Title
Collaboratively Adaptive Vibration Sensing System for High-fidelity Monitoring of Structural Responses Induced by Pedestrians

Authors
Pan, Shijia
Xu, Susu
Mirshekari, Mostafa
et al.

Publication Date
2022-08-15

DOI
10.3389/fbuil.2017.00028

Peer reviewed

Collaboratively Adaptive Vibration Sensing System for High Fidelity Monitoring of Structural Responses Induced by Pedestrians

Shijia Pan[a], Susu Xu[b], Mostafa Mirshekari[b], Pei Zhang[a], Hae Young Noh[b]

[a]*Electrical and Computer Engineering, Carnegie Mellon University*
[b]*Civil and Environment Engineering, Carnegie Mellon University*

Abstract

This paper presents a collaboratively adaptive vibration monitoring system that captures high fidelity structural vibration signals induced by pedestrians. These signals can be used for various human activity monitoring by inferring information about the impact sources, such as pedestrian footsteps, door open closing, dragging objects. Such applications often require high fidelity (high resolution and low distortion) signals. Traditionally, expensive high resolution and high dynamic range sensors are adopted to ensure sufficient resolution. However, for sensing systems that use low-cost sensing devices, the resolution and dynamic range are often limited; hence this type of sensing methods is not well explored ubiquitously. We propose a low-cost sensing system that utilizes 1) a heuristic model of the investigating excitations and 2) shared information through networked devices to adapt hardware configurations and obtain high fidelity structural vibration signals. To further explain the system, we use indoor pedestrian footstep sensing through ambient structural vibration as an example to demonstrate the system performance. We evaluate the application with three metrics that measure the signal quality from different aspects: the sufficient resolution rate to present signal resolution improvement without clipping, the clipping rate to measure the distortion of the footstep signal, and the signal magnitude to quantify the detailed resolution of the detected footstep signal. In experiments conducted in a school building, our system demonstrated up to 2X increase on the sufficient resolution rate and 2X less error rate when used to locate the pedestrians as they walk along the hallway, compared to a fixed sensing setting.

Keywords: Structural vibration sensing, indirect sensing, pedestrian monitoring

1. Introduction

Structural vibration sensing for pedestrian monitoring has been applied for various spatio-temporal information acquisition purposes. Works have been done on human information monitoring through vibration induced by their activities, including identity [1, 2, 3],

Email addresses: shijiapa@cmu.edu (Shijia Pan), susux@cmu.edu (Susu Xu), mmirshekari@cmu.edu (Mostafa Mirshekari), peizhang@cmu.edu (Pei Zhang), noh@cmu.edu (Hae Young Noh)
URL: https://users.ece.cmu.edu/~shijiapa/ (Shijia Pan)

gender [4, 5], location [6, 7, 8, 9], trajectory [10, 11], traffic [12, 13], activity [14], etc. The non-intrusive nature of this sensing system makes it a promising ubiquitous sensing method. Like other sensing systems, structural vibration sensing generally requires three steps in order to fulfill its purposes: signal acquisition, feature extraction, and information learning.

A large amount of research has been focusing on feature extraction and information learning for different vibration based applications [3, 4, 6, 9, 10, 13]. However, if the raw signals acquired are already distorted (signal clipping) or of low resolution, the learning can hardly compensate for such information loss. One way to improve the signal fidelity is to use sensors with high dynamic range and high resolution. These sensors are often expensive and impractical for large-scale deployment. On the other hand, our target signals induced by pedestrian vary in signal strength (amplitude) fast and significantly, hence existing adaptive hardware settings methods can hardly adapt fast enough to such changes.

Therefore, in this paper, we present a low-cost high-fidelity vibration signal acquisition system targeting at pedestrian induced structural vibration responses. Our system ensures high signal fidelity by predicting the pedestrian induced vibration signal strength and calculating the hardware configuration setting required. The predictions mainly are through two solutions: 1) for each sensor, it applies heuristic models of structural responses and adapts amplification settings dynamically to maximize signal resolution while minimizing clipping rate; and 2) for the networked sensors, the system models the structural variation through multiple locations to improve dynamic adaption of each local amplification setting. Finally, the system detects and outputs high fidelity pedestrian induced vibrations. In general, our paper provides the following contributions:

- We present a hardware system with low-cost off-the-shelf vibration sensors that adapts hardware configuration (e.g., amplification gains) to obtain high fidelity structural vibration responses induced by pedestrians.

