Dynamics-based estimation of Parkinson’s disease severity using Gaussian Processes

Satoshi Endo * Franz M. J. Pfister **, *** Jakob Fröhner *
Urban Fietzek ** Daniel Pichler **, **** Kian Abedinpour **, *****
Terry T. Um † Dana Kulić † Muriel Lang * Sandra Hirche *

* Chair of Information-oriented Control, Technical University of Munich, Munich, 80333 Germany (e-mail: {s.endo,muriel.lang,jakob.froehner,hirche}@tum.de)
** Department of neurology and clinical neurophysiology, Schön Klinik München Schwabing, Munich, Germany (e-mail: urban.fietzek@schoen-kliniken.de.)
*** Faculty of Mathematics, Informatics and Statistics, Ludwig-Maximilians-Universität München, Munich, Germany (e-mail: fmj.pfister@me.com)
**** Department of Neurology, School of Medicine, Technical University of Munich, Munich, Germany (e-mail: daniel.pichler@tum.de, kabedinpour@t-mobile.de)
† Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada (e-mail: terry.t.um@gmail.com, dana.kulic@uwaterloo.ca)

Abstract: The innovation potential of wearable technologies is very high, as continuous and non-invasive estimates of clinical symptoms during daily activities could provide valuable support for healthcare management. We propose a data-driven approach for monitoring patients with Parkinson’s disease (PD) using a wearable motion tracking device which predicts PD movement abnormalities and their severities during daily living. We consider severity evolution of PD symptom subsets, namely bradykinesia and dyskinesia as a dynamical system described by a set of movement features and the history of the motor state. Inside the dynamical system, we model the relationship between the severity of the symptom state and the measured sensor data using Gaussian processes (GPs). Our results show estimation of the PD severity was improved when the GP was modelled with our dynamical system formulation. Furthermore, the time-series characteristics of the GP estimates resemble the dynamics of the PD state fluctuations, resultantly leading to a stable and robust performance of the GP for this application.

Keywords: Dynamical systems, Gaussian processes, Inertial sensors, Medical applications

1. INTRODUCTION

Advancements of low-cost devices such as smartwatches have enabled continuous activity monitoring of patients for various health applications in the home environment. In order to reliably interpret the sensor readings, however, these technologies must be able to cope with noisy measurements arising from unscripted daily activities and environments. Here, we propose a novel control-oriented machine learning approach suitable for mobile health applications using Gaussian Processes (GPs) with a novel dynamical system formulation. As an example, we describe how movement abnormalities and their severities in people with Parkinson’s disease (PD) can be robustly estimated from inertial measurements of arm movements observed in daily activities.

PD is a neurodegenerative disorder with several characteristic movement abnormalities, such as bradykinesia and dyskinesia. Bradykinesia represents a general reduction in speed of the voluntary motion due to dopamine deficiency caused by PD, while dyskinesia is characterised by involuntary muscle movements induced by dopaminergic treatments of PD. Although PD treatments are designed to keep the patients at a balanced state between bradykinesia and dyskinesia, manifestation of movement abnormalities widely fluctuate within and between patients (Jankovic, 2005), partly due to a short-term efficacy of the medication and external factors such as nutrition (Seidl et al., 2014). Thus, healthcare management of PD patients can greatly benefit from continuous monitoring of the disease for optimising their treatment plan (Maetzler et al., 2016). Typically, wearable sensors for monitoring PD symptoms involve inertial sensors worn on the body (Um et al., 2017) with some variations of additional sensors such as electromyography (Cole et al., 2014). Performance matrices such as speed of motion are often calculated from

