Gait, Wrist, and Sensors: Detecting Freezing of Gait in Parkinson’s Disease from Wrist Movement

Sinziana Mazilu  
Wearable Computing Lab  
ETH Zürich  
sinziana.mazilu@ife.ee.ethz.ch

Ulf Blanke  
Wearable Computing Lab  
ETH Zürich  
ulf.blanke@ife.ee.ethz.ch

Gerhard Tröster  
Wearable Computing Lab  
ETH Zürich  
troester@ife.ee.ethz.ch

Abstract—We investigate the correlation between wrist movement and freezing of the gait in Parkinson’s disease. Detecting such freezes allows real-time monitoring to reduce the risk of falls in subjects with Parkinson’s. While most of research focuses on placing inertial sensors on lower limb, i.e., foot, ankle, thigh, lower back, we focus on the wrist as an alternative placement. Commonly worn accessories at the wrist such as watches or wristbands are easier to be accepted and worn by elderly users, in special subjects with motor problems. Experiments on data from 11 subjects show that freezing of gait episodes can be detected using the wrist movements, with a freeze hit-rate of 90% and 83% specificity in a subject-dependent evaluation scheme. This suggests that wrist sensors can be a feasible alternative to the cumbersome placement on the legs.

I. INTRODUCTION

Between 4 and 16.2 million people worldwide are suffering from Parkinson’s disease [1], a neurological degenerative disorder characterized by postural instability, reduced movement, postural rigidity and tremor of the limbs. From these, more than 70% develop freezing of gait (FoG) [2], a symptom characterized by the sudden incapacity of the subject to walk or move the lower limbs, even if she/he wishes so. FoG can last from few seconds up to 1-2 minutes [3], it is the main cause of falls in Parkinson’s disease [4], and has a strong impact of the social life of the patients [3].

Clinical findings [5] suggest that rhythmic auditory cues such as metronome sounds may help subjects with Parkinson’s disease to shorten or overcome the gait freeze and to resume walking. However continuous rhythrical cueing wears off with time, humans are getting used with the sound and are learning to ignore it. Thus it is important to start the rhythmic cueing for a limited period of time of 8 – 10 seconds, during a gait freeze or when the subject has gait difficulties that might lead to freeze. Wearable solutions ([6], [7], [8]) have been proposed to detect the FoG events in real-time and give a rhythmal cue to help subjects resume walking, using data from wearable sensors such accelerometers.

However, proposed wearable assistants use on-body motion sensors to be attached on the lower body of the user, e.g., on thighs, ankles or even on lower back. The acceptance of such on-body electronics by elderly subjects is still an issue for human-computer interaction in healthcare: The weight, bulkiness, and location of on-body sensors [9], in particular for people with motor deficits. Correct mounting of sensors can be challenging for subjects with mobility impairment. Stigmatization is also a barrier in accepting the wearable systems [10], as most technologies are visible on-body and can be observed very easily as different. Further issues are related to privacy, reliability or battery lifetime [11].

The emerging wrist bands or smartwatches, which incorporate sensors such accelerometers, gyroscopes and magnetometers, are promising to be easily accepted by such in-need users. Their design, the common on-body placed to be attached, the computational power and radio connectivity with the mobile phones make them a good candidate to be integrated in healthcare wearable solutions. A FoG wearable assistant composed from sensors integrated in bracelets or smart watches and the personal smartphone seem promising to be accepted by the elderly subjects in need, and easier to wear, as it is already integrated on-body.

Research until now focused on analyzing and detecting FoG from sensors attached on lower body limbs, as FoG happens at the gait level. Studies such as ([12], [6], [8], [13]) showed that by using movement data from ankles, thighs or lower back it is possible to detect gait freeze events. But during walking, humans tend to also move their arms in tandem. Moreover, humans and Parkinson’s disease subjects suffering from FoG in particular tend to not use their arms for other tasks during walking, as they pay attention on their gait and their next step.

This contribution is centered around the question whether arm movements during walking might be correlated with freezing of gait, and thus making it possible to detect FoG from wrist-attached wearable sensors? We make the following contributions:

(1) We investigate for the first time, up to our knowledge, if the wrist movements during walking are correlated with freezing of gait in Parkinson’s disease. We search whether the wrist movement shows typical properties during FoG which are different from the wrist movement during the rest of walking. We compute new features to describe FoG from wrist data. We use data collected from IMUs attached on both wrists of 11 subjects from the CuPiD dataset [14].