- We propose a prediction method that employs both a heuristic model to adapt hardware based on local signal change and a collaborative model to adapt hardware based on global variance.

- We apply the system to an application: pedestrian monitoring by footstep induced vibration and evaluate the system performance in this application.

To the best of our knowledge, this is the first work that investigates sensing signal quality for structural vibration monitoring.

The rest of the paper is organized as follows: In Section 2, we detail related work done on improving signal fidelity and what is the research gap between prior works and this work. Then, Section 3 presents the overview of the system. Next, in Section 4 and Section 5 , we introduced the optimization solution for hardware configuration, and the algorithm design for collaborative adaptation of the hardware. Then in Section 6, we present the system implementation. Section 7 evaluates the system modules and analyzes their abilities to preserve footstep induced structural vibrations with high fidelity. Then, in Section 8 we further discuss the system limitation, trade-offs, and usage. Finally, Section 9 presents the conclusions of this work.

2. Related Work

Prior works that focus on improving sensing signal quality mainly fall into three categories: 1) utilizing expensive enhanced sensors [15], 2) post-processing to restore signal shape [16, 17, 18], and 3) adaptive hardware settings to obtain high fidelity signals [19, 20]. The cost of **enhancing** sensing device to achieve high dynamic sensing range as well as high resolution could make large-scale deployment unrealistic. Previous methods for obtaining high-fidelity sensing data mainly fall into two categories: post- and pre-processing. **Post-processing** methods restore unknown or lost data after data collection [16, 17, 18]. These methods are usually used for audio data and evaluated by signal-to-noise ratio (SNR). Janssen et al. proposed an adaptive interpolation method to restore lost data, with the restrictions that the positions of the unknown samples are known [16]. Miura and his group introduced their clipping removal method through recursive vector projection [18]. Kitic et al. approached the problem from another perspective with iterative hard thresholding and evaluated the results using both signal-to-noise ratio and human listening [17]. However, for those feature-oriented applications such as identification [3] or TDoA-based localization [10], restored data is not dependable enough since it introduces signal artifacts.

Pre-processing methods utilize signal processing techniques to predict signal clipping and limit distortion of an amplified signal [21]. In addition, Zhang et al. proposed the robust taking pressure control (RPC) algorithm to adjust the system sensing configuration for better signal collection [20]. For pedestrian induced excitation, the rapid change and variation makes it difficult if not impossible to achieve high fidelity with those methods.

3. System Overview

The system goal is to capture high fidelity structural vibration signals induced by indoor pedestrians using low-cost low-dynamic-range sensors. It is achieved by maximizing the signal resolution while avoiding signal clipping. Figure 1 shows the relationship between the modules in the system. The vibration signal is obtained by the analog signal acquisition module, which specifies the sensing configuration used. Then the detected impact signals are sent to a collaborative adaptive prediction module where the sensing configuration is decided based on sensing data from the local device as well as from other networked devices.

The rest of the paper introduces the system based on the application of pedestrian monitoring through footstep induced vibration. The causes of variation in detected human footstep strength mainly fall into two categories: human and environmental. Human variation includes two aspects: 1) the personal level as inconsistencies of individual footstep-to-sensor distance within a series of steps (we refer it as a *trace* in the rest of the paper), and 2) the interpersonal level as variations between individuals. Environmental variation occurs when the sensors are placed at different locations, which have different impact response due to structural factors like beams and partitions.

To accommodate these variations, the system, first of all, needs to have a variety of applicable hardware configurations that support the signal variation range (Section 4). Then the system determines the hardware configuration settings to through the collaboratively adaptive algorithm (Section 5).

