Detecting Freezing of Gait with a Tri-Axial Accelerometer in Parkinson's Disease Patients

Claas Ahlrichs · Albert Samà · Michael
Lawo · Joan Cabestany · Daniel
Rodríguez-Martín · Carlos Pérez-López ·
Dean Sweeney · Leo Quinlan · Gearòid
Ò Laighin · Timothy Counihan ·
Patrick Browne · Lewy Hadas · Gabriel
Vainstein · Alberto Costa · Roberta
Annicchiarico · Sheila Alcaine · Berta
Mestre · Paula Quispe · Àngels Bayes ·
Alejandro Rodríguez-Molinero

Abstract Freezing of gait (FOG) is a common motor symptom of Parkinson's Disease (PD), which presents itself as an inability to initiate or continue gait. This paper presents a method to monitor FOG episodes based only on acceleration measurements obtained from a waist-worn device. Three approximations of this method are tested. Initially, FOG is directly detected by a support vector machine (SVM). Then, classifier's outputs are aggregated over time to determine a confidence value, which is used for the final classification

Claas Ahlrichs
neusta mobile solutions GmbH (NMS), Konsul-Smidt-Str. 24, D-28217 Bremen, Germany

Michael Lawo
Institute for Artificial Intelligence (AGKI), University of Bremen, Germany

Albert Samà · Joan Cabestany · Daniel, Rodríguez-Martín · Carlos Pérez-López
Technical Research Centre for Dependency Care and Autonomous Living (CETpD), Universitat Politcnica de Catalunya, Vilanova i la Geltr, Spain

Dean Sweeney · Leo Quinlan · Gearòid Ò Laighin · Alejandro Rodríguez-Molinero
Electrical & Electronic Engineering Department, NUI Galway, Galway, Ireland

Timothy Counihan · Patrick Browne
School of Medicine, NUI Galway, Galway, Ireland

Lewy Hadas · Gabriel Vainstein
Maccabi Healthcare Services, Tel Aviv, Israel

Alberto Costa · Roberta Annicchiarico
IRCCS Fondazione Santa Lucia, Rome, Italy

Sheila Alcaine · Berta Mestre · Paula Quispe · Àngels Bayes
Unidad de Parkinson y trastornos del movimiento (UParkinson), Barcelona, Spain

of freezing (i.e. second and third approach). All variations are trained with signals of 15 patients and evaluated with signals from another 5 patients. Using a linear SVM kernel, the third approach outperforms results in the literature with 98.7% accuracy and a geometric mean of 96.1%. Moreover, it is investigated whether frequency features are enough to reliably detect FOG. Results show that these features allow the method to detect FOG with accuracies above 90% and that frequency features enable a reliable monitoring of FOG by using simply a waist sensor.

1 Introduction

PD is a chronic, progressive, neurodegenerative disorder [2,14,30,10,16], with which a great number of motor and non-motor symptoms have been associated. The disorder was first described by James Parkinson in 1817 [26]. It affects the movement and it is typically characterized by a loss of (motor) function, increased slowness and rigidity. Presently the cause and origin of PD remain unknown [14,30,9,17] and it cannot be cured. Consequently, treatments aim at reducing severity and frequency of motor complications. The disease is generally associated with elderly people and is rarely diagnosed before the age of 40. It is estimated that the mean age of onset is about 65 years [30].

PD is a great burden as it considerably decreases the quality of life, due to a gradual loss of function and decreasing ability to take care of oneself. The World Health Organization (WHO) considers the burden of PD to be on the same disability level as an amputated arm, drug dependency, congestive heart failure, deafness and tuberculosis [20]. The cardinal symptoms are bradykinesia, rigidity, tremor and postural instability [14,30,10,16,17,32,1]. However, a number of non-motor related symptoms (e.g. sleep disturbances, depression, psychosis, autonomic and gastrointestinal dysfunction as well as dementia) may occur as well [14,30,10,16,18].

One of the motor symptoms is called FOG (also known as freezing or motor blocks). It is a form of akinesia, which presents itself as an inability to initiate or continue gait [30,16,12,31,24]. Motor blocks are a common symptom, experienced by people with Parkinson's (although it does not occur uniformly) and can affect various extremities (e.g. arms and legs) as well as the face [16]. Freezing greatly impairs the quality of life of those affected and is one the most disabling symptoms. It is usually attributed to medium and advanced stages of PD and it is a common cause of falls [30,16,6]. A single freezing episode typically lasts for several seconds. In severe cases, episodes can be apparent for as long as several minutes.

Continuous monitoring of FOG events can give neurologists information which is otherwise difficult to obtain. Clinical assessment of FOG at the doctor's office is considered to be problematic since symptoms are commonly not evident in this clinical environment [25]. Thus, a wearable device capable of ambulatory monitoring FOG could benefit patients in two ways. First, it could provide clinicians with complementary information of the disease that can be

used to improve treatment [31]. Second, since patients are capable of improving gait based on specific stimulations provided as haptic, visual or auditory cues [19], real-time FOG detection would allow patients to avoid some episodes and, consequently, avoid falls, such as the system presented in [3]. Consequently, many studies have attempted to develop wearable devices for the detection of FOG.

