# Analysis and Quantification of Repetitive Motion in Long-Term Rehabilitation

Loreen Pogrzeba, Thomas Neumann, Markus Wacker, and Bernhard Jung

***Abstract*—Objective assessment in long-term rehabilitation under real-life recording conditions is a challenging task. We propose a data-driven method to evaluate changes in motor function under uncontrolled, long-term conditions with the low-cost Microsoft Kinect Sensor. Instead of using human ratings as ground truth data, we propose kinematic features of hand motion, healthy reference trajectories derived by principal component regression, and methods from machine learning to analyze the progression of motor function. We demonstrate the capability of this approach on datasets with repetitive unrestrained bi-manual drumming movements in 3-dimensional space of stroke survivors, patients suffering of Parkinson's disease, and a healthy control group. We present processing steps to eliminate the influence of varying recording setups under real-life conditions and offer visualization methods to support clinicians in the evaluation of treatment effects.**

***Index Terms*—depth sensor, human motion, kinematic features, rehabilitation, movement quality assessment.**

## I. Introduction

NEUROLOGICAL deficits as a consequence of a stroke or Parkinson's disease have sustained impact on daily life. They entail symptoms such as reduced mobility, paralysis or rigidity of limbs, higher risk of falling and pain. The need for long-term rehabilitation is apparent, as stroke is "a major cause of long-term disability" [1] and Parkinson's disease as a chronic disease involves deterioration of symptoms.

The advent of low-cost, mobile, and easily applicable markerless motion recording systems like the Microsoft Kinect depth sensor (short: Kinect sensor) opens up new fields of application in therapy and rehabilitation, especially in elderly care, stroke rehabilitation, and exergaming [2], [3]. Current research focuses mainly on interdisciplinary short-term studies under controlled laboratory conditions, with motion analysis results being correlated with qualitative clinical assessment scales as gold standard. However assessment scales are dependent on the ratings and experience of the evaluators, thus can be subjectively distorted [4]. They may not coercively correlate with the results from motion analysis, because the chosen scales could be too coarse or too general, thus not responsive enough for long-term tracking of symptoms or motor changes [5], [6], [7].

L. Pogrzeba, T. Neumann and M. Wacker are with University of Applied Sciences Dresden, Germany (e-mail: pogrzeba@htw-dresden.de; tneumann@informatik.htw-dresden.de; wacker@informatik.htw-dresden.de.)

B. Jung is with University of Technology and Mining in Freiberg, Germany (e-mail: jung@informatik.tu-freiberg.de).

Manuscript received October 15, 2017; revised May 10, 2018. This research was supported by ESF (grant no. 100231931, TISRA, and grant no. 100098265, PhD scholarship).

In addition, previous studies in rehabilitation often focus on uni-directional trajectories, for example reaching movements with predefined start and end points in space. Real-life rehabilitation settings are usually much less constrained, containing unpredictable reaching targets in space and potentially asymmetric execution (forward and backward motion). Few studies have explored such a real-life rehabilitation setting. Here, we exemplarily study the quantification of repetitive motion from recordings of treatment sessions with function oriented music therapy (FMT). FMT is a non-verbal neuromuscular therapy based on repetitive drumming movements in changing setups of instrumentation [8], [9]. FMT is targeted to treat diverse neurological deficits, such as stroke (S) and Parkinson's disease (PD). In a long-term rehabilitation setting like this, the aim is not to detect diseases at an early stage, but instead to offer computational tools that help monitoring the rehabilitation progress as unobtrusively as possible.

Such a real-life scenario poses several technical and methodological challenges: we require a method of normalization that allows for an analysis not only invariant under varying recording conditions, but also invariant to changing motion tasks during therapy. Classical approaches record motion from an impaired patient group (PG) and compare it to data of a healthy control group (HG) [3], [10]–[13].

To monitor and quantify long-term rehabilitation progress, the quality (the "healthiness") of a given motion needs to be estimated from kinematic features. To obtain such a measure, we propose a model that predicts a probability between "healthy" and "impaired" from the kinematic features of a given motion. The model thus provides a continuous score of "healthiness" as a corridor of accepted motor function. Notably, this model is trained only from sets of healthy and impaired motion. It does not require subjective and potentially distorted therapist scores for calibration. Monitoring the model scores in an ongoing therapy allows us to estimate the recovery of the patient. We show that, both for stroke and Parkinson's patients, model scores successfully quantify the tendency of rehabilitation of a patient. A therapist in practice could thus use our model to quickly check whether symptoms improve or even disappear over the course of long-term therapy.

In summary, our contributions are:

1) We describe a framework for recording, automatic calibration, and analysis of repetitive motion in real-life conditions. We build a reference trajectory model to correct for varying setups and propose three kinematic features that quantify variability and consistency of a given repetitive reaching/drumming motion.

2) We show that a probabilistic model trained from a set of "healthy" and "impaired" motion can be used to monitor the recovery of patients towards "healthier" motion during long-term therapy.
3) Our study is the first to offer a computational, motion data based assessment of rehabilitation success of FMT, based on a novel dataset of drumming motion recorded in unconstrained therapy sessions. We quantify the motion of both stroke and Parkinson's patients.

## II. Related Work

Motion analysis, in particular quantification of motion quality, have been studied in various contexts: for personalized rehabilitation systems [6], [14], ergonomics [15], [16], for measuring motor symptoms of Parkinson [10], [17], [18] and stroke patients [5], [11]. For our specific use case in FMT, therapeutic observation criteria have been transferred [8], but not yet evaluated for automatic motion analysis. Current research in motion quantification is oriented towards establishing correlations between kinematic features and human ratings (e.g., the Wolf Motor Function Test (WMFT) [12] for stroke survivors) to build evaluative or predictive models. In contrast to such disease-specific motor performance scores, we analyze drumming motion during unconstrained, long-term music therapy by implicitely modeling a "healthyness" score without relying on human ratings. Note that drumming movements can be seen as compositions of multiple reaching tasks, therefore our framework also generalizes to motion analysis for reaching tasks and hopefully inspires future work also in this context.

To measure human motion, most studies rely on expensive marker-based systems. Recently, low-cost sensors such as the Kinect sensor have been shown to achieve comparable results [19]–[21] in various applications settings [2], [3], [22], [23]. We argue that the use of such a sensor in a real-life rehabilitation setting not only poses big challenges due to sensor noise and limited accuracy, but also causes problems due to uncontrolled recording conditions that have to be factored into the analysis framework. Our framework normalizes the data even in such uncontrolled setups.

To analyze the movements recorded from multiple subjects, many existing approaches explore the use of kinematic features for assessing movement quality: Venkataraman *et al.* [6] use curvedness, speed, and jerkiness; Das *et al.* [17] use frequency-domain features to measure tremor; Chen *et al.* [11] explore features such as temporal, velocity, and trajectory profiles; Adams *et al.* [24] analyze duration, normalized speed, and movement arrest period ratio. These kinematic features are usually combined to predict movement quality scores using machine learning. Leightley *et al.* [25] evaluate machine learning methods to first classify motion type, then compute deviations from a healthy control group to label movements as "good" or "poor". Mostafavi *et al.* [26] extensively analyze the relationships between kinematic features and clinical scores for reaching, matching, and object hit tasks in stroke survivors.

We extend the general idea of interpretable kinematic features in the case of drumming motion, for measuring long-term rehabilitation effects, and for a patient group with a very wide spectrum of characteristics poststroke and with Parkinson's disease.

The methods mentioned above typically learn a mapping directly from kinematic features to therapist ratings from a dataset of impaired and healthy patients that perform the same motion. Essentially, healthy motion is modeled in the kinematic feature space. As an alternative, some methods model healthy trajectories directly: For example, Olesh *et al.* [5] model motor function of the non-paretic (healthy) arm using Principal Component Analysis (PCA), reconstruct the other hand motion within this PCA space, and measure the difference (and vice versa). This gives a quantitative scale that works well for patients with hemiparesis, but it strongly fluctuates over movement types. Models that decompose motion into sparsely-activated motor primitives can also be used, e.g., to reveal problems in coordination [18]. Burget *et al.* [10] train a mathematical model of individual joint motion and show reduced activation of proximal joints for PD patients. Som *et al.* [27] generate an "optimal" trajectory synthetically as the shortest geodesic on a manifold that respects motion specific constraints of the human body. This allows for completely unsupervised modeling of motion, but cannot capture factors such as acceleration and energy efficiency, factors which are important for modeling natural human motion. Trajectories generated from recorded data overcome this limitation, for example by fitting Bezier curves to MoCap data [16] or by generating human gait trajectories based on variables such as gender [28] in a data-driven way. We argue that generative models like these can be used to factor external variables in reaching movements (such as start/end point), and show how such a model can be tied to the construction of kinematic features, thereby enabling real-life long-term rehabilitation analysis and monitoring without interfering with therapists.

