

III. A HUMAN MOTION PERFORMANCE MEASURE

Human motion is influenced by a number of different neurological and physiological factors and the challenge for a performance metric lies in the generalization among various different styles of motion like arm, leg or body movements. Even healthy subjects tend to move distinct from one to the other and they execute tasks in different ways.

A. Data

MoCap allows one to record the movement of the whole body. We make use of the XSens MoCap suit which stores the trajectory of a person by saving the position and orientation of each body segment in the global frame with $120\,\mathrm{Hz}$. There are a total of 23 body segments and 22 joints. The distribution of the test group is given in Table I. The whole testings involve timed up and go, functional reach, $10\,\mathrm{m}$ walk, 90 degree turn and a hand coordination task. However, we chose the $10\,\mathrm{m}$ walk which is a non-goal oriented whole body movement task with major impact on life quality.

Condition	Quantity
Healthy	25
PD w/o DBS, w/o med.	8
PD w DBS, w/o med.	7
PD w DBS, w med.	8
PD w and w/o DBS, w and w/o med.	6

TABLE I: The quantities of individuals in each condition group.

B. UPDRS

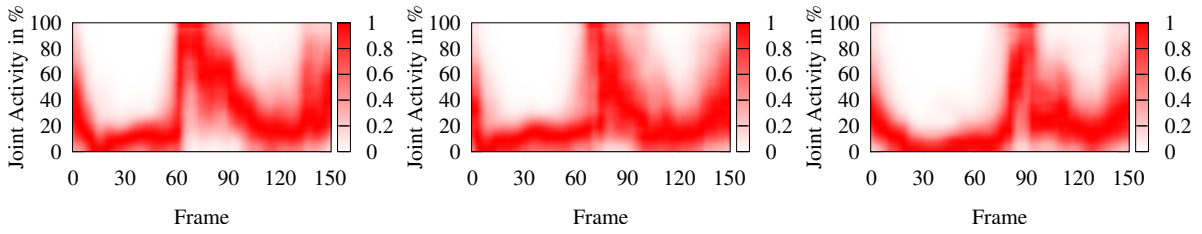
The disease state of PD subjects is usually measured with the UPDRS. It represents the rating of a human expert through a mainly visual examination. The score involves the stiffness in muscles, the quality of repetitive motions and the ability to perform multitasking with both hands. However, this evaluation is subjective and, thus, an objective performance measure is desirable, which is capable to identify the severeness of the current condition of a subject.

C. Joint Activity Metric

The Joint Activity metric represents the distance between a trajectory and a learned motion pattern. Each frame of each body segment is a probability distribution. In Fig. 2 we show distribution plots of normalized joint speeds during a stride of one segment. The first plot is about healthy subjects, the second one shows the group of PD with DBS and the third one PD patients without DBS. The color depicts the likelihood of a given speed at a certain frame during the stride. Each group has a slightly different speed pattern during walking. The result of the metric is the likeliness to be the trajectory of a healthy subject. Hence, healthy and PD subjects with DBS should achieve a high value while PD patients without DBS get a low score.

We model the disease state of the subject by the random variable $\mathbb{S}$, which takes the values $\{h, pd\} = \{\text{healthy}, \text{PD}\}$. In order to respect different walking speeds among the subjects, we normalize the joint speeds as follows. We scale the strides to equal length N and divide the joint speeds by the maximum velocity inside of the corresponding stride. For a given frame t and subject x let $\left\{v_{x,t}^k\right\}_{k=1}^{N_{S,x}}$ be the set of normalized joint speeds over all strides, where $N_{S,x}$ denotes the number of strides.