* This research is supported by the ERC Starting Grant ‘Control based on Human Models (con-humo)’ under grant agreement n°337654.
these measurements while the users perform standardised tests or scripted activities (e.g., Butt et al., 2017). In contrast, few studies have adopted ambient monitoring of non-specific motions, due to the presence of a greater number of uncontrollable variables. By large, these studies use supervised machine learning techniques which model the relationship between motion of PD patients and associated PD annotations (i.e., types and severities of PD symptoms) performed by experts in movement disorders (Um et al., 2017). However, estimation of the disease purely based on motion profiles remains a great challenge as uncontrollable factors in the measurements directly interfere with the results in these static modelling approaches. In contrast, there have been efforts to suppress transient measurement errors by explicitly learning longitudinal symptomatic fluctuations. Cole et al. (2014), for instance, explored several dynamical machine-learning frameworks including dynamic neural networks, dynamic support vector machines, and hidden Markov models, and demonstrated much improved prediction performance of the machine learning models when the dynamics of the symptom fluctuation was explicitly incorporated into these models. As an alternative, we propose to use GPs as they are well suited for modelling human movement behaviour due to their ability to generate smooth motion predictions for nonlinear dynamics. In addition, GPs can quantify the uncertainty of the state estimate introduced by the input noise and distance to the training data. In the present study, therefore, we propose a novel approach in formulating GP regression as a dynamical system to illustrate how the movement abnormality exhibited by PD patients can be reliably estimated from a low-cost inertia sensor in a smart watch. Although there are other canonical symptoms of PD such as speech disorders, as the study is designed to estimate the PD symptoms from measurements from an inertia sensor, we focus on severity estimation in two classes of movement symptoms (PD class); bradykinesia and dyskinesia.

2. MODELLING

2.1 GP in the dynamical framework

Given the smooth symptom transition over time (Jankovic, 2005), we assume the time-series of severity level in each PD class form a dynamical system, in which the next estimation depends on the current severity \(\xi_k\) with an unknown disturbance \(w_k\) such as medication intake. Here, the evolution is modelled as

\[
\xi_{k+1} = f(\xi_k, w_k),
\]

\[
\xi_0 \in \{0, \ldots, n\},
\]

\[
\lambda_k = g(\xi_k, w_k),
\]

where the functions \(f\) and \(g\) are unknown but smooth. The drift term \(f\) describes how the current severity state and the disturbance \(w_k\) influence the subsequent severity estimate. The measurable system output \(\lambda_k\) at time step \(k \geq 0\) depends on the discrete severity state \(\xi_k\) defined over \(n+1\) severity levels, and noise \(w_k\) which includes the sensor noise and concurrent motion of a daily activity. We have access to a set of kinematic metrics \(\lambda_k\) from the accelerometer and gyroscope (see Sec. 2.3 for details), but not the internal severity \(\xi_k\) in (1). Thus, we infer this state using the available kinematic metrics \(Z_k\) at time instance \(k \geq 0\), such that

\[
Z_k = \{\lambda_k, Z_{k-1}\},
\]

\[
Z_0 = \{\lambda_0\}.
\]

A state observer estimates the current severity level \(\xi_k\), given the available kinematic metrics \(Z_k\),

\[
\hat{\xi}_k = \mathbb{E}[\xi(\xi_{k-1})|Z_k].
\]

Due to the Markovian property of the underlying dynamics (1), the corresponding estimator has the structure

\[
\hat{\xi}_k = \mathbb{E}[\varphi(\xi_{k-1}, \lambda)],
\]

where \(\hat{\xi}_{k-1}\) contributes to the model-based prediction step and \(\lambda_k\) to the measurement based innovation. We then use a GP to estimate the function \(\varphi\) in (4), given the realization sets \(\{\xi_k\}_{k=0}^n\) and \(\{\lambda_k\}_{k=1}^n\). The training input \(\{x_k\}_{k=1}^n\) to the GP consists of the previous PD severity \(\{\xi_k\}_{k=0}^{n-1}\) and \(\{\lambda_k\}_{k=1}^n\) appended to the feature vector,

\[
x_k = (\lambda_k, \xi_k-1)^\top
\]

for training.

The process output describes a Gaussian distribution, which has the GP mean prediction as expected value \(\mathbb{E}[\cdot]\).