## III. Framework Design

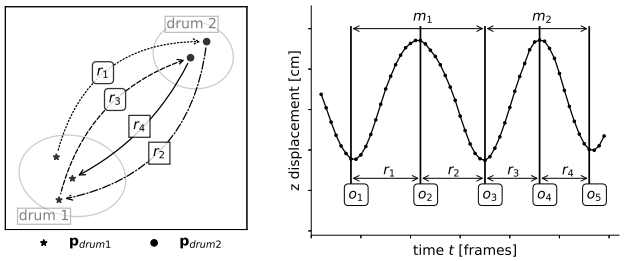
To quantify human movements over uncontrolled long-term treatment we present a framework which utilizes four steps. First, the trajectories are transformed into a uniform spatial representation to allow consistent analysis also under varied spatial recording conditions. Second, a reference model is built which synthesises healthy trajectories depending on the parameters of the recorded impaired movements. Third, kinematic parameters are calculated and forth, used for the estimation of treatment effect.

### A. Representation of Trajectories

Skeletal data consists of a time-ordered sequence of joint positions in Cartesian space, also named as *trajectories*. Let $\left\{\mathbf{p}_j(t) \in \mathbb{R}^3, t \in \mathcal{T}\right\}$ denote a set of joint trajectories at measured time points $\mathcal{T} \subset \mathbb{R}_+$ and for different joints $j \in \mathcal{J}$. For our use case we analyze the movements of two joints of interest, $\mathcal{J} = \{lh, rh\}$, namely the left ($lh$) and right hand ($rh$). We define repetitive drumming actions as composed movements, which are built from a number of repetitive reaching actions $r$ (cf. Fig. 1a). For a 2-drum-setup we arrange two reaching actions in a single *motion cycle* (cf. Fig. 1b) and combine ten motion cycles to a *set*.

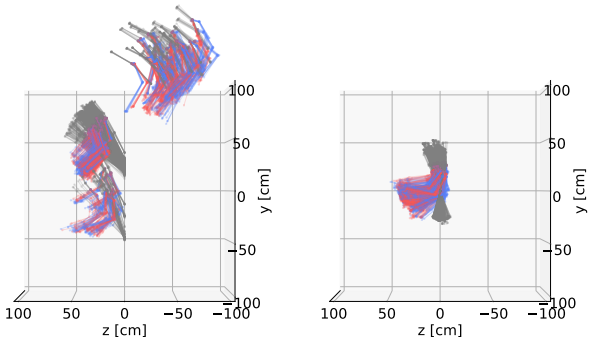
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2848103, IEEE Journal of Biomedical and Health Informatics





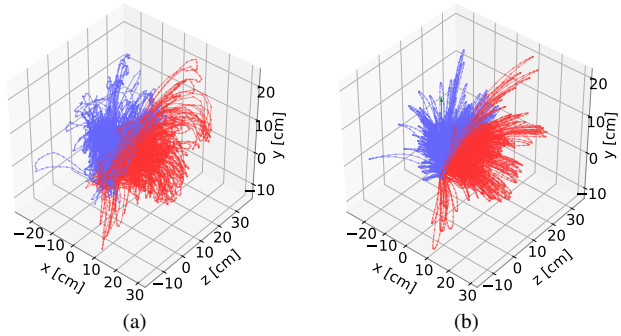

(a) (b)

Fig. 1. (a) Schematic representation of reaching trajectories $r_i$ ($i \in \mathbb{N}$) of one skeletal joint $j$ (here, one of the hands) in camera coordinates for a 2-drum-setup. (b) Exemplary position-time graph, with continuous reaching actions $r_i$ across time $t$ for a 2-drum-setup. The reaching actions start and end in points of time labeled with onsets $o_k$ ($k \in \mathbb{N}$) and are combined into motion cycles $m_l$ ($l \in \mathbb{N}$).




(a) Before Processing (b) After Processing

Fig. 2. Upper body joints in side view at odd onsets in (a) with different orientations due to varied recording conditions before and in (b) with matching orientation after processing. Spine joints in gray, left/right body side in blue/red, respectively.




(a) (b)

Fig. 3. (a) Captured (real) drumming trajectories of the left (blue) and right (red) hand joints of the patient group (PG) and (b) predicted trajectories of healthy subjects from the spatial coordinates of the PG in 3d.

### B. Registration of Onsets

Drumming movements are decomposed into $R \in \mathbb{N}_+$ reaching actions. Each reach starts (*onset*) and ends (*offset*) when a drum is hit, i.e., at the distinct points in time $\mathcal{O} \subset \mathcal{T}$, with $\mathcal{O} = \{o_1, o_2, \ldots, o_R\}$. Since drumming continues immediately after a hit, onsets and offsets coincide in our scenario. Consequently, in a 2-drum-setup where the patient hits each drum alternately, we have the "odd" onsets where the first drum is hit, $\mathcal{T}_{first} = \{o_1, o_3, \ldots, o_{R-1}\}$ (cf. Fig. 1b), and the "even" onsets when the second drum is hit $\mathcal{T}_{second} = \{o_2, o_4, \ldots, o_R\}$. Each hand can be modeled separately in this way, even if both hands participate in drumming. An example is shown in Fig. 5: although three drums are involved, each hand is alternating between the center and one of the outer drums, thus each hand still performs a 2-drum motion.

For our datasets (cf. Sec. IV-A) we register onsets manually based on the image data. We here look at symmetric drumming motion, so we select one onset for both hands: if the impaired hand (left or right) is known, we register the onset of the healthy hand. Otherwise (e.g., for healthy subjects) we take the frame where both hands are at minimal y-position. If the hands move asymmetrically, this will influence the trajectories' shape after processing (as described next), consequently making this asymmetry detectable by kinematic features (see III-E).

### C. Processing

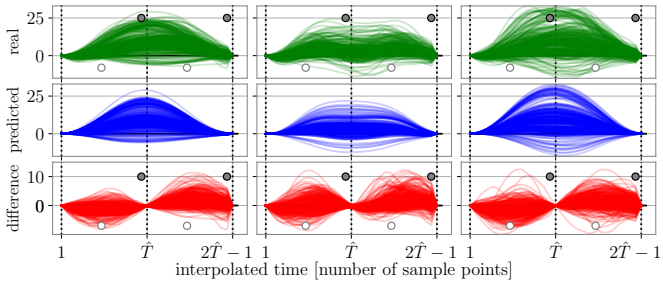
We use a five-step, fully automated processing and calibration routine to transform raw motion data into a unified trajectory representation. First, we use a Savitzky-Golay filter of order 3 following [29] to smooth the movements. Second, we correct for varying height of the sensor, which possibly arised during recording of different sessions: The height of the sensor influences the pitch angle in camera space, so we need to perform a rotation $\mathbf{R}_x$ around the $x$-axis by the angle $\theta$. To determine $\theta$ we measure the angle between the unit vector $\mathbf{y} = [0, 1, 0]^\top$ and the spine (gray lines in Fig. 2) at the start of the motion cycles, at $\mathcal{T}_{first}$, where we can assume the "most upright" posture. Averaging over all sets within one setup gives a robust estimate of the actual pitch of the sensor, cf. Fig. 2b, without requiring any manual calibration effort. Third, drumming trajectories for each joint are translated so that they start at the origin. Fourth, reaches are resampled to obtain $\hat{T} = 16$ equally spaced sample points from the raw trajectory, using cubic spline interpolation. This corresponds to the average sample rate of 29.8 Hz of the raw data and so the resampled trajectories reproduce the actual motion trajectory with high fidelity, cf. Fig. 4. Fifth, motion cycles are combined into sets. After this, all motion cycles start and end in the origin as depicted in, e.g., Fig. 3. This pipeline would also work for joint angles, but here we chose the trajectory representation of motion, as it is commonly used in the context of reaching and rehabilitation [6], [11], [14], [17]. Trajectories preserve the spatial conditions of the reaching actions, are easily visualized and lend themselves to application of scoring principles of the widely used WMFT [12] assessment for stroke survivors.

