



Characterizing the variability of footstep-induced structural vibrations for open-world person identification

Yiwen Dong^{a,*}, Jonathon Fagert^b, Hae Young Noh^a

^a Department of Civil and Environmental Engineering, Stanford University, 473 Via Ortega, Stanford, CA 94305, USA

^b Department of Engineering, Baldwin Wallace University, 275 Eastland Road, Berea, OH 44017, USA

ARTICLE INFO

Communicated by D. Bernal

Keywords:

Structural vibration

Gait biometrics

Open-world

Nonparametric Bayesian

Variability analysis

Feature transformation

ABSTRACT

Person identification is important in providing personalized services in smart buildings. Many existing studies focus on *closed-world* person identification, which only identifies a fixed group of people who have training data; however, they assume everyone has pre-collected data, which is not practical in real-world scenarios when newcomers are present. To overcome this drawback, *open-world* person identification recognizes both newcomers and registered people, which opens up new opportunities for smart building applications that involve newcomers, such as smart visitor management, customized retail, personalized health monitoring, and public emergency assistance. To achieve this, structural vibration sensing has various advantages when compared with the existing sensing modalities (e.g., cameras, wearables, and pressure sensors) because it only needs sparsely deployed sensors mounted on the floor, does not require people to carry devices, and is perceived as more privacy-friendly. However, one fundamental challenge in analyzing footstep-induced structural vibration data is its high variability due to the structural heterogeneity and the footstep variations. Therefore, it is difficult to distinguish different people given this high variability within each person, and it is more challenging to recognize a new person as that data is unobserved before.

In this paper, we characterize the variability in footstep-induced structural vibration to develop an open-world person identification framework. Specifically, we address three variability challenges in developing our method. First, the high variability within each person comes from multiple sources that are entangled in the vibration signals, and thus is difficult to be decomposed and reduced. Secondly, the distribution of features extracted from the vibration signals is irregularly shaped, and therefore is difficult to model. Moreover, the identity of the next person is correlated with the previous observations, which makes the identification process more complicated. To overcome these challenges, we first characterize multiple variability sources and design a transformation function that results in signal features that are less variable within one person and more separable between different people. We then develop a modified Chinese Restaurant Process (mCRP) for nonparametric Bayesian modeling to capture the irregularly shaped feature patterns both from local and global perspectives. Finally, we design an adaptive hyperparameter α that represents the prior probability of newcomers at each observation, which keeps updating depending on the time, location, and previous predictions. We evaluate our approach through walking experiments with 20 people across 2 different structures. With only 1 pre-recorded person at each structure, our method achieves up to 92.3% average accuracy with randomly appearing newcomers.

* Corresponding author.

E-mail address: ywdong@stanford.edu (Y. Dong).

1. Introduction

Structural vibrations induced by human footsteps contain gait-based biometric information of the footstep owner, which is an important indicator for providing personalized smart building services, such as person identification, health monitoring, and activity recognition [1–3]. Among them, person identification is essential because it is the premise for many personalized services in smart buildings. Many existing studies mainly focus on person identification in *closed-world* settings which only recognizes people who have pre-collected training data [4]; however, it is impractical to assume that everyone's data is pre-collected in real-world scenarios where newcomers may be present. This calls for *open-world* person identification, which identifies both new and existing people as they arrive. *Open-world* person identification opens up new opportunities for smart building applications that accommodate newcomers, such as visitor management, smart customer service for retail, personalized health monitoring, and public emergency assistance.

Footstep-induced structural vibration sensing has various advantages when compared with the existing sensing modalities for open-world person identification. While previous studies mainly use cameras, they require direct line-of-sight and have raised privacy concerns due to appearance exposures [5–7]. Other sensing modalities, such as radio frequency sensors, acoustic sensors, and pressure mats, can be used to reduce this privacy issue [8–13], but they have operational limitations such as walking trajectory, noise level, and dense deployment constraints. To overcome these limitations, our prior work explored footstep-induced floor vibration sensing, which only needs sparsely deployed sensors, is wide-ranged, non-intrusive, and is perceived as more privacy-friendly [1,2,14,15]. It has been successful in various applications in closed-world settings [1,3,16–20].

However, one fundamental challenge in open-world person identification using footstep-induced structural vibration is the multi-source, high variability in vibration signals, which causes high variability within a person's footstep data and low separability between people's gait-based features. Specifically, there are three unique challenges associated with this high variability issue. First, the variability within each person comes from multiple sources, including variations in footstep forces and structural heterogeneity across different locations, so it is difficult to understand and reduce the variability for each person. Secondly, the footstep features extracted from the signal have irregularly shaped distributions within each person — it is normally distributed around the mean, but has thick and asymmetrical tails at the boundary. Therefore, it is difficult to model both the global data distribution and the local patterns at the tails. Thirdly, the predictions of the previously observed persons affect the expectation of the next observed persons. Therefore, making independent predictions leads to bias accumulation in the long run.

In this paper, we characterize the variability of footstep-induced structural vibration to develop an open-world person identification framework. To reduce the variability, we first decompose its multiple sources and design a transformation function based on the dominant source by establishing the physical relationship between footsteps at various locations. We then formulate an optimization problem to find the transformation parameters that map features to a new space with minimum within-person variability and maximum between-people separability. To model the irregularly shaped feature distribution, we develop local and global modeling schemes. For local modeling, we introduce a modified Chinese Restaurant Process (mCRP) to accurately estimate the complex feature patterns at the decision boundaries between people; For global modeling, we leverage the Dirichlet Process Mixture Models (DPMM) to achieve open-world learning of overall feature patterns among the existing people and the newcomers. In order to keep track of the correlations between previous and future observed persons, we design an adaptive hyperparameter α that represents the prior probability of newcomers for each observation. This α alters after each observation based on the previous predictions, which significantly reduces the bias induced by the independent assumption between observations.

The core contributions of this paper are:

- We present the first open-world person identification framework based on footstep-induced structural vibration, which continuously identifies newcomers and registered people by modeling and learning their walking patterns on the fly.
- We address the challenges caused by the high variability in footstep-induced structural vibration data by variability analysis and feature transformation, global and local feature distribution modeling based on Nonparametric Bayesian models, and adaptive online updating for each newly observed sample.
- We evaluate the framework in a real-world setting with twenty people on two structures with different materials and structural configurations and conduct a sensitivity analysis to assess the framework performance under various conditions.

The remainder of the paper is organized as follows: We first review the existing work on person identification and structural vibration sensing to provide a background for this study (Section 2). Then, we present the variability characterization approach to understand the effect of structural and human influences in the footstep-induced structural vibration (Section 3). Then, we introduce the main components in our framework and present our solutions to address the high variability challenge (Section 4). Then, we present a real-world experiment and discuss the evaluation results in various test cases (Section 5). After that, a discussion of potential future studies (Section 6). Finally, a conclusion of our work is drawn (Section 7).

2. Related work

This section presents the related work for the paper. We first explore the existing studies on person identification frameworks, and then compare the biometrics and sensing modalities for person identification. Finally, we focus on footstep-induced structural vibration sensing for human gait characterization tasks.

2.1. Existing person identification frameworks

Person identification frameworks consist of multiple sub-branch tasks. Previous studies have explored occupant detection, verification, closed-world person identification, new person discovery (alternatively named as imposter detection), person re-identification, and open-world person identification [7,10,15,21–25]. Typically, different learning models are designed specifically for single tasks due to the difference in their scopes and objectives. For example, fixed-group identification is a multi-class supervised classification problem; newcomer discovery is a novelty detection problem; person re-identification is a semi-supervised learning problem; and online person identification is an unsupervised online learning problem.

2.2. Biometrics and sensing modalities for person identification

Different biometrics have been used in person identification frameworks, including fingerprints, face, iris, voice, footsteps, and so on [23,26,27]. They are widely applicable in government authorities, public infrastructures, and commercial security platforms. The choice of biometrics is based on criteria such as accuracy, security, and ease of collection [12,28,29]. Fingerprint-based identification is based on minutiae matching, which has sparse and highly distinguishable features among different people. Nevertheless, collecting fingerprints requires touching on optical/capacitive/ultrasound scanners, which raises sanitary concerns [30,31]. Non-contact biometrics are thus becoming popular for their ease of use, such as iris, face, and voice recognition. Despite the high accuracy in iris scanning, it requires close proximity to a specialized infrared radiation (IR) camera [28,32]. Face and voice recognition mimic the person recognition process in human brains and are widely used in personal devices [33]. However, more privacy concerns are raised towards these methods as they are also easily recognized/captured by unauthorized individuals [31,34].

2.3. Footstep-induced floor vibration sensing for human gait characterization

Floor vibration sensing was originally developed to ensure infrastructure safety and human comfort/serviceability. For example, the pedestrian-induced vibrations on the London Millennium Bridge have been thoroughly studied to understand the cause of the excessive vibration on the opening day and ensure structural safety for subsequent usage [35,36]. Other studies have explored the crowd-induced vibrations in public infrastructures such as sports stadiums, musical halls, shopping centers, and airports to understand the influence of pedestrian behavior on the structural systems [37,38]. Other sensing approaches have also been developed to detect pedestrians and their body dynamic parameters (e.g., damping, stiffness, mass), such as infrared and microwave sensors, and force plates [39,40]. However, they are limited in resolution and unable to produce accurate person identification.

Recent studies have leveraged footstep-induced floor vibrations for human gait characterization, which have been shown to demonstrate promising performances among various different tasks [41], including but not limited to: human activity monitoring [3,14], occupant localization and tracking [16,18,42–45], person identification [1,17], gait health monitoring [2,46–48], surface touch control [49], and so on.

For the topic of person identification through gait characterization, previous works developed various physics-informed machine learning techniques for more accurate, interpretable person identity prediction. To extract gait-related information, prior work on gait health monitoring found amplitudes from the low-frequency bands (0–200 Hz) are sensitive to the changes in gait patterns [1,50,51]. In addition, wavelet coefficients, power spectral density, and auto-regressive coefficients are also shown to be effective in predicting gait abnormally [2]. With these features, physics-informed machine learning models are developed to make predictions for gait disorders and person identity [47,48]. These models are based on support vector machines (SVM) [1], convolutional neural networks (CNN), Gaussian Mixture Models (GMM), etc. However, these closed-world learning approaches fail to accommodate previously unobserved people or behaviors. As a result of this restriction, these models are unable to discover new people or behaviors, leading to prediction errors when out-of-distribution samples are present.

3. Variability analysis of footstep-induced structure vibration

In this section, we analyze the variability in footstep-induced structural vibration in order to understand, model, and reduce the variability. We first characterize multiple variability sources in order to design a transformation function to reduce the variability within each person (Section 3.1). Then, we analyze the unique feature distribution patterns in order to develop a suitable model for identity prediction (Section 3.2).

3.1. Variability source characterization for within-person variability reduction

We characterize the sources of variability in footstep-induced structural vibration in order to develop a transformation function that reduces the variability within each person. These variability sources include natural variations in human gait (footstep variability) and the heterogeneity in structure across different locations (structural variability).

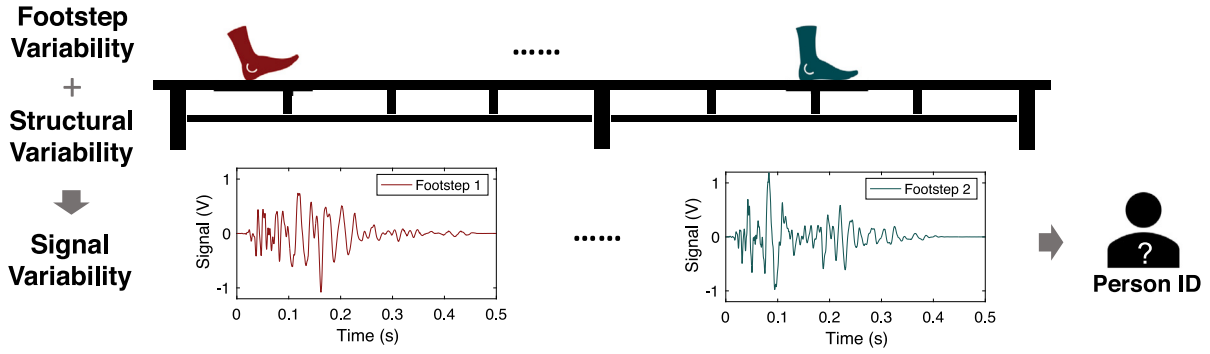


Fig. 1. Both the footstep variations (footstep variability) and the structure heterogeneity (structural variability) contribute to the overall variability of the structural vibration signals, leading to high variability within each person's data. Therefore, it is difficult to distinguish different people, and it is more challenging to recognize a new person whose gait pattern has not been observed before.