Thus, we consider the GP mean predictions as a process output \(y_k\), given the mean describes an estimate for the k-th severity rating, \(\xi_k\). In our approach, it comprises two GP estimations, one for dyskinesia class (the dyskinesia GP) and the other for bradykinesia class (the bradykinesia GP). Here, these GP classes are treated independently as they have different aetiology and clinical scales. Both PD classes are trained to approximate the severity of the respective class with its output

\[
y_{DK,k} \approx \lfloor \xi_k \rceil \quad \text{if } k \in J_{DK},\]

\[
0 \quad \text{else},
\]

\[
y_{BK,k} \approx \lfloor \xi_k \rceil \quad \text{if } k \in J_{BK},\]

\[
0 \quad \text{else},
\]

where the index sets \(J_{DK}\) and \(J_{BK}\) are defined as \(\{j \in \{1, \ldots, \nu\} \land \xi_j \in DK\}\) and \(\{j \in \{1, \ldots, \nu\} \land \xi_j \in BK\}\), respectively. The decision among the dyskinesia, bradykinesia or balanced (i.e., neither) classes is made based on the results from these two: when both GP models provide severity estimates \(y_{DK,k} < \tilde{c}\) and \(y_{BK,k} < \tilde{c}\) below a certain threshold \(\tilde{c}\), we consider the balanced state is the classification result, and this threshold is set as \(\tilde{c} \leq 0.5\) in the present study. Otherwise, the remaining PD class with higher severity prediction is selected.

Following the model training, we provide state estimates for unseen input data. As the available measurements only consist of the kinematic feature vector \(x_k\), \(k > 0\), we initialize the internal PD state \(\hat{\xi}_0\) by generating a uniformly drawn random integer from \([0, 4]\). Thus, we obtain a GP input as

\[
x_k^* = (\lambda_k^*, \xi_k^*-1)^\top
\]

for testing, where the previously predicted output \(y_{k-1}^*\) is used for the state estimate \(\hat{\xi}_{k-1}^* = [y_{k-1}^*]\). The function \([\cdot]\) rounds the GP mean prediction \(y_k^* \in \mathbb{R}\), which is obtained in continuous space, to the nearest integer, and in the unlikely case of predictions outside the interval \([-0.5, 4.5]\), maps the negative and positive values to 0 and 4, respectively.
2.2 PD data set

The PD data set was collected from thirty individuals with PD (20 male). The average age of the participants was 67±10 years old. The participants wore the smartwatch (Microsoft Band 2) on the wrist of the most affected side, and they were free to engage in any daily routine during recording. On average, the data was collected for 331.2±192.6 minutes per participant, and 9937 minutes in total. The linear acceleration and angular velocity of the wrist were measured by a 6-axis gyroscope/accelerometer module inside the watch at approx. 62.5 Hz. Annotations of PD class and severity were performed by a trained expert (D.P.) who passively monitored the participants every minute during the entire data collection period. The bradykinesia annotation was performed with a standard 5-level rating scale with the MDS-UPDRS (Goetz et al., 2008) item III.14. Dyskinesia was assessed using the abnormal involuntary movement scale (AIMS, item A2.5). In both assessments, the balanced state without any abnormal movement is annotated as 0, and the severity level of each PD class corresponds to 1 to 4. This study was approved by the local ethical board.

2.3 Data processing

A 124 dimensional kinematic feature vector was prepared from the scalar accelerometer and gyroscopic measurements, their derivatives, and wavelet decompositions with Daubechies wavelets (9 levels). Features of the time-series data include standard deviation, norm, maximum, root mean square, kurtosis and skewness per one-minute window. Furthermore, the maximum spectral power and average magnitude around it were added as they are indicative measures of PD (Griffiths et al., 2012). Absence of motion is quantified by calculating the time for which the signal magnitude was small. Multiple absolute thresholds was introduced to compensate for inter-patient and inter-activity variability and to cover all severity levels of the PD classes; 0.1, 0.15, 0.2, 0.25 and 0.3 $G$ for the scalar accelerometer and 1, 1.25, 1.5, 1.75 and 2 $dps$ for the scalar gyroscope. As severe bradykinesia could result in immobility during the whole one minute window or longer, morbidity over a 5-minute window is additionally calculated.