The literature indicates that these studies typically make use of multiple sensors (i.e. accelerometers, gyroscopes, etc.) at various body locations and they usually employ some form of supervised learning approach (e.g. SVM or neural network (NN)). Djurić-Jovičić et al. [11] achieved an error rate of up to 16% classifying "normal" (i.e. standing and regular steps) and pathological (i.e. festination, akinesia, shuffling and small steps) walking patterns of PD patients based on a NN (using multiple inertial measurement units). The approach by Cole et al. [7] yielded to 82.9% sensitivity and 97.3% specificity in detecting FOG (using acceleration and electromyograph (EMG) sensors) with a multi-staged algorithm that utilized a simple linear classifier and a dynamic neural network (DNN). Niazmand et al. [23] used an accelerometer based smart garment [22] to extract gait-related features. They achieved 88.3% sensitivity and 85.3% specificity (using multiple accelerometers). The approach by Bächlin et al. [5] yielded to 73.1% sensitivity and 81.6% specificity for detecting FOG events in real-time (using multiple accelerometers and gyroscopes).

In this work, the authors present a multi-staged approach based on an SVM and a single tri-axial acceleration sensor. Using a linear SVM kernel and the full feature set (see Table 5), an accuracy of 98.7% and a geometric mean of 96.1% have been achieved. The overall dataset (i.e. training and testing, together) includes signals from 20 PD patients. These results have been obtained with a patient-independent methodology and they exceed those results found in other publications. Furthermore, the algorithm can be configured toward a higher sensitivity or a higher specificity. The employed movement signals were collected for the REMPARK-project (Personal Health Device for the Remote and Autonomous Management of Parkinson's Disease) database [29]. This project aims to develop a closed loop system with the purpose of monitoring PD motor and non-motor symptoms, and responding to these symptoms in real-time using a series of actuators. Data collection of REMPARK's database inertial signals has taken place in 4 different countries (Spain, Italy, Israel and Ireland).

2 Methods

Firstly, the data acquisition is described. Then, the methodology and model selection of the proposed approach are outlined.

Data Acquisition and Labeling: All participants (aged between 50 and 75 years) had a clinical diagnosis of Idiopathic Parkinson's disease according to

	Training	Test
Number of freezing windows	93	45
Number of non-freezing windows	3883	2312
Recordings with freezing	6	2
Recordings without freezing	9	3
Overall number of recordings	15	5

Table 1: The number of windows (before aggregation) in each dataset that are used for signifying FOG.

Indexes	Feature
1	FFT (raw, no filtering)
2	Mean and standard deviation of amplitude (band: 0.5-3.0 Hz)
3	Entropy of signal in time domain (band: 0.5-3.0 Hz)
4	Peak amplitude and its frequency (band: 0.5-3.0 Hz)
5	Mean and standard deviation of amplitude (band: 3.0-8.0 Hz)
6	Entropy of signal in time domain (band: 3.0-8.0 Hz)
7	Peak amplitude and its frequency (band: 3.0-8.0 Hz)
8	Freezing Index

Table 2: The full set of features used for FOG detection. In contrast, the reduced feature set is only comprised of a fast Fourier transform (i.e. index 1).

the United Kingdom (UK) PD Society Brain Bank [15]. Clinical fluctuations were present in all patients as well as Hoehn and Yahr stage [13] above two (moderate-severe phase of PD). Furthermore, all patients gave their signed informed consent before their participation. The experimental protocol was approved by the corresponding local ethics review committee. For this paper, signals from 20 PD patients were used. The recordings are identical to those employed by Rodríguez-Martín et al. [27,28].

As part of the experiments, participants were recorded with an HD quality camera while wearing a set of sensors (i.e. accelerometer, gyroscope and magnetometer) as they performed a set of scripted activities. However, these activities are of a rather general nature (e.g. walking around the apartment and showing it to the researchers or carrying a full glass of water from the kitchen to another room) and they are much more variable in comparison to other typically scripted activities like hand-to-nose or similar gestures which are performed in a seated position. The recordings also include non-scripted activities that lead to FOG episodes (e.g. narrow places and turnings). The experiments took place at the participant's apartment and started in the morning. During the course of the day, two recording sessions took place: one in the "OFF" motor state and one while in the "ON" motor state. For the first session, participants were asked to skip their morning dose of medication, thus recordings started while the participant was in a clinically defined OFF state [29]. After finishing the first round, participants took their normal medication and the second recording session was started once the participant had reached a clinically defined ON state. During both recording sessions, participants performed a series of short controlled activities. The activities performed by

patients during their OFF state were an indoors walking test, a FOG provocation test and a gait test. During the ON state, a dyskinesia test, a dual task test and a set of activities of daily livings (ADLs) also were performed. ADLs included brushing teeth, shaking a deodorant, erasing with an eraser, writing with a pencil, typing on a computer keyboard, cleaning a window or furniture and drying a wet glass [29].