### D. Reference Trajectory Model

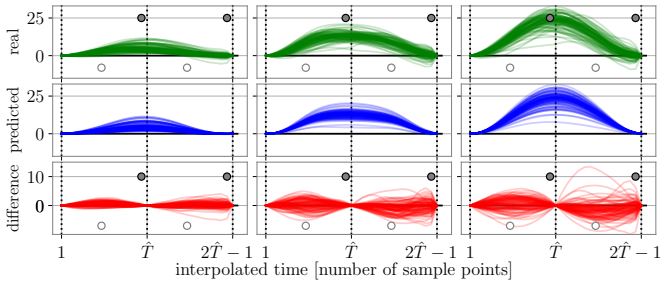
The preprocessing so far cannot sufficiently level out differences in the shape of the trajectories, which might be significantly different depending on the actual 3d location of the drums or, in general, of any reaching target [30], [31]. This is also clearly visible in Fig. 3a. Instead of forcing the therapists to place the drums exactly in the same 3d location for every patient to make patients comparable, we propose to train a model that synthesizes reference trajectories from an additional dataset, named "Setup Variation" (SV). In this

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2848103, IEEE Journal of Biomedical and Health Informatics

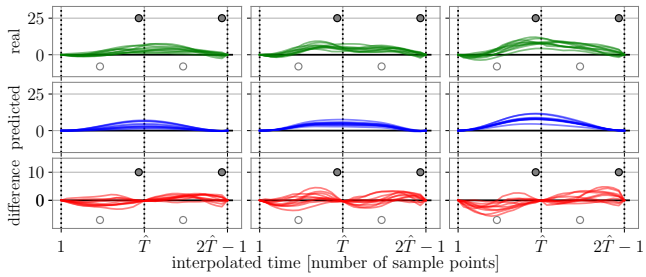




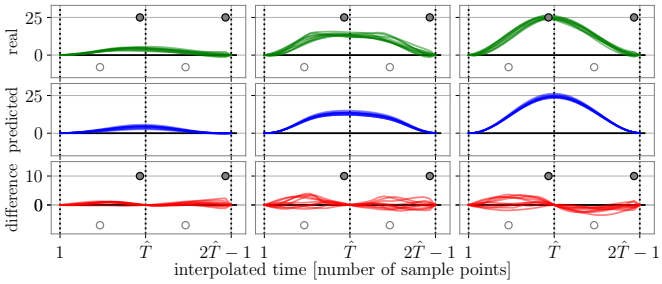

(a) All impaired subjects (PG).




(b) All healthy subjects (HG).




(c) Subject S8 from patient group (PG) after 19 weeks of treatment.




(d) Subject S28 from healthy group (HG).

Fig. 4. Real (top row, green), predicted (middle row, blue) and subtracted (bottom row, red) position data per patient group and for selected single subjects. Each row contains trajectories of two reaches in $x$, $y$ and $z$ direction, which were resampled to contain $\hat{T}$ time steps per reach. Vertical dotted lines indicate onsets of reaches. Circle markers indicate points in time, where variability features are calculated.

dataset, a part of variant features of a reaching or drumming setup is reproduced and systematically changed by healthy subjects.

From this, a model is learnt for a specific joint $j$ from $N_{SV}$ recorded trajectories that all went to different drum positions. As described previously, reach trajectories are preprocessed and resampled to contain $\hat{T}$ time steps. This allows us to collect all reaches into matrix $\mathbf{X}_{SV}^{(j)} \in \mathbb{R}^{N_{SV} \times 6\hat{T}-3}$, with each row containing the 3d positions of *two* reaches per motion cycle with forward *and* backward motion for the 2-drum-setup. The offset of 3 counts for onsets $\mathcal{T}_{second}$ (in Fig. 4 at times $\hat{T}$), that are part of multiple reaches, i.e., ending points of forward reaches and starting points of backward reaches. We then decompose the healthy trajectories into $K$ principal components $\mathbf{c}_k \in \mathbb{R}^{6\hat{T}-3}$ and weights $\mathbf{w}_k \in \mathbb{R}^{N_{SV}}$,

$$\mathbf{X}_{SV} \approx (\mathbf{x}_{mean})^\top + \sum_{k=1}^{K} \mathbf{w}_k \, (\mathbf{c}_k)^\top \quad . \tag{1}$$

Linear regression is used to model the relationship between PCA weights and variable parameters, so that $\mathbf{w}_k \approx \mathbf{Y}_{SV}\boldsymbol{\beta}_k$. In our case, $\mathbf{Y}_{SV} \in \mathbb{R}^{N_{SV} \times 4}$ contains the 3d joint positions at the even onsets, $\mathcal{T}_{second}$, as a proxy for the real drum position in 3d space, plus a constant to model the linear regression bias. To prevent overfitting and to increase robustness to outliers, we collect multiple (here, ten) motion cycles for each drum position in $\mathbf{Y}_{SV}$ and average the resampled trajectories in $\mathbf{X}_{SV}$. After learning $\boldsymbol{\beta}_k$, we can synthesize a trajectory $\bar{\mathbf{x}} \in \mathbb{R}^{6\hat{T}-3}$ for any given target drum location $\mathbf{y} \in \mathbb{R}^4$ (values for $x, y, z$, and a constant), thereby generating a trajectory that "simulates" healthy drumming to that target location by:

$$\bar{\mathbf{x}} = \mathbf{x}_{mean} + \sum_{k=1}^{K} \mathbf{y}^\top \boldsymbol{\beta}_k \mathbf{c}_k \quad . \tag{2}$$

Fig. 3 contrasts the real trajectories of the patient group (a) with predicted reference trajectories $\bar{\mathbf{x}}$ computed using Eq. (2) in (b). We can now subtract the reference trajectory from the patient data to even better reveal the irregularities visible in Fig. 3a, which is what we will show next. At first glance, this idea seems specific to our drumming use-case, but in fact, it can be easily extended simply by adding more columns to $\mathbf{Y}_{SV}$ (e.g., location of a second drum, walking speed for gait analysis, etc.). The reference model is also invariant under specific motion representation, joint angles in matrix $\mathbf{X}_{SV}$ would also work.

### *E. Kinematic Features*

We evaluate motor changes over long-term treatment with the help of three groups of kinematic features. The features $\mathbf{f}$ are calculated per set, for each joint $j \in \mathcal{J} = \{lh, rh\}$, for all subjects and sets of the patient group (PG) and healthy group (HG). They are collected in the matrix $\mathbf{F} \in \mathbb{R}^{N_{PG+HG} \times 2F}$, where $N_{PG+HG}$ denotes the number of sets contained in the datasets (PG, HG) and there are $F$ features for each joint. The features are inspired by criteria of the Functional Ability Scale, which is often used in stroke assessment as a part of WMFT. It rates amongst others completion time, precision and fine coordination of the upper extremity [32]. The first two features are also motivated by the research of Cirstea and Levin [33], who observed that pointing movements of stroke survivors involve increased movement variability and

more widely distributed end-point positions. In our work, we also assume higher variability in movements for impaired subjects, in particular: (i) at the end of the reaches supported by [33], [11], (ii) at the mid-time of the reaches, because this is the most undefined portion of the movement without any movements requirements concerning, e.g., reaching height, (iii) in total over the full trajectory. With these ideas in mind, we propose three groups of features:

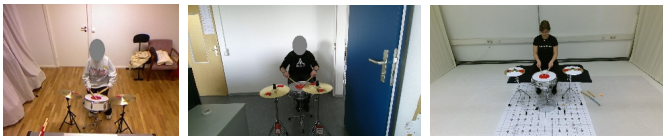
*1) Consistency with predicted healthy trajectories:* We measure how consistently our reference model can predict the impaired trajectories for every set from the spatial parameters $\mathbf{Y}_{PG,HG}$. Fig. 4c shows this procedure exemplarily for one subject of the PG, Fig. 4d for one subject of the HG. The measured and predicted trajectories are subtracted from each other and on the resulting difference the variance (VAR), mean absolute deviation (MAD), and median absolute deviation (MED) are calculated over the time frames mentioned in (i)-(iii) per joint. These features are summed over $x, y, z$-positions per joint and are normalized with the averaged spatial distance between instruments. We expect the differences between real and predicted trajectories to be higher and more fluctuating for impaired subjects than for healthy ones.