In the following, we derive the probability of the subject x to be healthy given the information at frame t. We assume that all information is given by the set of normalized joint speeds, modeled by the random variable V. Thus, we are interested in the probability of the subject x to be healthy given the set $\left\{v_{x,t}^k\right\}_{k=1}^{N_{S,x}}$ of normalized joint speeds at fixed



(a) Healthy subjects (b) PD with DBS (c) PD without DBS

Fig. 2: Normalized joint speeds during a stride at the foot segment. The stride length was normalized to $N = 150$ frames. The color shows the likelihood of the speed over a frame where red is the highest likelihood and white zero. The left picture depicts the profile of healthy subjects, the middle one the PD group with DBS and the right one the PD group without DBS.

frame t, which, according to Bayes theorem is

$$P_t\left(\mathbb{S} = h \mid \{v_{x,t}^k\}_{k=1}^{N_{S,x}}\right) = \frac{P_t\left(\{v_{x,t}^k\}_{k=1}^{N_{S,x}} \mid \mathbb{S} = h\right) P\left(\mathbb{S} = h\right)}{P_t\left(\prod_{k=1}^{N_{S,x}} V = \{v_{x,t}^k\}_{k=1}^{N_{S,x}}\right)}. \quad (1)$$

We assume that the normalized joint speeds are independent and that we have no prior knowledge of the probability distribution $P_t(V)$. Thus, we assume it to be uniform distributed in the interval $[0, 1]$, which leads to

$$P_t\left(\mathbb{S} = h \mid \{v_{x,t}^k\}_{k=1}^{N_{S,x}}\right) = \eta \prod_{k=1}^{N_{S,x}} P_t\left(v_{x,t}^k \mid \mathbb{S} = h\right), \quad (2)$$

where η denotes a normalization factor. Next, we model the probability distribution $P_t\left(v_{x,t}^k \mid \mathbb{S} = h\right)$. For each healthy subject $j \in H$ we get a set $\{v_{j,t}^i\}_{i=1}^{N_{S,j}}$ of normalized joint speeds over all $N_{S,j}$ strides. Thus, given the training set of healthy subjects, the optimal probability distribution would be given by a weighted sum of delta distributions

$$P_t\left(V \mid \mathbb{S} = h\right) = \frac{1}{|H|} \sum_{j \in H} \frac{1}{N_{S,j}} \sum_{i=1}^{N_{S,j}} \delta_{v_{j,t}^i}(V). \quad (3)$$

To avoid overfitting the training data, we add some Gaussian noise v through a convolution with a normal distribution with zero mean and through experiments tuned variance σ. Thus, we obtain the probability that at a given frame t the subject x is healthy, by

$$P_t\left(\mathbb{S} = h \mid \{v_{x,t}^k\}_{k=1}^{N_{S,x}}\right) = \eta \prod_{k=1}^{N_{S,x}} \left[\frac{1}{|H|} \sum_{j \in H} \frac{1}{N_{S,j}} \sum_{i=1}^{N_{S,j}} \mathcal{N}\left(v_{x,t}^k; v_{j,t}^i, \sigma\right)\right]. \quad (4)$$

Let $v^k = \left(v_1^k, \ldots, v_N^k\right)$ be the vector of normalized joint speeds of the k-th stride with N frames. Then the performance measure of this stride is given by the mean over the probabilities, i.e.,

$$L\left(v^k\right) = \frac{1}{N} \sum_{t=1}^{N} P_t\left(\mathbb{S} = h \mid V = v_t^k\right). \quad (5)$$

In the following we combine the Joint Activity metric with a variation of a Random Forest to obtain a probability distribution for a subject to be healthy.

D. Random Forest with Probability Distributions

The Random Forest (Breiman *et al.* [3]) is a method for classification and regression where multiple decision trees evaluate their own decision and combine them in a final step to a global prediction. Due to the nature of a decision tree a single one would overfit the training data whereas the Random Forest creates a good generalization. One generalization idea is to reduce the training set for each decision tree to a subset drawn with replacement from the original data. A second step is to consider a subset of features for each decision tree. One commonly takes the square root of the total number of features. In our case the feature set is given by $M = T \times B$, where T is the set of various types of metrics and B is the set of different body segments. For each node of the decision tree the metric-body segment pair which optimizes information gain is used to split the data into two subsets.

