Detecting Freezing of Gait with Earables Trained from VR Motion Capture Data

Nobuyuki Oishi
Wearable Technologies Lab, University of Sussex
Brighton, United Kingdom
n.oishi@sussex.ac.uk

Benedetta Heimler
Center of Advanced Technologies in Rehabilitation, Sheba Medical Center
Ramat Gan, Israel
benedetta.heimler@sheba.health.gov.il

Lloyd Pellatt
Wearable Technologies Lab, University of Sussex
Brighton, United Kingdom
lp349@sussex.ac.uk

Meir Plotnik
Center of Advanced Technologies in Rehabilitation, Sheba Medical Center
Ramat Gan, Israel
meir.plotnik@sheba.health.gov.il

Daniel Roggen
Wearable Technologies Lab, University of Sussex
Brighton, United Kingdom
daniel.roggen@ieee.org

ABSTRACT
Freezing of Gait (FoG) is a common disabling motor symptom in Parkinson’s Disease (PD). Auditory cueing provided when FoG is detected can help mitigate the condition, for which earables are potentially well suited as they are capable of motion sensing and audio feedback. However, there are no studies so far on FoG detection at the ear. Immersive Virtual Reality (VR) combined with video-based full-body motion capture has been increasingly used to run FoG studies in the medical community. While there are motion capture datasets collected in such an environment, there are no datasets collected from IMU placed at the ear. In this paper, we show how to transfer such motion capture datasets to IMU domain and evaluate the capability of FoG detection from ear position in an immersive VR environment. Using a dataset of 6 PD patients, we compare machine learning-based FoG detection applied to the motion capture data and the virtual IMU. We have achieved an average sensitivity of 80.3% and an average specificity of 87.6% on FoG detection using the virtual earable IMU, which indicates the potential of FoG detection at the ear. This study is a step toward user-studies with earables in the VR setup, prior to conducting research in over-ground walking and everyday life.

CCS CONCEPTS
• Applied computing → Consumer health.

KEYWORDS
Freezing of Gait; Virtual IMU; Earables; Parkinson’s Disease

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ACM Reference Format:

1 INTRODUCTION
Parkinson’s Disease (PD) is the second most widespread neurodegenerative disorder worldwide, affecting 1-2% of the general population and 1% of the population aged over 60 [24]. The number of people with PD is estimated to double from today to over 12 million by 2040 worldwide due to the ageing of the population [8].

Freezing of Gait (FoG) is a common disabling motor symptom in advanced PD and defined as “a brief, episodic absence or marked reduction of forward progression of the feet despite the intention to walk” [18]. It leads to increased fall risk, avoidance behaviour, wheelchair use, and poor quality of life.

Auditory cueing is a non-pharmacological method for mitigating FoG, which alleviates serious side effects of anti-FoG drugs. Auditory cueing can be provided continuously, synchronized with the gait, or on the detection of FoG, with the latter avoiding issues of acceptance and habituation [21]. Wearable devices are ideal for sensing and delivering biofeedback. Most works to date focus on detecting FoG with Inertial Measurement Units (IMUs) attached to the middle or lower part of the body [19]. Earable devices are potentially well suited for such an assistive application as they are capable of motion sensing and audio feedback in a single device [11]. However, there are no studies on FoG detection from IMU at the ears; hence, no such datasets are readily and publicly available to our knowledge. Though a dedicated dataset could be collected with an ear-worn IMU, it requires time, adds to costs, and has ethical implications, as vulnerable people would need to be brought in for data collection using a sensor placement that was unconfirmed to be suitable to detect FoG prior to our work.

An increasing number of biomechanics studies uses immersive VR with omnidirectional treadmills as it allows to run sophisticated experiments to understand contextual cues affecting FoG occurrences by running experiments in simulated indoor and outdoor scenarios in safe and reproducible settings [4]. FoG experiments in VR often contain physiological sensors and exhaustive video-based motion capture including head markers. Therefore, the data collected in such VR environments are expected to be useful to
develop methods for detecting FoG and providing biofeedback in wearable devices, which, however, requires translating the motion capture data into virtual IMU data to be sensed on the body. Such methods could be trialled in the VR environment first before being used in over-ground walking.

In this study, we show how to exploit the video-based motion capture data collected from PD patients in Motek’s CAREN (Computer Assisted Rehabilitation Environment) High-End [17] immersive VR environment to evaluate the capability of FoG detection from ear position. Figure 1 shows the setup of the environment.

The contributions of this paper are: 1) a mapping algorithm from motion-capture data to 6-axis virtual IMU to simulate earable IMU data with the novelty of including the treadmill motion into the mapping, 2) validation of the mapping on a healthy subject performing various head movements while standing and walking, 3) development of machine learning-based FoG detection pipeline with resampling to deal with imbalanced distribution of FoG and non-FoG samples, 4) evaluation on FoG detection performance from ear position using motion capture data and virtual earable IMU data at the ear and additionally at the left heel as baselines.