(2) We test and discuss the feasibility of detecting FoG using wrist-attached IMUs in both a subject-dependent and -independent evaluation schemes, using the FoG-detection methods based on supervised machine learning as in [15].

In the rest of the paper we survey the related studies (Section II), and present the dataset used in our investigation (Section III). In Section IV we detail and discuss our experiments and findings, and we conclude our work in Section V.
II. RELATED WORK

Several research groups have proposed wearable systems for the detection of FoG episodes which require on-body accelerometers and/or gyroscopes ([13], [16], [6], [7]). One standard feature which is extracted from acceleration signal is the freezing index, defined as the ratio between the power contained in the so-called gait freezing and locomotion frequency bands ([3-8] Hz and [0.5-3] Hz respectively) [12]. This feature is convenient since it requires only FFT-computation. Other feature extraction approaches involve entropy [13] or time-domain and statistical features such as mean, standard deviation, variance, together with FFT-features [8]. However, all the FoG-detection approaches except [13] require that the sensors are attached on the lower limbs, in order to analyze the gait properties. Tripoliti et al. [13] uses data from wrist sensors, but only in combination with data coming from sensors mounted on lower limbs.

A first evidence that freezing in Parkinson’s disease is present also in the upper limb is given in [17], where frequency analysis of wrist movements showed early-occurrences of manual motor blocks in Parkinson’s disease. Moreover, Vercruysse et al. [18] found evidence that upper limb freezing power spectra were broadened, with increased energy in the gait freezing band. Even if there is evidence found that there is freezing at the level of the arm, wrist or fingers in Parkinson’s disease, these two studies don’t look for a correlation between freezing of gait and wrist movement.

Findings in the study of Nieuwboer et al. [19] show that freezing episodes in the upper limb are correlated with the gait freeze episodes. Moreover, they argue that gait freeze may be also elicited by an upper limb task. Following this hypothesis, we make a first attempt to analyze the wrist movement during FoG episodes and compare with the movement of the wrist during walking, including straight line walking, turns, start or stop walking. We capture these movements with on-body attached inertial measurements units, which include an accelerometer, a gyroscope, and a magnetometer. We aim to find specific patterns in the wrist movement during FoG episodes, which are different from wrist movements during the rest of walking, and thus to detect them.

III. DATASET

To analyze whether the hand movements during walking correlate with freezing of gait episodes, we use the inertial measurement data from sensors attached on wrists in the CuPiD multimodal dataset [14]. The CuPiD dataset contains 24 hours of sensing data collected from 18 subjects with Parkinson’s disease which performed different walking protocols in a laboratory setting designed to provoke FoG, which included walking with 360- and 180-degrees turns, walking in straight lines and passing narrow corridors, or walking across the crowded hospital halls [14].

The data collection system contained 9 wearable ETHOS Inertial Measurement Units (IMU) [20] attached on different parts of the body, one electrocardiogram sensor, a galvanic skin response sensor and a near-infrared spectroscopy sensor. The ETHOS IMU sample 3D accelerometer, 3D gyroscope, and 3D magnetometer data at 128 Hz. We use in this work the data collected from the IMUs attached on both wrists of the subjects, as shown in Figure 1.

The walking protocol performed by each subject was video recorded and the IMU data was synchronized with the videos. Offline, two clinicians labeled the freezing of gait episodes and other walking events such as start walking or turns, using stopwatch annotations and videos synchronized with the sensor datastream. Labels were updated by also taking into account sensor data visualizations synchronized with videos. Clinicians considered the moment of arrested gait pattern, i.e., stop in alternating left-right stepping, as start of a FoG episode, and the instant when the patient resumed a regular gait pattern as the end of FoG.

In total, clinicians labeled 182 FoG episodes from 11 out of 18 subjects, with a duration between 0.2 seconds and 98.8 seconds (average: 9 seconds, standard deviation: 15 seconds). The rest of 7 subjects did not encounter any gait freezing event during the in-the-lab protocol.

