
Human Gait Monitoring Using Footstep-Induced Floor Vibrations Across Different Structures

Mostafa Mirshekari

Carnegie Mellon University
Pittsburgh, PA 15213 USA
mmirshekari@cmu.edu

Pei Zhang

Carnegie Mellon University
Moffett Field, CA 94035 USA
peizhang@cmu.edu

Jonathon Fagert

Carnegie Mellon University
Pittsburgh, PA 15213 USA
jfagert@cmu.edu

Hae Young Noh

Carnegie Mellon University
Pittsburgh, PA 15213 USA
noh@cmu.edu

Amelie Bonde

Carnegie Mellon University
Moffett Field, CA 94035 USA
amelie@cmu.edu

Abstract

In this paper, we present a structure-adaptive approach for monitoring human gait using footstep-induced floor vibrations. Human gait information is critical for timely and accurate assessment and diagnosis of many health conditions. Footstep-induced vibration monitoring has been introduced in prior works as an accurate and cost-effective mean to provide continuous gait monitoring in indoor environments. Prior works in this field typically rely on inverse modeling approaches to map signal characteristics to gait information, but are limited due to varying signal characteristics in each new deployment structure. These changing characteristics introduce errors when attempting to use a model trained in one structure in another structure. To overcome this challenge, we propose a structure-adaptive approach that enables the transfer of a model trained in one structure to a new structure where no training data is present. To this end, we first find a lower-dimension space in which vibration responses share similarity across various structures and label the data in the target structure. Then, we use the labeled data to train a new model for footstep detection and monitoring in that structure. We evaluated our approach through real-world experiments in three types of structures (wood, concrete, and steel). Our evaluation results show that our approach achieves a footstep detection accuracy of between 85 and 97 percent and F1-score of between 86

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 ACM 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.
UbiComp/ISWC'18 Adjunct, October 8–12, 2018, Singapore, Singapore
© 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-5966-5/18/10...\$15.00
<https://doi.org/10.1145/3267305.3274187>

and 97 percent, which represents 3-10X and 2.2-15X improvements over the baseline approach, respectively.

Author Keywords

Transfer Learning; Gait Monitoring; Footstep Detection; Structural Vibration

Introduction

Obtaining occupant gait information (such as balance, step frequency, number of steps, etc.) is a critical component of overall health assessment. For example, timely information about a patient's gait in a medical setting can enable caretakers to monitor and make timely decisions regarding treatment of common diseases such as chronic obstructive pulmonary disease (COPD), muscular dystrophy, and dementia [9, 17]. An emerging sensing approach for obtaining occupant gait information passively and sparsely monitors human activity by utilizing their footstep-induced structural vibrations [7, 13].

Prior works in this field have leveraged footstep-induced structural/floor vibrations to extract gait characteristics and information with applications ranging from footstep location and walker's identity [7, 14, 13], to step frequency and left/right balance [8, 5, 15, 2]. Some other examples of applications of vibration-based sensing are human-computer interaction [11], hand-washing monitoring [3], and activity monitoring [10, 6]. These works primarily focus on supervised machine learning or inverse modeling of the recorded vibration signatures to obtain occupant information. While effective, this is limited in wide-spread applications due to the need to "re-train" the model for each new structure in which the sensing system is deployed. This re-training is necessary because of changes in the structural response characteristics in various structures which results in different temporal and spectral vibration signal characteristics [8]. Thus, the prior works will likely incur increases

in deployment costs or lower performance for typical real-world scenarios (i.e. deployments in multiple buildings).

To address these deployment challenges, we present an approach for modeling footstep-induced vibrations that adapts to different structures. This modeling technique is then utilized to distinguish footstep events from other impulsive excitations. We accomplish this through a model transfer method that considers the shared principal components of vibration events (e.g., footsteps) across a source structure (in which labelled data is available) and a target structure (in which no labelled data is available).