*2) Variability of trajectories:* With similar motivation as above, we calculate the variability features (VAR, MAD, MED) on the measured trajectories directly. This captures the variability within multiple repetitions of the same motion.

*3) Deviation from bell-shaped speed profile:* Flash and Hogan observed in [34] a symmetrical bell-shaped speed profile for reaching actions with a peak velocity in the mid-time of the movement. Chen *et al.* [11] fit a Gaussian curve to the speed profiles of trajectories and measure the fitting error. It turned out that this procedure is not robust enough for our data, as the fit is sometimes not possible for too divergent speed profiles especially in the reversal movements, so we assess the deviation from the bell-shaped speed profile by two methods: We either fit a Gaussian curve and take the peak on the fitted curve, following [11], setting a constant value if this is not successful (e.g., a flat trajectory curves where the center of the Gaussian cannot be estimated). Alternatively, the peak velocity is directly estimated from the trajectory. To construct the feature from this, the peak velocity (either from Gaussian fit or from the direct method) is summed for odd and even reaches over all motion cycles of a set.

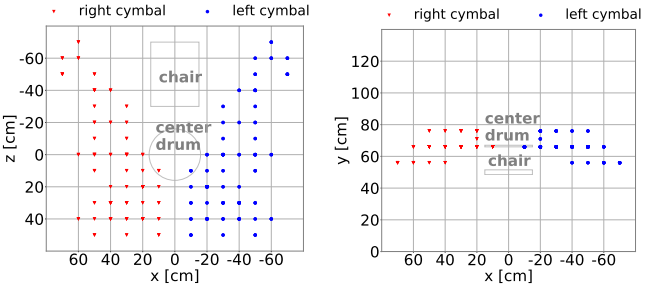
### F. Estimation of Treatment Effect

To estimate a treatment effect, we predict a "healthiness" score from the kinematic features. Instead of correlating features with human ratings to obtain the score, we train a model that estimates the probability $p(y = PG|\mathbf{F})$ of a motion to be "healthy" or "impaired" (i.e., belonging to the PG) given features $\mathbf{F}$ of that motion. Depending on the kind and severity of the disease, we expect that the probabilities of the PG change over the progress of the treatment. This change in probability is used to quantify the treatment effect of a patient: If the probabilities of belonging to the patient group, $p(y = PG|\mathbf{F})$, do not change, we would deduce that treatment had no effect on the motor function. If the probability decreases, we assume positive effects of the treatment on the motor function. An increasing probability hints at a deterioration of motor function, e.g., due to degenerative processes from the disease and/or ineffective treatment.



(a) Patient group (b) Healthy group (c) Setup variation

Fig. 5. Representative images acquired by Kinect sensor in different datasets. Please note the variation in the sensor placement, subject's location and distribution of instrumentation.




(a) Top view (b) Front view

Fig. 6. Recording area in dataset with controlled setup variation (SV), with varying positions of outer cymbals.

For training the model that predicts $p$, we use motion of the HG ($y \neq PG$) and only the first treatment sessions of the PG ($y = PG$). In these first sessions, treatment effect did not yet kick in and motion can definitely be rated as "impaired" due to a verified diagnosis that lead to the therapy in the first place. For prediction of $p$ from $\mathbf{F}$, non-probabilistic methods such as SVMs would be possible, but require additional calibration of probabilities on a separate validation set (cf. [35], [36]), which we lack due to the small size of our datasets. Decision trees are another alternative, but provide accurate probability calculation only for very large datasets. For these reasons, probabilities $p$ are modeled using a linear model with logistic sigmoid function [37].

## IV. Experimental Results

We now demonstrate results of the proposed motion analysis framework in the context of long-term FMT treatment. We first present the captured datasets (IV-A) and evaluate the ability of the reference trajectory model to predict realistic healthy trajectories under varied spatial conditions (IV-B). We then explore the importance of kinematic features and tune the selection of best features (IV-C). Then, we describe the results of our main contributions in Sec. IV-D, where we illustrate the suitability of our model to assess changes in motor function over long-term treatment and analyze the results per subject group and disease. We also test how the model responds to individual features (IV-E) and demonstrate the robustness of our model to substantial setup variations (IV-F).

TABLE I
DEMOGRAPHICS OF PATIENT GROUP AND INFORMATION ABOUT INVOLVED TREATMENT SESSIONS

| Name | Disease | Affected | Time (years) | Treatment duration (weeks) | | | | Drumming speed (hpm) | | | | Therapists' Scores | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | A | B | C | D | A | B | C | D | A | B | C | D |
| S5 | S | Left | 2 | – | 4 | 19 | 20 | – | 97 | 96 | 102 | – | 3 | 1 | 1 |
| S6 | S | Right | 14 | 1 | 3 | 18 | 20 | 168 | 185 | 139 | 112 | 4 | 3 | 1 | 1 |
| S7 | S | Right | 12 | 6 | 7 | 14 | 17 | 124 | 146 | 116 | 140 | 4 | 4 | 1 | 1 |
| S8 | S | Left | 2 | 1 | 2 | 19 | 20 | 73 | 92 | 118 | 118 | 4 | 3 | 0 | 0 |
| S10 | S | Left | 1 | 1 | 2 | 18 | 20 | 66 | 71 | 105 | 113 | 2 | 2 | 0 | 0 |
| S15 | PD | Right | 9 | 1 | 4 | 12 | 20 | 128 | 161 | 156 | 220 | 5 | 5 | 3 | 4 |
| S16 | PD | Left, Right | 10 | 1 | – | – | 20 | 114 | – | – | 134 | 4 | – | – | 2 |
| S18 | PD | Left | 5 | 2 | 5 | 18 | 19 | 99 | 133 | 154 | 150 | 2 | 1 | 1 | 1 |
| S19 | PD | Right | 12 | 1 | 2 | 19 | 20 | 134 | 116 | 92 | 113 | 2 | 3 | 1 | 1 |

Patients with stroke (S) or Parkinson's disease (PD), more affected left or right body side and time since stroke or onset, in sessions A to D, described by treatment duration in weeks, drumming speed in hits per minute (hpm) and therapists' scores for total improvement of bodily functions. The symbol – indicates missing sessions that did not fit the selection criteria described in Sec. IV-A1.

### A. Data Acquisition

We captured 3 datasets to investigate the influence of impairment, number of treatment, and spatial distribution of instrumentation in bi-manual repetitive drumming.

*1) Patient group (PG):* We recorded 20 subjects (of which 11 female) in 5 to 20 FMT sessions of approx. 20 min. length under real-life conditions (cf. Fig. 5a). 10 of these patients were diagnosed with Parkinson's disease (PD), 10 were stroke survivors (S). The motion data was acquired weekly with varying setups of instrumentation. Table I presents an overview of data used in this paper, more patient information is given as in [9].

For the analysis, robustly tracked session parts in a 2-drum-setup, performed in a self-chosen speed, with the desired drumming-pattern and with comparable drum sticks were selected in 2 to 4 sessions (cf. Table I, A to D). Data was excluded, if (i) the patients were repeatedly recorded with unstable skeletal tracking, (ii) did not perform the exercise correctly in the required number of sessions and with the minimal number of ten motion cycles and (iii) the subjects' age was below 18 years. In the selected drumming samples, each hand is alternating between the center drum and the one outer drum, which is located at the corresponding body side (cf. Fig. 9). Two skilled therapists with a working and teaching experience of at least ten years in FMT used our self-written software [38] (based on the Kinect for Windows SDK 1.5 Version) to record the motion data and rated the drumming performance visually on a FMT-specific 5-point scale.

*2) Healthy group (HG):* The motion data of 10 healthy subjects (of which 3 female, 1 left-handed, mean±SD age of 31.4±2.54 years) was acquired by the Kinect for Windows SDK 2.0 under lab conditions (cf. Fig. 5b) in 3 setups of instrumentation with the same drumming pattern as in the PG group in a drumming speed of 122-126 hits per minute (hpm). The recording process was conducted with a self-written software and was initialized and ended by the instructor.

*3) Controlled setup variation (SV):* Data of 1 healthy subject (female, 31 years) was acquired by the the Kinect for Windows SDK 2.0 with controlled variation of the positions of the instruments (cf. Fig. 5c) with the same drumming pattern as in the PG group in a drumming speed of 122 hpm. The position of the chair and the $XZ$-position of the center drum was fixed, the positions of the outer cymbals were changed as displayed in Fig. 6.