IV. ARM MOVEMENT VS. FOG

To observe whether wrist movements change during FoG, we plot in Figure 2 the raw accelerometer and gyroscope data from an IMU mounted on the wrist of a subject with Parkinson’s disease from CuPiD dataset. During the approx. 70 seconds of walking with turns, two freezing of gait episodes occurred. We can visually spot and distinguish the FoG episodes from the wrist movements – both accelerometers and gyroscopes have different patterns during FoG episodes, compared with the rest of walking.

A. Wrist Movement Features to Describe FoG

Previous visualization of the raw IMU data suggests that wrist movement during walking might have different and specific properties when a subject encounters a gait-freeze. Thus it might be possible to detect the freeze at the level of the gait using movement features captured at the wrist. Having in mind a real-time application for FoG-detection with wrist-mounted sensors, we need to extract features from sensing data which describe FoG. Gait-specific features extracted from acceleration data such as statistical features (mean, standard deviation) [8] or FFT-based features such as freezing index [12]...
are used to describe and detect in real-time the gait freezing episodes. But up to now, all these features require to have the sensors mounted on the lower body limbs, such as ankles or thighs, in order to capture the FoG characteristics. The first contribution of this work is to find and analyze new specific features from wrist mounted IMUs, which are related to the lower limb FoG episodes.

In Table I we list all the features we extract in a sliding-window manner from the wrist accelerometer and gyroscope data. We use a window size of $W = 3$ seconds, with a sliding-window step of $S = 0.25$ seconds, similar to the methodology applied on the same dataset from [15]. Prior to feature extraction, we compute the magnitude vectors from acceleration and gyroscope data from each window. We extract statistical features such as mean and standard deviation from both acceleration and rotation data, and FFT-based features from the acceleration magnitudes.

In Figure 3 we illustrate the extracted features from the same sequence of 70 seconds of walking in Figure 2. We observe that both sets of statistical features from accelerometer and gyroscope have higher mean and standard deviation values during the FoG events, compared with the rest of walking. Moreover, in Figure 3(c) we observe there are high values of the power on $[0-1] \text{Hz}$, and on $[8-13] \text{Hz}$ during FoG. Similarly, acceleration power on $[5-8]$, $[9-12]$, and $[13-16] \text{Hz}$ bands are higher during FoG compared with the other walking events.

**B. FoG-Detection from Wrist Movement**

To detect FoG episodes from wrist movement, we will employ the FoG-detection chain based on supervised machine learning methods with data from ankle mounted IMUs [15]. Instead of ankle data, we extract features from both right and left wrist-attached IMUs in the CuPiD dataset. Raw IMU readings are separated into overlapping windows from which wrist movement features are extracted, as detailed in the previous paragraph, together with the FoG or walking labels, as set by the clinicians. A total of 46 features from both right and left wrist together with the label create a feature vector.
Combinations of the feature vectors are then used to train C4.5 classification models, in order to automatically distinguish between FoG and the rest of walking events.

**Evaluation scheme.** We consider first a (1) subject-dependent evaluation model. That means we evaluate the wrist movement features for each of the 11 subjects in the CuPSD dataset separately. For each subject we consider a leave-one-FoG-out cross-validation evaluation scheme: We split the sensing data into sessions which contain in the center a FoG episode, and the remaining of the data in the session is composed from walking. Each of these sessions is then used as testing data, while the classification model is trained on the rest available sessions. We repeat this procedure for all the sessions in the dataset. Second, we consider a (2) subject-independent cross-validation scheme. That means each subject dataset is considered as testing data, while the rest of subjects data are used to train the classification model. We repeat the procedure for all the 11 subjects.

**Evaluation measures.** For both evaluation schemes we report the hit rate FoG-detection measure, the number of the false positive events, and the specificity with respect to the whole walking period for each subject dataset. The hit rate represents the number of correctly detected FoG events divided by the number of total FoG events. Different from previous work [8], which reports the window-based sensitivity, we consider FoG hit-rate a better measure, as it gives exactly the statistics we are interested in case of a wearable assistant – how many FoG events can be actually detected, and not how accurate it is on a window-basis comparison. The false positive events represent how many times a false FoG event was labeled, with respect to the whole period of walking. The specificity measures the proportion of correctly detected walking windows to all reference walking windows. Before reporting the hit-rate and number of false positives, we first pre-process the window-based output of the FoG-detection method: If the difference between two consecutive windows in which the classifier detected FoG is less than the window size $N = 3$ seconds, then the whole period between these two consecutive detections is considered to be part of the same FoG event.