This approach assumes that 1) the physical phenomena (the footstep force transfer) remains the same, whereas 2) the underlying structure is different. Due to the similarity of physical phenomena, there exists a lower-dimension space in which the vibration responses still share some similarities across the source and target structures. Hence, in this space, the footstep models in each structure are similar and the model trained in the source structure can be utilized for footstep detection in the target structure. However, due to the structural differences, this principal component-based model transfer alone is not sufficient for estimating the detailed footstep model in the target structure, and errors may still be present. To overcome this challenge, using the labeled data from the source structure, we train a model for the target structure in the lower-dimensional space and use this model to label the data for the target structure. Then, we select the labeled data that has the highest prediction confidence to train a new model in the target structure. By learning the final model using the data from the target structure, we improve the model accuracy. We validate our approach by conducting real-world experiments in three locations: 1) a wood-framed elder care facility, 2) a concrete framed campus building, and 3) a steel-framed elder care facility.

The core contributions of this work are as follows:

1. We present a footstep modeling technique that can adapt to varying structures, enabling gait health tracking through footstep-induced floor vibrations without re-training.
2. We characterize the common physical properties of each deployment structure by mapping into a lower-dimension space composed of the principal shared components to achieve model transfer between structures without additional labeled data.
3. We validate our approach using real-world structural vibration data in multiple structures.

In the paper that follows we describe our model transfer approach in detail in the Model Transfer for Gait Analysis section. Next, we evaluate the performance of our method in the Evaluation section. Lastly, we provide conclusions and discuss future work utilizing this approach.

Model Transfer for Gait Analysis

Our model transfer approach for structure-adaptive gait analysis using floor vibrations has three primary modules: 1) an impulse detection module which monitors incoming floor vibration signals and isolates the signals with larger variation as potential footstep events using the technique presented in the "Impulse Detection" section below, 2) a structure-informed model transfer module that labels detected impulses in the new structure using the source/original structure through the process discussed in the "Structure-Informed Model Transfer" section below, and 3) a labelled training module which trains a new footstep classification model for the target/new structure with the labeled data from module 2 (as described in the "Labeled Training" section below). Figure 1 provides an overview of our approach

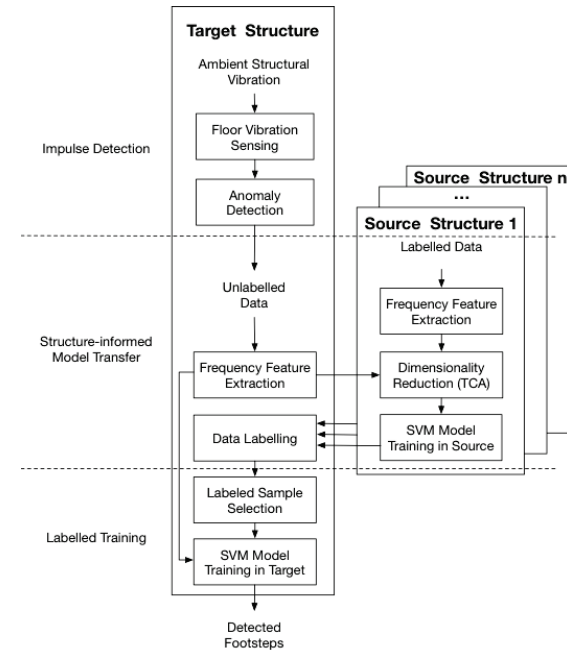


Figure 1: Approach Overview.

and each module is described in more detail in the sections that follow.

Impulse Detection

Our impulse detection module continuously monitors vibration activity and isolates the portions of the signal that contain "impulse-like" excitations. This is accomplished by establishing a baseline vibration measurement level (corresponding to ambient conditions), and detecting times when the vibration signal variance exceeds that of these baseline ambient conditions. To achieve this we first measure the ambient vibrations of the floor structure (i.e. when no other

activity is taking place) using a geophone sensor [16]. In our application, these geophone sensors are placed directly on the floor surface and are characterized as being low-cost and low-distortion sensing devices which record the vertical (uniaxial) vibration velocity of the surface they are attached to (in this case the floor structure), and convert this velocity to a voltage that can be collected with an analog-to-digital converter (ADC). To improve the signal resolution, we amplify it using an operational amplifier (op-amp). An example of the sensing system employed is shown in Figure 3.