The resulting decision tree has on every node a threshold for a specific metric and body segment. All results, combined in a majority count or mean, predict the class or a score. However, we are not interested in the precision of the classification but separation of data in each tree node. The idea of a Random Forest with probability distributions is that on each node the probability distribution of being a healthy subject and being a PD w/o DBS are computed. Hence, the probability of a data value x_m of a metric-body segment pair m to be of a healthy subject is

$$P_\tau\left(\mathbb{S} = h \mid x_m\right) = \frac{P\left(x_m \mid \mathbb{S} = h\right) P\left(\mathbb{S} = h\right)}{\sum_{s \in \{h, pd\}} P(x_m \mid \mathbb{S} = s) P(\mathbb{S} = s)}, \quad (6)$$

where τ is one of the decision trees.

Given the data value x_m of a stride of one person $\widehat{j}$, we assume the probabilities $P\left(x_m \mid \mathbb{S} = h\right)$ and $P\left(x_m \mid \mathbb{S} = pd\right)$ to be normally distributed plus uniform noise. We use leave-one-out cross-validation to compute the means $\left(\mu_{H_m}, \mu_{O_m}\right)$ and variances $\left(\sigma_{H_m}^2, \sigma_{O_m}^2\right)$ of each normal distribution. Let $H_m = H_{m,\widehat{j}}$ be the values of healthy subjects and let $O_m = O_{m,\widehat{j}}$ be the values of PD subjects without Deep

Brain Stimulation for m, except all values which correspond to person $\widehat{j}$. For the computation of the variances we use the corrected empirical variance, i.e.,

$$\sigma^2_{H_m} = \frac{1}{N_{H_m} - 1} \sum_{i=1}^{N_{Hm}} \left(H^i_m - \mu_{H_m}\right)^2, \tag{7}$$

where N_{H_m} is the number of elements in H_m. We calculate the same for the set O_m and obtain $\mathcal{N}(\mu_{H_m}, \sigma^2_{H_m})$ and $\mathcal{N}(\mu_{O_m}, \sigma^2_{O_m})$. Thus, we get

$$P(x_m|\mathbb{S} = h) \propto \alpha \mathcal{N}\left(\mu_{H_m}, \sigma^2_{H_m}\right) + (1-\alpha)\mathcal{U}(0,1)$$
$$\text{and} \quad P(x_m|\mathbb{S} = pd) \propto \beta \mathcal{N}\left(\mu_{O_m}, \sigma^2_{O_m}\right) + (1-\beta)\mathcal{U}(0,1),$$

where $\alpha, \beta \in [0,1]$ denote weight parameter and $\mathcal{U}(0,1)$ is a uniform distribution in the interval $[0,1]$, which takes movement outliers into account. In the experiments we choose $\alpha, \beta = \frac{2}{3}$. In each node $n \in \mathcal{N}$ the metric-body segment pair $m = m(n) = m_n$ is chosen which optimizes the information gain. Thus, one might assume that the metric-body segment pair of a subsequent node is independent of its root node. With this assumption we end up with

$$P_\tau(\mathbb{S} = h \mid x) = \frac{P(\mathbb{S} = h) \prod\limits_{n \in \mathcal{N}} P(x_{m_n} \mid \mathbb{S} = h)}{\sum\limits_{s \in \{h, pd\}} \left(P(\mathbb{S} = s) \prod\limits_{n \in \mathcal{N}} P(x_{m_n} \mid \mathbb{S} = s) \right)}, \tag{8}$$

where $x = (x_k)_{k \in M}$ denotes the total feature-vector.

Each decision tree τ computes an individual probability $P_\tau(\mathbb{S} = h \mid x)$. The mean over all probabilities yields the final score for a stride

$$P(\mathbb{S} = h \mid x) = \frac{1}{N_T} \sum_{\tau=1}^{N_T} P_\tau(\mathbb{S} = h \mid x), \tag{9}$$

where N_T denotes the number of decision trees.