Experienced clinicians labeled the videos based on the activities that patients performed and the symptoms shown during the video. The clinicians who performed the labeling were also physically present during the recording sessions. Each of the clinical sides (one for each country) had two clinicians with several years of experiences with PD patients (i.e. ≥ 5 years). Prior to the recording sessions, all clinicians received a training session on setting baselines for labeling of symptoms (including FOG). The group that performed the labeling is disjoint from the group that performed the analysis.

Video and inertial signals were synchronized based on the procedure described in [29]. FOG labels provided by clinicians have been treated with an automatic relabeling procedure in order to consider specific peculiarities of FOG. Recordings of freezers were cut to the point where only FOG labels remained. This reduced the overall amount of data for recordings of freezers but ensured that no freezing episodes (which might not have been properly labeled) were used. On the other hand, those patients without any freezing episodes were relabeled in such a way that all available data were used. Consequently, sensitivity was determined by using data from patients with freezing episodes while specificity was determined by using data from non-freezing patients. Overall, this procedure allowed using larger portions of the recordings.

As far as the actual labeling is concerned, the presence of any type of freezing (e.g. start, turn, end, etc.) was considered to be an episode of FOG. The detection of individual types of freezing requires additional contextual information which is not contained within the database (DB). Furthermore, such a fine granularity might not provide an additional value (e.g. to a PD monitoring system). The fact that a freezing episode is happening is more relevant than the actual type of episode (e.g. for rhythmic cueing purposes). Consequently, freezing episodes are detected rather than individual types of freezing.

Methodology: The general methodology is such that acceleration signals from a waist-mounted sensor are split into equally sized windows (i.e. a sliding window is applied to the time series). Features are extracted from those windows and fed to an SVM for training or classification. The classification output of n consecutive windows $s_1, \ldots, s_n$ is then aggregated over time t to achieve higher accuracies. However, the volatile nature of FOG must be considered during the development of an algorithm for detecting such episodes. In contrast to resting tremor (or dyskinesia for that matter) episodes of FOG do not last for prolonged periods of time which may emphasize the importance of the chosen window size ws. In any case, the contents of the database are split into two datasets (i.e. training and testing) that are used for training an SVM as well as

optimizing additional parameters and testing, respectively. Datasets stay the same for all approaches (details are listed in Table 1). The individual datasets hold 15 and 5 patients for the training and testing dataset, respectively.

Two feature sets are evaluated: a reduced feature set with only the fast Fourier transform (FFT) and a full feature set with various additional features (see Table 2). The effect of adding these additional features is quantified in section 3. These features are comprised of the freezing index [4] as well as some frequency related features for differing frequency ranges [21].

At first varying window sizes ws were evaluated such that freezing of gait detection was optimized. The comparison of different window sizes was done on an episode level (rather than a window level). An episode of FOG was detected when at least one window within an actual FOG episode was classified as such. As far as non-freezing episodes were concerned, an aggregation of windows over a period of time that corresponds to the average length of a FOG episode plus twice the standard deviation is performed. The acceleration data are resampled to 40Hz and split into unisized chunks of data $s_1, \ldots, s_m$ with a certain length ws that overlapped to 50%. These windows are then used to extract features which in turn were fed to an SVM for training and classification. This resembles the first and naive approach, where $freezing_j^1$ represents the jth window in the series $s_1, \ldots, s_m$ and whether FOG is present in that window.

$$freezing_j^1 = \begin{cases} 0 \text{ no freezing if } f_{SVM} \leq 0 \\ 1 \text{ freezing } \quad \text{if } f_{SVM} > 0 \end{cases} \tag{1}$$

where $f_{svm} = \sum_{i=1}^{l} y_i \alpha_i K(\mathbf{x}_i, \mathbf{f}) + b$, $\mathbf{x}_1, \ldots, \mathbf{x}_l$ are the support vectors (SVs), y_i, α_i are the corresponding label and lagrange multiplier of each SV and b is the bias [8]. The number in the superscript (here: 1) indicates the variation. The second and third variation will use 2 and 3, respectively.

The second variation aggregates the SVMs' outputs over a time period t and calculates the degree of confidence c_j. If the confidence value exceeds a threshold th then the aggregated time frame t is considered to be an episode of FOG, otherwise not. Here, $freezing_j^2$ covers a time frame t (starting at the jth window and covering n windows) and determines whether FOG is apparent in that time frame.

$$t = \frac{ws(n+1)}{40 * 2} \tag{2}$$

$$c_j = \sum_{i=j}^{j+n-1} \frac{freezing_i^1}{n} \tag{3}$$

$$freezing_j^2 = \begin{cases} 0 \text{ no freezing if } c_j < th \\ 1 \text{ freezing } \quad \text{if } c_j \geq th \end{cases} \tag{4}$$

where $c_j, th \in [0, 1]; n, j \in \mathbb{N}; n, j > 0; t \in \mathbb{R}^+$

The third variation introduces a second threshold. The lower threshold th_l and upper threshold th_u can be used to tune sensitivity and specificity separately. The lower threshold th_l sets the maximum confidence value for "no

freezing" periods and the upper threshold th_u sets the minimum confidence value for freezing episodes. By not requiring that these thresholds need to be equal (which would essentially be variation two), the final output of the algorithm may indicate the presence of freezing as well as "undefined". This is the case when the confidence value is between the two thresholds. Consequently, some aggregated windows may be ignored and data usage is lowered.

$$freezing_j^3 = \begin{cases} 0 & \text{no freezing if} \quad c_j < th_l \\ 1 & \text{freezing} \quad \text{if} \quad c_j \geq th_u \\ -1 & \text{undefined} \quad \text{if} \quad th_l \leq c_j < th_u \end{cases} \tag{5}$$

where $c_j, th_l, th_u \in [0, 1]; th_l \leq th_u; j \in \mathbb{N}; j > 0$

Model Selection: The individual SVM models are trained with the features that were extracted from the training dataset. For the second and third variation, the individual parameters t, th, th_l and th_u are also optimized on the training dataset. The final results are obtained from the testing dataset.