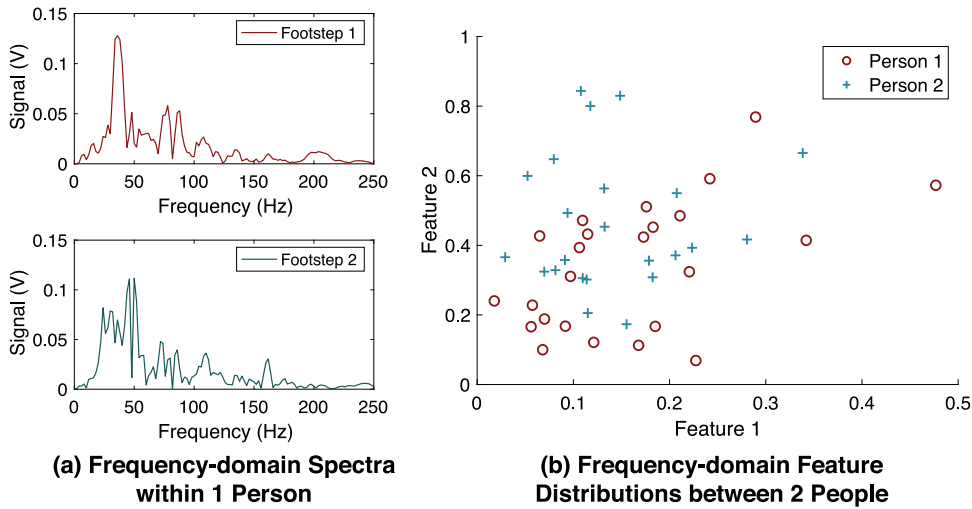


Fig. 2. Visualization of the variability in footstep-induced structural vibration: (a) frequency-domain spectra of two single footsteps within 1 person, (b) frequency-domain feature distributions between 2 people (these 2 people's footstep features are overlapping, making it difficult to identify each person).

3.1.1. Identifying variability sources in footstep-induced structural vibration

The identified variability sources are: (1) footstep variability and (2) structural variability, as illustrated in Fig. 1. They are identified through the process of wave generation and propagation. When people walk, their footstep impact forces are the inputs to the structure, so the variations in the footstep forces affect the resultant signals. These input forces generate vibration waves that propagate through the structure and are then received by the sensors, so the differences in structural properties across various locations also affect the resultant vibration signals. The formal definitions of these two sources are summarized below:

- **Footstep variability** is defined as the natural variation of footsteps when a person is walking normally at a relatively constant speed without changing shoes. These variations may come from the minor adjustments of balance when a person is walking in the same direction at a relatively constant speed [46].
- **Structural variability** is defined as the signal variations caused by the same footstep force stepping at different locations of the structure [45]. It represents the mixed effect from various aspects of the structure heterogeneity, including material heterogeneity, spatial variations of the structure layout, and so on.

To show the combined effect of both the footstep and structural variability, Fig. 2(a) describes the discrepancy in frequency-domain features between two footsteps from the same person. To further demonstrate the effect of this multi-source, high variability in our person identification problem, Fig. 2(b) shows the feature distributions between 2 people (after the t-SNE dimensionality compression for 2D visualization purpose [52]). We observe that each person's data has a high variability, making it difficult to separate these two people.

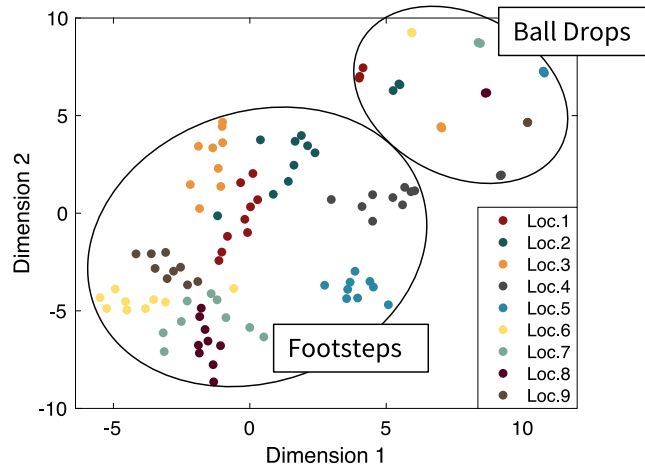


Fig. 3. Feature patterns of vibration responses from 2 different types of impulses: (1) ball drops, (2) footsteps. Different colors indicate different excitation locations, numbered 1–9. The ball drops are on the right and the footsteps are on the left. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.1.2. Decomposing variability in footstep-induced structural vibration

After identifying variability sources, we decompose the overall variability in footstep-induced structural vibration signals into (1) footstep variability and (2) structural variability, in order to understand their properties and develop a method for variability reduction (as described in Fig. 1). To isolate the structural and footstep variability, we compare the vibration responses from footstep impact forces and the ball drop forces from the same height at the same excitation and sensing location. Since the impact forces from ball drops at the same height are almost identical, it removes the variability contribution from footstep impact forces. In addition, the same excitation and sensing location allows the wave generation and propagation paths to be almost the same, which removes the structural variability in the vibration signal.

Fig. 3 shows the variability patterns induced by a walking person and a repeated dropping ball, visualized in a 2-dimensional plot using t-distributed stochastic neighbor embedding (t-SNE) [52]. For the same excitation location (i.e., dots with the same color), the ball drop-induced vibration features significantly overlap, meaning that their resultant signals are almost the same; while the footstep-induced vibration features vary, indicating the variability across different footsteps. Among different excitation locations (i.e., dots with different colors), we observe that both the ball drop and the footsteps induce different vibration signals, which represents structural variability. In addition, we observe that the structural variability has a much larger variability than the footstep variability because the average Euclidean distance between different excitation locations in the feature space is larger than that of different footsteps at the same excitation location. The detailed experiment setup and the quantitative evaluation for variability decomposition will be described in Section 5.3.

3.1.3. Quantifying the footstep and structural variability within one person

To understand how much structure heterogeneity and footstep variations contribute to the overall variability, we quantify these decomposed variability sources based on the Analysis of Variance (ANOVA) of the footstep features. Since the footstep impact forces are generated from repeated trials of the same person with a consistent body mechanism, they are assumed to be independent and identically distributed (i.i.d.). Therefore, we model the footstep variability as a statistically random factor incurred during sampling (i.e., walking in our problem), which is defined as follows:

$$\Sigma_{footstep} = \frac{1}{K} \sum_{k=1}^K \frac{1}{N_k} \sum_{i=1}^{N_k} (x_i - \mu_k)(x_i - \mu_k)^T \quad (1)$$

The structural variability, on the other hand, is estimated by sampling uniformly distributed excitation locations over the structure and then computing the covariance among the sample means at each location. As observed in Fig. 3, the mean of footstep features at each location are scattered around the feature space and can be modeled as an overall normal distribution. They also satisfy the i.i.d. assumption because they are generated through the same structure that is typically stationary without damage. Therefore, it is quantified as follows:

$$\Sigma_{structure} = \frac{1}{K} \sum_{k=1}^K (\mu_k - \mu)(\mu_k - \mu)^T \quad (2)$$

where K denotes the total number of excitation locations, N_k denotes the number of footsteps at location k . x_i means the features of i th footstep sample at location k , μ is the mean among all footstep features and μ_k is the mean of footstep features at location k .

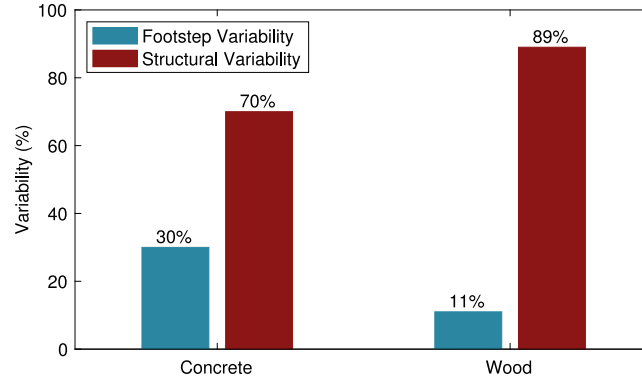


Fig. 4. Overall variability proportion for wood and concrete structures. The wooden structure has higher structural variability.

Based on the above quantification of structural and footstep variability, we compare the variability proportions on the wood and concrete structures and found that the structural variability dominates (See Fig. 4). The detailed analysis and discussion are included in Section 5.3.

To this end, we aim to reduce the structural variability and keep the footstep variability. This is because the identity of the person is independent of the structure properties, and the footstep variability may be useful in reflecting a person's unique walking habits.

3.1.4. Designing physics-guided transformation function to reduce structural variability

In order to reduce the structural variability, we design a data transformation function to develop more separable footstep features between people. Assuming that the cross-sections of our structure along the walking path can be simplified as a simply-supported beam within the linear elastic range, the dynamic response of the structure can also be simplified, in which we can compute the analytical solution of the structural vibration given the input force and location [53]. Let $P(t)$ be the input force with respect to time, $w(x, t)$ be the vertical deflection of the beam at location x , the person's walking speed as v , the governing equation for vertical vibration can be written as:

$$EI \frac{\partial^4 w(x, t)}{\partial x^4} + \rho A \eta \frac{\partial w(x, t)}{\partial t} + \rho A \frac{\partial^2 w(x, t)}{\partial t^2} = P(t) \delta(x - vt) \quad (3)$$

where E , ρ , A , I are material and geometry constants, η is the damping coefficient, $\delta(t)$ is the Dirac delta function.

Since our features are the frequency domain magnitudes from the footstep signal, it represents the distribution of frequency content of a footstep at each location. Therefore, we associate the vertical responses of the structure with the frequency features of our signal $\phi(x, \omega)$ through Fourier Transform:

$$\phi(x, \omega) = \int_{-\infty}^{\infty} w(x, t) e^{-i\omega t} dt \quad (4)$$

The closed form solution of Eq. (3) is solved as:

$$\phi(x, \omega) = \sum_{k=1}^M \phi_k(x, \omega) \quad (5)$$

where M is the number of modes, $\phi_k(x, \omega) = C_k(\cosh \lambda x - \cos \lambda x) + D_k(\sinh \lambda x - \sin \lambda x) + Q_k(A_1 \cosh \lambda x - A_2 \sinh \lambda x + A_3 \cos \lambda x - A_4 \sin \lambda x - A_0 e^{-i\omega x/v})$. Note that C_k , D_k , A_0 , A_1 , A_2 , A_3 , A_4 are constants, Q_k represents the force in frequency domain, and $\lambda^4 = \rho A(\omega^2 - i\eta\omega)/EI$. If we take the k th mode of the structure, then this solution can be approximated as a nested exponential function with respect to ω : $\phi_k(x, \omega) \approx e^{\lambda(\omega)x}$. Therefore, the structural responses at 2 different excitation location x_1 and x_0 can be associated as follows:

$$\begin{bmatrix} \phi_k(x_1, \omega_1) \\ \vdots \\ \phi_k(x_1, \omega_d) \end{bmatrix} = C \begin{bmatrix} e^{\lambda(\omega_1)(x_1-x_0)} & \dots & 0 \\ 0 & \ddots & 0 \\ 0 & \dots & e^{\lambda(\omega_d)(x_1-x_0)} \end{bmatrix} \begin{bmatrix} \phi_k(x_0, \omega_1) \\ \vdots \\ \phi_k(x_0, \omega_d) \end{bmatrix} \quad (6)$$

where $\omega_i, i = 1, 2, \dots, d$ means different frequency bands, and C is a constant value.

Therefore, we can apply a linear transformation matrix to approximate the relationship of structural responses across different excitation locations, represented as follows:

$$X_{transformed} = W^T X_{original} \quad (7)$$

where $X_{original}$ represents the extracted footstep features before the transformation, and $X_{transformed}$ denotes the transformed features after the optimization.

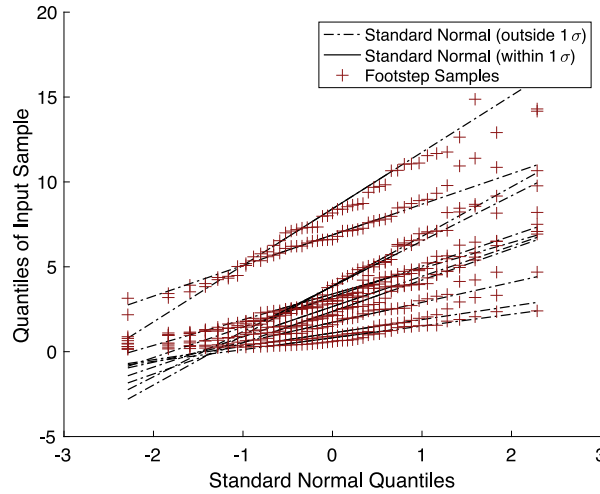


Fig. 5. Quantile-quantile plot of the footstep features versus the standard normal distribution in all feature dimensions. The features are normally distributed within 1 standard deviation while drifting significantly from standard normal at both ends.