As mentioned above we use the Joint Activity metric as a feature set for our proposed variation of the Random Forest. Besides this, we use several metrics defined by Balasubramanian *et al.* [1] as underlying feature sets. This includes the Spectral Arc Length, the Root Mean Square Jerk, the Speed Arc Length, the Dimensionless Jerk, the Log Dimensionless Jerk and finally the Normalized Mean Absolute Jerk. Additionally, we apply the Random Forest with probability distributions to the union of all the used feature sets above. In the following, we call the obtained combinations measures and denote them by the name of the used metric.

IV. EXPERIMENTAL EVALUATION

In our context a good performance metric has to fulfill multiple aspects. First, the metric has to be able to classify each subject into its respective class. Second, the information content of the metric should be high, and third, the metric should be close to state of the art PD rating scales evaluated by human experts like the UPDRS.

We first look at the classification results before we discuss the quality of a measure to distinguish between healthy subjects and PD patients without stimulation.

Metric	Precision in %	Recall in %	F-Score in %
JA	91.5	85.5	88.4
SAL	98.5	**95.9**	**97.2**
RMSJ	87.3	86.1	86.7
SpAL	95.7	74.7	83.9
DJ	85.9	72.6	78.7
LDJ	85.9	71.7	78.2
NMAJ	95.7	73.1	82.9
Combined	**100.0**	91.0	95.9

TABLE II: Precision, recall and F-Score of all metrics combined with a Random Forest and a Random Forest over all metrics. Hereby, JA denotes the Joint Activity, SAL the Spectral Arc Length, RMSJ the Root Mean Square Jerk, SpAL the Speed Arc Length, DJ the Dimensionless Jerk, LDJ the Log Dimensionless Jerk and NMAJ the Normalized Mean Absolute Jerk.

A. Classification Rates

Fernandez *et al.* [5] compared various classification and regression algorithms. The Random Forest came out to be the best among all training sets. Table II displays precision, recall and F-score for Random Forests trained on different metrics as feature set. Each Random Forest is either trained on all dimensions of a metric or all metrics combined and consists of 100 trees without depth restriction where we train each tree with 40 percent of the data and the square root of the total number of features. Except for the Root Mean Square Jerk, the Dimensionless Jerk and the Log Dimensionless Jerk all other metrics have precision rates above 90 % where the Spectral Arc Length and the combination of all measures have recall rates above 90 %. The Spectral Arc Length has with 97.2 % the highest F-Score.

However, our main focus is not the classification but the definition of a objective measure for the performed movements. Overall, the metric should distinguish between Parkinson subjects with and without stimulation while maintaining a smooth transition between both groups.

B. Entropy

Our goal of a objective measure is not a binary classification but a one-dimensional measure which spreads over the entire metric space. Not every subject inside a group performs equally well. Hence, we want that the distribution of metric values is spread over the entire interval to have the possibility of a smooth transition between off and on state. Fig. 4 shows the normalized entropy values

$$\eta = -\sum_{n=1}^{N} \frac{P(\mathbb{S} = h \mid x_n) \log P(\mathbb{S} = h \mid x_n)}{\log N} \tag{10}$$

for each metric over all samples (Parkinson patients with and without stimulation and healthy subjects). Hereby, $\eta = 1$ would represent a uniform distribution and $\eta = 0$ a delta distribution. On top of each bar is the F-Score and we see that metrics with an high F-Score (bold) achieve a higher normalized entropy than their classic Random Forest counterparts. Overall, the variation of a Random Forest creates a more distributed metric space than the classic one