TABLE II
MEAN AND STANDARD DEVIATION (SD) IN CM AND EXPLAINED VARIANCE OF REFERENCE MODEL IN DEPENDANCE OF NO. OF PCA COMPONENTS AND SUBJECT GROUP. BEST RESULTS DENOTED IN **BOLD**.

| No. PCA Comp. | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| **Patient Group (PG)** | | | | | | |
| *Left Hand* | | | | | | |
| Mean [cm] | 5.60 | 3.03 | 3.00 | 2.99 | **2.98** | 3.01 |
| SD [cm] | 3.45 | 2.25 | 2.24 | 2.28 | **2.28** | 2.30 |
| *Right Hand* | | | | | | |
| Mean [cm] | 4.47 | 3.04 | 3.04 | **3.00** | 3.01 | 3.01 |
| SD [cm] | 2.89 | 2.36 | 2.35 | 2.35 | **2.34** | 2.34 |
| **Healthy Group (HG)** | | | | | | |
| *Left Hand* | | | | | | |
| Mean [cm] | 2.90 | 2.20 | **2.20** | 2.21 | 2.21 | 2.20 |
| SD [cm] | 2.23 | 1.91 | **1.91** | 1.94 | 1.94 | 1.94 |
| *Right Hand* | | | | | | |
| Mean [cm] | 2.58 | 2.11 | **2.10** | 2.10 | 2.11 | 2.11 |
| SD [cm] | 1.96 | 1.82 | **1.81** | 1.82 | 1.83 | 1.83 |
| **Explained Variance** | | | | | | |
| Left Hand [%] | 0.966 | 0.985 | **0.990** | 0.992 | 0.994 | 0.996 |
| Right Hand [%] | 0.966 | 0.984 | **0.990** | 0.993 | 0.995 | 0.996 |

### B. Validation of Reference Trajectory Model

We can use the reference model (Sec. III-D) to predict a reference trajectory for real trajectories of impaired or healthy subjects. In order to measure the precision of this model we calculated the average Euclidean distance and standard deviation (SD) in $\mathbb{R}^3$ over all trajectories of the PG and HG and the explained variance ratio depending on the numbers of used PCA components and separately for joints of the left and right body side. Table II reveals the results.

As expected, the healthy trajectories are closer to the reference trajectories (lower average distance and SD): while

healthy drumming contains style variations, impaired drumming apparently contains aberrations from the healthy reference due to the disease. Differences between left and right hand are negligible in PG, presumably because the body side affected by the disease is balanced in our dataset, cf. Table I. With increasing number of PCA components, the model overfits to peculiarities of the training set (here: SV), which results in higher error of the model on the HG and PG dataset. Therefore, we select $K = 4$ PCA components.

### C. Feature Tuning and Selection

We further optimize the probabilistic model that computes the treatment effect (cf. III-F) by performing leave-one-subject-out cross validation (LOOCV) using Scikit-learn [39]. This simulates performance of the model on a subject that was not used for training, which is repeated and averaged over all subjects to obtain an expected model accuracy. We first tested different variants per feature group (VAR vs MAD vs MED, cf. III-E): MAD and MED achieved the best accuracy of $0.88$. Speed profile features worked best when using peak velocity (III-E3) instead of using the Gaussian fit. Each of our three kinematic features actually provide several sub-features (e.g., deviation of speed at forward and backward reach, cf. III-E). To further reduce the number of features, we systematically select one sub-feature in each of the three feature groups. The variability around the end of reaches and the deviation from the speed profile in the even (reversal) reaches contributed the most per group. The importance of variability around the end of reaches coincides with findings of [33] for stroke survivors. The feature tuning and selection process leaves us with an optimal feature set of two scalars per feature (= 6 features in total) that we collect in $\mathbf{F}^{best}$.

### D. Analysis of Treatment Effect

The classification model from Sec. IV-C, which was trained on the best sub-feature combination $\mathbf{F}^{best}$ of both body sides from early treatment sessions, was used to predict probabilities for later treatment sessions for the patient group (PG). Please note that the classification is based on noisy class labels, because we used no information about the impaired body side of the patients in the training procedure and conducted no medical assessment of the healthy group (HG) about the quality of motor function.

We compare these model predictions with therapists' ratings. Two experienced FMT therapists described the total coordination of subjects, including motor function of both hands, with a score from 0-"no disability" to 5-"severe disability", for the *full* treatment session. The focus of the human evaluation can be followed in [8], the ratings are displayed in Table I. Fig. 7 shows the probabilities $p(y = PG|\mathbf{F}^{best})$ of subjects to be labeled as belonging to the disabled PG over the duration of treatment and the corresponding ratings of the therapists. Thus, each data point per subject represents a treatment session. The therapists' ratings and model probabilities mostly match and show similar trends. This is remarkable since the model was never trained/calibrated on the therapists' ratings and, on top of that, the therapists rated the full treatment session (about 20 minutes) while our method only analyzes a single exercise of that same session (20 reaches during at most 20 seconds of motion). The therapists rated the total coordination of all S patients as improved. In agreement, we see that the probabilities decrease over the duration of treatment and seem to converge to the range of the HG (cf. Fig. 7a), which could be a signal of recovery. In PD (cf. Fig. 7b), except for subject 16, a converse development in the probabilities can be observed, suggesting a deterioration of motor function over the treatment. This may correspond to the usual course of PD. Additionally, the model does not fully agree with the therapists, who saw a small improvement. A reason could be that the model looks at kinematic features on the hand joints while a therapist considers additional criteria (e.g., total coordination, breathing, cf. [8]), thus, depending on the disease, model and humans inherently focused on different functions. In both groups, some probabilities fluctuate per subject, which could be a consequence of, e.g., daily condition, fatigue, or medication. The HG was not rated by the therapists, but the probabilities are clearly smaller than $0.5$ (cf. Fig. 7c), hence would be labeled correctly as "healthy". The variability within the HG can be seen as usual phenomenon in an untrained subject group, which perform the desired exercise for the first time.

We now compare if the model and the therapist both detect an improvement after treatment by defining a positive effect when: (i) for human scores when they improve by at least 2, (ii) for classification probabilities when they show an improvement of at least 0.2. Table III shows the results. The therapists and our model agree for S subjects 6, 8, 10 and in all PD subjects. In the majority of PD subjects the therapists assessed only cautious improvements, which were too low to

TABLE III
IMPROVEMENTS OF MOTOR FUNCTION FROM DIFFERENCES BETWEEN 1ST AND LAST SESSION: BOTH THE THERAPIST AND OUR MODEL AGREE IN MOST CASES.

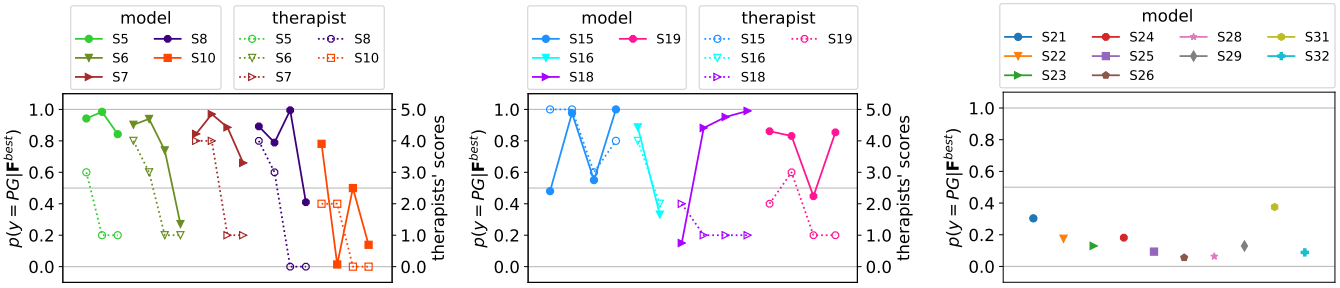
| Name | Disease | Therapists' Scores | | Probabilities | | Both agree? |
|---|---|---|---|---|---|---|
| | | Diff. | Improved | Diff. | Improved | |
| S5 | S | 2.0 | ✓ | 0.10 | ✗ | – |
| S6 | S | 3.0 | ✓ | 0.63 | ✓ | ✓ |
| S7 | S | 3.0 | ✓ | 0.18 | ✗ | – |
| S8 | S | 4.0 | ✓ | 0.48 | ✓ | ✓ |
| S10 | S | 2.0 | ✓ | 0.64 | ✓ | ✓ |
| S15 | PD | 1.0 | ✗ | -0.52 | ✗ | ✓ |
| S16 | PD | 2.0 | ✓ | 0.55 | ✓ | ✓ |
| S18 | PD | 1.0 | ✗ | -0.84 | ✗ | ✓ |
| S19 | PD | 1.0 | ✗ | 0.01 | ✗ | ✓ |

