Feature Learning for Detection and Prediction of Freezing of Gait in Parkinson's Disease

Sinziana Mazilu[1], Alberto Calatroni[1], Eran Gazit[2], Daniel Roggen[1], Jeffrey M. Hausdorff[2], and Gerhard Tröster[1]

[1] Wearable Computing Laboratory, ETH Zürich, Switzerland
{sinziana.mazilu, alberto.calatroni, daniel.roggen, troester}@ife.ee.ethz.ch
[2] Laboratory of Gait and Neurodynamics, Tel Aviv Sourasky Medical Center
{erang, jhausdor}@tasmc.health.gov.il

Abstract. Freezing of gait (FoG) is a common gait impairment among patients with advanced Parkinson's disease. FoG is associated with falls and negatively impact the patient's quality of life. Wearable systems that detect FoG have been developed to help patients resume walking by means of auditory cueing. However, current methods for automated detection are not yet ideal. In this paper, we first compare feature learning approaches based on time-domain and statistical features to unsupervised ones based on principal components analysis. The latter systematically outperforms the former and also the standard in the field – Freezing Index by up to 8.1% in terms of F1-measure for FoG detection.
We go a step further by analyzing FoG prediction, i.e., identification of patterns (pre-FoG) occurring before FoG episodes, based only on motion data. Until now this was only attempted using electroencephalography. With respect to the three-class problem (FoG vs. pre-FoG vs. normal locomotion), we show that FoG prediction performance is highly patient-dependent, reaching an F1-measure of 56% in the pre-FoG class for patients who exhibit enough gait degradation before FoG.

Keywords: Unsupervised feature learning, Freezing of Gait, Parkinson's disease

1 Introduction

Freezing of gait (FoG) is a common gait impairment among patients with Parkinson's disease (PD), defined as a "brief, episodic absence or marked reduction of forward progression of the feet despite the intention to walk" [18]. Patients describe FoG as the feeling of having the feet glued to the ground and being temporarily unable to re-initiate gait. According to a survey of 6620 PD patients, 47% of the subjects reported regular freezing and 28% experienced FoG daily [13]. FoG is associated with falls [12], has substantial clinical and social consequences [6, 15] and is often resistant to pharmacological treatment [2].

Rhythmic auditory stimulation (RAS) was introduced as an assistive tool for FoG treatment [8]. RAS can be applied to produce a rhythmic ticking sound upon

detection of a FoG episode, to help the patient resume walking. Wearable systems based on motion sensors have been proposed for the detection and treatment of FoG with auditory stimulation [1, 11]. While RAS upon detection helps to shorten the duration of FoG episodes [1], it cannot avoid them altogether due to the latency of the detection, which is at best on the order of hundreds of milliseconds [11]. A step further is to predict when a patient is *about to* experience FoG, thus enabling preemptive RAS, with the goal of avoiding the FoG episodes. We call this *FoG prediction* as opposed to *FoG detection*.

There are some known specific properties that differentiate the sensor data during FoG episodes from normal walking (e.g., a large increase in the signal energy in the 3-8Hz frequency band [9,15]) and the gait of patients with FoG also differs between freezing episodes, compared to patients who do not experience FoG [10]. There are even suggestions of a characteristic change in the gait pattern just prior to the occurrence of a FoG episode; however, currently, there is no way of automatically identifying the prodromal state, when the normal gait pattern is about to transform into FoG.

The lack of physiological understanding of the gait deterioration preceding FoG makes it difficult to come up with a model or with problem-specific features based on expert knowledge. Moreover, walking styles of PD patients differ across subjects (including diverse motor anomalies) [17]. Thus eventual patterns in the data just before a FoG event will also likely be highly subject-specific. Nevertheless, previous work suggests that there is a deterioration of the normal gait before FoG, although this deterioration can be expressed in various ways [17–19].

In this work, we first formulate the *FoG detection* problem as a two-class classification problem: FoG versus normal locomotion. Similarly, we treat *FoG prediction* problem as a three class classification problem. Beside FoG and normal locomotion, we consider the walking periods before FoG episodes as a third class called *pre-FoG*. We hypothesize that there is a detectable deterioration of gait in this phase which precedes FoG. We assume different durations of the pre-FoG events, since these cannot be labeled by an expert, but can rather only be retrieved through data mining from segments of data preceding FoG events. We focus on the analysis of different feature extraction approaches that lead to a meaningful representation of both the FoG and the new pre-FoG class. The feature extraction approaches that we investigated are the following:

(a) Extraction of standard frequency-based features, namely Freezing Index and total energy in the frequency band 0.5-8 Hz. This is the current standard in the field and serves as a baseline [15].
(b) Extraction of various *hand-crafted* time-domain and statistical features, which are used in pattern recognition problems involving motion or human activity recognition.
(c) Unsupervised feature learning [20]. This method involves extraction of information from the raw data, without relying on domain specific knowledge, or on the availability of ground truth annotations. We evaluate the use of principal component analysis for extracting a compact representation of the structure of the signals.