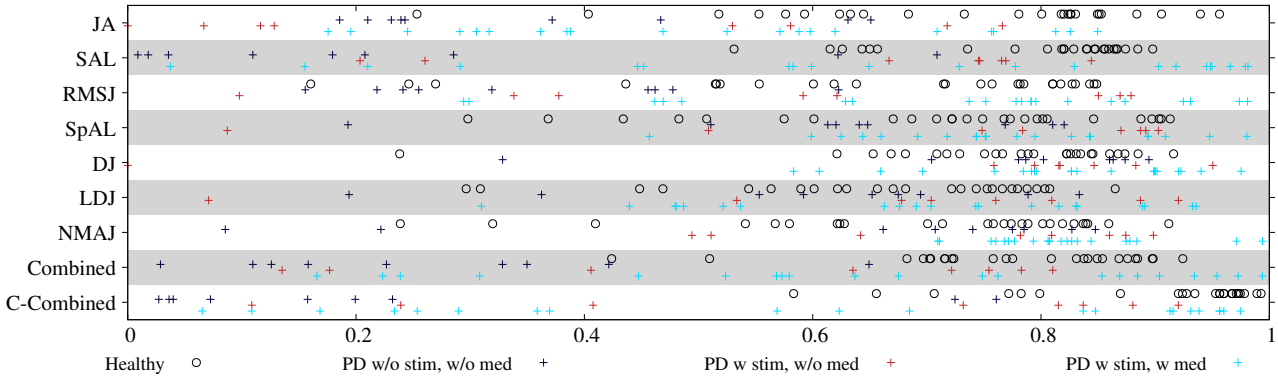


Fig. 3: Each measure consists of either 22 or 23 dimensions upon we train a Random Forest. The probabilities which we calculate with the structure of the Random Forest represent the shown values. The combination measure involves all dimensions from each single measure. Hence, it has 160 dimensions to train a Random Forest with.

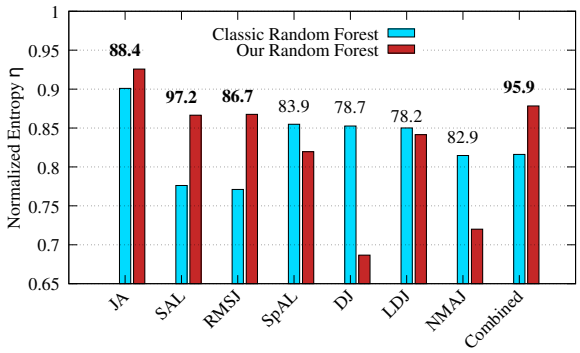


Fig. 4: The entropy for each metric either with our proposed alternation of a Random Forest or the certainty of a classic Random Forest. The measure notations are as in Table II. On top of each bar is the F-Score in %.

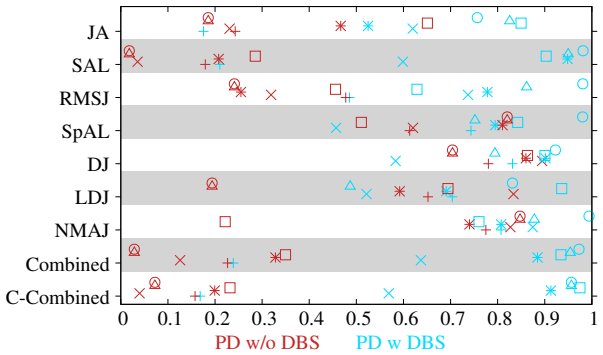


Fig. 5: Shift of patients with both active (blue) and inactive DBS (red). Overall, the performance increases with an active stimulation. The measures are notated as in Table II.

(0.875 against 0.821). The most uniform distributed metrics are the Joint Activity with our variation of a Random Forest and with the classic Random Forest.

This reveals that the combination of all metrics with a Random Forest distributes all measures over the entire metric space. Hence, it is possible to see meaningful differences between stimulation activated and stimulation turned off.