3.2. Distribution analysis for footstep feature modeling

After designing the transformation function, we analyze the distribution of the transformed features to develop models for open-world person identification. We observe that the transformed footstep features form a unique distribution: it is normally distributed around the mean, while it has asymmetrical tails outside the first standard deviation (see Fig. 5).

To validate our observation of data patterns, we conduct different normality tests on the transformed footstep features, such as the Kolmogorov–Smirnov test and the goodness-of-fit test. We found the transformed data rejected the majority of the tests with less than 0.05 p-values, which means that it is not valid to assume the data within each person are normally distributed. However, when we take the middle 68% percentile of the data and conduct these tests again, they pass the majority of the tests with more than 0.2 p-values.

We also interpret this data distribution based on the sources of variability. The Gaussian distribution forms because the gait naturally varies when walking. For example, the impact forces, contact angles, and the location of contact typically vary around an inherent mean value represented by body configurations and walking habits. The irregular outliers may result from gait anomaly (e.g., dragging, kicking, temporary loss of balance), environmental noises, and/or distortions during the wave propagation.

3.3. Heuristics on newcomer patterns

In real-life application scenarios, the probability of the presence of newcomers depends on the time and location. For example, companies may expect visitors during scheduled business hours while households may welcome guests during the weekends. This means that the decision boundary between the known and unknown people changes in different user scenarios — companies should set the decision boundary prone to the unknown people during the weekdays while households do so on weekends.

In addition to the changing newcomer probability over different time and locations, the previous observation also affects the future decision. For example, if a guest and a resident are expected to arrive in total, the arrival of the resident leads to a much higher expectation for observing the guest to come. In this case, the prior probability of the newcomers changes from 50% to 100% after observing either person.

Therefore, we need to track the prior probability of the newcomers based on their patterns of presence. The method to keep track of the above-mentioned newcomer patterns will be introduced in Section 4.3.

4. Open-world person identification framework based on nonparametric Bayesian modeling

We develop an open-world person identification framework based on nonparametric modeling of footstep-induced structural vibration. Our framework consists of three modules: (1) footstep data transformation for variability reduction, (2) global and local feature distribution modeling for identification, and (3) adaptive online learning, as illustrated in Fig. 6. The framework mimics the learning process of humans, who first refine the raw information, then organize it, and finally remember it [54,55]. The three modules correspond to the three steps people take to acquire new knowledge, which are designed to achieve continual knowledge expansion without much initial data.

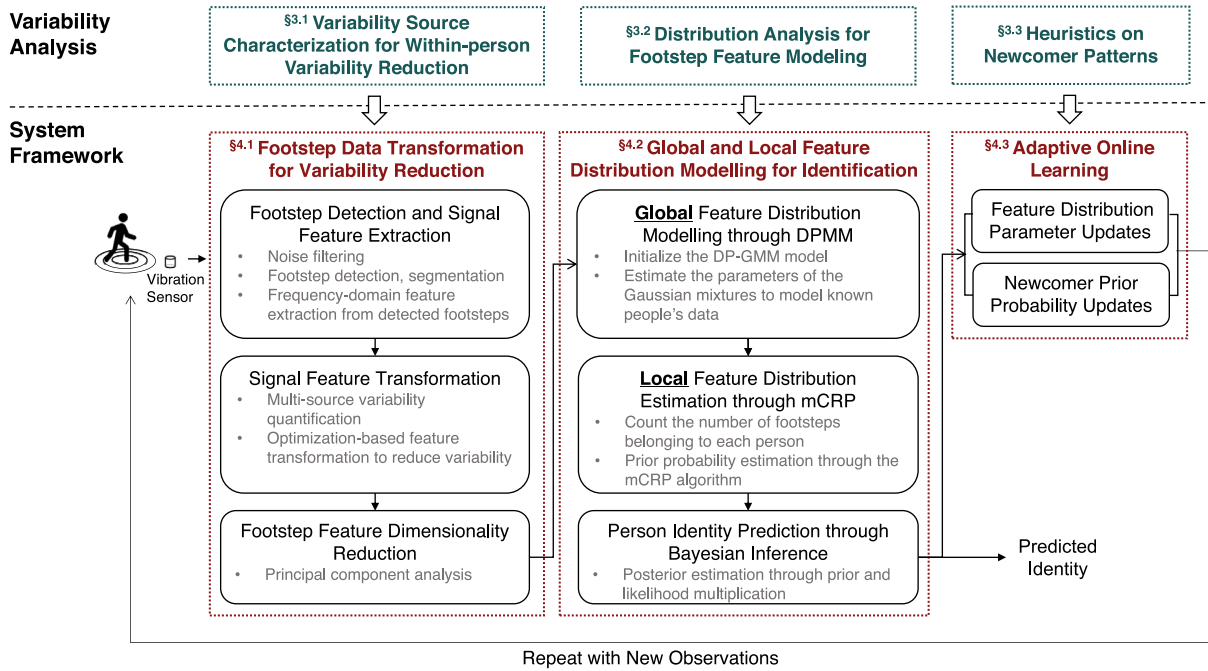


Fig. 6. An overview of our open-world person identification framework. The framework consists of 3 modules. In module 1, we develop footstep features that are less variable within each person through feature transformation. In module 2, we model the irregularly shaped footstep feature distributions, where we combine global and local feature distribution models to achieve more accurate identification. In module 3, we update the feature distribution parameters and the newcomer's prior probability to prepare for the new observations.

4.1. Footstep data transformation for variability reduction

When a person's footsteps are detected, we first reduce the variability in footstep-induced structural vibration and develop features that are more separable between different people.

4.1.1. Footstep detection and signal feature extraction

The first stage of our approach involves detecting the occurrence of footsteps in the sensing area and extracting frequency-domain signal features from the detected footsteps. In this work, we use SM-24 geophone sensors to sense the vertical velocity of the structural vibration induced by people's footsteps. This is because the sensitivity of geophones is relatively consistent across the frequency range of footstep-induced floor vibrations and the sensing mechanism is less sensitive to noises due to temperature changes or lateral disturbances on the floor surface [56,57]. While this work focuses on geophone sensors, the developed approach is applicable to other vibration-based sensors (e.g., accelerometers, displacement meters). When a person walks by a vibration sensor, their footsteps generate impulsive vibration waves that propagate through the floor, which are sensed by the vibration sensors. These vibration signals are pre-processed first through a lowpass filter in order to filter out the high-frequency sensory noise (≥ 200 Hz) and then through an adaptive Wiener filter that reduces the environmental noises from ambient vibrations from the environment. The choice of the 0–200 Hz frequency range is based on the empirical observation that 90% of the frequency components in footstep-induced floor vibrations are within this range. This is consistent with the findings from the prior work, which demonstrates that using the 0–200 Hz range leads to promising performance in person identification [1,58]. In order to capture footsteps in a time series, we first differentiate the footstep impulses from the other environmental impulses (e.g., a ball dropping, a door closing, etc.) based on the duration and the number of consecutive events observed, which was developed in prior work [50]. Impulsive events detected as footsteps represent the presence of a person and trigger our system. The triggering criteria are that the average amplitude of the signal within a 0.1-second window is larger than $10\times$ mean amplitude of the background noise.

With the detected footstep events, we isolate the consecutive footsteps as a footstep trace using an anomaly detection algorithm developed in prior works [50]. Footstep features that represent different walking characteristics of each individual are then extracted from these traces. Specifically, the features are the mean amplitudes at each 5 Hz frequency band across the frequency domain (0–200 Hz), which are shown to be effective for person identification in prior works [1,58].

The frequency-domain footstep features are extracted from each detected footstep instead of an entire walking trace to mitigate the effect of walking speed on features. These features are then used as inputs for variability characterization and reduction.

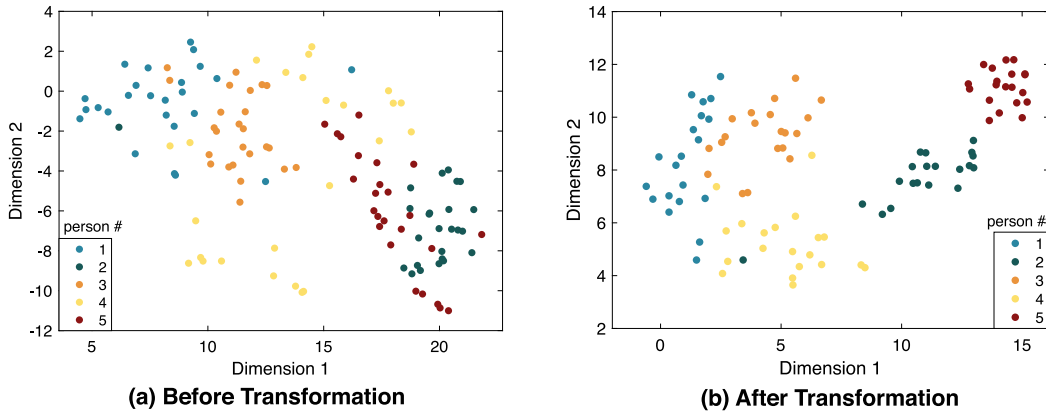


Fig. 7. 2D t-SNE visualization of footstep features before and after variability reduction among 5 people: (a) before and (b) after. Footstep features from these 5 people are more distinguishable after the transformation.

4.1.2. Signal feature transformation to reduce variability in each person

With the designed transformation function in Section 3.1.4, we formulate an optimization problem to separate different people. The optimization goal is consistent with the new person identification objective, which is to distinguish different people through their footsteps. In order to approach this goal, we translate this objective into an optimization problem that aims to minimize the total variability within-person and maximize the between-person variability. In this case, we simplify Eq. (2) and Eq. (1) by setting $N_k = 1$ and $K = N$. This is because the footstep locations are relatively random when the person is walking, and thus it is unlikely that two footsteps fall at the exact same location. With this simplification, the total variability within one person is then expressed as:

$$\Sigma_{total} = \Sigma_{structure} + \Sigma_{footstep} = \frac{1}{N} \sum_{n=1}^N (x_n - \mu)(x_n - \mu)^T \quad (8)$$

where N represents the total number of footsteps within one person.

After computing the total variability within one person, the optimization problem is formulated to minimize the sum of within-person covariance and maximize the between-people covariance, expressed as follows:

$$\text{maximize } J(w) = \frac{w^T S_B w}{w^T S_W w} \quad (9)$$

where the between-person covariance matrix is calculated as $S_B = \sum_{i=1}^C N_i (\mu_i - m)(\mu_i - m)^T$, and the within-person covariance matrix as $S_W = \sum_{i=1}^C \sum_{n=1}^{N_i} (x_n^i - \mu_i)(x_n^i - \mu_i)^T$. i represents the person number that ranges from 1 to C ; n represents the sample number within each person i that ranges from 1 to N_i ; x_n^i denotes the n th data sample in the i th person; m and μ_i are the mean of all data and the mean of data in person i respectively. Since the optimization goal matches the objective of Fisher's Discriminant Analysis parameterized by coefficients w [59,60], the optimal coefficients w are estimated through a closed-form solution in the Fisher's formulation [61]. This formulation assumes Gaussian distributions of the overall data, which is consistent with our observation that the data samples around the mean are normally distributed in Section 3.2.

After we obtain the coefficients w , we then transform the data into the targeted feature space. As shown in Fig. 7, before the transformation, footstep features from the 5 people overlap with each other with high variability, while the footsteps between different people are more distinguishable after the data transformation.

4.1.3. Footstep feature dimensionality reduction

After the variability reduction algorithm, we further compress the feature dimensions to reduce the overfitting effect for data modeling and learning. This is because the features after the transformation are high-dimensional and are negatively affected by the *curse of dimensionality*. This means the high dimensional data leads to sparse sample distribution in the feature space, resulting in the overfitting of learning models to the local patterns rather than the overall patterns of footstep features.

To overcome this limitation, we reduce the feature dimension through Principal Component Analysis (PCA). We choose PCA because it maintains the within-person continuity and between-person separability of the footstep features by taking the direction of maximum spread of the data as principal components [62]. In order to incorporate at least 90% of the overall footstep variance, the first 10 principal components are kept as footstep features. With a set of preliminary data of 100 footsteps (which represents 3–4 known people), the coefficients w remain relatively stable (within a 5% change) with different sets of training-test splits after reducing the dimension of the features. Since the model is expected to be more stable as more people are known, the preliminary result with a relatively low number of known people shows that PCA is effective in reducing the variance of our data transformation model.