![Figure 1: The VR environment provides a 360° field of view, an omnidirectional treadmill, a safety harness, and full body motion capture with 41 markers. It simulates a corridor setting including narrow passages and turns, which are known as common triggers of FoG.](image)

2 RELATED WORK

A review of wearable FoG detection shows that sensor placement is mostly on the lower-body [19]. Initial work defined a “freeze index” based on the high-frequency components of leg movements [16]. The “freeze index” is represented by the power in the “freeze” band (3-8 Hz) divided by the power in the “locomotor” band (0.5–3 Hz). This signal-processing approach was later extended for real-time operation [1], showing a detection accuracy of 78.1% sensitivity and 86.9% specificity. Recently, machine learning and deep learning methods could be trialled in the VR environment first before being evaluated in over-ground walking.

As common triggers of FoG.

3 MAPPING FROM VICOM TO IMU

Our mapping algorithm consists of three steps: adjust Vicon marker positions using the treadmill’s belt speed data, calculate acceleration and angular velocity at the left ear position, rotate the calculated values to the target sensor’s coordinate system. The first step is for simulating virtual IMU data that would be collected overground.

The CAREN system reports the positions of the attached markers in the Vicon’s coordinate system and the belt speed data at time \( t \). We use the 4 markers (LFHD, LBHD, RFHD, and RBHD) placed on the head, assuming that an IMU is positioned between LFHD and LBHD. L, R, F, B, and HD of the marker names stand for left, right, front, back, and head, respectively. (See Figure 2)

![Figure 2: The head markers, the ear positions, and the coordinate system of the virtual earable.](image)

First, marker positions are adjusted by adding the distance the belt has moved up to time \( t \) to the walking direction as in (1). Here, \( p_{xt} \) is a marker’s position along the X axis, and \( fs \) is the sampling rate of the Vicon system.

\[
p'_{xt} = p_{xt} + \sum_{i=0}^{t} \frac{b_{xt}}{fs} \quad (1)
\]

To get acceleration and angular velocity of the virtual earable IMU, we need to first calculate the orientation of the virtual earable \( (\phi_{y}, \Theta_{y}, \psi_{y}) \) at each timestamp \( t \) with respect to the Vicon’s coordinate system as in (2). Here, for example, \( p_{LEARy} \) is the y-axis component of the left ear position at time \( t \).

Figure 2 also visually explains how the orientation is obtained.
Angular velocity of the virtual earable $\omega_v$ is then calculated by differentiating the obtained orientation $\omega_e$ calculated as in (3). Here, $a_v$ is the virtual earable’s acceleration along the X, Y and Z, $g$ is the gravitational acceleration.

$$a_v = R(-\phi_t, -\theta_t, -\psi_t)(av_t + g) \tag{3}$$

Finally, we rotate this virtual earable IMU data to align it with the target sensor’s coordinate system. We calculate the rotation matrix $R$ from the virtual earable IMU to the target IMU sensor using the Kabsch algorithm [10]. The Kabsch algorithm finds the optimal rotation matrix that minimizes the root mean squared error between a set of vectors. We use virtual IMU data and real IMU data collected concurrently to find the rotation matrix. The target sensor’s virtual IMU is obtained by $a_{tar} = Ra_e, \omega_{tar} = Ra_e$.

To validate this method, we collected data from the four Vicon head markers and an OPAL IMU sensor [9] attached near the left ear of a healthy subject. This validation is about ensuring the correctness of physical and geometrical transformations, independent of the health status of the user. The dataset consists of four scenarios: 1) standing while looking in various directions, 2) standing with quick head movements, 3) walking while looking in various directions, 4) walking with quick head movements. The total duration of the dataset is 206 seconds. Vicon data and OPAL data were collected at 120 Hz and 128 Hz, respectively. The Vicon data was upsampled to 128 Hz to match the sampling frequency of the OPAL data. During the recording, the subject simultaneously tapped the RFHD marker and the OPAL IMU multiple times; synchronization was performed using these reference points. We used the data collected during scenario 3 for the input of the Kabsch algorithm.

Figure 3 shows the error histograms of reconstructed data. Overall, the 95% values of the reconstructed accelerometer and gyroscope values were within [-154.177, +134.859] mG and [-13.724, +18.612] degree/s, respectively. This method estimates the target sensor’s values with smaller error compared to the previous work [20].

Noise in the target sensor and suboptimal Vicon marker placement are considered to be the possible causes of the errors. In this experiment, we did not perform any calibration or filtering process on the target sensor. In addition, our method expects the line between LFHD and LBH HD markers and the line between LEAR and REAR to be orthogonally aligned to correctly calculate the orientation. However, it is unlikely that the actual markers are attached so that these lines are perfectly orthogonal. Thus, some further calibration may be needed in the future.
trained with the resampled data, and the performance is evaluated using the test data. We use a random forest classifier.