2.4 Modelling and testing procedure

We split the patient cohort, consisting of 30 participants, into two disjoint sets; one for training the model and the other for evaluating the model. Specifically, we perform a leave-one-out approach, where we repeat the training and testing procedure 30 times. In each of the independent runs, the test group consists of data from one participant and the training is performed on the remaining data from the 29 participants. The hyperparameters are chosen so the model accuracy of the training set is balanced over occurrences of false positive/negative predictions across the training sets. Finding suitable GP model hyperparameters that represent the PD class characteristics, not the state frequency, becomes more difficult when the states are non-uniformly distributed in the training data set. Therefore, we do not train the GP models on the full data set, but on data subsets that only comprise the balanced data and the data where the respective PD class are present.

3. RESULTS

3.1 Dynamics of PD annotations

The descriptive analysis of the severity labels collected by the clinical expert, shows 46.3% belongs to the balanced class, and the rest are distributed between the bradykinesia class with 29.4% and dyskinesia class with 24.3%. We first study the dynamics of the PD annotations by loosely grouping the participants in terms of the balanced (ON), dyskinetic (DYS), bradykinetic (OFF) and fluctuating (FL) groups in terms of the average and standard deviation of the PD annotations (Fig. 1).

For each group, we analysed the time-series of the PD annotations in terms of autocorrelation and partial autocorrelations to assess the temporal evolution within each PD class (Fig. 2). The mild reduction of the correlation coefficient over a lag of 15 minutes indicate a generally mild transition of the PD severity across all the patient groups. In contrast, the partial autocorrelation coefficients indicate a drop of the coefficients after one minute lag, suggesting relatively simple linear dependence of the PD annotations over time regardless of the dominant types of movement abnormality or fluctuation patterns of particular patients.

![Fig. 1. Examples of the time-series and histograms of the PD annotations of a single participant from each patient group. For simplicity, the severity of bradykinesia is inverted, resulting in 9 severity levels.](image1)

![Fig. 2. (Partial) autocorrelation coefficients of the PD annotations. The dotted lines indicate approximate confidence bounds.](image2)
Fig. 3. (a) An example of GP prediction with and without dynamical system formulation and (b) the resultant (partial) auto correlation coefficients. Shaded area indicates 2 standard deviations of the GP prediction and the severity level of bradykinesia is inverted.

3.2 GP prediction performance

The accuracy reports in all GP models are normalized by the amount of data available in each run, to prevent biased results due to the distortion of repetitions, where the data set for a certain PD class differs strongly in size. In addition, we provide the ±1 accuracy, defining the percentage when the predicted state severity is less than one level off of the severity assessed by the expert rater to account for potential inter-rater variability. Training accuracy of dyskinesia GP and bradykinesia GP in terms of the mean (standard deviation) are 61.94 (6.24)% and 59.99 (2.66)%, respectively. For comparison, we present the results for pure regression where the GP input for both training and testing is the observed PD annotations (Fig. 2). Our preliminary analyses have shown promising results, showing estimation of PD-related movement abnormalities using a low-cost inertia sensor.

We proposed a novel approach for estimating PD-related movement abnormalities using a low-cost inertia sensor. A key property of the approach is the dynamical system formulation employed to model the temporal evolution of the movement abnormality. Our preliminary analyses have produced promising results, showing estimation of PD severity markedly improved when the GP was modelled with our dynamical system approach. As a next step, we will investigate the robustness of our approach, for example, against the sample size of the training set by studying the change in the GP regression accuracy and confidence level. These results will strengthen the evidence for our approach in modelling and predicting PD symptoms during unscripted activities using inertia measurements.

REFERENCES