The contributions of this work are summarized as follows:

1. We go beyond state of the art by explicitly introducing and performing a first step toward FoG prediction, as opposed to mere detection, thereby potentially allowing for the possibility of applying preemptive RAS;
2. We compare three methods for feature extraction in the FoG detection and FoG prediction problems and show that unsupervised feature learning outperforms on average standard feature extraction schemes in our real-life dataset;
3. We show that removal of pre-FoG sequences from the training data for FoG detection improves classification performance;
4. For FoG prediction, we show that, for some patients, gait anomalies associated with the upcoming onset of FoG can be detected, thereby allowing for an early intervention with RAS.

2 Related Work

FoG detection. Several research groups have proposed wearable systems for the detection of FoG episodes [1, 4, 5, 7, 11, 14–16, 21, 23]. Most sensor setups involve accelerometers and/or gyroscopes [1, 5, 11, 16, 21], extended with electroencephalography (EEG) [7] or electromyography (EMG) [4]. One standard feature which is extracted from the raw signals is the Freezing Index (FI), defined as the ratio between the power contained in the so-called *freezing* and *locomotion* frequency bands (3-8 Hz and 0.5-3 Hz respectively) [1, 11, 15]. This feature is convenient since it requires only FFT-computation. Other feature extraction approaches involve mixed time-frequency features [23] and entropy [21]. In [14], the authors investigated the use of time-domain and statistical features, together with FFT-features. Various classifiers have been used for the two-class classification problem (FoG versus no-FoG), including Decision Trees, Random Trees/Forests, Naive Bayes [21] as well as rule-based classifiers [5] and simple thresholds on the FI [1]. Overall, the different proposed approaches reach detection sensitivities that often exceed 80%, but the detection is performed with at best a latency of a few hundred milliseconds. Handojoseno and colleagues [7] make use of wavelet decomposition to analyze the dynamics of EEG signals during the onset and the freezing periods. Their aim was to achieve an early detection of FoG from brain activity that could, potentially, help patients to avoid an impending FoG episode. To our best knowledge, no attempts have been yet made at tackling the FoG prediction problem using just motion sensors. We therefore perform a first analysis in this direction.

Unsupervised feature learning. Automatic (unsupervised) feature extraction has been proposed in the context of human activity recognition based on motion sensors. Plötz et al. [20] argued that instead of using the explicit knowledge to select specific features, one can extract the core signal characteristics by means of principal components analysis. This allows one to uncover meaningful, low-dimensional representations of raw data without relying on domain-specific knowledge. The results on public activity recognition datasets showed that the

features learned in this unsupervised manner are more discriminative than state of the art representations based on time- and frequency-domain features. We propose to apply this method for detection and prediction of FoG, since the properties of the FoG and pre-FoG signals are subject-dependent and difficult to model.

3 Feature Extraction for FoG Detection and Prediction

The general process that we adopt for signal processing and classification is depicted in Figure 1. The set of operations is standard in pattern recognition problems involving motion data from on-body 3-dimensional accelerometers: sensor signals are sampled and sliced into partially overlapping windows. In each window, features are extracted and the resulting vectors are classified according to a pretrained model. In this work, we empirically set the window length to $1s$ (64 samples) with $0.25s$ of overlap (16 samples). We choose a Decision Tree classifier,

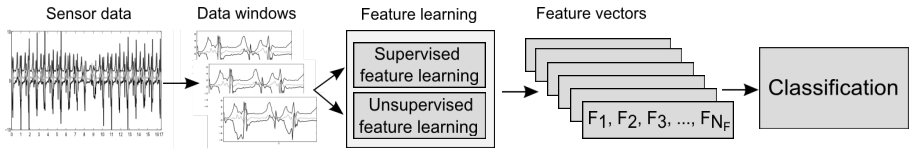


Fig. 1. Signal processing and classification for the detection and prediction of FoG.

because of its low computational cost when deployed. In this work, we focus on the selection of the appropriate features for detection and prediction of FoG, so the optimization of the classifier parameters is out of our scope. In the training phase of the system, we use feature ranking based on Mutual Information (MI) to rank the top discriminant time-based and statistical features [3]. We denote with N_F the number of top-ranked features retained in the classification process.

3.1 Feature Extraction Schemes

We choose three groups of features, the first of which has been already used in the context of FoG detection and is used here as a baseline. We call *supervised* the first two feature extraction approaches, since they involve features manually selected due to expert knowledge. The features are computed for each window.

Supervised: Domain-specific Feature Extraction. The first feature group contains the Freezing Index and the sum of energy in the freezing (3-8 Hz) and locomotory (0.5-3 Hz) frequency bands. These features are obtained by computing the FFT, followed by binning, in order to compute the spectral distribution of the energy in the desired bands.

Supervised: Feature Extraction of Time-domain and Statistical Features. The second group of features is often used in activity recognition [22].

Until now, only a small subset of these has been also applied to FoG detection [14]. We list the used features in Table 1. We extracted 18 features for each of the three accelerometer axes (x, y, z) and six features using data from all three axes. **Unsupervised Feature Learning.** For learning the implicit structure of

Table 1. Computed statistical features and their brief descriptions.