TABLE IV
ACCURACY OF CLASSIFICATION MODEL DEPENDING ON SELECTED BEST SUB-FEATURES

| Speed | Consistency w. Model | Variability | Accuracy |
|---|---|---|---|
| ✓ | ✓ | ✓ | 0.88 |
| ✓ | | | 0.85 |
| | ✓ | | 0.77 |
| | | ✓ | 0.77 |

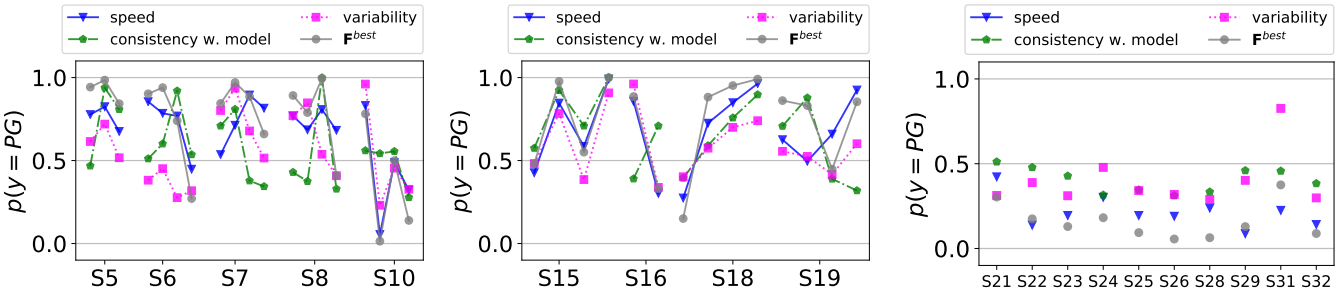
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2018.2848103, IEEE Journal of Biomedical and Health Informatics






(a) Stroke survivors (b) Subjects with Parkinson's disease (c) Healthy subjects

Fig. 7. Estimation of treatment effect per patient group: Probabilities of subjects of being labeled as belonging to disabled patient group (PG) over the duration of treatment (solid) and corresponding ratings of the therapists (dotted). For better readability, the time between sessions is uniformly scaled.




(a) Stroke survivors (b) Subjects with Parkinson's disease (c) Healthy subjects

Fig. 8. Estimation of treatment effect for different aspects of motor function: Probabilities of subjects of being labeled as belonging to disabled patient group (PG) over the duration of treatment, whereas model was trained with single features. "speed" stands for the speed profile in the even (reversal) reaches, "consistency w. model" and "variability" for the variability around the end of reaches both for predicted and real trajectories, $\mathbf{F}^{best}$ for the best sub-features. For better readability, the time between sessions is uniformly scaled.

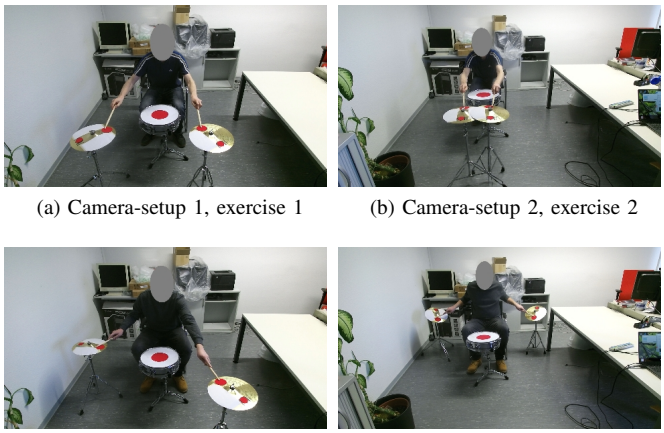
count for a positive treatment effect and leads to this high agreement. In the S group, the motor function of subjects 5 and 7 were not evaluated as "improved" by our model as the differences in probabilities are not high enough, but a tendency is clearly there.

### *E. Quantification of Different Aspects of Motor Function*

The model evaluated above combines all three kinematic features to produce a single output $p(y = PG|\mathbf{F}^{best})$. We can also train the model on just a single kinematic feature in order to assess if that feature responds to different aspects of a patient's motor function, which might provide further insights for a therapist. Table IV shows classification accuracies obtained for models trained on a single feature (again, computed from LOOCV). The features are informative on their own, especially "speed", but cannot achieve accuracy of a model trained on all three features. More importantly, models trained on singular features respond differently to different patients, which is what is visualized in Fig. 8. For example, S8 is able to quite precisely reproduce motion (also illustrated in Fig. 4c, top row) in session C after 19 weeks of treatment, which is depicted in low probabilities for the feature describing the variability around the ends of real reaches ("variability" in Fig. 8, 3rd data point belonging to S8). However, the higher probabilities concerning the variability around the ends of the predicted reaches ("consistency w. model" in Fig. 8) indicate, that the executed motion still differs substantially from the predicted motion of a healthy subject (cf. Fig. 4c, middle and bottom row), thus is repeatedly performed in a not optimal, "unhealthy" way. In summary, while sub-features might reveal such different syndromes in individual patients, only a combination of all features characterizes treatment effect robustly.

### *F. Validation of Model Invariance to Setup Changes*

To ensure that our model is invariant to setup changes, we recorded two additional healthy subjects in two substantially different camera setups while they are performing 4 different drumming exercises (cf. Fig. 9). Subjects as well as camera setups were not part of any training set. Fig. 10 shows the probabilities of subjects to be labeled as belonging to the disabled PG in both camera-setups. The probabilities remain unaffected in both setups, thus were not influenced by the variation of the camera-setup. Both subjects are correctly classified as "healthy" with probabilities that are clearly smaller than 0.5. Although the exercises were also changed and demand different motion strategies from the subjects, the probabilities stay stable. This indicates a versed, well adapted motor coordination in the healthy subjects and harmonizes with the low inter-subject variability in healthy subjects in the HG as depicted in Fig. 4b.



(a) Camera-setup 1, exercise 1 (b) Camera-setup 2, exercise 2

(c) Camera-setup 1, exercise 3 (d) Camera-setup 2, exercise 4

Fig. 9. Representative images acquired by Kinect sensor in dataset with subjects 33 (top row) and 34 (bottom row) drumming in 2 camera-setups with 4 exercises each.

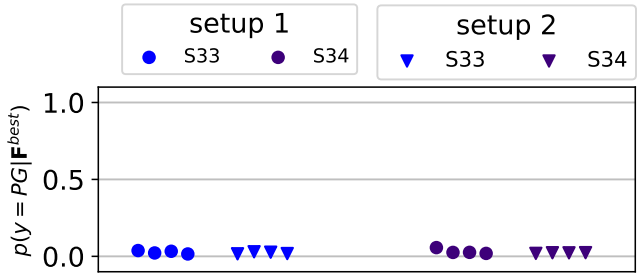



Fig. 10. Estimation of treatment effect depending on setup changes: Probabilities of subjects of being labeled as belonging to disabled patient group (PG) in 2 different camera setups with 4 different exercises each. The similar probabilities indicate that our processing and reference trajectory model successfully eliminate the influence of setup variations on motion.

## V. Discussion and Conclusion

In this paper, we address the problem of assessing motor function in long-term rehabilitation with varied spatial setups and without using subjective therapists' scores. We describe a framework to process, normalize, and compare unrestrained trajectories located in 3d and recorded from uncontrolled conditions. Clinical studies usually implicate strongly controlled and thus restricted circumstances with extensive technical assistance and expensive equipment. On the contrary, our method allows low-cost data collection, analysis of real-life motion recordings, and is more robust against user negligence.

We propose and evaluate a reference trajectory model to predict healthy hand trajectories from variable 3d joint positions of variable given trajectories. We demonstrate the robustness of our model to substantial setup variations. So it allows the comparison of movements which were recorded from uncontrolled conditions, as they often occur in long-term, real-life treatment. By considering left and right joints separately, it is also suitable for trajectory synthesis in case of diseases that affect whole body motion, without a necessary determination of, e.g., an affected hand or body side. Similar to other trajectory models, ours can only learn and predict variances contained in the healthy, controlled training data, namely varied motion target locations. Currently it lacks the ability to predict speed-relevant features like overall movement duration in dependance on different given or self-paced speeds, peak speed, and similar attributes. Decreased speed in relation to a healthy control group [33] is an interesting attribute of impairment and would be helpful to monitor in the trajectory model. Given sufficient data, our framework also allows learning effects like these via the healthy reference model.