The window size ws is determined before any of these parameters are evaluated. For each of the proposed window sizes ws (see below), the naive algorithm is applied to the training dataset. The window size that yields to the best combination of accuracy and geometric mean is chosen.

During training, varying settings for kernel, weighting, cost and gamma were considered. The weighting parameters were used to balance both classes "FOG" and "non-freezing". The cost and gamma parameters were systematically evaluated (i.e. $10^q, q \in \{-3, -2, ..., 2, 3\}$) depending on the chosen kernel (i.e. radial basis function (RBF) kernel or linear kernel). Additionally, a ten-fold cross-validation is performed on the training dataset. However, instead of averaging the accuracy of the training set, the geometric mean of sensitivity and specificity is used (i.e. $\sqrt{sensitivity * specificity}$) to identify those parameters combinations with high sensitivity and specificity. The maximum geometric mean is used to select the optimal parameters and obtain the final SVM model to be used with the test dataset. The geometric mean was chosen as it does treat both classes equally as opposed to accuracy which implicitly weights the classes. The weighting of latter measure can be a problem if the classes have (very) different priors.

The following discrete values have been evaluated: $ws \in 2^{\{5,6,7,8\}}$; $t \in \{10, 15, 20, 25, 30, 45, 60\}$; $th, th_l \in \{0, 0.05, 0.1, \ldots, 0.95, 1.0\}$; $th_u \in (th_u \geq th_l | th_u \in \{0, 0.05, 0.1, \ldots, 0.95, 1\})$. The appropriate values and parameters were evaluated for each of the four conditions (two kernels and two feature sets).

3 Results

The average length of a FOG episode in our dataset was 3.48 [$\pm$ 3.29] seconds (total: 209 freezing events). Figure 1 shows several measures for varying window sizes (i.e. sensitivity, specificity, geometric mean and accuracy). The
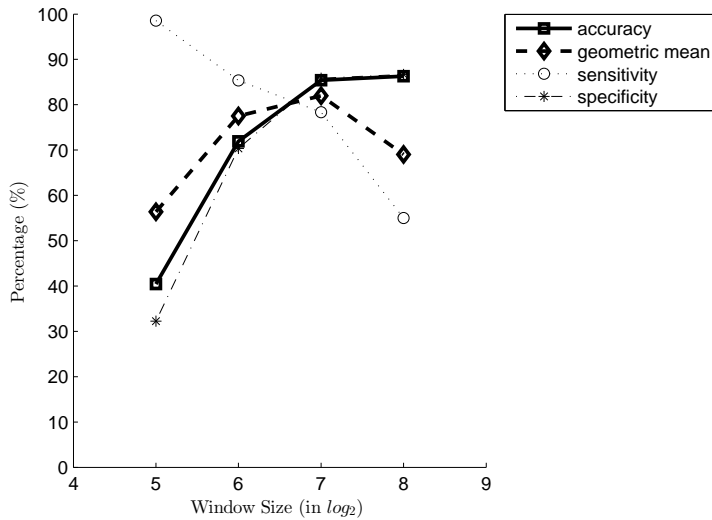
Fig. 1: Results of an evaluation for varying window sizes with respect to freezing episodes. For each window size, an SVM has been trained on the training dataset and evaluated on the test dataset.

best values for those measures were achieved with a window size of 128 samples (i.e. 2^7 samples). Accuracy and geometric mean were closest at this level. Consequently, this window size was utilized in all further analysis.

Table 3 presents the results obtained for the first variation. It was observed that, on the training dataset, both full and reduced feature sets yield a similar geometric mean regardless of the employed SVM kernel. This, however, diverges on the test dataset. The RBF kernel seems to benefit from the reduced feature set while the linear kernel favors the full feature set. Acceptable levels of specificity are consistently achieved on the test dataset, while sensitivity was reduced by false negatives (FNs). The latter may be counteracted when windows are aggregated. Nonetheless, accuracies above 90% were consistently reached.

The impact of window aggregation t and threshold th are highlighted in Figure 2 for all four conditions. The subfigures indicate that a threshold close to 50% works best in all cases. Furthermore, it is observed that the geometric mean increases with the aggregation level.