Axis Features		
No.	**Feature**	**Description**
1,2	Min, Max	Minimum and maximum of the signal
3	Median	Median signal value
4,5	Mean, ArmMean	Average value, and the harmonic average of the signal
6	Root Mean Square (RMS)	Quadratic mean value of the signal
7	GeoMean	Geometric average of the signal
8	Variance	Square of the standard deviation
9	Standard Deviation (STD)	Mean deviation of the signal compared to the average
10	Kurtosis	The degree of peakedness of the sensor signal distribution
11	Skewness	The degree of asymmetry of the sensor signal distribution
12	Mode	The number that appears most often in the signal
13	TrimMean	Trimmed mean of the signal in the window
14	Entropy	Measure of the distribution of frequency components
15	Asymmetry coefficient	The first moment of the data in the window divided by STD over the window
16	Range	The difference between the largest and smallest values of the signal
17	Zero Crossing Rate (ZCR)	Total number of times the signal changes from positive to negative or back, normalized by the window length
18	Mean Crossing Rate (MCR)	Total number of times the signal changes from below average to above average, normalized by the window length
Sensor Features		
No.	**Feature**	**Description**
55	Signal Magnitude Vector (SMV)	Sum of the euclidean norm over the three axis over the entire window normalized by the window length
56	Normalized Signal Magnitude Area (SMA)	Acceleration magnitude summed over three axes normalized by the window length
57,58,59	Eigenvalues of Dominant Directions (EVA)	Eigenvalues of the covariance matrix of the acceleration data along x, y, and z axis
60	Averaged Acceleration Energy (AAE)	Mean value of the energy over three acceleration axes

the data, each data window containing 64 samples for the three accelerometer axes is arranged into a 192-dimensional vector (the first three entries correspond to the first samples from the x, y and z axes, and so on). In the training phase, principal component analysis (PCA) is then applied to the whole training data matrix, obtained by stacking all the 192-dimensional vectors in the training set and disregarding class labels. This yields a projection matrix, which is then used in the testing phase to project the single data frames.

3.2 Assumptions about pre-FoG Events: From FoG Detection to Prediction

We assume that the gait cannot enter in the FoG state directly from walking state. Rather, we assume that prior to FoG, there is a gait deterioration that eventually leads to the FoG. This is represented by a transition period of variable duration T_{Prefog} that we refer to as the pre-FoG state. An example of FoG episode, with supposed pre-FoG, is shown in Figure 2. The optimal value of T_{Prefog} will be patient-dependent. The identification of segments of pre-FoG data is valuable both for FoG detection and prediction. **Making Detection**

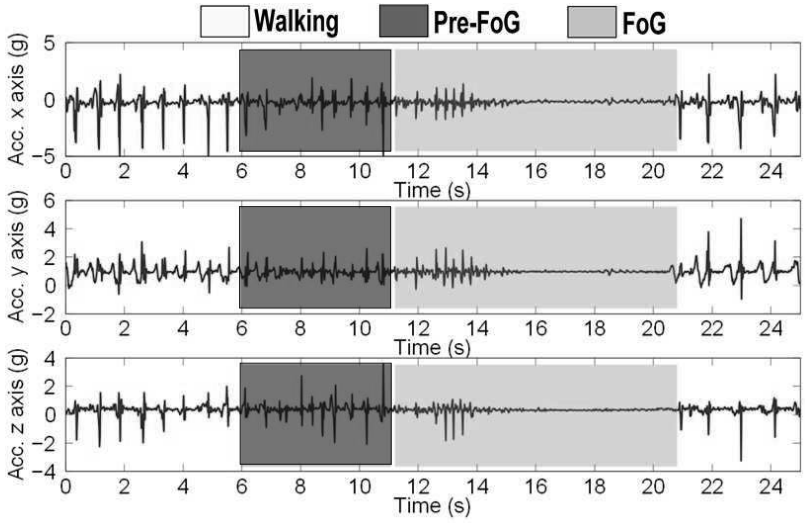


Fig. 2. An example of an accelerometer signal, on the three acceleration axes, that captures the motor variations in the gait of a patient with Parkinson's disease. The sequence contains normal gait, a FoG episode, preceded by a assumed pre-FoG period.

More Robust. For the detection problem, we set up a two-class classification problem. We name the two classes WALK (which includes instances of normal locomotion, including walking, standing, turning, etc.) and FoG, which represents the freezing episodes. In the training phase, we remove data for a duration of T_{Prefog} before each FoG event contained in the training set. This aims at having a more precise classifier model for the FoG and for the WALK classes.

Towards FoG Prediction. For prediction, we set up a three-class classification problem. Besides the two classes described above (WALK and FoG), we use the segments assumed to be in a pre-FoG state to build the model for the third class.

4 Dataset

We validated the proposed approach on the public available DAPHNet dataset[3] [1], which contains data collected from eight PD patients that experienced regular FoG in daily life. Data were recorded using three 3D accelerometers attached to the shank (above the ankle), the thigh (above the knee) and to the lower back of each subject. For our experiments here we focused on movement data recorded from the ankle, as the data from the other two sensors generally behave similarly. Subjects completed sessions of 20-30 minutes each, consisting of three walking tasks: (1) Walking back and forth in a straight line, including several 180-degrees turns; (2) Random walking with a series of initiated stops and 360 degrees turns; (3) Walking simulating activities of daily living, which included entering and leaving rooms, walking to the kitchen, getting something to drink and returning to the starting room with a cup of water.