Furthermore, we considered FoG-detection latency, defined as the delay between the start of a FoG episode as labeled by clinicians, and the start of a detected FoG episode by the algorithm.

**Subject-dependent cross-validation.** In Figure 4 we plot the FoG-hit rate and specificity metrics for each of the 11 subjects, and the averaged metrics across all subjects. The FoG-hit rate varies from 0.4 to 1, with an overall average of 0.9, and the specificity varies between 0.62 and 0.96, averaged to 0.83 across all subjects. Table II completes the FoG-detection statistics, with the reported number of detected FoG per subject, detection latency, and the number of false FoG events detected by the algorithm. Overall, 164 out of 182 FoG episodes and a total of 164 false detections across all datasets. FoG is correctly detected with an average detection-latency of 1.53 seconds, thus during the onset of the freeze, given that the average FoG duration in the dataset is of 9 seconds.

![Fig. 4. The FoG hit rate and specificity for each of the 11 subjects, and their average values across over all 11 datasets.](image)

<table>
<thead>
<tr>
<th>Subject</th>
<th># FoG</th>
<th># FoG detected</th>
<th>Latency (seconds)</th>
<th># False positives</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>19</td>
<td>16</td>
<td>3.79</td>
<td>31</td>
</tr>
<tr>
<td>S2</td>
<td>11</td>
<td>11</td>
<td>0.34</td>
<td>14</td>
</tr>
<tr>
<td>S3</td>
<td>22</td>
<td>22</td>
<td>0.14</td>
<td>4</td>
</tr>
<tr>
<td>S4</td>
<td>2</td>
<td>2</td>
<td>1.25</td>
<td>2</td>
</tr>
<tr>
<td>S5</td>
<td>5</td>
<td>2</td>
<td>0.12</td>
<td>5</td>
</tr>
<tr>
<td>S6</td>
<td>37</td>
<td>36</td>
<td>0.52</td>
<td>18</td>
</tr>
<tr>
<td>S11</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>S12</td>
<td>27</td>
<td>26</td>
<td>1.23</td>
<td>26</td>
</tr>
<tr>
<td>S17</td>
<td>24</td>
<td>23</td>
<td>1.02</td>
<td>20</td>
</tr>
<tr>
<td>S18</td>
<td>26</td>
<td>19</td>
<td>4.26</td>
<td>39</td>
</tr>
<tr>
<td>S19</td>
<td>6</td>
<td>5</td>
<td>4.24</td>
<td>4</td>
</tr>
</tbody>
</table>

The FoG episodes are not distributed equally among the subjects: 4 out of 11 datasets contain between 2 and 6 FoG episodes. The low number of FoG did not seem to affect the FoG-recognition performances, except in case of S5, with only 2 out of 5 FoG events detected. Moreover we observe the overall high FoG-detection results for 5 subjects where all the FoG events were detected. However the high FoG detection rate comes at the cost of false positives: for 4 out of 11 subjects (S1, S2, S5, and S18) the number of false detected FoG is equal or higher than the number of actual detected freezing episodes. These suggests that wrist movement data related to FoG might not happen only during FoG events, but also during other walking events, thus not being specific to FoG only.

**Subject-independent cross-validation.** In case of subject-independent evaluation, the wrist movements were useful to detect the same number of 164 out of 182 FoG episodes, but with a lower average latency of 0.98 seconds. However, the number of false positives increased considerably to 259 false events. The FoG-hit rate is 0.90, but the overall specificity drops down to 0.70, from 0.83 in the subject-dependent evaluation. This drop in performance recognition is expected when using data from different subjects. Similar to the gait characteristics, also the wrist movements might be specific for each subject. Also, each subject might have a different reaction during gait freeze, thus one could have very specific wrist movements during FoG.

**Feature statistics.** To have a deeper understanding of the FoG-detection performances, we visualized how the
features extracted from wrist movement differ during walking events and during FoG periods. We consider the same S6 dataset given as an example until now, from which we compute the mean and standard deviation for all the features during walking sessions. We then extract the same statistics from all the data representing the FoG events.