To establish the ambient vibration threshold, we consider a Chi-squared-based anomaly detection algorithm [1]. The collected vibration signal typically consists of a mixture of impulsive events (e.g. footsteps, doors closing, objects falling, etc.), periodic excitations (e.g. HVAC systems, machinery, etc.), and white background noise (e.g. measurement/sensor noise). Our intuition in using the Chi-squared anomaly detection algorithm is that the impulsive events have a larger variance than the periodic and white noise excitations (i.e. the ambient conditions). Using this intuition, we first establish the variance of the ambient noise, and then continuously monitor the variance of the signal with a sliding window to determine if the variance exceeds the ambient conditions. Specifically, we accomplish this by evaluating the null hypothesis for each window (H_0) against its alternative hypothesis (H_1). In this case, H_0 is defined as: $\sigma^2 w = \sigma^2 n$ (i.e., signal is noise), and H_1 is defined as $\sigma^2 w > \sigma^2 n$ (i.e., signal is an impulsive event), where $\sigma^2 n$ and $\sigma^2 w$ represent the variance of the noise and the windowed signal, respectively. Figure 2 shows an example of the detected signals using this approach. The red boxes in the figure represent the time windows where an impulsive event is detected, which could include excitations such as footsteps, objects falling, and doors closing. These detected

impulses are then used as the input for the following modules.

Structure-Informed Model Transfer

Our task in the structure-informed model transfer module is to use the labeled data from the source structure(s) along with unlabeled data from the target structure to train a model that can effectively predict footstep labels for the target structure. This involves a three-step approach: 1) find a low-dimensional latent space across the different structures represented by a set of shared principal components, 2) use the labeled source structure data in that space to train a footstep model, and 3) predict the sample labels in the target structure with the model. We use the physical insight that the frequency representation of the vibration signals represents the dynamic response of the structure. This allows us to use the frequency representation as a feature for estimating the footstep models.

To find the principal components, we adapt the domains of our source structures to our target structure domain via transfer component analysis (TCA) [12]. This projects the source and target feature spaces into a latent space which attempts to both preserve the variance of the data and minimize the distance between the data distributions of the different structures. Mathematically, TCA solves the following optimization problem,

$$\begin{aligned} & \underset{W}{\text{minimize}} && \text{tr}(W^T K L K W) + \mu \text{tr}(W^T W) \\ & \text{subject to} && W^T K H K W = I \end{aligned}$$

where K is the kernel matrix of the source and target data, L is a coefficient matrix, and H is a centering matrix, and W is the projection matrix. Overall, $\text{tr}(W^T K L K W)$ represent the maximum mean discrepancy (MMD) which is the distance between the distributions in source and target, $\mu \text{tr}(W^T W)$ is the regularization term to prevent over-

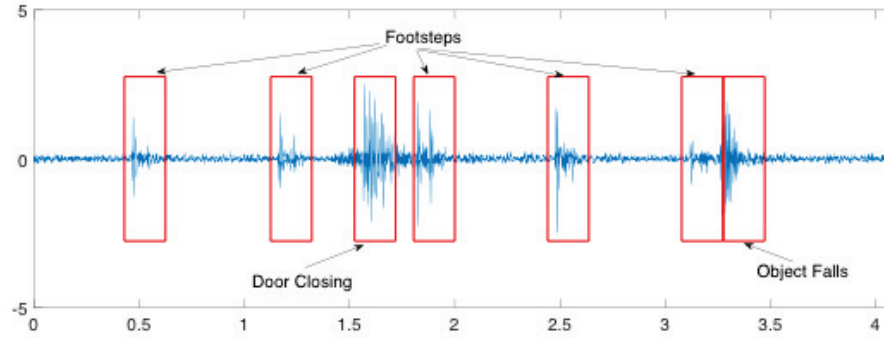


Figure 2: An example of the impulse detection.

fitting, and $W^T K H K W$ is the covariance matrix of the data. Interested readers can refer to [12] for more information about TCA.

Let Y_s and Y_t be the label spaces in the source and target structures, and X_s and X_t be the feature spaces in the source and target structures, respectively. Let $P(X_s)$ and $Q(X_t)$ be the marginal distributions of X_s and X_t , respectively. The key assumption behind TCA is that while $P(X_s)$ and $Q(X_t)$ are different distributions, $P(Y_s|\phi(X_s)) \approx P(Y_t|\phi(X_t))$ where ϕ is the mapping function. Based on this assumption, in the latent space, the footstep model trained in the source structures can be utilized to predict the labels in the target structure.