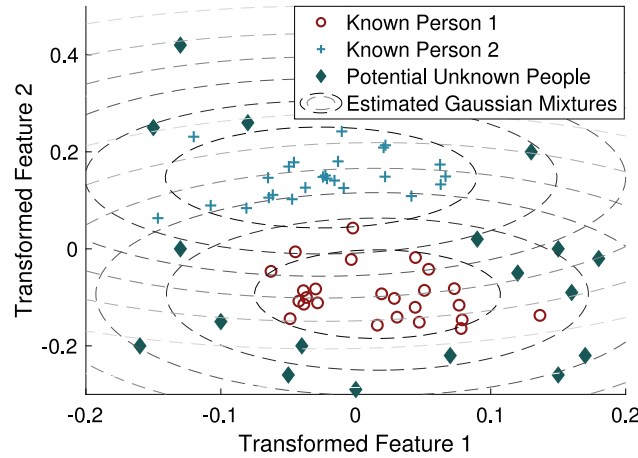


Fig. 8. Dirichlet Process Mixture Model (DPMM) with two Gaussian mixture components: footstep features from two known people are modeled as Gaussian mixture components globally. Footsteps from unknown people (i.e., potential mixture components) could be present anywhere in the feature space.

4.2. Global and local feature distribution modeling for identification

In this section, we aim to model the transformed footstep features and predict the identity of each new observation. To deal with the irregularly shaped feature distributions analyzed in Section 3.2, we introduce mCRP with DPMM, which captures both local and global distributions of the features and significantly improves the accuracy when predicting people's identity.

4.2.1. Global feature distribution modeling through Dirichlet Process Mixture Model (DPMM)

For global modeling, we describe the feature distributions among different people as a Dirichlet Process Mixture Model (DPMM) where each mixture component is a distribution that represents footstep features from one person. Since the data samples in each person's footstep features near the mean are normally distributed based on analysis from Section 3.2, we assume each mixture is a Gaussian distribution in global feature modeling. DPMM is a generalized form of finite mixture models, which has an unlimited number of mixture components each representing the feature distribution within a person [63,64]. To form the global model, we estimate both the within-person and between-people model parameters. The within-person parameters include the mean and covariance of features of each person, and the between-people parameters include the number of overall samples, the sample proportions between people, etc.

Once we formulate our features as a DPMM, we infer the identity of each person by comparing the posterior probability of a footstep trace x_i belonging to each person k (i.e., $P(z_i = k|x_i)$). It can be obtained by Bayes' Rule, where posterior $\propto$ prior $\times$ likelihood [65,66] (see Fig. 8).

4.2.2. Local feature distribution estimation through modified Chinese Restaurant Process (mCRP)

While the global feature distribution is modeled as DPMM with Gaussian mixtures, the local feature patterns are highly complex in shape as discussed in Section 3.2. To locally capture such complexity, we introduce a modified Chinese Restaurant Process (mCRP) that focuses on the proximity of the newly observed footsteps to estimate the prior probability of each person. The proximity is quantified by feature values that are within three standard deviations of the mean (which includes around 99.7% of the data for Gaussian distributions). This is because we want to include most people who might be the owner of newly observed footsteps. The mCRP is formulated as below:

$$\text{known person } k : p(z_i = k|z_{-i}, \alpha) = (1 - \alpha) \frac{n'_k + n_{K+1}}{\sum n'_k + n_{K+1}} \quad (10)$$

$$\text{unknown person } K + 1 : p(z_i = K + 1|z_{-i}, \alpha) = \alpha \quad (11)$$

where z_i represents the class label for i th footstep trace; n_{K+1} represents the number of footsteps of the incoming trace; n'_k denotes the number of footsteps from person k that are in the proximity of the current observation. We develop the mCRP in a way such that it also has a new hyperparameter α - it represents the prior expectation of the unknown people as shown in Eq. (11), which is discussed in greater detail in Section 4.3).

There are two key advantages of mCRP: (1) it does not require a distribution assumption, and (2) it filters out the noisy information from people who are irrelevant to the prediction by focusing on the area around the new observations. Similar to the philosophy of Support Vector Machines (SVM), mCRP looks at the data samples around the decision boundaries (i.e., margins) to make predictions.



Fig. 9. Decision boundary changes with the hyperparameter α : a larger α means the incoming footsteps are more likely to belong to an unknown person.

In addition to the prior estimation, we also estimate the likelihood of the known and unknown people. We assume that the newly observed data either join one of the known person's footstep distribution θ_k or form a new distribution θ_{K+1} . With the above assumptions, we approximate the likelihood of the unknown class using the Markov property and the Bayes' Rule [63]:

$$p(x_i|\lambda) = \int F(x_i|\theta_{K+1})G_0(\theta_{K+1}|\lambda)d\theta_{K+1} = \frac{p(x_i|\theta_{K+1}, \lambda)p(\theta_{K+1}|\lambda)}{p(\theta_{K+1}|x_i, \lambda)} \quad (12)$$

where $p(\theta_{K+1}|\lambda)$ means the probability density of θ_{K+1} in between-people distribution G_0 and $p(\theta_{K+1}|x_i, \lambda)$ means the probability density of θ_{K+1} in that distribution after adding the new footstep trace x_i ; $p(x_i|\theta_{K+1}, \lambda)$ can be approximated by the centroid of footstep features of x_i given that x_i forms a new distribution component itself.

4.2.3. Person identity prediction through Bayesian inference

After the prior and likelihood estimation above, we use Bayesian inference to predict the identity of the new observation. The identity of a newly observed person is then determined by comparing the posterior probability of a footstep trace belonging to each person. According to the Bayes' Rule, posterior is estimated through the multiplication of prior and likelihood [65–67], presented in the formula below:

$$\text{known person } k : p(z = k|\alpha)p(x_i|\theta_k) \quad (13)$$

$$\text{unknown person } K + 1 : p(z = K + 1|\alpha)p(x_i|\lambda) \quad (14)$$

Finally, we choose the person that has the largest posterior probability computed by Eqs. (13) and (14) as the predicted identity.

4.3. Adaptive online learning

In this module, the model is updated from two different aspects, including (1) the feature distribution parameter updates (these parameters include the coefficients w in data transformation and the distribution parameters in DPMM), and (2) the newcomer prior probability (i.e., α) updates based on the predicted identity in the previous module.

4.3.1. Feature distribution parameter updates

After the owner of the newly observed footsteps is predicted in Section 4.2, the distribution of the footstep features within the predicted person changes correspondingly. To incorporate these footsteps, the footstep features of the corresponding person are updated by adding rows and columns before computing the covariances. The updates of DPMM model parameters can be separated into two different cases depending on whether the incoming person is a known person or not. In the former case, we update the mean and covariance matrix in the mixture component that describes the footstep feature distribution of this known person. If the incoming person is predicted as unknown, we create a new mixture component (i.e., a distribution that describes the footsteps of this unknown person) in the DPMM model. One significant advantage of DPMM is that it is inherently incremental, meaning that adding a new component does not affect the distribution parameters of other components. In addition to within-person distribution, the parameter λ that describes the base distribution among all footstep data in Eq. (14) is updated accordingly.

4.3.2. Newcomer prior probability updates

To keep track of the prior changes in different situations discussed in Section 3.3, we introduce a customizable, adaptive hyperparameter $\alpha(t, l)$ that alters over different time t and location l . As shown in Eq. (11), the α is the prior probability of the unknown people (i.e., $p(z_i = K + 1 | z_{-i}, \alpha)$), which is defined as the ratio of the expected number of visitors over the expected number of all incoming people, ranging between 0 and 1. The lower α is, the less likely we would expect another unknown person. The influence of α on the decision boundary is shown in Fig. 9.

To account for the influence of previous observations on future predictions, we alter the prior probability of known and unknown people by changing the hyperparameter α every time we make an observation. If the incoming person is unknown, the number of expected unknown people (i.e., the numerator in α) decreases by 1; otherwise, the expected number of all incoming people (i.e., the denominator in α) decreases by 1. The automatic change in the hyperparameter α provides real-time adjustments to the decision boundary so that it matches the real situation better. The effectiveness of such adaptive updates will be shown in Section 5.2.

5. Evaluation

We evaluate our framework through a real-world experiment in two different structures with 20 people. The evaluation contains three parts. First, the experiment setup and the overall performance of our framework are presented (Sections 5.1 and 5.2). Then, the results from within-person variability characterization and between-people footstep data modeling are presented (Section 5.3). Finally, a sensitivity analysis is conducted to explore the influence of different factors, including the number of known people, the sequence of the incoming people, the order of footstep traces, and the hyperparameter α (Section 5.4). This allows us to gain a systematic understanding of the advantages and limitations of our framework.

5.1. Experiment setup

We conducted experiments at 2 deployment sites: (1) a wood-framed platform and (2) a concrete corridor in an educational building. For the setup, 4 SM-24 geophone sensors recording at a sampling frequency of 25.6 kHz were mounted along the two edges of the corridor [68], spaced apart by 2 meters (as shown in Fig. 10). The sensor layout is chosen to minimize the bias due to noise and structural layout and to maximize the area of coverage along the walking path. The sensors at both sides are more robust to environmental noises and structural heterogeneity when compared with placing sensors at one side only. The 2-meter offsets allow the sensors to cover a longer walking path than aligning the sensors horizontally. In addition, the sensor placement was significantly denser than required to provide redundancy in data collection and analysis for research purposes. Preliminary test results show that the coverages of these sensors are significantly overlapping and reducing the number of sensors to one only impairs the performance by 5% to 10%. Amplifiers were used to improve the signal-to-noise ratio (SNR). The amplification rate we used is 100-1000 $\times$, which was selected empirically by conducting preliminary walking experiments to maximize the vibration amplitude while minimizing the clipping of the signal. The amplification of the signal increases the sensing range to up to 20 meters in diameter [14]. A National Instruments NI-DAQ was used to acquire and convert the analog signal to the digital signal [1].

The experiment includes two steps: (1) study introduction and participation consent, and (2) walking trials on the wooden and concrete structures. First, we introduce the high-level goal of our study (i.e., human gait characterization) and obtain the participant's consent after explaining the procedures, risks, and data handling process. Then, each participant was asked to walk across each floor structure for 10–15 trials. During the walking trials, the participants are asked to walk freely with their natural gait, following the direction of the walkway. The sequence of floor types for walking is randomly selected. The rest time between trials is around 15 s. These are designed to mitigate the bias due to physical fatigue or mental familiarity of the participants. All experiments were conducted in accordance with the approved IRBs.

As shown in Fig. 10, the wooden structure is an elevated wood-framed platform constructed in the laboratory. The structural components are connected through metal plates fixed with bolts. The concrete structure is part of the existing concrete building (second floor) with concrete floorings. The concrete slab is supported by three secondary beams underneath its span.

5.2. Overall performance in open-world person identification

Our framework achieves an average of 92.3% accuracy in open-world person identification with 1 pre-recorded person only, as shown in Fig. 11. The initial setup of the framework is to randomly pick 1 known person and record 15 footsteps from this person into our database. The rest of the people are permuted into a random sequence with arbitrary repetitions to mimic the unpredictable arrival of known and unknown people. Each test set consists of 200 footsteps from 10 people (1 known person and 9 unknown people), with an average of 3 consecutive footsteps per sample trace.

In order to show the effectiveness of individual modules, we evaluate the performances of our framework without each module: within-person variability reduction (Module 1), modified mCRP (Module 2), adaptive α (Module 3). As shown in Fig. 11, the accuracy of our framework drops by 12.7%, 21%, and 12.5% when each of these three modules is removed, respectively. Among them, the accuracy drops the most when the modified Chinese Restaurant Process (mCRP) is replaced with traditional CRP. This means that our mCRP plays an important role in identifying a person correctly because it captures the irregular local distributions at the decision boundaries between people. To evaluate our framework in sub-branch person identification tasks, we also compare the performance of our non-parametric Bayesian approach with other new person discovery and closed-world person identification algorithms. After the footstep features are transformed by the variability control module, they are input to the existing algorithms proposed by the



Fig. 10. Walking experiment for evaluation. The experiments were conducted in a real-world setting with human subjects (right). The sensor layout, walking paths, and the cross-sectional structural layout of the wooden structure (top) and the concrete structure (bottom) are shown in the diagram. The floor dimensions are marked in the diagram.