C. Performance Improvement Between Off and On Condition

In Fig. 5 we see the performance increase among PD subjects. For the Spectral Arc Length six out of six subjects show an improvement, where the effect for one is negligible. Five patients show an performance increase if we look at the Joint Activity, the Root Mean Square Jerk and the combination of all measures. The Speed Arc Length, Dimensionless Jerk, Log Dimensionless Jerk and Normalized Mean Absolute Jerk have patients with increased performance, too, but the values are closer to each other.

Overall, the combination of all metrics with our altered Random Forest and the classic Random Forest show the correct shift of performance between off and on condition as well as a data distribution over the entire interval.

D. Measure Scores

Fig. 3 displays the value distribution of each measure. It contains four groups of points for each measure: one for the healthy subjects (circles) and three for the PD patients (crosses). Joint Activity, Spectral Arc Length and combined show the expected outcome that PD subjects get a low score, healthy subjects are close to one while PD patients with DBS are somewhere in between. Root Mean Square Jerk's tendency is correct but distributes the healthy scores. The other measures mix all four groups and are not reliable.

Fig. 6 shows the probabilities to be not healthy such that a comparison to the UPDRS mean score is possible. Joint Activity, Spectral Arc Length, Root Mean Square Jerk, and combination of all measures compute for most healthy subjects (black) a low probability and a high one for PD subjects without DBS (blue). The mean UPDRS is 44, where 0 is a perfect score and 100 is the worst possible outcome. However, 61 is the worst value in our data set. Hence, we use it as a threshold in Fig. 6. Joint Activity, Spectral Arc Length, Root Mean Square Jerk and the combination behave similar to the UPDRS. Interestingly, PD subjects with DBS (red) offer values between that of healthy subjects and PD without DBS. In comparison, the mean UPDRS

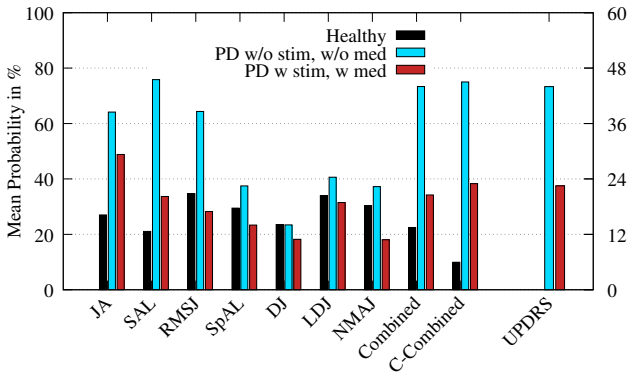


Fig. 6: The mean probability to be not healthy over all subjects separated by their condition - healthy (black), PD without DBS (blue) and PD with DBS (red). As ground truth we have the mean values of the UPDRS on the interval between $[0, 61]$ where 61 is the worst score of one of our PD patients. Measure notations are as in Table II.

is 22.5. Hence, for both, the measures and UPDRS, the performance increased from inactive to active stimulation. Dimensionless Jerk, Log Dimensionless Jerk, Speed Arc Length and Normalized Mean Absolute Jerk have slight differences between the off condition and healthy subjects and as such are not reliable.

The classic Random Forest behaves similar than our version. The mean values for either Parkinson's patient with and without stimulation are close to the UPDRS value. However, the mean value of healthy subjects is closer to zero than the mean of the altered Random Forest.

We have seen that our combination of metrics with an altered Random Forest delivers a possibility to separate data of different subjects into their respective groups as well as present a shift between different states like with and without stimulation.

V. CONCLUSIONS

The paper proposes a new metric termed Joint Activity to quantify the performance of a PD patient in comparison to the motor behavior of healthy subjects. The Joint Activity and multiple other metrics are the input for a Random Forest where we exploit the tree structure to compute probability distributions to be either a healthy subject or a PD patient. Our results show that these probability distributions can be used to form a performance measure to quantify the progress during a therapy or to evaluate the efficacy of a set of DBS stimulation parameters by comparing motions before and after DBS is turned on. In contrast to the UPDRS our metric is objective and not prone to intra- and inter-expert variability errors. Furthermore, it is the first step towards a closed-loop system to automatically adjust DBS parameters.