Numerical results for the second variation are shown in Table 4. All conditions yielded to a threshold close to the intuitive border of 50%, which is consistent with the observations in Figure 2. Moreover, the aggregation period t is the same across all four conditions. Having optimized parameters t and th on the training dataset, the results on the testing dataset show an increase by 9.4% (on average). All conditions achieve a high specificity of 98% or greater and, furthermore, most conditions also reached a sensitivity of 90% or above for an aggregation period of 60 seconds.

| Kernel | RBF | Linear | RBF | Linear |
Features	Freq.	Freq.	All	All
Sensitivity (train)	0.946	0.903	0.946	0.925
Specificity (train)	0.860	0.903	0.901	0.932
Data Usage (train)	1.000	1.000	1.000	1.000
Geometric Mean (train)	0.902	0.903	0.924	0.928
Accuracy (train)	0.862	0.903	0.902	0.932
True Positives	37	30	32	37
False Positives	167	117	124	101
True Negatives	2145	2195	2188	2211
False Negatives	8	15	13	8
Sensitivity (test)	0.822	0.667	0.711	0.822
Specificity (test)	0.928	0.949	0.946	0.956
Data Usage (test)	1.000	1.000	1.000	1.000
Geometric Mean (test)	**0.873**	**0.796**	**0.820**	**0.887**
Accuracy (test)	**0.926**	**0.944**	**0.942**	**0.954**

Table 3: Results in signifying FOG with the naive approach (i.e. variation 1). Various measures are listed for both datasets.

Results in Table 5 are those of the third variation. Most conditions still favor an aggregation level of 60 seconds. The lower and upper thresholds (i.e. th_l and th_u) were consistently found to enclose the previously found thresholds th in the second approach (see Table 4). Allowing for two thresholds increased sensitivity and specificity values on the test dataset for the linear kernel. However, the RBF kernel did not benefit from this approach in terms of geometric mean. The average change in geometric mean from variation two to three was -1.2% and 3.7% for the RBF and linear kernel, respectively. Nonetheless, all conditions yield to a sensitivity of roughly 90% and a specificity well above 90%. However, this was at the cost of a slightly reduced data usage although still above 90% for the most part.

4 Discussion

The presented FOG detection methods result in a geometric mean of 88.7%, 96.1% and 96.1% for each of the three proposed approaches (linear kernel with full feature set). Thus, the meta-analysis used in the second and third variation is shown to enable a better recognition of FOG episodes since it improves the overall performance (geometric mean) by 8%. Regarding the feasibility of detecting FOG uniquely by means of frequency features, it was observed that a geometric mean of 96.1% (one sided approach with RBF kernel) can be achieved based on them. This way, it is concluded that frequency features enable a reliable monitoring of FOG.

The results previously obtained by Niazmand et al. [23], Cole et al. [7] and Bächlin et al. [5] were consistently outperformed by the presented approach in its third variation, which has yielded to an average sensitivity and specificity above 94%. Niazmand et al. [23] achieved a sensitivity of 88.3% and a specificity of 85.3%. Compared to the results from Bächlin et al. [5], both sensitivity
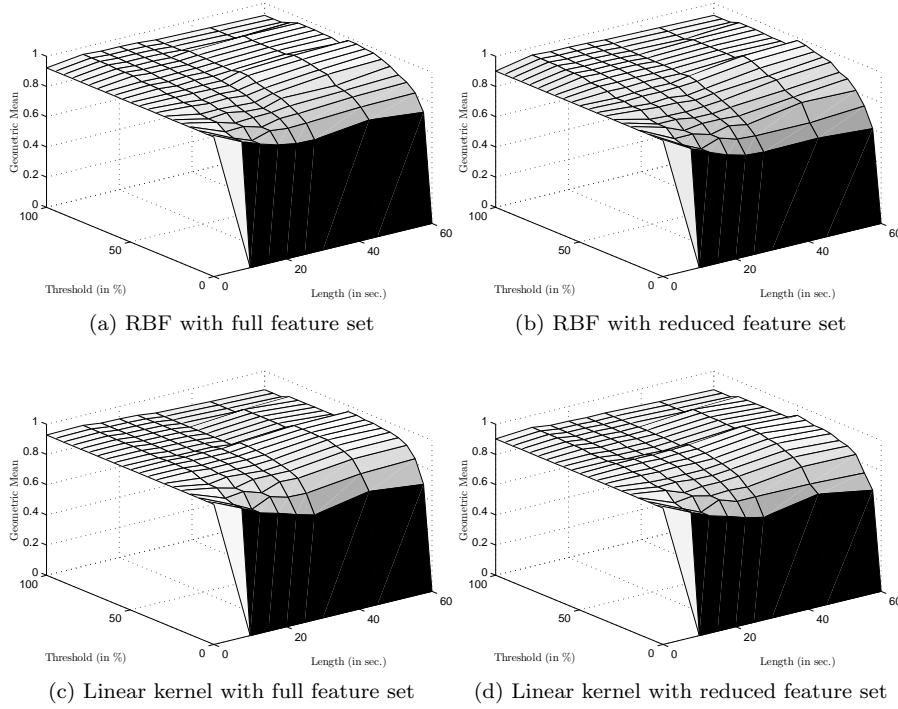
(a) RBF with full feature set   (b) RBF with reduced feature set

(c) Linear kernel with full feature set   (d) Linear kernel with reduced feature set

Fig. 2: Effect of window aggregation t and threshold th on geometric mean. The results are shown for all four conditions.