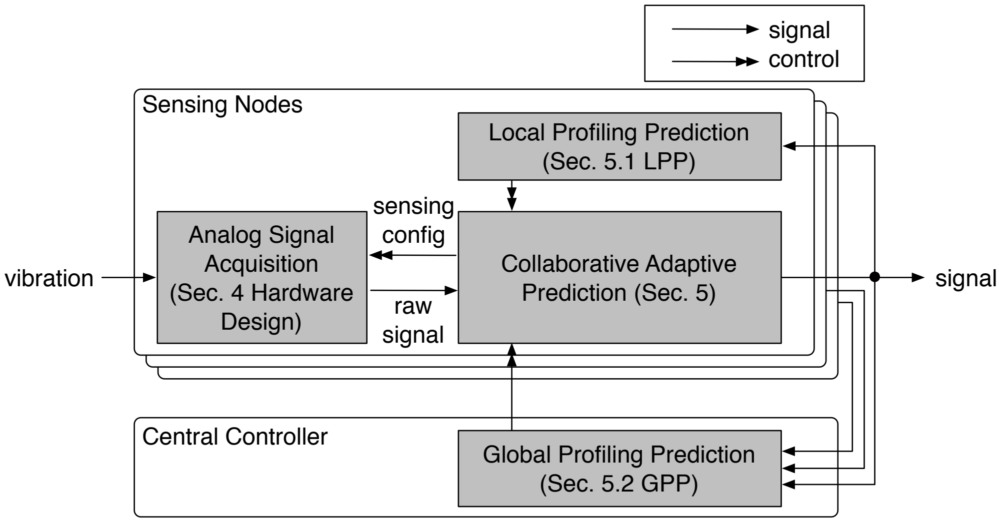

Figure 1: System overview.

4. Hardware: Sensing Configuration Optimization

The hardware configuration selection is the foundation of the sensing system. The goal of the selection is to use minimum number of amplifiers to satisfy the sensing requirement, which we solve through an optimization problem. We define an amplified footstep signal that is represented by a range of integer values as of 'sufficient resolution' when that range is over a selected threshold. For a different system or application, this threshold can be defined differently. The goal of optimization is to maximize the probability that a detected signal falls in the sufficient resolution interval after amplification with a limited number of amplifiers. How do we select amplification gain so that amplifiers allow a step signal on a surface to have sufficient resolution? First, we explain the relation between the concept of amplification and signal resolution (Amplification and Signal Resolution). Next, since the optimization mainly targets footstep strength change in a trace due to footstep-to-sensor distance variation, we model the distribution of the signal amplitude at different locations on a floor plane (Signal Amplitude Distribution). We form the optimization problem (Objective Function) to maximize the probability that a signal with the modeled distribution falls in the sufficient resolution range with limited amplification settings and obtain the optimal solution (Optimal Solution). Finally, the hardware design using the optimal solution is discussed.

Before we form the optimization problem, we list the notations used in the following sections in Table 1. In order to model the amplitude distribution of footstep impulses measured by a sensor, we represent the floor with a 2-dimensional X-O-Y Cartesian Coordinate plane. Since the calculation depends only on the relative locations of the footsteps and the sensor, we simplify the computation by taking the sensor's location as the origin on the plane without losing generality. We make four assumptions to form the optimization problem:

Assumption 1. *The sensing area A is a circular area with the sensor at the origin (0,0).*

Table 1: Notations

Notation	Descriptions of Notations
$X-O-Y$	Cartesian coordinates of sensor/footstep position
T	Input signal amplitude
T_1	Output signal threshold 1
T_2	Output signal threshold 2
k	Signal amplitude measured 1 unit distance from sensor
d	Distance between sensor and footstep
(L_1, L_2)	Footstep location in X-O-Y
g_i	ith amplifier; where $\forall\ i,\ 1 < g_i < g_{i+1}$
n	Number of amplifiers
A	Sensing area
R	Radius of A
$F_T(t)$	Cumulative distribution function of T
$F_D(d)$	Cumulative distribution function of d in A