We performed a window-based evaluation. Feature selection and the average FoG detection performance with the virtual IMU at the ear was 6pp lower than with the virtual IMU at the heel, and the average sensitivity with the motion capture data splitting, feature selection, resampling, and model training were performed for each patient individually.

The average number of selected features for each patient is 150.7 ± 32.5. Common frequency domain features across all PD patients were the FFT coefficients between 1-3 Hz for Acc-YM and Gyro-X and 6-7.5 Hz for Acc-YM and Gyro-Z, which correspond to the “locomotion” band and the upper half of the “freeze” band.

Representative FoG detection results are shown in Table 1 with the numbers of FoG and voluntary stop feature vectors. The results are the average of five attempts of running the evaluation, including data splitting, feature selection, resampling, and model training. The average FoG detection performance with the virtual IMU at the ear was 80.3% sensitivity and 87.6% specificity. The resampling improved the average sensitivity by about 6 pp but decreased the average specificity by about 3 pp. The average sensitivity with the virtual IMU at the ear was 6pp lower than with the virtual IMU at the heel, and the average sensitivity with the motion capture coordinates at the ear was 8.5pp lower than with the motion capture coordinates at the heel.

### 5 DISCUSSION

About 80% or more of the FoG were detected from the ear while keeping high specificity for most patients. However, for P3, we can see large gaps in sensitivity between the ear and the heel, which may indicate that FoG detection at the ear may be particularly difficult for certain users. We have not looked at user-independent performance because a user-specific model may be preferable in a medical context; it does not necessarily mean that a lot of effort is required to create a personalized model. For example, Bächlin et al. showed a very high benefit of a user-specific model with just 2 settings, low and high sensitivity [1]. Further improvements of the pipeline to further tune it to the particular characteristics of FoG at the ear could include: extract more ear position-specific features, consider longer time dependencies and gait deterioration before FoG onsets, and design for easier personalization.

The challenges faced by this dataset are mainly due to the small amount of data and the short FoG events. The models were trained and tested using only a few FoG events for some subjects. Also, we limited the window length to 2.0 seconds because the majority of the FoG events were shorter than this, even though FoG detection accuracy was reportedly improved by increasing the window size up to 4.5 seconds [1]. We performed binary classification (Fog/non-Fog) following existing studies [1, 16, 23], but it should be noted that the total duration of the voluntary stop events accounted for only 2.6% of the dataset, which is less than one-third of the total duration of the FoG events in the dataset. It might have caused overfitting to the dataset. While results show the high sensitivity and specificity, there is nevertheless a need for further evaluation including a closer look at voluntary stops on larger-scale datasets.

### 6 CONCLUSION

In this paper, we proposed the mapping method from motion capture to virtual earable IMU to exploit valuable motion-capture datasets collected from PD patients in VR and evaluated the capability of FoG detection at the ear. Based on our investigation, this is the first study on FoG detection at the ear. The mapping method estimated the target sensor values with smaller errors. The experiments using the virtual IMU demonstrated the potential of FoG detection with earable IMU for most users. While this work was evaluated in VR, it did account for the belt speed of the treadmill. Therefore, it is expected that the FoG detection models trained with the virtual IMU would work with the real IMU data in overground walking. This must be verified through future experiments. These findings provide evidence to support future research and indicate that there may be a pathway to realise a wearable FoG assistant providing biofeedback upon FoG detection from a single earable device, which would significantly enhance comfort compared to the multi-device systems reviewed by Sweeney et al. [21]. Given that users have different preferences, continuation of this work would provide a new option for FoG assistance.

### Table 1: Sensitivity and specificity for FoG detection with virtual IMU and motion capture at the ear and the heel

<table>
<thead>
<tr>
<th>Pt</th>
<th>Virtual IMU</th>
<th>Motion capture</th>
<th>No. of Stop</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ear</td>
<td>Heel</td>
<td>Ear</td>
</tr>
<tr>
<td>1</td>
<td>0.886</td>
<td>0.889</td>
<td>0.863</td>
</tr>
<tr>
<td>2</td>
<td>0.898</td>
<td>0.901</td>
<td>0.896</td>
</tr>
<tr>
<td>3</td>
<td>0.529</td>
<td>0.736</td>
<td>0.698</td>
</tr>
<tr>
<td>4</td>
<td>0.855</td>
<td>0.767</td>
<td>0.759</td>
</tr>
<tr>
<td>5</td>
<td>0.890</td>
<td>0.918</td>
<td>0.923</td>
</tr>
<tr>
<td>6</td>
<td>0.805</td>
<td>0.970</td>
<td>0.839</td>
</tr>
<tr>
<td>Avg</td>
<td>0.803</td>
<td>0.862</td>
<td>0.817</td>
</tr>
</tbody>
</table>
ACKNOWLEDGMENTS

This study was supported in part by the Israel Science Foundation (ISF) grant #1657-16.

REFERENCES