REFERENCES

[1] S. Balasubramanian, A. Melendez-Calderon, and E. Burdet. A robust and sensitive metric for quantifying movement smoothness. *IEEE Transactions on Biomedical Engineering*, 59(8):2126–2136, 2012.

[2] M. Boczarska-Jedynak, S. Kwiek, R. Sordyl, K. Kubicki, M. Humeniuk, L. Przeklasa, M. Stawarz, A. Polanski, K. Wojciechowski, M. Arkuszewski, P. Bazowski, and G. Opala. Motion capture method in the assessment of resting tremor in patients with Parkinson's disease treated with deep brain stimulation (dbs). In *Movement Disorders*, volume 28, pages S460–S460, 2013.

[3] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[4] S. Das, L. Trutoiu, A. Murai, D. Alcindor, M. Oh, F. De la Torre, and J. Hodgins. Quantitative measurement of motor symptoms in Parkinson's disease: A study with full-body motion capture data. In *International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'11)*, pages 6789–6792, 2011.

[5] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim. Do we need hundreds of classifiers to solve real world classification problems? *The Journal of Machine Learning Research*, 15(1):3133–3181, 2014.

[6] 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 *International Conference on Wearable and Implantable Body Sensor Networks (BSN'14)*, pages 115–120, 2014.

[7] 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 of Things Journal*, 2(1):41–51, 2015.

[8] R. Krupicka, Z. Szabo, S. Viteckova, and E. Ruzicka. Motion capture system for finger movement measurement in Parkinson disease. *Radioengineering*, 23(2):659–646, 2014.

[9] M. D. Lewek, R. Poole, J. Johnson, O. Halawa, and X. Huang. Arm swing magnitude and asymmetry during gait in the early stages of Parkinson's disease. *Gait & Posture*, 31(2):256–260, 2010.

[10] T. Mera, J. L. Vitek, J. L. Alberts, and J. P. Giuffrida. Kinematic optimization of deep brain stimulation across multiple motor symptoms in Parkinson's disease. *Journal of Neuroscience Methods*, 198(2):280–286, 2011.

[11] Movement disorder society task force on rating scales for Parkinson's disease. The Unified Parkinson's Disease Rating Scale (UPDRS): status and recommendations. *Movement disorders: official journal of the Movement Disorder Society*, 18(7):738, 2003.

[12] I. Pachoulakis and K. Kourmoulis. Building a gait analysis framework for Parkinson's disease patients: Motion capture and skeleton 3D representation. In *International Conference on Telecommunications and Multimedia (TEMU'14)*, pages 220–225, 2014.

[13] B. Pogorelc and M. Gams. Discovery of gait anomalies from motion sensor data. In *IEEE International Conference on Tools with Artificial Intelligence (ICTAI'10)*, volume 2, pages 331–336, 2010.

[14] M. Richards, K. Marder, L. Cote, and R. Mayeux. Interrater reliability of the Unified Parkinson's Disease Rating Scale motor examination. *Movement Disorders*, 9(1):89–91, 1994.

[15] V. Rincon Montes, Y. Quijano, J. E. Chong Quero, D. Villanueva Ayala, and J. Perez Moreno. Comparison of 4 different smoothness metrics for the quantitative assessment of movement's quality in the upper limb of subjects with cerebral palsy. In *Pan American Health Care Exchanges (PAHCE)*, pages 1–6, 2014.

[16] M. Stawarz, A. Polański, S. Kwiek, M. Boczarska-Jedynak, Ł. Janik, A. Przybyszewski, and K. Wojciechowski. A system for analysis of tremor in patients with Parkinsons disease based on motion capture technique. In *Computer Vision and Graphics*, pages 618–625. Springer, 2012.