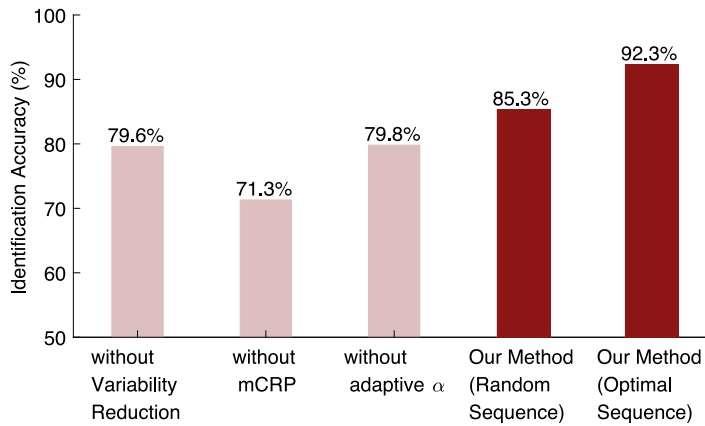


Fig. 11. Overall performance and the effectiveness of individual modules: average person identification accuracy of our framework compared with the framework without addressing the challenges.

prior studies [1,58,69–71]. For new person discovery, these models include Mahalanobis Distance-based Thresholding (D-THRE) and One-class SVM (OC-SVM). For closed-world person identification, we select the classification models that are shown to perform better in footstep-based person identification, including weighted K-nearest neighbors (KNN) and support vector machine (SVM) with radial basis function (RBF) kernels. Both (a) and (b) in Fig. 12 show our nonparametric Bayesian method (i.e., DPMM with mCRP) performs significantly better than the existing models on our dataset — we achieved an average of 88.4% for new person discovery (6× error reduction) and 95.6% closed-world person identification (1.4× error reduction). Unlike existing separate model approaches where the error from the first model propagates to the next model, our method achieves high accuracy for both tasks by simultaneously conducting them in one model only.

Fig. 13 shows the robustness of our method in a wood and concrete structure. The figure shows similar accuracy across these 2 structures under different person identification scenarios (92.9% and 91.7% test accuracy for wood and concrete structures respectively). The wooden structure has a slightly higher accuracy because it has lower damping and stiffness than the concrete structure, leading to less attenuation when the vibrations are recorded a distance away from the footstep location. Especially, the attenuation effect of higher frequency components (100–200 Hz) is more significant on the concrete structure due to its higher

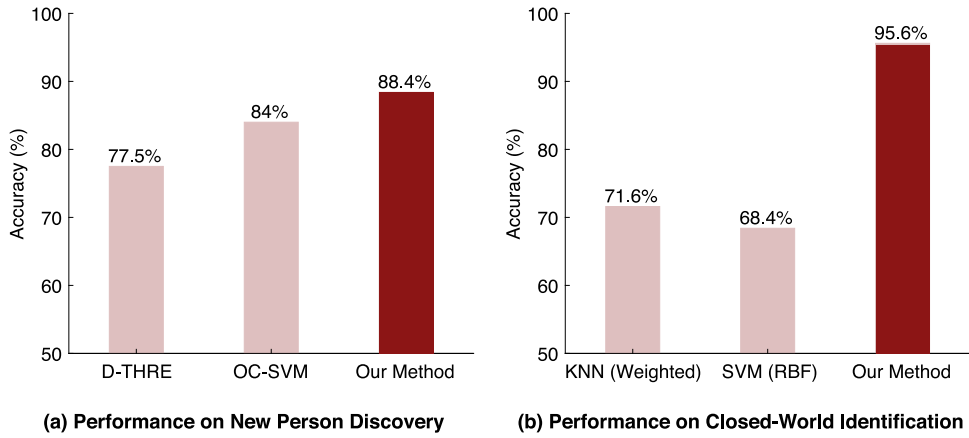


Fig. 12. Performance of our framework in (a) new person discovery and (b) closed-world person identification, compared with the existing models.

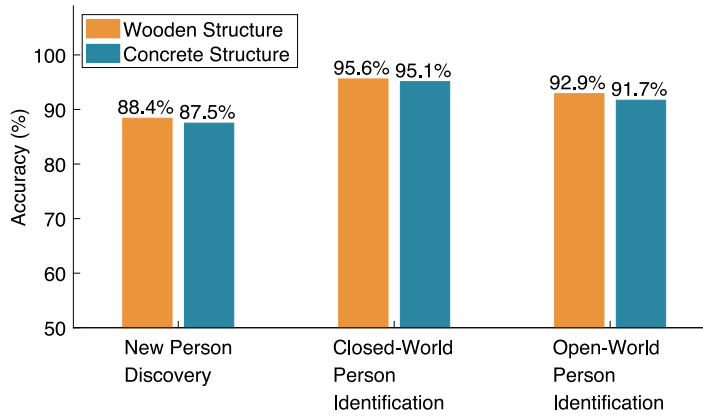


Fig. 13. Performance of our framework across wooden and concrete structures for (1) newcomer detection, (2) closed-world person identification, and (3) open-world person identification.

stiffness and damping ratio, leading to a lower signal-to-noise ratio within that range. As a result, less footstep information is preserved on the concrete floor.

The computational time of person identity prediction is near real-time during our evaluation, which is on average within 1 s after a person passes by using a commercial laptop with a quad-core CPU that is connected to the sensor nodes. The distribution of computational time across three modules for a 10-person identification are around 80%, 15%, and 5% for module 1, 2, and 3, respectively. This is because module 1 requires feature extraction and transformation, each contains multiple loops over all the existing data. For comparison, module 2 and 3 only requires modeling and updating the new data sample by comparing it with the existing data's distribution parameters.

5.3. Evaluation on footstep variability analysis and modeling

We present the evaluation for the variability analysis and modeling to further discuss the effectiveness of our framework.

5.3.1. Variability characterization on wood and concrete structures

The experiment for variability characterization involves two parts: (1) ball drops and (2) people walking, to analyze the two sources of variability (i.e. the structural variability and footstep variability, respectively). For both experiments, the structure/corridor is divided into 10 uniform parts with 9 markers representing the fixed locations of ball drops and footsteps (marked as footprints in Fig. 14). These locations are marked on the floor using colored duct tape. For the ball drop experiments, a tennis ball drops from 1.2 m above each designated location on the floor and is repeated 8 times. For each time of dropping, the signals are recorded by all 4 sensors. For the walking experiment, three participants were asked to walk across the sensing area more than 8 times using their natural gait; for each time of walking, a series of consecutive footstep impulses were recorded. The ball drop and walking experiment allow variability composition as discussed in Section 3.1.2.

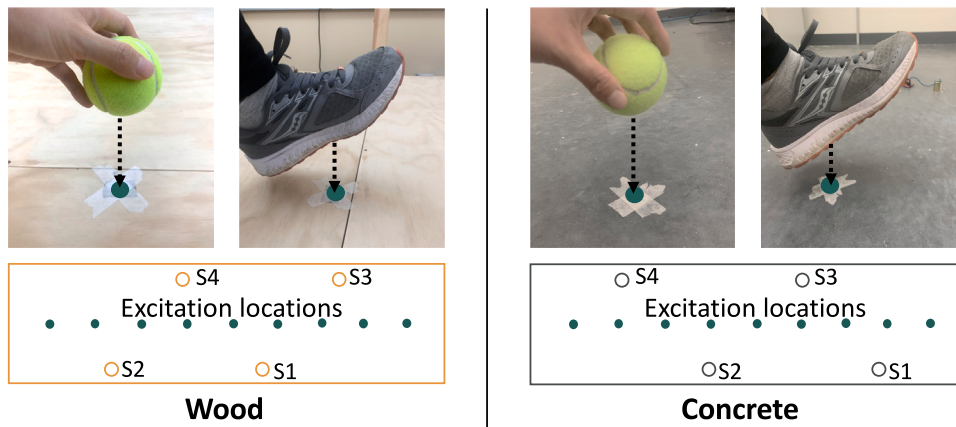


Fig. 14. Experiment setup for within-person variability quantification and decomposition: ball drops and footsteps as excitation forces on 9 uniformly distributed excitation locations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

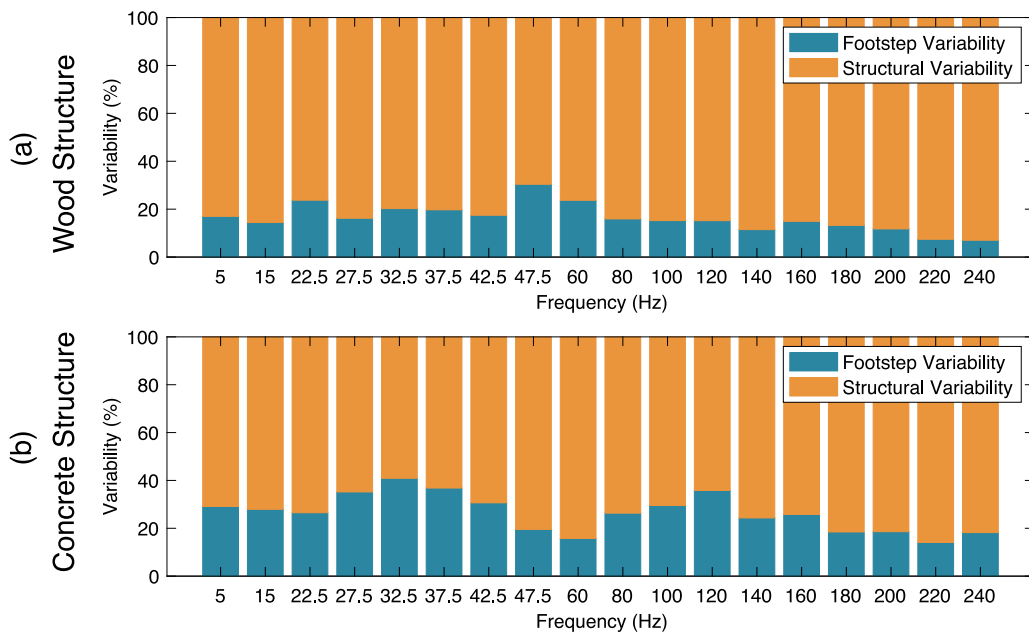


Fig. 15. Variability decomposition for wood and concrete structure across the sensing bandwidth (0–240 Hz). The variations in proportions are due to the natural frequency of the structure and the frequency content in footstep impact forces.

Comparing the variability proportion of the wood and concrete floor, the structural variability in the wooden floor is significantly higher than that of the concrete floor (see Fig. 15). This is because these two structures are different in material, layout, joint connections, etc. Comparing the material types, concrete is more homogeneous than wood because the concrete is a man-made mixture with controlled strength and ductility but wood is a naturally obtained material with higher variations of physical properties. In addition, the configuration of the wooden structure is often more heterogeneous than the concrete because it requires more secondary beams to support the slab and more complex joints to connect separate structural elements together. In contrast, the concrete floor, due to its higher strength, does not require as many secondary supports and often has fixed joints bounded by steel reinforcements.

Variability decomposition results on wooden structure. Fig. 15(a) shows the variability distribution across the bandwidth of the sensor (i.e., 240 Hz). Overall, the footstep variability gradually decays as the frequency increases, except for the frequency band centered at 22.5 Hz and 50 Hz. To explore the reason behind it, we recorded free vibration of the structure under different levels of impulse forces and found that the natural frequency of the structure is around 23 Hz. This means that the frequency magnitude at 23 Hz mainly represents the free vibration of the structure after the footsteps and thus has a much lower structural variability compared to the other bands. For the magnitude at around 50 Hz, the footstep variability also contributes more than that of the

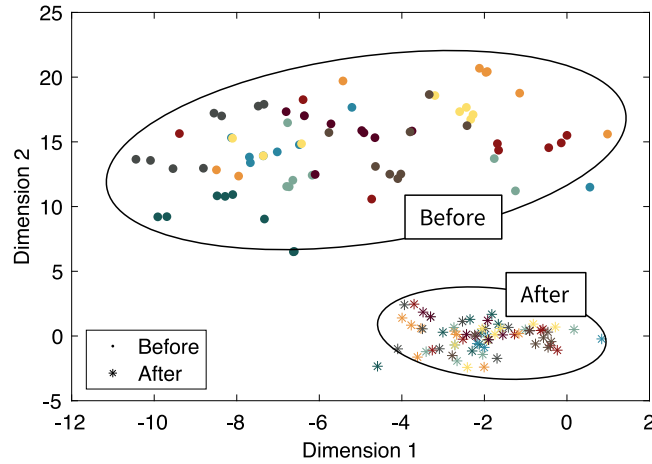


Fig. 16. Visualizing footstep features through 2D t-SNE visualization. Overall, the footstep features are less variable after transformation. The structural variability is reduced significantly while there is a slight increase in footstep variability because we aim to enhance the footstep influence to distinguish different people.

Table 1

Consistency and separability of footstep features before and after the variability reduction.

Category	Avg. cohesion $a(i)$	Avg. separation $b(i)$	Avg. Silhouette $s(i)$
Before	12.6735	12.6110	0.0193
After	2.0478 (more cohesive)	13.4148 (more separable)	0.8405

adjacent bands, this may result from the frequency change in the footstep impact forces or the structural frequency from the second mode.

Variability decomposition results on concrete structure. Fig. 15(b) shows the distribution of variability in each frequency band on the concrete structure. As observed, the footstep variability is more significant at 5–10 Hz, 30–40 Hz, and 110–130 Hz. Similar to the wood structure, we found that the natural frequency of the structure under various impulses is around 9 Hz, leading to a lower structural variability from 5–10 Hz. Interestingly, the observation at the later bands echoes the findings from the vibration-based gait health monitoring [2,47], where 30–40 Hz and 110–130 Hz also have the highest feature importance in detecting the presence of muscular dystrophy in individuals. This suggests that these two bands contain more information on people's footstep forces than the other frequency bands.