Motor performances varied strongly among the participants. While some subjects maintained regular gait during nonfreezing episodes, others walked slowly and were very unstable. The DAPHnet dataset contains 237 FoG episodes; the duration of FoG episodes is between 0.5s and 40.5s (7.3±6.7s). 50% of the FoGs lasted for less than 5.4s and 93.2% were shorter than 20s. FoGs were labeled by physiotherapists using synchronized video recordings. The start of a FoG event was defined as the point when the gait pattern (i.e., alternating left-right stepping) was arrested, and the end of a FoG was defined as the point in time when the pattern was resumed.

5 Experiments and Evaluation

We performed two sets of experiments using the DAPHnet dataset described in Section 4: one for *FoG detection* and one for *FoG prediction*. For FoG-detection we ignored the pre-FoG sequences. For both sets of experiments and for two of the three groups of features introduced in Section 3.1, we varied the number of selected features N_F from 5 to 60 in steps of 5. This cannot be done for the domain-specific features, since they are only two - FI and total energy. We further characterized the influence of different choices of the pre-FoG duration on both the two-class and three-class problem, by sweeping the assumed pre-FoG duration in the range $T_{Prefog} \in \{1s, 2s, ..., 11s\}$.

The evaluation was performed on a patient-dependent basis. Since in each patient dataset the WALK class was over-represented compared to the FoG class, we chose to balance the data by having $size(\text{WALK}) = X * size(\text{FoG})$, where $X \in \{1.5, 2, ..., 10\}$. We performed an $N = 10$-fold cross validation, in which the training data contains N-1 parts from the FoG data, N-1 parts from normal locomotion data, and the testing data the rest. The data were split for each fold in such a way as to avoid having time-correlated chunks of the same FoG, WALK, or pre-FoG events in the training and testing data.

[3] www.wearable.ethz.ch/resources/Dataset

We report results in terms of overall patient datasets average sensitivity and average specificity of the FoG class, and F1-measures for FoG, WALK and pre-FoG classes, in a window-to-window comparison.

6 Results

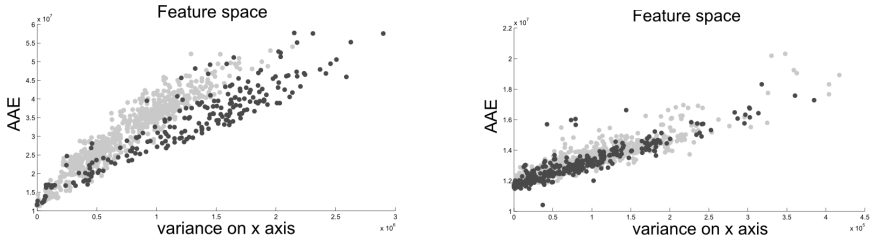
In the following, we analyze the performance of the different feature extraction strategies for the FoG-detection and FoG-prediction problems.

6.1 Time-domain and Statistical Features

The top ranked features based on MI for FoG detection are AAE, eigenvalues of dominant directions, range, variance, root mean square, and standard deviation (some features, like standard deviation and variance are of course strictly related). The top ranked features according to their MI are those computed from the entire data window (all axes), followed by those on x-axis. Table 2 shows the top k ranked features, for k ranging from 5 to 20. Note that selecting features

Table 2. Average top ranked features with Mutual Information.

Top k	**x axis**	**y axis**	**z axis**	**sensor**
Top 5	variance	-	variance	EVA (2 directions), AAE
Top 10	RMS, variance, range	variance, range	variance	EVA (3 directions), AAE
Top 15	variance, range, RMS, min, STD	variance, range, RMS	variance, range, RMS	EVA (3 directions), AAE
Top 20	max, RMS, variance, STD, min, range	variance, range, RMS, max, min, STD	RMS, variance, range, min	EVA (3 directions), AAE



(a) Patient 2 (b) Patient 8

Fig. 3. AAE vs. variance on x-axis for (a) Patient 2 data and (b) Patient 8 data.

that are ranked highly by the MI does not automatically guarantee that they are also discriminative enough. Figure 3 contains an example of distribution of the top ranked features AAE and variance on x-axis. For some patients these features are enough to distinguish between FoG and WALK – the two classes form two distinct clusters when represented by these two features. Still, this does not work for all the patients. For example, in the case of Patient 8, even if the

top ranked features with MI are the same, their discriminative power is lower – FoG is easily confused with WALK, when represented by the same two features.

6.2 FoG Detection

We compare the different feature learning approaches in terms of producing discriminative feature sets for distinguishing FoG and normal gait. In Table 3 the sensitivity, specificity and F1 measures for two feature-extraction methods are depicted as a function of number of features N_F used for classification, for a fixed duration of ignored data $T_{Prefog} = 3s$ before each FoG. The classification results are significantly improved when performing unsupervised feature learning compared to the results for the standard feature set, for values of $N_F < 30$. Classification based on PCA features also outperforms the one using FFT-based features, when using small number of PCA features (see Figure 4). In Figure 4,

Table 3. Average of the sensitivity, specificity and F1-measure for the FoG class for supervised and unsupervised feature extraction methods, in the two-class classification problem. The pre-FoG duration is fixed as $T_{Prefog} = 3s$.