Kernel	RBF	Linear	RBF	Linear
Features	Freq.	Freq.	All	All
t (in seconds)	60	60	60	60
th (in %)	0.400	0.500	0.400	0.450
Sensitivity (train)	1.000	0.885	1.000	0.923
Specificity (train)	0.911	0.991	0.946	1.000
Data Usage (train)	1.000	1.000	1.000	1.000
Geometric Mean (train)	0.954	0.936	0.973	0.961
Accuracy (train)	0.928	0.971	0.957	0.986
True Positives	12	10	12	12
False Positives	0	1	0	0
True Negatives	66	65	66	66
False Negatives	1	3	1	1
Sensitivity (test)	0.923	0.769	0.923	0.923
Specificity (test)	1.000	0.985	1.000	1.000
Data Usage (test)	1.000	1.000	1.000	1.000
Geometric Mean (test)	**0.961**	**0.870**	**0.961**	**0.961**
Accuracy (test)	**0.987**	**0.949**	**0.987**	**0.987**

Table 4: Results in signifying FOG with the one-sided approach (i.e. variation 2).

| Kernel | RBF | Linear | RBF | Linear |
Features	Freq.	Freq.	All	All
t (in seconds)	60	45	60	60
th_l (in %)	0.150	0.250	0.300	0.350
th_u (in %)	0.900	0.800	0.800	0.500
Sensitivity (train)	1.000	0.923	1.000	0.923
Specificity (train)	1.000	1.000	1.000	1.000
Data Usage (train)	0.696	0.891	0.906	0.986
Geometric Mean (train)	1.000	0.961	1.000	0.961
Accuracy (train)	1.000	0.987	1.000	0.985
True Positives	9	8	9	12
False Positives	0	0	0	0
True Negatives	55	82	65	65
False Negatives	1	1	1	1
Sensitivity (test)	0.900	0.889	0.900	0.923
Specificity (test)	1.000	1.000	1.000	1.000
Data Usage (test)	0.823	0.919	0.949	0.987
Geometric Mean (test)	**0.949**	**0.943**	**0.949**	**0.961**
Accuracy (test)	**0.985**	**0.989**	**0.987**	**0.987**

Table 5: Results in detecting FOG with the two-sided approach (i.e. variation 3).

(73.1%) and specificity (81.6%) were clearly outperformed. However, Cole et al. [7] achieved a similar level of specificity (97.3%), but with lower sensitivity (82.9%).

A limitation of the presented work relies in its applicability to real-time detection in order to provide rhythmic cues. In this case, a short lag between the appearance of a FOG episode and its detection is desired. The meta-analysis in the second and third variations may add a delay that could reach the aggregation time (60 s.), which would not allow to be used for this particular purpose, although remaining useful in monitoring tasks. The first variation, however, could be employed, since the lag provided is roughly 3.2 s. (128 samples at 40 Hz).

Besides the performance comparative, the proposed approach has with respect to [5] the advantage of being patient-independent, given that the same classifier can be used by any patient. Moreover, we only use a single tri-axial accelerometer at the waist (e.g. Cole et al. [7]: three tri-axial accelerometers and surface EMG, Niazmand et al. [23]: five accelerometers, Bächlin et al. [5]: three accelerometers and three gyroscopes). The presented approach offers configuration capabilities since the algorithm can be tuned toward high sensitivity and high specificity by adjusting the thresholds. Finally, the optimal window size has been determined by evaluating the performance of the algorithm at episode level, as opposed to window level used in previous works, which may have increased specificity.

5 Conclusion

This work has evaluated three approaches to detecting FOG in Parkinson's patients based on a waist-worn accelerometer. The optimal window size was determined and it has been analyzed whether frequency features are sufficient to reliably detect FOG.

Although the linear and RBF kernel do not benefit equally from the third approach, combining the results from both variations (i.e. second and third variation) shows promising results. While the RBF kernel achieved a geometric mean greater than 95% and an accuracy greater than 98% with the second approach, the linear kernel reached similar levels (close to 95% geometric mean and 98% accuracy) with the third approach. However, in latter case the data usage is slightly penalized. The findings suggest that the full feature set is not required for satisfactory results. Instead, a linear kernel that has been trained with an FFT alone can accurately detect FOG episodes. Finally, the optimal window size has been found to be 128 samples (at 40 Hz).

The present approach works in a patient-independent way and requires only a single tri-axial accelerometer. The results outperform related work.