Currently, we account for hand motion only, both within the trajectory model and for the kinematic features. The wrist and hand joints prepare and guide the movement in drumming [40] as in reaching. However, shoulder, elbow, and torso joints also contribute to reaching tasks [7], with increased trunk displacement for stroke patients in comparison to healthy controls as found by [33]. Hence, we plan the integration of additional joints into the framework as future work.

We motivated and evaluated the influence of different kinematic features to serve as indicators for impairment over long-term treatment. We could confirm that end-point positions are more widely distributed for impaired subjects, as stated by [33]. Additionally, larger deviations from a symmetrically bell-shaped velocity profile in forward and especially reversal movements were pronounced in the disabled group. A shifting of the peak velocity outside of the mid-time between start and end point of movement is usually associated with requirements for speed and accuracy as described by Fitts's law [41]: The peak velocity occurs earlier if the movement needs to be performed very accurately, e.g., towards a small target size, and it occurs later if the subject has to move very fast. We assume that the requirements for accuracy are higher for disabled subjects, also with a comparable target size in a healthy and patient group. In FMT, the therapists continually urge the patients to carry out new drum movements that go beyond their current motor skills. Due to their impairment they are more challenged to hit the target accurately and adapt their movements strategies accordingly. A second possible reason could be that the subjects were especially rushed in the reversal motion to prepare the next forward motion in time and put more emphasis on the speed of their movements. Third, the requirements and strategies in unrestrained drumming could be generally different for forward and backward motion, e.g., as observed for upward and downward motion by Atkeson and Hollerbach [30]. Future research needs to investigate the relationship between impairment, movement direction, and requirement for accuracy and speed in reaching and repetitive drumming. We will focus on the extension of the amount of kinematic features and application to our patient groups, e.g., hand path length and initial movement direction error as investigated by [13]. We can easily integrate these features as additional indicators for the evaluation of the treatment effect.

We presented a probabilistic model which allows statements about long-term progression of impaired motor function in relation to a healthy control group and compared it to therapists' scores. To the best of our knowledge, this paper is the first to reveal an objective analysis of therapy effect in FMT and for the first time investigates long-term changes in drumming from

real-life treatment sessions. The distribution of the kinematic features per subject group and the accordance between subject-related improvements in model probabilities and human ratings indicate that the proposed framework is appropriate to evaluate motor function of patients after stroke and with Parkinson's disease. Hence, our model can be helpful to assist therapists in the objective assessment of therapy success or encourage changes in treatment if used concomitantly to the therapy. Arguably, the comparison of our quantitative model with the therapists' scores in Sec. IV-D demands further investigation and discussion, because the therapists' ratings are given for the whole treatment session and for total movement, not only hand motion. And, while our model does not use any subjective human scores, this comparison does. In the future, additional methods for objective treatment estimation may be worth implementing. However, we think a basic accordance with human scores of multiple, experienced raters will support clinical appliance of our model. In this study, the sample size of the subject groups and the number of events per variable (EPV) is too small to produce stable estimates of the treatment effect. If, e.g., $4$ out of $9$ patients experience a treatment effect (cf. improved probabilities in Table III), the EPV of the initial model with $16$ features is only $0.25$, but should be $10$ to $15$ EPV following [42].

So, in the future, a more detailed study would be a profit for the validation of the model and FMT in general: (i) with a higher amount of involved patients and sessions, (ii) joint-wise scored motor function by different therapists, (iii) a physical examination of the patients by specialists to compare bodily functions with kinematic features, and (iv) comparative measurements with a more precise movement sensor (e.g., a Vicon system) to evaluate the suitability of the Kinect sensor for tracking reaching and drumming motion and to validate the model predictions. Especially the involvement of more treatment sessions as well as metadata like medication or parallel rehabilitation programs could be helpful to find reasons for strong fluctuations in model probabilities. We hope that motion quantification is increasingly applied in real-life rehabilitation, also outside of clinical studies, and that the methods of this paper drive this development forward in the analysis of drumming and reaching movements.

## ACKNOWLEDGMENT

The authors would like to sincerely thank the subjects and patients for their participation, Margareta Ericsson and Karina Larsson for support and consultations about FMT therapy, Åsa Rosin for provided demographic data of the patient group, Sven Hellbach for methodological advice, and team and student members Mark Schramm, Martin Wolff and Jens Friedrich for technical support and labeling.

## REFERENCES

[1] World Health Organization, "Neurological disorders: Public health challenges," 2006.

[2] D. Webster and O. Celik, "Systematic review of kinect applications in elderly care and stroke rehabilitation." *J. Neuroeng. Rehab.*, vol. 11, 2014.

[3] H. M. Hondori and M. Khademi, "A review on technical and clinical impact of microsoft kinect on physical therapy and rehabilitation," *J. Med. Eng.*, pp. 1–16, 2014.

[4] B. H. Bornstein and A. C. Emler, "Rationality in medical decision making: a review of the literature on doctors decision-making biases," *J. Eval. Clin. Pract.*, vol. 7, no. 2, pp. 97–107, 2001.

[5] E. V. Olesh, S. Yakovenko, and V. Gritsenko, "Automated assessment of upper extremity movement impairment due to stroke," *PLOS ONE*, vol. 9, no. 8, pp. 1–9, 2014.

[6] V. Venkataraman, P. Turaga, M. Baran, N. Lehrer, T. Du, L. Cheng, T. Rikakis, and S. L. Wolf, "Component-level tuning of kinematic features from composite therapist impressions of movement quality," *IEEE J. Biomed. Health Inform.*, vol. 20, no. 1, pp. 143–152, 1 2016.

[7] P. M. McCrea, J. J. Eng, and A. J. Hodgson, "Biomechanics of reaching: Clinical implications for individuals with acquired brain injury," *Disability and Rehab.*, vol. 24, no. 10, pp. 534–541, 2002.

[8] L. Pogrzeba, M. Ericsson, K. Larsson, M. Wacker, and B. Jung, "Towards a vocabulary for describing 3d motion data in functionally oriented music therapy," in *Proc. 13th Int. Conf. Music Percept. Cogn., Proc. 5th Conf. Asia-Pacific Soc. Cogn. Sci. Music*, 2014, pp. 222–228.

[9] Å. Rosin, M. Ericsson, and K. Larsson, "The effects of functionally oriented music therapy on body function and quality of life in chronic stroke survivors and on patients with parkinson's disease," *Music & Medicine*, vol. 7, no. 2, pp. 14–19, 2015.

[10] F. Burget, C. Maurer, W. Burgard, and M. Bennewitz, "Learning motor control parameters for motion strategy analysis of parkinsons disease patients," in *IROS Conf. Proc.*, 2015, pp. 5019–5025.

[11] Y. Chen, M. Duff, N. Lehrer, H. Sundaram, J. He, S. L. Wolf, and T. Rikakis, "A computational framework for quantitative evaluation of movement during rehabilitation," in *AIP Conf. Proc.*, 2011, pp. 317–326.

[12] S. L. Wolf, P. A. Catlin, M. Ellis, A. L. Archer, B. Morgan, and A. Piacentino, "Assessing wolf motor function test as outcome measure for research in patients after stroke," *Stroke*, vol. 32, no. 7, pp. 1635–1639, Jul. 2001.

[13] A. M. Coderre, A. A. Zeid, S. P. Dukelow, M. J. Demmer, K. D. Moore, M. J. Demers, H. Bretzke, T. M. Herter, J. I. Glasgow, K. E. Norman, S. D. Bagg, and S. H. Scott, "Assessment of upper-limb sensorimotor function of subacute stroke patients using visually guided reaching," *Neurorehab. and Neural Repair*, vol. 24, no. 6, pp. 528–541, 2010.

[14] M. Baran, N. Lehrer, D. Siwiak, Y. Chen, M. Duff, T. Ingalls, and T. Rikakis, "Design of a home-based adaptive mixed reality rehabilitation system for stroke survivors," in *IEEE EMBS Conf. Proc.*, 2011, pp. 7602–7605.