Sensitivity (%)					
Features	5	10	15	20	25
Unsupervised	77.15	77.7	76.29	76.86	76.86
Supervised	67.8	68.53	69.42	66.65	67.58

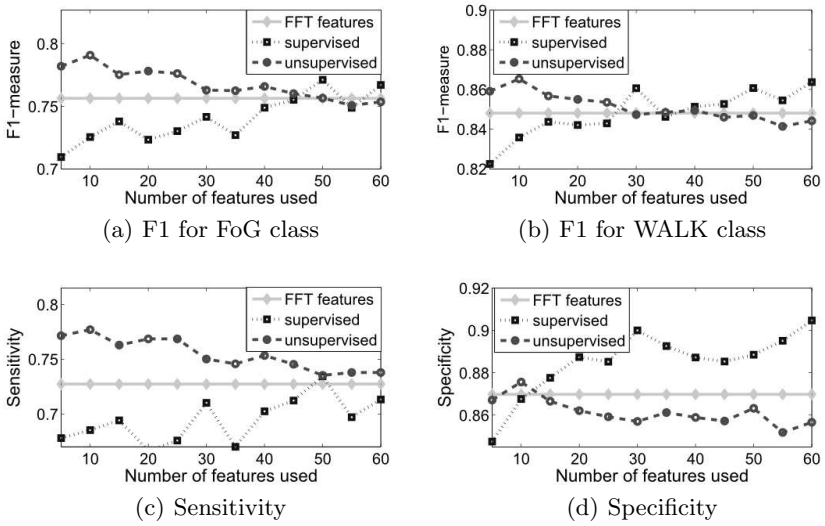
F1 (FoG)(%)					
Features	5	10	15	20	25
Unsupervised	78.2	79.09	77.53	77.62	76.29
Supervised	70.94	72.54	73.79	72.33	73.02

Specificity (%)					
Features	5	10	15	20	25
Unsupervised	86.71	87.56	86.65	86.21	85.52
Supervised	84.75	86.76	87.76	88.74	88.52

F1 (WALK) (%)					
Features	5	10	15	20	25
Unsupervised	85.91	86.53	85.67	85.5	85.35
Supervised	82.25	83.58	84.37	84.21	84.29

we observe that for larger values of N_F, the classification performances tend to decrease for unsupervised extracted features. PCA concentrates the variability and the useful information from the raw data in the first features. The usefulness of a feature decreases with its rank. However, our target is to use as few features as possible, as noted above.

In Figure 5, we present the classification results with $N_F = 10$ features, when varying the amount of discarded data before each FoG episode in the range $T_{Prefog} \in [1s, 11s]$, in steps of 1s. We observe that for both supervised and unsupervised features, FoG detection performance increases with the increase of T_{Prefog}, until reaching a plateau value at $T_{Prefog} = 5 - 6s$. This suggests that these discarded portions of data could contain properties that are different both from FoG and normal locomotion. In the next set of experiments, we analyze whether this dataset has specific properties that will lead to prediction of FoG episodes.



(a) F1 for FoG class (b) F1 for WALK class

(c) Sensitivity (d) Specificity

Fig. 4. Sensitivity, Specificity and F1 measures for FoG detection when using different values for N_F unsupervised and supervised extracted features. The amount of discarded data is fixed to $T_{Prefog} = 3s$ before each FoG episode.

6.3 Towards FoG Prediction.

In the previous experiments, we observed that discarding T_{Prefog} data preceding each FoG episode improved the FoG detection results for all types of feature extraction. We now present the results for the three-class classification problem, where we use the discarded chunks as examples of the pre-FoG class. As a first step, we analyzed the impact of the addition of this third class to the mutual information between the various features and the classes.

Mutual Information. We compare the mutual information of the features in the FoG-detection and FoG-prediction problems. Figure 6 shows an example of MI values computed for both supervised and unsupervised features, on the same Patient 2 dataset, in case of FoG detection and FoG prediction problems. We observe that all MI values improve for top ranked features, when adding the third class. This suggests that the pre-FoG data can indeed be representative.

Performance of FoG Prediction In the next experiments, we set $N_F = 10$, and we varied $T_{pre-FoG}$ from 1s to 6s, in steps of 1s. We stopped at $6s$ because a further increase did not improve the FoG prediction results.

Figure 7 shows the variation of F1-measures for all the three classes versus the value of T_{Prefog}. We first observe that, like in the two-class classification problem (FoG detection), the unsupervised features perform better than the supervised ones. Second, the F1-measures for the FoG and WALK classes are smaller than in the two-class problem. This is expected since we are trying to solve a classification task with one extra class – pre-FoG – which is identified using an assumption on its presence and duration, which leads inevitably to a