References

1. Andlin-Sobocki, P., Jnsson, B., Wittchen, H.U., Olesen, J.: Cost of Disorders of the Brain in Europe. European Journal of Neurology **12**, 1–27 (2005). DOI 10.1111/j.1468-1331.2005.01202.x. URL http://dx.doi.org/10.1111/j.1468-1331.2005.01202.x
2. Armstrong, R.A.: Visual Signs and Symptoms of Parkinson's Disease. Clinical and Experimental Optometry **91**(2), 129–138 (2008). DOI 10.1111/j.1444-0938.2007.00211.x. URL http://dx.doi.org/10.1111/j.1444-0938.2007.00211.x
3. Bächlin, M., Plotnik, M., Roggen, D., Giladi, N., Hausdorff, J.M., Tröster, G.: A Wearable System to Assist Walking of Parkinson's Disease Patients. Methods of Information in Medicine **49**(1), 88–95 (2010). DOI 10.3414/ME09-02-0003. URL http://www.schattauer.de/en/magazine/subject-areas/journals-a-z/methods/issue/special/manuscript/12447/show.html
4. Bächlin, M., Plotnik, M., Roggen, D., Maidan, I., Hausdorff, J.M., Giladi, N., Troster, G.: Wearable Assistant for Parkinson's Disease Patients with the Freezing of Gait Symptom. IEEE Transactions on Information Technology in Biomedicine **14**(2), 436–446 (2010). DOI 10.1109/TITB.2009.2036165
5. Bächlin, M., Roggen, D., Troster, G., Plotnik, M., Inbar, N., Meidan, I., Herman, T., Brozgol, M., Shaviv, E., Giladi, N., Hausdorff, J.M.: Potentials of Enhanced Context Awareness in Wearable Assistants for Parkinson's Disease Patients with the Freezing of Gait Syndrome. In: 2009 International Symposium on Wearable Computers (ISWC), pp. 123–130 (2009). DOI 10.1109/ISWC.2009.14
6. Bloem, B.R., Hausdorff, J.M., Visser, J.E., Giladi, N.: Falls and Freezing of Gait in Parkinson's Disease: A Review of Two Interconnected, Episodic Phenomena. Movement Disorders **19**(8), 871–884 (2004). DOI 10.1002/mds.20115. URL http://dx.doi.org/10.1002/mds.20115

7. Cole, B.T., Roy, S.H., Nawab, S.H.: Detecting Freezing-of-gait During Unscripted and Unconstrained Activity. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 5649–5652 (2011). DOI 10.1109/IEMBS.2011.6091367

8. Cristianini, N., Shawe-Taylor, J.: An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Cambridge University Press, New York, NY, USA (2000)

9. Dauer, W., Przedborski, S.: Parkinson's Disease: Mechanisms and Models. Neuron **39**(6), 889–909 (2003). DOI 10.1016/S0896-6273(03)00568-3. URL http://www.sciencedirect.com/science/article/pii/S0896662703005683

10. Davie, C.A.: A Review of Parkinson's Disease. British Medical Bulletin **86**(1), 109–127 (2008). DOI 10.1093/bmb/ldn013. URL http://bmb.oxfordjournals.org/content/86/1/109.abstract

11. Djurić-Jovičić, M., Jovičić, N.S., Milovanović, I., Radovanović, S., Kresojević, N., Popović, M.B.: Classification of Walking Patterns in Parkinson's Disease Patients Based on Inertial Sensor Data. In: 2010 10th Symposium on Neural Network Applications in Electrical Engineering (NEUREL), pp. 3–6 (2010). DOI 10.1109/NEUREL.2010.5644040

12. Giladi, N.: Freezing of Gait: Risk Factors and Clinical Characteristics. Parkinsonism and Related Disorders **12**(Supplement 2), S52 (2006). DOI 10.1016/j.parkreldis.2006.05.015

13. Hoehn, M.M.: Parkinsonism: Onset, Progression, and Mortality. Neurology **17**, 427–442 (1967)

14. Hou, J.G.G., Lai, E.C.: Non-motor Symptoms of Parkinson's Disease. International Journal of Gerontology **1**(2), 53–64 (2007). DOI 10.1016/S1873-9598(08)70024-3. URL http://www.sciencedirect.com/science/article/pii/S1873959808700243

15. Hughes, A.J., Daniel, S.E., Kilford, L., Lees, A.J.: Accuracy of Clinical Diagnosis of Idiopathic Parkinson's Disease: A Clinico-pathological Study of 100 Cases. Journal of Neurology, Neurosurgery & Psychiatry **55**(3), 181–184 (1992). DOI 10.1136/jnnp.55.3.181. URL http://jnnp.bmj.com/content/55/3/181.abstract

16. Jankovic, J.: Parkinson's Disease: Clinical Features and Diagnosis. Journal of Neurology, Neurosurgery & Psychiatry **79**(4), 368–376 (2008). DOI 10.1136/jnnp.2007.131045. URL http://jnnp.bmj.com/content/79/4/368.abstract

17. Korczyn, A.D.: Parkinson's Disease. In: E. in Chief: Kris Heggenhougen (ed.) International Encyclopedia of Public Health, pp. 10–17. Academic Press, Oxford (2008). DOI 10.1016/B978-012373960-5.00028-9. URL http://www.sciencedirect.com/science/article/pii/B9780123739605000289

18. Krenz, A.: The Pathological Role of Synphilin-1 and the Therapeutic Potential of Hsp70 in Models of Parkinson's Disease Using Viral Vectors. Ph.D. thesis, Universität Tübingen, Wilhelmstr. 32, 72074 Tübingen (2010). URL http://tobias-lib.uni-tuebingen.de/volltexte/2010/4620