5.3.2. Effectiveness of variability reduction

We evaluate the effect of variability control through a metric named *silhouette* that describes the consistency of the same person's footsteps and the separability of different people's footsteps [72,73]. It is an integrated metric that measures (1) how similar a footstep is to the other footsteps of the same person (i.e., within-person cohesion), and (2) how separable footsteps are among different people (i.e., between-person separation). *Silhouette* ranges from -1 to 1 , a larger value indicates better within-person cohesion and between-person separation. The metric is defined as

$$Silhouette(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \quad (15)$$

where $a(i)$ is the average distance between the i th footstep and footsteps that belong to the same person; a smaller value indicates that the footsteps are more similar within this person, meaning that they are more cohesive. $b(i)$ is the average distance between the i th footstep and footsteps that belong to people that are different from i ; a larger value indicates the larger distance between the different people's footsteps and thus being more separable.

The average Silhouette over all points measures the cohesiveness and separability of the overall dataset, which matches with the goal of minimizing within-person covariance and maximizing between-person covariance. Table 1 shows the average of cohesion, separation, and Silhouette among the entire dataset before and after reducing the feature variability. It achieves a more than 80× increase in silhouette, resulting from a 6× increase of within-person cohesion and a slight increase of between-person separation. As a result of such changes, the distances between different people's footsteps become larger while footsteps belonging to each person are more densely distributed as shown in the previous Fig. 7. This leads to overall performance improvement in person identification because the features are more separable after the optimization-based data transformation.

Fig. 16 shows the effect of the variability reduction method in the shared feature space before and after transformation. For both structures, our transformation method reduces 65% of variability compared with the original features. The method reduced 70% of structural variability while the footstep variability increased by 30%. This is because the method aims to enhance the footstep differences between people. In this case, the footstep variability within the same person is also slightly enhanced to ensure better separability between people.

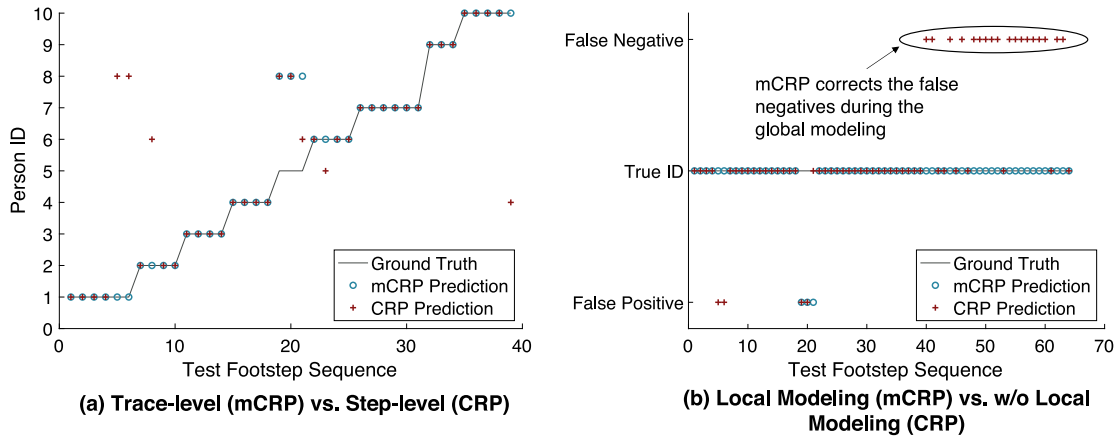


Fig. 17. Our mCRP performs better than CRP in both (a) closed-world identification and (b) new person discovery. In (a), mCRP smoothens the prediction results by accommodating footstep traces as inputs; In (b), mCRP significantly reduces false negatives.

5.3.3. Effectiveness of combined global and local modeling

As mentioned in Section 4.2.2, our local modeling not only extends the scope of CRP by making footstep trace-level predictions instead of using step-level predictions, but also focuses more on local feature patterns instead of global distributions. The effectiveness of our method is shown in Fig. 17.

In Fig. 17(a), we observe that the mCRP is “smoothing” the prediction results of the step-level prediction in each footstep trace. As a result, mistakes made by step-level predictions are corrected by the majority of the footsteps in a trace, which reduces the effect of outlier footsteps. Our evaluation results show that an average of 2.6× error reduction rate is achieved in person identification by leveraging trace-level predictions in mCRP.

As shown in Fig. 17(b), mCRP achieves 7× overall error reduction by significantly reducing the false negative rate. This is because the known people who have very different footsteps from the incoming person are eliminated before further identification, which reduces the estimation bias towards people who have more footsteps. Instead of making decisions among every known person with any footsteps, mCRP shortens the candidate list by considering people that have “similar” footsteps only. Since there are zero known footsteps belonging to the unknown people, conventional CRP prediction favors the known people (especially those with many recorded footsteps) much more than the unknown group, resulting in a high false negative rate (i.e., the probability that an unknown person is predicted as known). In the long run, when more and more known people are recorded in the database, mCRP improves the scalability by shortlisting a small subset of candidates out of a large number of known people.

5.4. Sensitivity analysis

In this section, we conduct a sensitivity analysis under various conditions to demonstrate the robustness of our framework. These conditions include the changing number of known people, the effect of the initial known people, the size of the footstep traces and their sequences, and the expected probability of unknown people.

5.4.1. Effect of the number of known people

In typical semi-supervised learning problems, an increased number of known labels makes the problem easier to address as we have more information during training. In extreme cases, knowing all labels during training makes a semi-supervised learning problem supervised while no label information means the problem is unsupervised. Therefore, in our problem, more known person labels help to reduce the uncertainty in person identification, so the model performance is expected to increase.

Fig. 18(a) shows the increasing trend of online person identification accuracy as more people are known to our database. It is worth noting that the model accuracies are almost the same across 1–5 initial known people, and then a significant increase is observed as we know more than 50% of the potential people that could appear in our framework. The slight fluctuation of prediction accuracy from 1–4 known people can be interpreted as a stage of high uncertainty — even though the number of known people is increasing, the accuracy may not be improving due to the influence of other factors such as the specific set known person, the sequence of the footstep traces, etc., which will be discussed in the following sections.

5.4.2. Effect of the characteristics of initial known people

In addition to the initial number of known people, we explore the characteristics of the initially known people that affect the model performance. Assuming there is only 1 known person in our database, Fig. 18(b) shows the change of prediction accuracy for the different initial known persons. Among them, person IDs 4, 9, and 7 have the highest average accuracies, and, interestingly, the distributions of their footsteps have very different characteristics but they lead to the same consequence. Specifically, the centroid

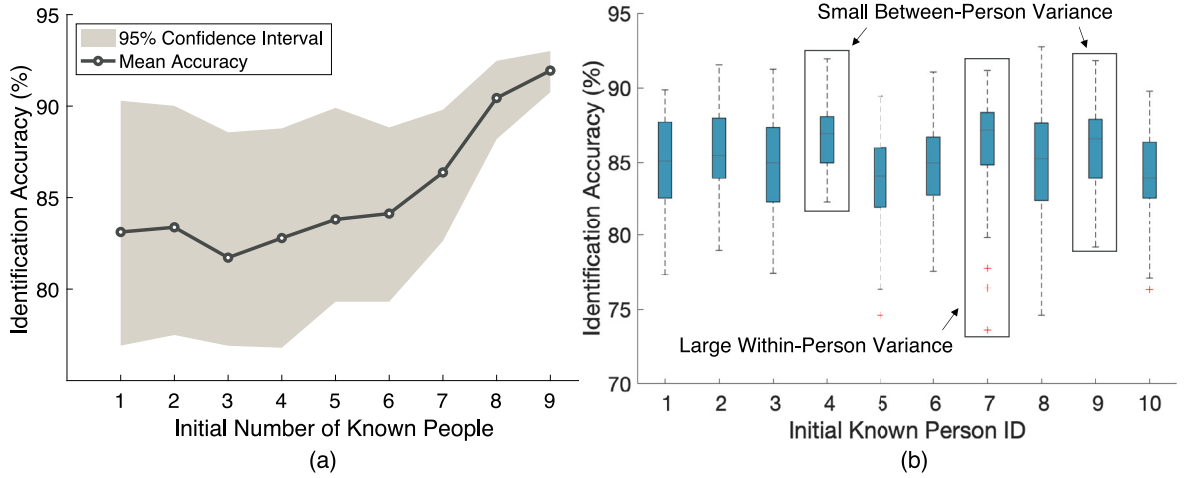


Fig. 18. Sensitivity analysis on (a) the initial number of known people and (b) the initial known person's ID. Higher prediction accuracy is observed with more initial known people, especially when they have small between-person variance and large within-person variance.

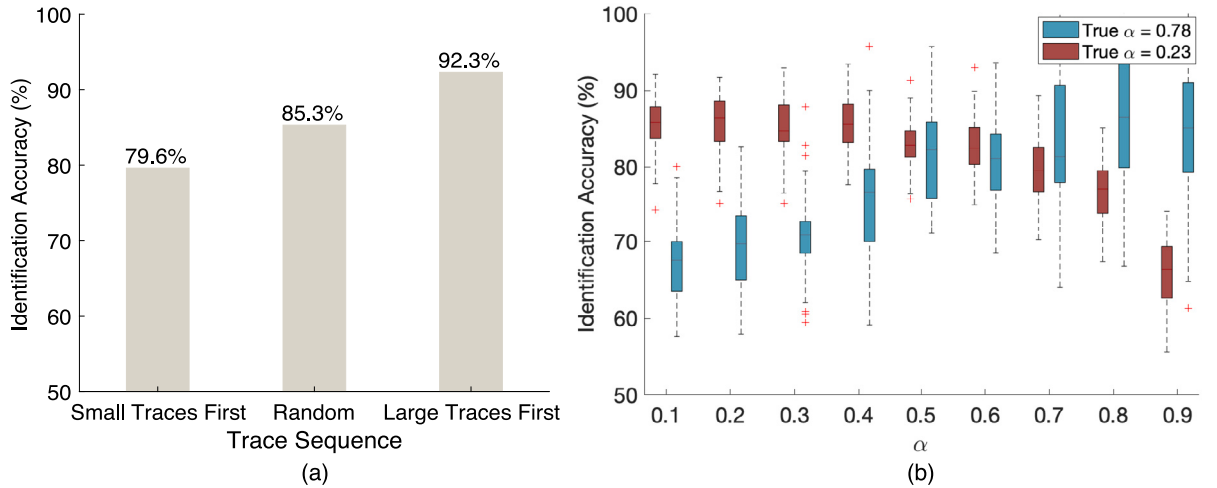


Fig. 19. Sensitivity of the framework performance due to (a) different sequences of the footstep traces based on their sizes, and (b) the mismatch between the true and user-defined α (i.e., expected probability of unknown people).

of person 4's and person 9's footstep features have the smallest Mahalanobis distance to their nearby clusters, making them hard to distinguish from the other people. In contrast, the centroid of footstep features from person 7 has large distances to the nearby clusters instead. However, person 7's footstep features have a high within-person covariance, resulting in a flat distribution in the feature space that is easily mixed up with the unknown footsteps. As discussed in Section 4.1, the separability of people's footsteps is influenced by the between-person distance and within-person variance. The characteristics of footstep feature distributions for persons 4, 7, and 9 show that their separability is low, and thus, setting one of them as an initially known person allows us to first increase their separability via variability control. As a result, following people's footsteps are less likely to be misclassified, which increases the overall accuracy of our framework. The initial knowledge of the people who have low separability helps with the model performance.

5.4.3. Effect of footstep trace sizes

As mentioned in Section 4.2.2, accommodating multiple footsteps from each trace allows us to accommodate their within-person variability. As the size of the footstep trace increases, the distribution of a person's footsteps has more sample points so that the distribution parameters are estimated more accurately. Typically, a large footstep trace has an average of 6 footsteps while smaller ones have only 2 footsteps in our dataset. The sequence of these traces feeding into our model influences the model's performance in the long run. As we observe in the increasing trend from Fig. 19(a), the identification accuracy increases with the increasing size of the trace that is known first. Large traces give higher prediction accuracy because of the increase in the number of samples at the initial stage of online learning.

In real-life applications, we are unable to control the size of footstep traces in our framework since it depends on people's walking patterns such as stride length, the strength of footstep impact forces, and so on. The evaluation results based on a random mixture of different trace sizes are more realistic. Alternatively, more sensors can be deployed along the walkway in order to collaboratively increase the number of footsteps in a trace.