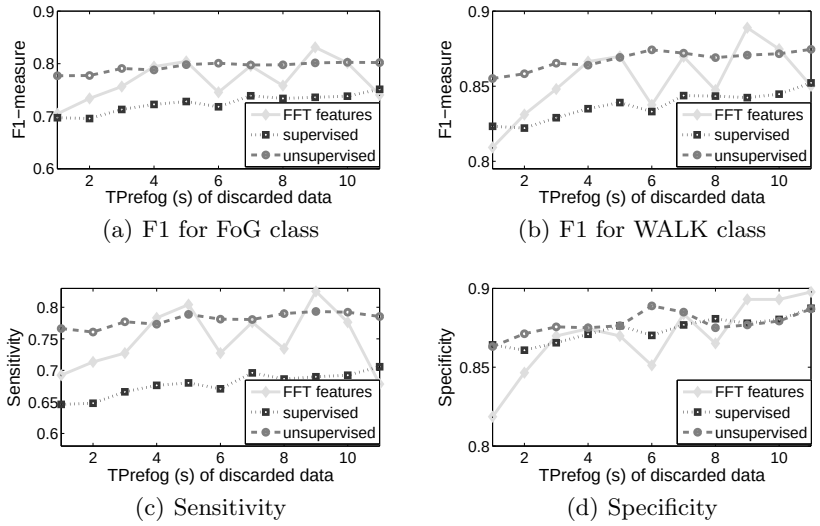


Fig. 5. Sensitivity, Specificity and F1-measures for FoG detection when using $N_F = 10$ unsupervised and supervised extracted features. The amount of discarded data varies between $T_{Prefog} = 1s$ and $T_{Prefog} = 11s$.

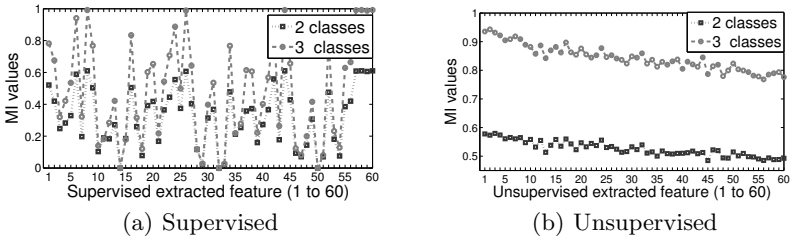
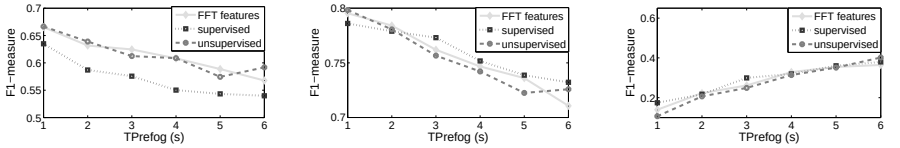


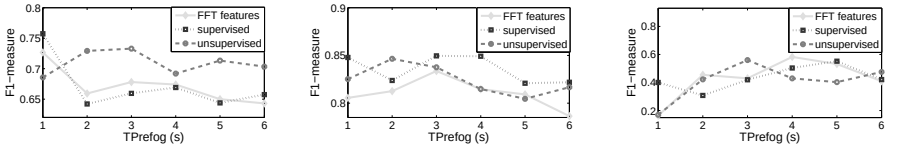
Fig. 6. MI values for all $N_{total} = 60$ computed features, both with supervised and unsupervised methods, for FoG detection and FoG prediction. $T_{Prefog} = 5s$.

noisy training. Instances of the pre-FoG class will indeed not always be radically different from WALK or FoG instances, which will introduce confusion. Nevertheless, we identify a trade-off: we can use the unsupervised feature extraction to perform FoG prediction at the expense of performance on detection. Too small values of T_{Prefog} lead to poor results on the F1-measures, because the pre-FoG class is not representative enough. On the contrary, indefinitely increasing T_{Prefog} improves the F1-measure only for this class, but dramatically decreases that of the other classes. This is due to the fact that the pre-FoG and WALK classes become more and more similar.

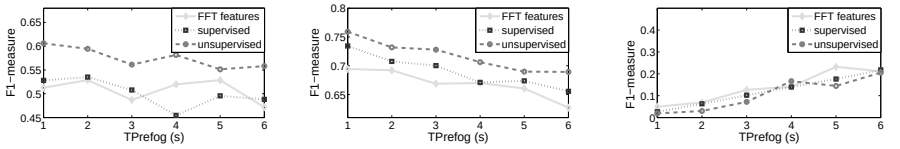


(a) F1 for FoG class (b) F1 for WALK class (c) F1 for pre-FoG class

Fig. 7. F1-measures for FoG prediction when using $N_F = 10$ unsupervised and supervised extracted features, with $T_{Prefog} \in [1s, 6s]$.



(a) FoG class – PD3 (b) WALK class – PD3 (c) pre-FoG class – PD3



(d) FoG class – PD8 (e) WALK class – PD8 (f) pre-FoG class – PD8

Fig. 8. F1-measures for FoG prediction on Patient 3 and Patient 8 data when using $N_F = 10$ unsupervised and supervised extracted features, with $T_{Prefog} \in [1s, 6s]$.