Assumption 2. *Attenuation model* $T \propto \frac{1}{\sqrt{d}}$ *[22, 11, 23]. When* $d > R$*, the impulse is outside the sensing area* A*, so we assign* $T = 0$*.*

Assumption 3. *The probability distribution of* $(L_1, L_2) \in$ *X-O-Y is a uniform distribution, that is, the probability that a footstep falls on any point in the sensing area is the same.*

Assumption 4. *The number of amplifiers is smaller than the least number needed to properly amplify the raw signal over the whole input signal range* $(\frac{k}{\sqrt{d}}, T_2)$*, i.e., amplification ranges do not overlap.*

Table 2: Amplitude and Resolution

Amplitude	Resolution
$(0, T_1)$	Insufficient
$[T_1, T_2]$	Sufficient
$(T_2, +\infty)$	Clipping (distorted)

4.1. Amplification and Signal Resolution

The analog-to-digital converter using limited number (resolution) of values to describe a signal within a specific voltage range; hence, for each impulsive vibration signal investigated, the amplification that maximize the resolution is different. For an analog-to-digital converter of the specific resolution, a signal that is represented with a large enough number of different values is defined as sufficient resolution. This indicates that the amplified signal falls into a designated voltage range of $[T_1, T_2]$. For different applications requirements, the optimal range of $[T_1, T_2]$ can be different. For example, human identification may require higher resolution signal to achieve high accuracy compared to the application of presence detection. Thus, identification application may have a higher optimal value for T_1 than that of presence detection. We quantify the relation between signal amplitude and resolution level as shown in Table 2. If a signal is amplified by the gain of g and its output falls into the range of $[T_1, T_2]$, then the original range of the signal is $[\frac{T_1}{g}, \frac{T_2}{g}]$. In that case, the sufficient resolution interval for input signal amplitude is expanded to $[\frac{T_1}{g}, \frac{T_2}{g}] \cup [T_1, T_2]$. With multiple available amplification gains, say $1 = g_0 < g_i < g_n$ $(0 < i < n)$, the system can cover sufficient resolution intervals within the full expected signal range. Although the method is applicable for any g values, considering the footstep signal range, it is practical to assume that the signal does not need to be amplified down, therefore we have $g_0 = 1$ here.

$$SigRange = [\frac{T_1}{g_0}, \frac{T_2}{g_0}] \cup [\frac{T_1}{g_1}, \frac{T_2}{g_1}] \cup \cdots \cup [\frac{T_1}{g_n}, \frac{T_2}{g_n}] \tag{1}$$

With this definition of $SigRange$, we further interpret the optimization goal as follows. Given the number of amplification configurations (amplifier gain) n, find a set of amplification gains $1 = g_0 < g_i < g_n$ $(0 < i < n)$ so that the probability of the input signal amplitude that belongs to the $SigRange$ is maximized.

4.2. Signal Amplitude Distribution

To select the optimal amplification setting combination, we need to understand the possible signal amplitudes (T) and their distribution. To simplify the model, we consider an ideal surface described by Assumption 1 and 3 as a start. On an ideal surface, the distance (d) between the footstep and the sensor affects this distribution. Therefore, we can estimate the probability of obtaining a signal of amplitude T from the probability of a step falling on a point of d away from the sensor, where a relationship between d and T as $T = \frac{k}{\sqrt{d}}$, $(k > 0)$ can be specified. Based on Assumption 2, the value k is derived from the absolute value of the impulse strength, which is caused by interpersonal level difference and not modeled in the optimization problem.