5.4.4. Effect of expected probability of newcomers

In this subsection, we evaluate how an initial value of α affects the identification accuracy. Specifically, we explore the change of accuracy due to the deviation of the assumed α from its true value. Section 4.3 presents the physical meaning of true α and discusses its influence on the decision boundary — the increase of alpha value indicates increases in the expected probability of unknown people, resulting in a decision boundary that is prone to the unknown class. α is a user-defined hyperparameter with a default value of 0.5, meaning that the probability of observing an unknown person is 50%. However, there is no guarantee that the user will always estimate the α value correctly. The true probability of unknown people at a given location and time may be different from the user's expectation. For example, invited guests may not show up and unexpected visitors may occur. As a result, it is important to understand how much deviation of α is tolerable in terms of overall accuracy.

In Fig. 19(a), we test two scenarios when the true probabilities of unknown people are 0.23 and 0.78 respectively, which has a relatively large deviation from the default value. As observed in this figure, slight deviations from the true α only resulted in 2%–3% of accuracy drop for both cases. Without any action required from the user, a default value of 0.5 is relatively robust to the change of user scenarios, resulting in a 5% decrease in accuracy on average. However, large deviations (e.g., when true α is 0.23 and we specify it to be 0.9) significantly impair the performance up to 20%, meaning that our framework is sensitive to α for large deviations. The aforementioned sensitivity pattern of α is favorable because it means that the framework tolerates small deviations and meanwhile stays sensitive to large differences, which allows the framework to be robust to manual errors. Also, it also validates our idea of improving the framework performance by adjusting its value adaptively. This is because the true α may shift significantly after several observations, and fixing its value may lead to large errors in prediction.

6. Discussion and future work

The evaluation demonstrates the promising performance of our framework in open-world person identification. To further improve our approach for real-world implementation, we identified several directions for future work, including (1) time-evolving variability control, (2) scalability enhancement with more people, and (3) person re-identification across various sensing locations.

First, since our current variability control method is based on previously known data, the data transformation outcome mainly depends on previous footsteps among different people. However, human walking patterns change over time as people's physical and psychological status changes, such as height, weight, emotion, and motivation. These changes can be reflected in the change in walking speed, direction, and magnitude of the footstep forces. As a result, those previous footsteps could be no longer representative of the current and future footsteps of the same person. To incorporate such changes, we plan to explore time-evolving variability control methods that could predict the potential footsteps from each person based on observed trends of their footsteps.

Secondly, the scalability of the system when more people are being observed remains challenging. Due to the variable nature of human footsteps, there is a higher chance that different people walk similarly when more people are known to the database. As footsteps encrypt body information, associations between this information and the features in footstep-induced floor vibration are important to understanding distinguishable features among different people's footsteps. Therefore, we will explore the physical meanings behind footstep-induced vibration signals so that our system is sensitive to minor differences between different people. Also, we will extend our dataset to more people over a longer period to better observe the characteristics of their footsteps.

Finally, re-identifying the same person at a new sensing location with different physical properties of the floor is important for occupancy tracking and monitoring. However, the same footstep impact force at various floor structures results in very different vibration signals due to the discrepancies among these wave propagation mediums. Also, noises in different environments distort the vibration signals further. In order to re-identify a person despite such changes, we will design a robust identification system that adapts to various physical conditions.

The development of this framework enables new opportunities in person identification and smart buildings from three main aspects: (1) potential applications on privacy-friendly smart home systems, (2) performance boosting for other sensing modalities, and (3) advancements in continual learning. To begin with, the structural vibration-based approach provides a privacy-friendly alternative for person identification, which enables a wide spectrum of personalized smart home services such as activity and health monitoring; Secondly, the developed variability reduction approach can also be adapted to other sensing modalities. For example, cameras, wearables, and radio-frequency-based approaches also share the variability challenge because a person may change clothes and movement patterns from day to day. Yet, this challenge has not been addressed thoroughly in the existing person identification literature, especially under-explored in the open-world setting. This provides opportunities to improve their performance using our algorithm. Thirdly, the non-parametric Bayesian model we developed overcomes the well-known issue of “catastrophic forgetting” in the field of continual learning [74]. This means that our model will not “forget” the observed people as more people are observed by storing them as feature distribution parameters. Furthermore, the dataset we collected in this paper can be utilized for analyzing the pedestrian-structure interaction mechanism on different structural types, which can advance the field for better design of human-centered buildings, footbridges, and sidewalks.

In addition to using floor vibration for person identification, it is also possible to infer the structural characteristics from human-induced floor vibration, which can lead to potential applications in structural identification and structural health monitoring. The

model developed in this work can be extended to distinguish the structural layout and loading conditions based on the locations of the footsteps. Opposite to the purpose of person identification, structural identification requires reducing the footstep variability caused by personal differences, and knowing a person's identity is helpful to ensure the consistency of the excitations. On the other hand, incorporating multiple people's gait forces for structural identification provides diverse group excitations, which may induce a larger variety of modes. To achieve this, we will further analyze and characterize the footstep-induced floor vibrations in terms of the structural layout/loading conditions.

7. Conclusion

In this paper, we present an open-world person identification framework, and develop a variability analysis and data modeling approach to overcome the high variability challenge when dealing with footstep-induced structural vibration data. To cope with the multi-source variability within each person, we first develop a feature transformation method based on variability decomposition results, which leads to features that are more separable between people. To model the irregularly shaped features, we model the global distribution based on DPMM and introduce an mCRP to estimate the local non-Gaussian feature distributions. As more people are observed and identified, we update our model parameters and design a hyperparameter to adapt to the changing expectations of the newcomers. The performance of our framework is evaluated through real-world walking experiments on 2 different structures. Our framework achieves up to 92.3% test accuracy on average. It also has consistent results across different structures. The sensitivity analysis further shows the robustness and flexibility of our framework under various conditions.

Declaration of competing interest

The authors have no competing interest to declare.

Data availability

Data will be made available on request.

Acknowledgments

This work was funded by the U.S. National Science Foundation (under grant numbers NSF-CMMI-2026699). The views and conclusions contained here are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either express or implied, of any University, the National Science Foundation, or the United States Government or any of its agencies.

References

- [1] S. Pan, T. Yu, M. Mirshekari, J. Fagert, A. Bonde, O.J. Mengshoel, H.Y. Noh, P. Zhang, FootprintID: Indoor pedestrian identification through ambient structural vibration sensing, *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1 (3) (2017) 1–31, <http://dx.doi.org/10.1145/3130954>.
- [2] Y. Dong, J.J. Zou, J. Liu, J. Fagert, M. Mirshekari, L. Lowes, M. Iammarino, P. Zhang, H.Y. Noh, MD-Vibe: Physics-informed analysis of patient-induced structural vibration data for monitoring gait health in individuals with muscular dystrophy, in: *UbiComp/ISWC 2020 Adjunct - Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*, No. Md, 2020, pp. 525–531, <http://dx.doi.org/10.1145/3410530.3414610>.
- [3] A. Bonde, S. Pan, M. Mirshekari, C. Ruiz, H.Y. Noh, P. Zhang, OAC: Overlapping office activity classification through iot-sensed structural vibration, in: *Proceedings - 5th ACM/IEEE Conference on Internet of Things Design and Implementation, IoTDI 2020*, 2020, pp. 216–222, <http://dx.doi.org/10.1109/IoTDI49375.2020.00028>.
- [4] J. Parmar, S.S. Chouhan, V. Raychoudhury, S.S. Rathore, Open-world machine learning: Applications, challenges, and opportunities, 2021, arXiv. <http://dx.doi.org/10.48550/ARXIV.2105.13448>. URL: <https://arxiv.org/abs/2105.13448>.
- [5] Y. Lu, A. Fleury, J. Boonaert, S. Lecoeuche, S. Ambellouis, Online person identification and new person discovery using appearance features, in: *2015 IEEE International Conference on Evolving and Adaptive Intelligent Systems, EAIS 2015*, 2015, <http://dx.doi.org/10.1109/EAIS.2015.7368794>.
- [6] A. Iosifidis, A. Tefas, I. Pitas, Activity-based person identification using fuzzy representation and discriminant learning, *IEEE Trans. Inf. Forensics Secur.* 7 (2) (2012) 530–542, <http://dx.doi.org/10.1109/TIFS.2011.2175921>.
- [7] Y. Iwashita, R. Baba, K. Ogawara, R. Kurazume, Person identification from spatio-temporal 3D gait, in: *Proceedings - EST 2010 - 2010 International Conference on Emerging Security Technologies, ROBOSEC 2010 - Robots and Security, LAB-RS 2010 - Learning and Adaptive Behavior in Robotic Systems, IEEE*, 2010, pp. 30–35, <http://dx.doi.org/10.1109/EST.2010.19>.
- [8] L. Huang, C. Wang, Unobtrusive pedestrian identification by leveraging footstep sounds with replay resistance, *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5 (4) (2022) <http://dx.doi.org/10.1145/3494963>.
- [9] Y. Chen, W. Dong, Y. Gao, X. Liu, T. Gu, Rapid: A multimodal and device-free approach using noise estimation for robust person identification, *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1 (3) (2017) <http://dx.doi.org/10.1145/3130906>.
- [10] D. Gafurov, E. Snekenes, P. Bours, Gait authentication and identification using wearable accelerometer sensor, in: *2007 IEEE Workshop on Automatic Identification Advanced Technologies - Proceedings*, 2007, pp. 220–225, <http://dx.doi.org/10.1109/AUTOID.2007.380623>.
- [11] J. Suutala, J. Rönning, Methods for person identification on a pressure-sensitive floor: Experiments with multiple classifiers and reject option, *Inf. Fusion* 9 (1) (2008) 21–40, <http://dx.doi.org/10.1016/j.inffus.2006.11.003>.
- [12] P. Connor, A. Ross, Biometric recognition by gait: A survey of modalities and features, *Comput. Vis. Image Underst.* 167 (June 2017) (2018) 1–27, <http://dx.doi.org/10.1016/j.cviu.2018.01.007>.
- [13] C. Feng, J. Xiong, L. Chang, F. Wang, J. Wang, D. Fang, RF-Identity: Non-intrusive person identification based on commodity RFID devices, *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5 (1) (2021) 1–23, <http://dx.doi.org/10.1145/3448101>.