Figure 8 displays the F1-measure variations for the datasets of Patient 3 (PD3) and Patient 8 (PD8). In the case of PD3, when using unsupervised extracted features, for T_{Prefog} periods of $2s$ and $3s$, the F1-measures increase for all the classes. The F1-measures for the pre-FoG class are 0.42 for $T_{Prefog} = 2s$ and 0.56 for $T_{Prefog} = 3s$. So there are common patterns in the 2s or 3s before FoG episodes that are distinct from WALK and FoG. The same behavior of F1-measures is observed for supervised extracted features, but with a delay compared to using unsupervised features. For $T_{Prefog} = 1s$ supervised features even outperform the unsupervised ones. The likely reason is that with such short pre-FoG durations, PCA in unable to capture the structure of that class. On the other hand, for PD8, an increase of T_{Prefog} leads to a constant decrease in performance for the detection of the FoG and WALK classes, while having a small increase for pre-FoG (along with a decrease of the WALK F1-measure). That shows that WALK and pre-FoG are similar, thus using two distinct classes just leads to confusion in the classification. So, for this patient, we cannot extract specific patterns that could differentiate pre-FoG from the global WALK class.

6.4 Discussion

The classification performance for FoG detection is not as high as that reported in other works. We believe this is mainly due to a less optimistic evaluation scheme, where we selected the training and testing data in each fold to avoid having training and testing data chunks coming from the same FoG episodes at once. This should lead to a more realistic estimate of the performance of a real-world deployed system. Furthermore, FoG-prediction performances vary considerably across subjects. We can claim that for some PD patients, like Patient 3, there are patterns, visible in the accelerometer data, that are characteristic of the pre-FoG class, making it different from the normal locomotion class. These patients exhibit a deterioration of the walk just before FoG episodes.

There are some limitations related to the assumtpions on the pre-FoG class:

- Different nature of the FoG episodes. Some FoG events occur when the subject starts walking, meaning there is no gait before the FoG, so that the gait deterioration that we assume to exist in the pre-FoG phase does not exist, rendering FoG prediction especially challenging for those cases.
- The duration of the pre-FoG class T_{Prefog} is considered to be fixed for each patient. Nevertheless, the pre-FoG pattern duration will probably vary even for different FoG episodes for the same patient. This means that an optimal training set for the pre-FoG class for a single patient might need to contain segments having different values for T_{Prefog}. In order to determine the correct value for each single instance, an approach could involve a direct monitoring of the variation of the features, to detect when they start changing from the normal status to the FoG status.

7 Conclusion

In this work, we analyzed the performance of three feature extraction approaches for detecting freezing of gait in patients with Parkinson's disease. Features based on time-domain and statistical features where compared to unsupervised ones based on principal components analysis, while Freezing Index (FI) was used as a baseline reference. We tested the approaches on acceleration data collected at the ankle from patients that experienced FoG in daily-life. Unsupervised feature learning outperformed FI by up to 7.1% and the time-domain and statistical features by up to 8.1% in terms of F1-measure for FoG detection.

We went a step further by analyzing FoG prediction, i.e. identification of patterns (pre-FoG) occurring before FoG episodes, based only on acceleration data. The purpose is to predict FoG so to assist patients in avoiding freezing periods altogether. For this, we assume that walking sequences of a fixed length T_{Prefog} just previous to a FoG episode have different characteristics compared to normal locomotion patterns and to FoG. On the three-class problem (FoG vs. pre-FoG vs. normal locomotion) we obtained results highly patient-dependent, reaching an F1-measure of 56% in the pre-FoG class for one patient. The identification of pre-FoG patterns is also beneficial for the simple FoG detection: when pre-FoG

data are discarded from the training set, performance on FoG detection increases for all the feature extraction methods.

The use of unsupervised features is a promising avenue, since these capture important variations in the data, without the bias of an expert choosing features manually and without any prior knowledge of the class labels. In order to improve the results, other, more complex unsupervised methods for feature learning will be tested (PCA using nonlinear kernels, deep learning). Furthermore, additional unobtrusive sensing modalities could be considered (e.g. gyroscopes). Finally, our assumption on a fixed duration of the pre-FoG class for all FoG events might need to be revised. To this end, methods monitoring directly changes in the extracted features could be beneficial for identifying the actual start of the pre-FoG phases, where present.

Acknowledgments. The research leading to these results has received funding from the European Union - Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 288516 (CuPiD project).