19. Lim, I., van Wegen, E., de Goede, C., Deutekom, M., Nieuwboer, A., Willems, A., Jones, D., Rochester, L., Kwakkel, G.: Effects of External Rhythmical Cueing on Gait in Patients with Parkinson's Disease: A Systematic Review. Clinical Rehabilitation **19**(7), 695–713 (2005). DOI 10.1191/0269215505cr906oa. URL http://cre.sagepub.com/content/19/7/695.abstract

20. Mathers, C., Fat, D.M., Boerma, J.T., WHO: the Global Burden of Disease : 2004 Update. World Health Organization, Geneva, Switzerland (2008). URL http://www.who.int/healthinfo/global_burden_disease/GBD_report_ 2004update_full.pdf

21. Moore, S.T., MacDougall, H.G., Ondo, W.G.: Ambulatory Monitoring of Freezing of Gait in Parkinson's Disease. Journal of Neuroscience Methods **167**(2), 340–348 (2008). DOI http://dx.doi.org/10.1016/j.jneumeth.2007.08.023. URL http://www.sciencedirect.com/science/article/pii/S0165027007004281

22. Niazmand, K., Somlai, I., Louizi, S., Lueth, T.C.: Proof of the Accuracy of Measuring Pants to Evaluate the Activity of the Hip and Legs in Everyday Life. In: J.C. Lin, K.S. Nikita, O. Akan, P. Bellavista, J. Cao, F. Dressler, D. Ferrari, M. Gerla, H. Kobayashi, S. Palazzo, S. Sahni, X.S. Shen, M. Stan, J. Xiaohua, A. ZoMaya, G. Coulson (eds.) Wireless Mobile Communication and Healthcare, *Lecture Notes of the Institute for*

Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 55, pp. 235–244. Springer Berlin Heidelberg (2011). URL http://dx.doi.org/10.1007/978-3-642-20865-2_30. 10.1007/978-3-642-20865-2_30

23. Niazmand, K., Tonn, K., Zhao, Y., Fietzek, U.M., Schroeteler, F., Ziegler, K., Ceballos-Baumann, A.O., Lueth, T.C.: Freezing of Gait Detection in Parkinson's Disease Using Accelerometer Based Smart Clothes. In: 2011 IEEE Biomedical Circuits and Systems Conference (BioCAS), pp. 201–204 (2011). DOI 10.1109/BioCAS.2011.6107762

24. Nieuwboer, A., Giladi, N.: Characterizing Freezing of Gait in Parkinson's Disease: Models of an Episodic Phenomenon. Movement Disorders **28**(11), 1509–1519 (2013). DOI 10.1002/mds.25683. URL http://dx.doi.org/10.1002/mds.25683

25. Nieuwboer, A., Weerdt, W.d., Dom, R., Lesaffre, E.: A Frequency and Correlation Analysis of Motor Deficits in Parkinson Patients. Disability and Rehabilitation **20**(4), 142–150 (1998). DOI 10.3109/09638289809166074. URL http://informahealthcare.com/doi/abs/10.3109/09638289809166074

26. Parkinson, J.: An Essay on the Shaking Palsy. 1817. The Journal of neuropsychiatry and clinical neurosciences **14**(2), 223–236; discussion 222 (2002). URL http://www.ncbi.nlm.nih.gov/pubmed/11983801

27. Rodríguez-Martín, D., Samà, A., Pérez-López, C., Cabestany, J., Català, A., Rodríguez-Molinero, A.: Enhancing FoG Detection By Means of Postural Context Using a Waist Accelerometer. First International Freezing of Gait Congress (IFOG 2014) (2014)

28. Rodríguez-Martín, D., Samà, A., Pérez-López, C., Cabestany, J., Català, A., Rodríguez-Molinero, A.: Posture Transition Identification on PD Patients Through a SVM-based Technique and a Single Waist-worn Accelerometer. (2015). Accepted for publication in Neurocomputing

29. Samà, A., Peréz, C., Rodríguez-Martin, D., Cabestany, J., Moreno Aróstegui, J.M., Rodríguez-Molinero, A.: A Heterogeneous Database for Movement Knowledge Extraction in Parkinson's Disease. In: European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (2013)

30. Samii, A., Nutt, J.G., Ransom, B.R.: Parkinson's Disease. The Lancet **363**(9423), 1783–1793 (2004). DOI 10.1016/S0140-6736(04)16305-8. URL http://www.sciencedirect.com/science/article/pii/S0140673604163058

31. Schaafsma, J.D., Balash, Y., Gurevich, T., Bartels, A.L., Hausdorff, J.M., Giladi, N.: Characterization of Freezing of Gait Subtypes and the Response of Each to Levodopa in Parkinson's Disease. European Journal of Neurology **10**(4), 391–398 (2003). DOI 10.1046/j.1468-1331.2003.00611.x. URL http://dx.doi.org/10.1046/j.1468-1331.2003.00611.x

32. Sian, J., Gerlach, M., Youdim, M.B.H., Riederer, P.: Parkinson's Disease: A Major Hypokinetic Basal Ganglia Disorder. Journal of Neural Transmission **106**, 443–476 (1999). URL http://dx.doi.org/10.1007/s007020050171. 10.1007/s007020050171