- [14] S. Pan, M. Mirshekari, J. Fagert, C.G. Ramirez, A.J. Chung, C.C. Hu, J.P. Shen, P. Zhang, H.Y. Noh, Characterizing human activity induced impulse and slip-pulse excitations through structural vibration, *J. Sound Vib.* 414 (2018) 61–80, <http://dx.doi.org/10.1016/j.jsv.2017.10.034>.
- [15] J. Han, S. Pan, M.K. Sinha, H.Y. Noh, P. Zhang, P. Tague, Smart home occupant identification via sensor fusion across on-object devices, *ACM Trans. Sensor Netw.* 14 (3–4) (2018) <http://dx.doi.org/10.1145/3218584>.
- [16] S. Drira, S.G. Pai, I.F. Smith, Uncertainties in structural behavior for model-based occupant localization using floor vibrations, *Front. Built Environ.* (2021) 13.
- [17] S. Drira, S.G. Pai, Y. Reuland, N.F. Olsen, I.F. Smith, Using footstep-induced vibrations for occupant detection and recognition in buildings, *Adv. Eng. Inform.* 49 (2021) 101289.
- [18] S. Drira, Y. Reuland, S.G. Pai, H.Y. Noh, I.F. Smith, Model-based occupant tracking using slab-vibration measurements, *Front. Built Environ.* 5 (2019) 63.
- [19] J. Fagert, M. Mirshekari, P. Zhang, H.Y. Noh, Recursive sparse representation for identifying multiple concurrent occupants using floor vibration sensing, *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 6 (1) (2022) 1–33.
- [20] Y. Dong, A. Bonde, J.R. Codling, A. Bannis, J. Cao, A. Macon, G. Rohrer, J. Miles, S. Sharma, T. Brown-Brandl, et al., Pigsense: Structural vibration-based activity and health monitoring system for pigs, *ACM Transactions on Sensor Networks* (2023).
- [21] Y. Sun, Y. Chen, X. Wang, X. Tang, Deep learning face representation by joint identification-verification, *Adv. Neural Inf. Process. Syst.* 3 (January) (2014) 1988–1996, [arXiv:1406.4773](https://arxiv.org/abs/1406.4773).
- [22] Z. Li, S. Chang, F. Liang, T.S. Huang, L. Cao, J.R. Smith, Learning locally-adaptive decision functions for person verification, in: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2013, pp. 3610–3617, <http://dx.doi.org/10.1109/CVPR.2013.463>.
- [23] R. Brunelli, D. Falavigna, Person identification using multiple cues, *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (10) (1995) 955–966, <http://dx.doi.org/10.1109/34.464560>.
- [24] L. Zheng, Y. Yang, A.G. Hauptmann, Person Re-identification: Past, Present and Future, Vol. 14, No. 8, 2016, pp. 1–20, [arXiv:1610.02984](https://arxiv.org/abs/1610.02984).
- [25] Y. Wu, Y. Lin, X. Dong, Y. Yan, W. Ouyang, Y. Yang, Exploit the unknown gradually: One-shot video-based person re-identification by stepwise learning, in: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5177–5186, <http://dx.doi.org/10.1109/CVPR.2018.00543>.
- [26] A.K. Jain, P. Flynn, A.A. Ross, *Handbook of Biometrics*, first ed., Springer Publishing Company, 2010, Incorporated.
- [27] J. Anil, H. Lin, S. Pankanti, Biometric identification, *Commun. ACM* 43 (2) (2000) 90–98.
- [28] E. Chirchi, Vanaja, Roselin, M. Wagmare, L., R. Chirchi, E., Iris biometric recognition for person identification in security systems, *Int. J. Comput. Appl.* 24 (9) (2011) 975–8887, URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.259.342&rep=rep1&type=pdf>.
- [29] O. Costilla-Reyes, R. Vera-Rodriguez, P. Scully, K.B. Ozanyan, Analysis of spatio-temporal representations for robust footstep recognition with deep residual neural networks, *IEEE Trans. Pattern Anal. Mach. Intell.* 41 (2) (2019) 285–296, <http://dx.doi.org/10.1109/TPAMI.2018.2799847>.
- [30] M.M. Ali, V.H. Mahale, P. Yannawar, A.T. Gaikwad, Fingerprint recognition for person identification and verification based on minutiae matching, in: *Proceedings - 6th International Advanced Computing Conference, IACC 2016, IEEE*, 2016, pp. 332–339, <http://dx.doi.org/10.1109/IACC.2016.69>.
- [31] S.J. Elliott, S.A. Massie, M.J. Sutton, The perception of biometric technology: A survey, in: *2007 IEEE Workshop on Automatic Identification Advanced Technologies - Proceedings*, 2007, pp. 259–264, <http://dx.doi.org/10.1109/AUTOID.2007.380630>.
- [32] R. Roizenblatt, P. Schor, F. Dante, J. Roizenblatt, R. Belfort, Iris recognition as a biometric method after cataract surgery, *BioMed. Eng. Online* 3 (2004) 1–7, <http://dx.doi.org/10.1186/1475-925X-3-2>.
- [33] S.R. Schweinberger, N. Kloth, D.M. Robertson, Hearing facial identities: Brain correlates of face-voice integration in person identification, *Cortex* 47 (9) (2011) 1026–1037, <http://dx.doi.org/10.1016/j.cortex.2010.11.011>.
- [34] S. Prabhakar, S. Pankanti, A.K. Jain, Biometric recognition: Security and privacy concerns, *IEEE Secur. Priv.* 1 (2) (2003) 33–42, <http://dx.doi.org/10.1109/MSECP.2003.1193209>.
- [35] P. Dallard, T. Fitzpatrick, A. Flint, A. Low, R.R. Smith, M. Willford, M. Roche, London millennium bridge: Pedestrian-induced lateral vibration, *J. Bridge Eng.* 6 (6) (2001) 412–417.
- [36] S.H. Strogatz, D.M. Abrams, A. McRobie, B. Eckhardt, E. Ott, Crowd synchrony on the Millennium Bridge, *Nature* 438 (7064) (2005) 43–44.
- [37] P. Reynolds, A. Pavic, Z. Ibrahim, Changes of modal properties of a stadium structure occupied by a crowd, in: *22nd International Modal Analysis Conference (IMAC XXII)*, Dearborn, Detroit, USA, Vol. 26, 2004, p. 29.
- [38] D.P. Thambiratnam, N.J. Perera, C.M. Abeysinghe, M.-H. Huang, S.S. De Silva, Human activity-induced vibration in slender structural systems, *Struct. Eng. Int.* 22 (2) (2012) 238–245.
- [39] S. Turner, D.R. Middleton, R. Longmire, M. Brewer, R. Eurek, et al., Testing and Evaluation of Pedestrian Sensors, Technical Report, Southwest Region University Transportation Center (US), 2007.
- [40] C. Gaspar, E. Caetano, C. Moutinho, J.G.S. da Silva, Active human-structure interaction during jumping on floors, *Struct. Control Health Monit.* 27 (3) (2020) e2466.
- [41] M. Valero, F. Li, L. Zhao, C. Zhang, J. Garrido, Z. Han, Vibration sensing-based human and infrastructure safety/health monitoring: A survey, *Digit. Signal Process.* 114 (2021) 103037.
- [42] Y. Reuland, S.G.S. Pai, S. Drira, I.F.C. Smith, Vibration-based occupant detection using a multiple-model approach, in: J. Caicedo, S. Pakzad (Eds.), *Dynamics of Civil Structures*, Vol. 2, Springer International Publishing, Cham, 2017, pp. 49–56.
- [43] M. Mirshekari, J. Fagert, S. Pan, P. Zhang, H.Y. Noh, Obstruction-invariant occupant localization using footstep-induced structural vibrations, *Mech. Syst. Signal Process.* 153 (2021) 107499, <http://dx.doi.org/10.1016/j.ymssp.2020.107499>.
- [44] L. Shi, M. Mirshekari, J. Fagert, Y. Chi, H.Y. Noh, P. Zhang, S. Pan, Device-free multiple people localization through floor vibration, in: *DFHS 2019 - Proceedings of the 1st ACM Workshop on Device-Free Human Sensing*, 2019, pp. 57–61, <http://dx.doi.org/10.1145/3360773.3360887>.
- [45] S. Pan, A. Bonde, J. Jing, L. Zhang, P. Zhang, H.Y. Noh, BOES: Building Occupancy Estimation System using sparse ambient vibration monitoring, in: *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2014*, Vol. 9061, 2014, p. 906110, <http://dx.doi.org/10.1117/12.2046510>.
- [46] J. Fagert, M. Mirshekari, S. Pan, L. Lowes, M. Iammarino, P. Zhang, H.Y. Noh, Structure- and sampling-adaptive gait balance symmetry estimation using footstep-induced structural floor vibrations, *J. Eng. Mech.* 147 (2) (2021) 04020151, [http://dx.doi.org/10.1061/\(asce\)em.1943-7889.0001889](http://dx.doi.org/10.1061/(asce)em.1943-7889.0001889).
- [47] Y. Dong, J. Liu, H.Y. Noh, GaitVibe+ Enhancing Structural Vibration-Based Footstep Localization Using Temporary Cameras for in-Home Gait Analysis, in: *Proceedings of the 20th ACM Conference on Embedded Networked Sensor Systems*, 2022, pp. 1168–1174.
- [48] Y. Dong, M. Iammarino, J. Liu, J. Codling, J. Fagert, M. Mirshekari, L. Lowes, P. Zhang, H. Noh, Ambient Floor Vibration Sensing Advances Accessibility of Functional Gait Assessment for Children with Muscular Dystrophies, 2023.
- [49] S. Pan, C.G. Ramirez, M. Mirshekari, J. Fagert, A.J. Chung, C.C. Hu, J.P. Shen, H.Y. Noh, P. Zhang, SurfaceVibe: Vibration-based tap & swipe tracking on ubiquitous surfaces, in: *Proceedings - 2017 16th ACM/IEEE International Conference on Information Processing in Sensor Networks, IPSN 2017*, ACM, 2017, pp. 197–208, <http://dx.doi.org/10.1145/3055031.3055077>.
- [50] M. Mirshekari, J. Fagert, S. Pan, P. Zhang, H.Y. Noh, Step-level occupant detection across different structures through footstep-induced floor vibration using model transfer, *J. Eng. Mech.* 146 (3) (2020) 1–18, [http://dx.doi.org/10.1061/\(ASCE\)EM.1943-7889.0001719](http://dx.doi.org/10.1061/(ASCE)EM.1943-7889.0001719).
- [51] M. Mirshekari, J. Fagert, S. Pan, P. Zhang, Y. Noh, Physics-guided model transfer for human gait monitoring using footstep-induced floor vibration, in: *Structural Health Monitoring 2019: Enabling Intelligent Life-Cycle Health Management for Industry Internet of Things (IIOT) - Proceedings of the 12th International Workshop on Structural Health Monitoring*, Vol. 2, 2019, pp. 2037–2044, <http://dx.doi.org/10.12783/shm2019/323337>.

- [52] L. Van der Maaten, G. Hinton, Visualizing data using t-SNE, *J. Mach. Learn. Res.* 9 (11) (2008).
- [53] N.T. Khiem, P.T. Hang, Frequency response of a beam-like structure to moving harmonic forces, *Vietnam J. Mech.* 38 (4) (2016) 223–238, <http://dx.doi.org/10.15625/0866-7136/6235>.
- [54] R.A. Bjork, Positive forgetting: The noninterference of Items intentionally forgotten, *J. Verb. Learn. Verb. Behav.* 9 (3) (1970) 255–268, [http://dx.doi.org/10.1016/S0022-5371\(70\)80059-7](http://dx.doi.org/10.1016/S0022-5371(70)80059-7), URL: <https://www.sciencedirect.com/science/article/pii/S0022537170800597>.
- [55] G. Miller, Human memory and the storage of information, *IRE Trans. Inf. Theory* 2 (3) (1956) 129–137, <http://dx.doi.org/10.1109/TIT.1956.1056815>.
- [56] MetrixInstrument, Accelerometers versus velocity sensors - What's the difference? 2023, URL: <https://www.metrixvibration.com/applications/accelerometer-or-velocity-sensor>.
- [57] J. Wei, Comparing the MEMS accelerometer and the analog geophone, *Lead. Edge* 32 (10) (2013) 1206–1210.
- [58] Y. Dong, J. Fagert, P. Zhang, H.Y. Noh, Stranger detection and occupant identification using structural vibrations, in: *European Workshop on Structural Health Monitoring*, Springer, 2022, pp. 905–914.
- [59] B. Scholkopf, K.-r. Mullert, Fisher discriminant analysis with kernels, 1999, pp. 41–48.
- [60] J. Yang, A.F. Frangi, J.Y. Yang, D. Zhang, Z. Jin, KPCA plus LDA: A complete kernel fisher discriminant framework for feature extraction and recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (2) (2005) 230–244, <http://dx.doi.org/10.1109/TPAMI.2005.33>.
- [61] A.J. Izenman, Linear discriminant analysis, in: *Modern Multivariate Statistical Techniques*, Springer, 2013, pp. 237–280.
- [62] S. Wold, K. Esbensen, P. Geladi, Principal component analysis, *Chemometr. Intell. Labor. Syst.* 2 (1–3) (1987) 37–52.
- [63] X. Yu, Gibbs Sampling Methods for Dirichlet Process Mixture Model: Technical Details, Technical Report, 2009, pp. 1–18.
- [64] Y.W. Teh, Dirichlet process, in: *Encyclopedia of Machine Learning*, 2011, pp. 280–287.
- [65] A. Ghosal, Subhashis and Van der Vaart, *Fundamentals of Nonparametric Bayesian Inference*, Vol. 44, Cambridge University Press, 2017.
- [66] P. Mitra, Riten and Muller, *Nonparametric Bayesian Inference in Biostatistics*, Springer, 2015.
- [67] A.J. Izenman, Review papers: Recent developments in nonparametric density estimation, *J. Amer. Statist. Assoc.* 86 (413) (1991) 205–224, <http://dx.doi.org/10.1080/01621459.1991.10475021>.
- [68] Sparkfun, SM-24 Geophone Element, Input/Output, Inc, 2006.
- [69] S. Anchal, B. Mukhopadhyay, S. Kar, Person identification and imposter detection using footstep generated seismic signals, *IEEE Trans. Instrum. Meas.* 70 (2021) <http://dx.doi.org/10.1109/TIM.2020.3022486>.
- [70] M. Markou, S. Singh, Novelty detection: A review - Part 1: Statistical approaches, *Signal Process.* 83 (12) (2003) 2481–2497, <http://dx.doi.org/10.1016/j.sigpro.2003.07.018>.
- [71] E. Poongothai, A. Suruliandi, Person re-identification using kNN classifier-based fusion approach, *Int. J. Adv. Intell. Paradigms* 16 (2) (2020) 113–131.
- [72] P.J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, *J. Comput. Appl. Math.* 20 (C) (1987) 53–65, [http://dx.doi.org/10.1016/0377-0427\(87\)90125-7](http://dx.doi.org/10.1016/0377-0427(87)90125-7).
- [73] D. Zhang, Y. Wang, Investigating the separability of features from different views for gait based gender classification, in: *Proceedings - International conference on pattern recognition*, 2008, pp. 15–18, <http://dx.doi.org/10.1109/icpr.2008.4761872>.
- [74] T.L. Hayes, K. Kafle, R. Shrestha, M. Acharya, C. Kanan, Remind your neural network to prevent catastrophic forgetting, in: *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VIII* 16, Springer, 2020, pp. 466–483.