To model the clipping of amplifiers, we define a threshold T_2: when $T > T_2$, the amplitude is too large and exceeds the upper bound output, meaning the signal is clipping. The amplitude in the clipped range $(T_2, +\infty)$ will always be sensed as the value T_2. In that case, according to Assumption 2, given the circular area A around a sensor, we formulate the

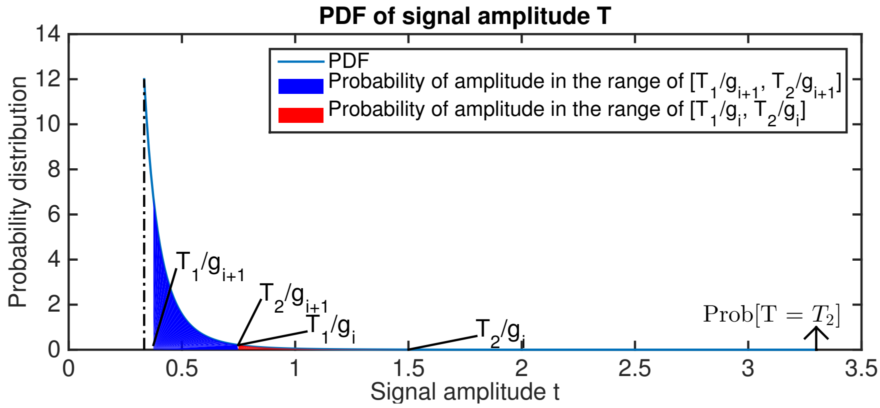

Figure 2: Probability distribution of the signal amplitude T. Note that the distribution is continuous in interval $[\frac{k}{\sqrt{R}}, T_2)$ with the function $f_T(t) = \frac{4k^4}{R^2t^5}$, while discrete at $T = T_2$ due to clipping. This figure shows the scenario where $k = 1$, $R = 9$ and $T_2 = 3.3$. The red region is smaller than the blue region, which means when g_i increases, the probability that the signal amplitude lies in the sufficient resolution interval also increases. Considering $d \leq R$, we can derive that $\frac{k^2}{T^2} \leq R$, which leads to $\forall t \in [\frac{T_1}{g_i}, \frac{T_1}{g_i}]$, $\frac{T_2}{g_i} > \frac{T_1}{g_i} \geq \frac{k}{\sqrt{R}}$, which is the constraint shown in Eq. 7. In order to prevent overlapping of the red region and the blue region, the constraint in Eq. 8 should be satisfied.

amplitude T as a function of distance (d) and the impulse strength (k):

$$T = \begin{cases} T_2 & d \in (0, \dfrac{k^2}{T_2^2}) \\ \dfrac{k}{\sqrt{d}} & d \in [\dfrac{k^2}{T_2^2}, R] \\ 0 & d \in (R, +\infty) \end{cases} \tag{2}$$

Once we understand the relation between d and T, in order to derive the distribution of T, we first calculate the distribution of d. Assumption 1 defines $O = (0, 0)$, so the distance between the sensor and the footstep can be represented as $d = \sqrt{L_1^2 + L_2^2}$. Assumption 3 defines the probability distribution of (L_1, L_2), which can be applied here to derive the probability distribution of d as Eq. 3.

$$F_D(d) = \begin{cases} \dfrac{d^2}{R^2} & 0 \leq d \leq R \\ 1 & d > R \end{cases} \tag{3}$$

Then we can derive the cumulative distribution function (CDF) of the signal amplitude from Eq. 2 and 3, and formulate it in Eq. 4

$$F_T(t) = P(T \leq t) = \begin{cases} 0 & t \in [0, \dfrac{k}{\sqrt{R}}) \\ 1 - \dfrac{k^4}{R^2t^4} & t \in [\dfrac{k}{\sqrt{R}}, T_2) \\ 1 & t \in [T_2, +\infty) \end{cases} \tag{4}$$

Figure 2 indicates that the probability distribution of amplitude is continuous in the interval $[\frac{k}{\sqrt{R}}, T_2)$, while discrete at $T = T_2$. For the continuous part, the probability density function of amplitude (PDF) $f_T(t)$ decreases when t increases. Together with Assumption 4, this implies that in the optimal solution, the sufficient resolution intervals of different amplifiers should not overlap unless we have more than enough amplifiers to cover the entire input signal range, which violates Assumption 4. That is, $\forall g_i < g_j$, if $\frac{T_2}{g_j} > \frac{T_1}{g_i} > \frac{T_1}{g_j}$, there must be $g_i' < g_i$ and $\frac{T_2}{g_j} = \frac{T_1}{g_i'}$, such that probability that amplitude lies in $[\frac{T_1}{g_j}, \frac{T_2}{g_j}] \cup [\frac{T_1}{g_i'}, \frac{T_2}{g_i'}]$ is greater than probability that in $[\frac{T_1}{g_j}, \frac{T_2}{g_j}] \cup [\frac{T_1}{g_i}, \frac{T_2}{g_i}]$ (i.e., $F(\frac{T_2}{g_i}) - F(\frac{T_1}{g_j}) < F(\frac{T_2}{g_i'}) - F(\frac{T_1}{g_j})$).