In Figure 5 we show the differences between the statistics for the main features extracted from the S6 data. We observe that statistical features extracted from the wrist acceleration data does not give discriminative information during large periods of time of walking and FoG. Gyroscope features have higher overall values during the FoG episodes compared with the rest of walking periods, but the difference is not major, and the values from these two classes of features overlap. Acceleration power on [8−9], [9−10], and [10−11] Hz bands have distinguishable higher values during FoG compared with walking.

Thus the FFT-based features extracted from acceleration seem to be the most informative to describe the wrist-movements correlated with FoG during gait-freeze episodes.

C. Discussion

Subject-dependent movements. The previous analysis shows that wrist movement patterns are different from subject to subject during FoG. We could split the 11 subject datasets we used in two groups, based on the wrist movement type during the gait freeze: The first group has similar wrist movement reactions during FoG as those illustrated from subject S6 in Figure 3. While the second group is characterized by an opposite wrist movement type during freeze episodes – people in this group tend to freeze all the movement of their arms during a gait freeze, meaning that their wrists seem totally blocked and they don’t perform any movement. Thus all the features extracted from IMUs during FoG have lower values compared with the rest of walking, even during turning or starting walking. An example of such pattern is given in Figure 6, where both acceleration and gyroscope statistical features decrease during the FoG event. Moreover, the power from acceleration on the 16 frequency bands have values close to 0, lower and different from the rest of walking. In this case, the features extracted during FoG are similar with the one extracted during short periods in which the subjects are standing, e.g., in between different walking sessions, or the periods just before staring walking. This might be one cause of the false detections, in which FoG events share similar features with standing movements, thus a confusion between them. For a real-time FoG-detection scenario [15], sitting or standing activities can be detected using phone’s internal sensors, and as a result eliminating such FoG false alarms.
Sensor placement. We presented the FoG-detection results obtained using combination of features from both wrists’ IMUs. However, we performed experiments in which we considered combinations of features coming from either the left, either the right wrist. The results obtained are comparable, suggesting that the correlated wrist movement patterns during a gait-freeze are similar on both of the right and left limbs.

The FoG-detection performances based on wrist data are slightly decreased in terms of FoG hit-rate – 0.9 – than when using the data from the IMUs mounted on the ankle – 0.94 [15], in a subject-independent cross-validation scheme. However, the specificity drops from 0.9 when using ankle IMU data [15] to 0.7 when using wrist IMUs. With subjects reacting very differently during a gait freeze, we require a model that takes into account this variety. However, although a user-independent model yielded a higher false positive rate when using movement data from wrists, it may not be critical in application: In this particular use case less missed FoG events are favored to a high precision (few false positives), as FoG are high-risk events during walking in Parkinson’s disease.

For a real-time FoG-detection system, we need to lower the sensors’ sampling rate to save battery. The FFT features we extract from wrist are using data up to 16 Hz, thus sensors can send information at 16 Hz sampling rate, or up to 32 Hz as in [15].

V. CONCLUSION

We investigated the correlation between wrist movement and freezing of gait in Parkinson’s disease. Our motivation is the possibility of building a FoG-detection assistant using sensors integrated in smartwatches or wristbands, as they are already accepted as on-body electronics, thus minimally intrusive as wearable. We found evidence that there are correlated and specific wrist movements during the gait-freeze episodes: Experiments on data from 11 subjects show that freezing of gait episodes can be detected using wrists’ movement features and supervised machine learning methods, with a hit-rate of 0.90 (164 out of 182 FoG detected), and a specificity of 0.83 (164 false FoG alarms) in a subject-dependent evaluation scheme. For subject-independent experiments, the hit-rate remains the same (164 out of 182 FoG), but with a decrease in specificity (259 false alarms). These results suggest that FoG can be detected only by using data from IMUs attached on the wrist.

We investigated specific statistical and FFT-based features extracted from the wrist-mounted accelerometers and gyroscopes, in order to capture and describe a correlated-with-FoG movement. However, subjects tend to have specific wrist movements during FoG. This and the high number of false positives obtained in the evaluation motivate the investigation of different features from the wrist-attached IMUs, in order to refine the FoG-correlated patterns.

ACKNOWLEDGMENT

The research leading to these results has received funding from the European Union - Seventh Framework Programme (FP7/2007-2013) under grant agreement n288516 (CuPID project).

REFERENCES