We use binary support vector machine (SVM) classifiers to train the footstep models; a common approach which does not require us to make assumptions about the data distributions. SVM maximizes the distance of the sample points from a dividing hyperplane, and outputs a class label and a score [4]. The score represents the distance of the

sample point from the dividing hyperplane and can be used as a prediction confidence measure, with larger distances corresponding to higher confidences. The sign of our final score for each sample point corresponds to the label of that point. We train a separate model for each source structure, and combine the predictions of multiple source structures by combining the final scores. By adding the SVM scores from different structures, we give more weight to labels with higher confidence scores. We use the physical insight that source structures that are more similar to the target structure will provide more accurate labeling, which is reflected in higher confidence scores.

Labeled Training

The labeled training module selects a subset of the newly-labeled data with highest confidence to train a new model for the target structure. The data distribution in the target structure differs from the data distributions of the source structures, causing inaccuracies in the newly-labeled data. Therefore, to get the most accurate labeling, we choose a subset of our newly-labeled data with high confidence

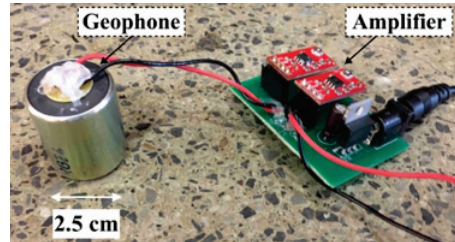


Figure 3: An example of the sensing node.

scores, with a heuristically chosen confidence threshold. Higher confidence scores reflect target structure data that is more similar to the source structure data we trained on in the previous step, and for which the TCA assumption is most likely to hold. We now have a subset of labeled data for the target structure, and can use a standard supervised learning algorithm now that we have achieved domain adaptation. We train our final footstep model using SVM in the original feature space of frequency representation (i.e. not the principal component space). This new model for the target structure adapts to the changes between the source and target structures, increasing the overall footstep detection accuracy.

Evaluation

To understand the performance of the model transfer approach across multiple real-world structures, we have conducted a set of experiments with human participants in three different structures. We first introduce the experimental setup, then we present and analyze the accuracy of gait analysis using our approach in the "Footstep Classification Robustness in Different Buildings" section.



Figure 4: Experiment locations.

Implementation and Experimental Setup

To evaluate our approach, we have utilized a sensing system which measures the floor vibration with a geophone. Geophone is a sensor which converts the vertical velocity of the floor to voltage [16]. Figure 3 shows an example of the sensing node. The collected signals are amplified on the order of 200-2000X. After amplification and depending on the structure type and footstep strike energy, the effective sensing range of our system for footstep detection is up to 20 meters in diameter. Amplified signals are then digitized and transferred to a server using a 24-bit A/D converter. To evaluate how robust our approach is in different structures, we have conducted experiments with one person walking in three different structures in Pittsburgh, PA as shown in Figure 4. The locations considered are: 1) a carpeted wooden floor in the third floor of the Baptist Homes senior care building, 2) a carpeted concrete hallway on the ground level of a campus building at Carnegie Mellon University, 3) and a carpeted metal deck floor on the second floor of the Vincentian Home senior care facility. The first observed natural frequency of these structures are 16.02, 23.83, and 14.84 Hz, respectively.

Figure 5 presents the sensing configuration used for all three structures. The footstep data from each structure includes 20 traces of 7 steps in each structure. Furthermore, to evaluate how accurate our approach is in distinguishing footsteps from other impulsive excitations, we have included door closings, dropping objects (ball and keychain) on the floor, and hammer strikings. These are injected locations within the 3.5 meters around the sensor. We have included 50 of each of these excitations.

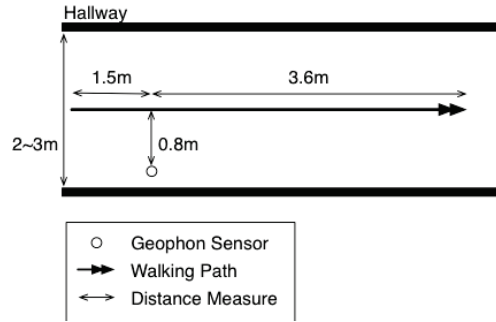


Figure 5: The sensing configuration.