References

1. M. Bächlin, M. Plotnik, D. Roggen, I. Maidan, J. M. Hausdorff, N. Giladi, and G. Tröster. Wearable Assistant for Parkinson's Disease Patients with the Freezing of Gait Symptom. *IEEE Transactions on Information Technology in Biomedicine*, 14:436–446, 2010.
2. B. R. Bloem, J. M. Hausdorff, J. E. Visser, and N. Giladi. Falls and Freezing of Gait in Parkinson's Disease: A Review of Two Interconnected, Episodic Phenomena. *Movement Disorders*, 19:871–884, 2004.
3. R. Cilibrasi and P. M. B. Vitányi. Clustering by Compression. *IEEE Transactions on Information Theory*, 51:1523–1545, 2005.
4. B.T. Cole, S.H. Roy, and S.H. Nawab. Detecting Freezing of Gait During Unscripted and Unconstrained Activity. In *Engineering in Medicine and Biology Society (EMBC)*, pages 5649–5652, 2011.
5. M. Djurić-Jovičić, N.S. Jovičić, I. Milovanović, S. Radovanović, N. Kresojević, and M.B. Popović. Classification of Walking Patterns in Parkinson's Disease Patients Based on Inertial Sensor Data. In *10th Symposium on Neural Network Applications in Electrical Engineering*, pages 3–6, 2010.
6. N. Giladi and J. M. Hausdorff. The Role of Mental Function in the Pathogenesis of Freezing of Gait in Parkinson's Disease. *Journal of the Neurological Sciences*, 248:173–176, 2006.
7. A.M. Ardi Handojoseno, James M. Shine, Tuan N. Nguyen, Yvonne Tran, Simon J.G. Lewis, and Hung T. Nguyen. The Detection of Freezing of Gait in Parkinson's Disease Patients Using EEG Signals Based on Wavelet Decomposition. In *Engineering in Medicine and Biology Society (EMBC)*, pages 69–72, 2012.
8. T. Hashimoto. Speculation on the Responsible Sites and Pathophysiology of Freezing of Gait. *Parkinsonism & Related Disorders*, 12:55–62, 2006.
9. J. M. Hausdorff, Y. Balash, and N. Giladi. Time Series Analysis of Leg Movements During Freezing of Gait in Parkinsons Disease: Akinesia, Rhyme or Reason? *Physica A: Stat Mechanics & Appl*, 321:565–570, 2003.

10. J. M. Hausdorff, J. Schaafsma, Y. Balash, A. Bartels, T. Gurevich, and N. Giladi. Impaired Regulation of Stride Variability in Parkinsons Disease Subjects with Freezing of Gait. *Experimental Brain Research*, 149:187–194, 2003.
11. E. Jovanov, E. Wang, L. Verhagen, M. Fredrickson, and R. Fratangelo. deFog - a Real Time System for Detection and Unfreezing of Gait of Parkinson's Patients. In *31th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2009.
12. M. Latt, S. Lord, J. Morris, and V. Fung. Clinical and Physiological Assessments for Elucidating Falls Risk in Parkinson's Disease. *Movement Disorders*, 24:1280–1289, 2009.
13. M. Macht, Y. Kaussner, J. C. Möller, K. Stiasny-Kolster, K. M. Eggert, H.-P. Krüger, and H. Ellgring. Predictors of Freezing in Parkinson's Disease: A Survey of 6,620 Patients. *Movement Disorders*, 22:953–956, 2007.
14. S. Mazilu, M. Hardegger, Z. Zhu, D. Roggen, G. Tröster, and J. M. Hausdorff. Online Detection of Freezing of Gait with Smartphones and Machine Learning Techniques. In *6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, pages 123–130, 2012.
15. O. Moore, C. Peretz, and N. Giladi. Freezing of Gait Affects Quality of Life of Peoples with Parkinson's Disease Beyond its Relationships with Mobility and Gait. *Movement Disorders*, 22:2192–2195, 2007.
16. K. Niazmand, K. Tonn, Y. Zhao, U.M. Fietzek, F. Schroeteler, K. Ziegler, A.O. Ceballos-Baumann, and T.C. Lueth. Freezing of Gait Detection in Parkinson's Disease Using Accelerometer based Smart Clothes. In *Biomedical Circuits and Systems Conference (BioCAS)*, pages 201–204, 2011.
17. A. Nieuwboer, R. Dom, W. De Weerdt, K. Desloovere, L. Janssens, and V. Stijn. Electromyographic Profiles of Gait Prior to Onset of Freezing Episodes in Patients with Parkinson's Disease. *Brain*, 127:1650–1660, 2004.
18. J. G. Nutt, B. R. Bloem, N. Giladi, M. Hallett, F. B. Horak, and A. Nieuwboer. Freezing of Gait: Moving Forward on a Mysterious Clinical Phenomenon. *The Lancet Neurology*, 10:734–744, 2011.
19. M. Plotnik, N. Giladi, and J. M. Hausdorff. Is Freezing of Gait in Parkinsons Disease a Result of Multiple Gait Impairments? Implications for Treatment. *Parkinson's Disease*, 2012:doi:10.1155/2012/459321, 2012.
20. T. Plötz, N.Y. Hammerla, and P. Olivier. Feature Learning for Activity Recognition in Ubiquitous Computing. In *22nd International Joint Conference on Artificial Intelligence (IJCAI)*, 2011.
21. E. Tripoliti, A. Tzallas, M. Tsipouras, G. Rigas, P. Bougia, M. Leontiou, S. Konitsiotis, M. Chondrogiorgi, S. Tsouli, and D. Fotiadis. Automatic Detection of Freezing of Gait Events in Patients with Parkinson's Disease. *Computer Methods and Programs in Biomedicine*, 110(1):12–26, 2013.
22. M. Zhang and A. Sawchuk. A Feature Selection-Based Framework for Human Activity Recognition Using Wearable Multimodal Sensors. In *International Conference on Body Area Networks (BodyNets)*, 2011.
23. Y. Zhao, K. Tonn, K. Niazmand, U.M. Fietzek, L.T. D'Angelo, A. Ceballos-Baumann, and T.C. Lueth. Online FOG Identification in Parkinson's Disease with a Time-Frequency Combined Algorithm. In *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, pages 192 –195, 2012.