[15] R. Luo and D. Berenson, "A framework for unsupervised online human reaching motion recognition and early prediction," in *IEEE/RSJ Intern. Conf. on Intell. Robots and Systems (IROS)*, 2015, pp. 2426–2433.

[16] J. J. Faraway, M. P. Reed, and J. Wang, "Modeling 3d trajectories using bezier curves with application to hand motion," *Applied Statistics*, vol. 56, no. 1, pp. 571–585, 2007.

[17] 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 *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, Aug 2011, pp. 6789–6792.

[18] Y. Li, C. Fermuller, Y. Aloimonos, and H. Ji, "Learning shift-invariant sparse representation of actions," in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Jun. 2010, pp. 2630–2637.

[19] R. A. Clark, Y.-H. Pua, K. Fortin, C. Ritchie, K. E. Webster, L. Denehy, and A. L. Bryant, "Validity of the microsoft kinect for assessment of postural control," *Gait & Posture*, vol. 36, no. 3, pp. 372–377, 2012.

[20] Q. Wang, G. Kurillo, F. Ofli, and R. Bajcsy, "Evaluation of pose tracking accuracy in the first and second generations of microsoft kinect," in *IEEE Intern. Conf. on Healthcare Inform.*, 2015, pp. 380–389.

[21] K. Otte, B. Kayser, S. Mansow-Model, J. Verrel, F. Paul, A. U. Brandt, and T. Schmitz-Hübsch, "Accuracy and reliability of the kinect version 2 for clinical measurement of motor function," *PLOS ONE*, vol. 11, no. 11, pp. 1–17, 2016.

[22] L. Pogrzeba, H. Orth, B. Maurer-Burkhard, M. Wacker, and B. Jung, "Application of computer-assisted neutral zero method of the shoulder joint in vojta therapy," *J. Rehab. Med.*, vol. 47, no. suppl. 54, p. 391, 2015.

[23] M. Eltoukhy, C. Kuenze, J. Oh, and J. Signorile, "Validation of static and dynamic balance assessment using microsoft kinect for young and elderly populations," *IEEE J. Biomed. Health. Inform.*, 2017, advance online publication. doi:10.1109/JBHI.2017.2686330.

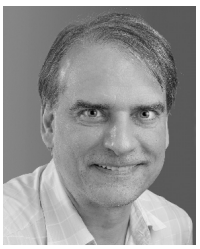
[24] R. J. Adams, M. D. Lichter, E. T. Krepkovich, A. Ellington, M. White, and P. T. Diamond, "Assessing upper extremity motor function in

practice of virtual activities of daily living," *IEEE Transactions on Neural Systems and Rehab. Eng.*, vol. 23, no. 2, pp. 287–296, 2015.
[25] D. Leightley, M. H. Yap, and J. McPhee, "Automated analysis and quantification of human mobility using a depth sensor," *IEEE J. Biomed. Health Inform.*, 2016, advance online publication. doi:10.1109/JBHI.2016.2558540.
[26] S. M. Mostafavi, P. Mousavi, S. P. Dukelow, and S. H. Scott, "Robot-based assessment of motor and proprioceptive function identifies biomarkers for prediction of functional independence measures," *J. NeuroEng. Rehab.*, vol. 12, no. 1, p. 105, 2015.
[27] A. Som, R. Anirudh, Q. Wang, and P. Turaga, "Riemannian geometric approaches for measuring movement quality," in *Proc. IEEE Conf. on Comp. Vis. Pattern Recogn. Workshops*, 2016, pp. 43–50.
[28] N. F. Troje, "Decomposing biological motion: a framework for analysis and synthesis of human gait patterns," *J. Vis.*, vol. 2, no. 5, pp. 371–387, 2002.
[29] M. Azimi, "Skeletal joint smoothing white paper," 2012, accessed: 2016-06-07. [Online]. Available: https:////msdn.microsoft.com/en-us/library/jj131429.aspx
[30] C. Atkeson and J. Hollerbach, "Kinematic features of unrestrained vertical arm movements," *J. Neurosci.*, vol. 5, no. 9, pp. 2318–2330, 1985.
[31] M. D. K. Breteler, R. G. J. Meulenbroek, and S. C. A. M. Gielen, "Geometric features of workspace and joint-space paths of 3d reaching movements," *Acta Psychol.*, vol. 100, pp. 37–53, 1998.
[32] M. Woodbury, C. A. Velozo, P. A. Thompson, K. Light, G. Uswatte, E. Taub, C. J. Winstein, D. Morris, S. Blanton, D. S. Nichols-Larsen, and S. L. Wolf, "Measurement structure of the wolf motor function test: Implications for motor control theory," *Neurorehab. and Neural Repair*, vol. 24, no. 9, pp. 791–801, 2010.
[33] M. C. Cirstea and M. F. Levin, "Compensatory strategies for reaching in stroke," *Brain*, vol. 123, no. 5, pp. 940–953, 2000.
[34] T. Flash and N. Hogan, "The coordination of arm movements: An experimentally confirmed mathematical model," *J. Neurosci.*, vol. 5, no. 7, pp. 1688–1703, 1985.
[35] J. C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," *Advances in large margin classifiers*, vol. 10, no. 3, pp. 61–74, 1999.
[36] A. Niculescu-Mizil and R. Caruana, "Predicting good probabilities with supervised learning," in *Proc. 22nd Intern. Conf. on Machine Learn.*, 2005, pp. 625–632.
[37] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer-Verlag, 2006, p. 139.
[38] L. Pogrzeba, M. Wacker, and B. Jung, "Potentials of a low-cost motion analysis system for exergames in rehabilitation and sports medicine," in *E-Learn. and Games for Train., Educ., Health and Sports*, ser. LNCS, vol. 7516. Springer, 2012, pp. 125–133.
[39] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Machine Learn. Research*, vol. 12, pp. 2825–2830, 2011.
[40] S. Dahl, "Striking movements: A survey of motion analysis of percussionists," *Acoust. Sci. & Tech.*, vol. 32, no. 5, pp. 168–173, 2011.
[41] P. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *J. Experim. Psych.*, vol. 47, pp. 381–391, 1954.
[42] P. Peduzzi, J. Concato, E. Kemper, T. R. Holford, and A. R. Feinstein, "A simulation study of the number of events per variable in logistic regression analysis," *J. Clin. Epidemiol.*, vol. 49, no. 12, pp. 1373–1379, 1996.

**Loreen Pogrzeba** received the diploma degree in media computer science from the University of Applied Sciences (HTW) Dresden, Germany, in 2009. She is currently working towards the Ph.D. degree at HTW Dresden and University of Technology and Mining in Freiberg, Germany. Her research interests include motion analysis and quantification in rehabilitation and sports, as well as the development and design of user-friendly tools for computer-aided motion assessments in real-life conditions.

**Thomas Neumann** is a postdoctoral researcher and leads the junior research group "TISRA" at Hochschule für Technik und Wirtschaft (HTW) Dresden. He studied media and computer science at HTW Dresden and received his PhD in 2016 from the Technical University of Braunschweig in collaboration with HTW Dresden and Max-Planck-Institut Saarbrücken with a thesis on "Reconstruction, Analysis, and Editing of dynamically deforming 3D-Surfaces". His research interests concern visual computing and machine learning with a focus on image based 3D reconstruction, geometry processing, statistical shape modelling, data-driven animation, and sparse modeling.

**Markus Wacker** is a Professor of computer graphics at the University of Applied Sciences in Dresden, Germany, since 2004. His research covers several applications of computer science and interdisciplinary projects, which are bundled in the Dresden Matrix Computer Graphics (DREMATRIX) team since 2010. His main research areas are motion capture and analysis, media stations for museums and visitor interaction, and digital reconstruction of historical buildings. In the field of motion capture and analysis he is mainly interested in markerless motion capture and decision supporting tools for therapists and physicians.



**Bernhard Jung** studied computer science and computer linguistics at the University of Stuttgart, Germany, and the University of Missouri, Saint Louis. He received his doctorate degree from the University of Bielefeld with a thesis on dynamic knowledge representation as well as a Habilitation degree for a thesis on intelligent virtual environments. From 2003 to 2005 he was full professor for Media Informatics at University of Lübeck's International School of New Media. Since October 2005 Bernhard Jung chairs the Virtual Reality and Multimedia group of the University of Technology and Mining in Freiberg, Germany. Prof. Jung's research interests are in the fields of Virtual Reality, Large Data Visualization, Human-Computer Interaction, and Advanced Robotics.