Footstep Classification Robustness in Different Buildings

In this section, we evaluate the accuracy and robustness of our classification approach which employs footstep-induced floor vibrations for gait analysis and step counting. To this end, we discuss and compare the robustness of our transfer learning-based approach with the baseline approach in the three structures. The baseline approach trains the footstep model in one structure and then tests the model in a different structure. Two metrics are utilized for this comparison: 1) classification accuracy which is the ratio of the correctly labelled samples to all the samples; and 2) F1 score which is the harmonic mean of the precision and recall rate and can be found through the following equation. The results of

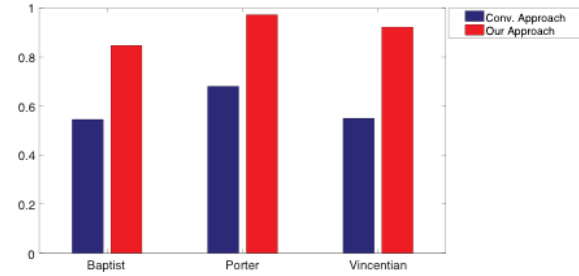


Figure 6: Accuracy in three structures using raw and our approach.

our approach are compared with a baseline approach which trains a model using the signals in the source structures and then utilizes this model for predicting the samples in the target structure. Further, the utilized source structures depend on the target structure (e.g., if Baptist is the target, Vincentian and Porter are used as the source.)

$$F_1 = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (1)$$

Where,

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives} \quad (2)$$

$$Recall = \frac{TruePositives}{TruePositives + FalseNegatives} \quad (3)$$

Figure 6 and Figure 7 show the comparison of the classification accuracy and F1 score in each structure, respectively. Based on these results, in the first structure, our approach exhibits 0.85 accuracy and 0.86 F1-score which

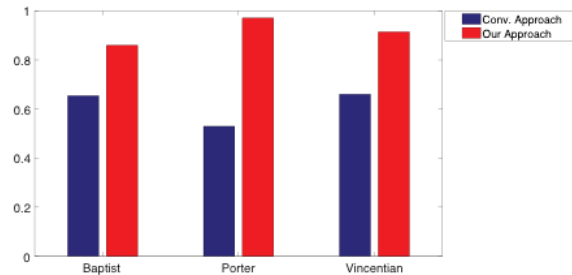


Figure 7: F1-Score in three structures using raw and our approach.

is equivalent to 3X and 2.2X reduction in classification error over the baseline approach which results in 0.54 accuracy and 0.65 F1 score. Similarly, in the second structure our approach results in 0.97 accuracy and 0.97 F1-score compared to 0.68 accuracy and 0.53 F1-score using the baseline approach (equivalent to 10X and 15X reduction in error). Finally, in the third structure, our approach results in accuracy of 0.92 and F1-score of 0.91 compared to the accuracy of 0.55 and F1-score of 0.66 using the baseline (equivalent to 5.6X and 3.8X reduction in error). Higher frequency noise components in Baptist homes due to its wooden floor structure result in less similarity with the other structures and reduces the classification accuracy. However, consistent improvement in all the three structures shows that our approach is more robust to the changes in the structure.

Discussion and Future Work

In this paper, we have focused on utilizing two structures as the source to find the footstep model in the target structure. However, we expect that the accuracy can be further improved by considering more structures as the source.

As part of our future work, we plan to explore the relationship between the amount of available data (e.g., number of structures) and the transfer model performance. Additionally, as discussed in the Evaluation section, the similarity of the source and target structures is an important factor in the model performance. As part of our future work, we plan to explore different metrics of similarity (e.g., cross-correlation of the frequency representation of ambient noise) and study their effect on the performance of the model.

Conclusion

In this paper, we presented a structure-adaptive approach for monitoring human gait using footstep-induced structural vibrations. Our approach transfers a trained model in one structure to a new structure by considering a projection into a lower-dimension latent space based on the spectral principle components in the source structure. Using this latent space, a footstep model is trained in the source structure and used to label data in the target structure. The labeled data in the target structure is then used to build a new model for that structure, which can then be used for detecting footstep events. Through real-world experimental evaluation, our approach achieves up to 15X improvement over a baseline approach. Our structure-adaptive gait monitoring approach enables continuous monitoring of human gait in indoor environments without the need for timely and costly re-training for each new structure.

Acknowledgements

This research was partially supported by NSF (CNS-1149611 and CMMI-1653550), Google, Intel, and Highmark. The authors would also like to acknowledge Vincentian Nursing Home and Baptist Nursing Home for providing deployment sites to conduct experiments and collect data.

REFERENCES

1. M. Baron. 2013. *Probability and statistics for computer scientists*. CRC Press.
2. J. Fagert, M. Mirshekari, S.a Pan, P. Zhang, and H.Y. Noh. 2017a. Characterizing left-right gait balance using footstep-induced structural vibrations. In *SPIE 10168, Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems*, Vol. 10168. 10168 – 10168 – 9.
3. Jonathon Fagert, MOSTAFA MIRSHEKARI, SHIJIA PAN, PEI ZHANG, and HAE YOUNG NOH. 2017b. Monitoring Hand-Washing Practices Using Structural Vibrations. *Structural Health Monitoring 2017 shm* (2017).
4. Jerome Friedman, Trevor Hastie, and Robert Tibshirani. 2001. *The elements of statistical learning*. Vol. 1. Springer series in statistics New York, NY, USA:.
5. Mike Lam, Mostafa Mirshekari, Shijia Pan, Pei Zhang, and Hae Young Noh. 2016. Robust occupant detection through step-induced floor vibration by incorporating structural characteristics. In *Dynamics of Coupled Structures, Volume 4*. Springer, 357–367.
6. Ramin Madarshahian, Juan M Caicedo, and Nicholas Haerens. 2019. Human Activity Benchmark Classification Using Multilayer Artificial Neural Network. In *Dynamics of Civil Structures, Volume 2*. Springer, 207–210.
7. M. Mirshekari, S. Pan, J. Fagert, E.M. Schooler, P. Zhang, and H.Y. Noh. 2018. Occupant localization using footstep-induced structural vibration. *Mechanical Systems and Signal Processing* 112 (2018), 77–97.
8. M. Mirshekari, P. Zhang, and H.Y. Noh. 2017. Calibration-free footstep frequency estimation using structural vibration. *Dynamics of Civil Structures 2* (2017), 287–289.
9. T. Oberg, A. Karsznia, and K. Oberg. 1993. Basic gait parameters: reference data for normal subjects, 10-79 years of age. *Journal of Rehabilitation Research and Development* 30, 2 (1993), 210.
10. Shijia Pan, Mostafa Mirshekari, Jonathon Fagert, Ceferino Gabriel Ramirez, Albert Jin Chung, Chih Chi Hu, John Paul Shen, Pei Zhang, and Hae Young Noh. 2018. Characterizing human activity induced impulse and slip-pulse excitations through structural vibration. *Journal of Sound and Vibration* 414 (2018), 61–80.
11. Shijia Pan, Ceferino Gabriel Ramirez, Mostafa Mirshekari, Jonathon Fagert, Albert Jin Chung, Chih Chi Hu, John Paul Shen, Hae Young Noh, and Pei Zhang. 2017. Surfacevibe: vibration-based tap & swipe tracking on ubiquitous surfaces. In *Information Processing in Sensor Networks (IPSN), 2017 16th ACM/IEEE International Conference on*. IEEE, 197–208.
12. S. Pan, Tsang, I. W., J. T. Kwok, and Q. Yang. 2011. Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks* 22, 2 (2011), 199–210.
13. S. Pan, T. Yu, M. Mirshekari, J. Fagert, A. Bonde, O. J. Mengshoel, and P. Zhang. 2017. FootprintID: Indoor Pedestrian Identification through Ambient Structural Vibration Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 89.
14. Jeffrey D Poston, R Michael Buehrer, and Pablo A Tarazaga. 2017. Indoor footstep localization from structural dynamics instrumentation. *Mechanical Systems and Signal Processing* 88 (2017), 224–239.

15. Y. Reuland, Pai, S. G., S. Drira, and I. F. Smith. 2017. Calibration-free footstep frequency estimation using structural vibration. *Dynamics of Civil Structures* 2 (2017), 49–56.
16. I/O sensor Nederland. 2006. SM-24 Geophone Element. (2006). P/N 1004117.
17. H. Visser. 1983. Gait and balance in senile dementia of Alzheimer's type. *Age and Ageing* 12, 4 (1983), 296–301.