4.3. Objective Function

We use an optimization problem to describe the goal of our amplification setting selection, which is to maximize the probability that the vibration signal amplitude lies in the sufficient resolution interval. We formulate the optimization problem into Eq. 5.

$$\max_{g_1,\cdots,g_n} \sum_{i=0}^{n} F(\frac{T_2}{g_i}) - F(\frac{T_1}{g_i}) \tag{5}$$

$$s.t. \quad 1 < g_i < g_{i+1} \quad \forall i \in \{1, \cdots, n-1\} \tag{6}$$

$$\frac{T_1}{g_i} \geq \frac{k}{\sqrt{R}} \quad \forall i \in \{1, \cdots, n\} \tag{7}$$

$$\frac{T_2}{g_{i+1}} \leq \frac{T_1}{g_i} \quad \forall i \in \{1, \cdots, n-1\} \tag{8}$$

Three constraints are applied to the optimization problem:

1. Constraint in Eq. 5. We simplify the calculation by define the order of amplification gain g_i is monotone increasing with i. We consider g_0 to represent the scenario where there is no amplifier applied, therefore the gain is $g_0 = 1$, and $[\frac{T_1}{g_0}, \frac{T_2}{g_0}]$ is the sufficient resolution interval of the raw signal.
2. Constraint in Eq. 7. Assumption 2 asserts that $d \leq R$, which leads to $\frac{k^2}{T^2} \leq R$, therefore we can derive that $\forall t \in [\frac{T_1}{g_i}, \frac{T_2}{g_i}]$, $\frac{T_2}{g_i} > \frac{T_1}{g_i} \geq \frac{k}{\sqrt{R}}$.
3. Constraint in Eq. 8. Because $\forall i, j \in \{1, \cdots, n\}$, $(\frac{T_1}{g_i}, \frac{T_2}{g_i})$ can not overlap with $(\frac{T_1}{g_j}, \frac{T_2}{g_j})$ and $g_i < g_{i+1}$, the signal that gets clipping when g_{i+1} is used should not be of insufficient resolution when the next level of gain g_i is applied.

4.4. Optimal Solution

To solve the optimization problem (Section 4.3) using the cumulative distribution function of signal amplitude from Eq. 4, the objective function can be rewritten as

$$S = \frac{k^4}{R^2}(\frac{1}{T_1^4} - \frac{1}{T_2^4}) \cdot (1 + \sum_{i=1}^{n} g_i^4) \tag{9}$$

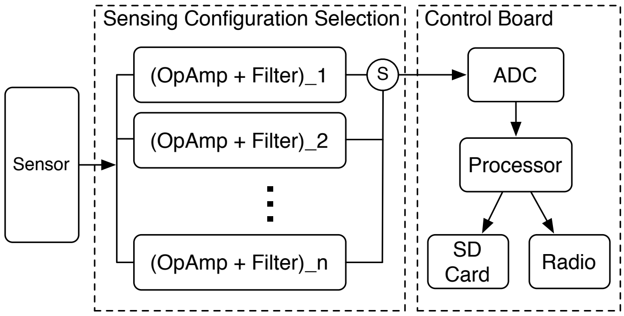

Figure 3: Adaptive amplification module. n levels of the OpAmp are designed to allow the analog signal to be amplified to different ranges. For each iteration, a level of OpAmp is selected (S), and sent to Analog-to-Digital Convertor (ADC). Then the digitized signal is sent to processor for further analysis.

where $\frac{k^4}{R^2}(\frac{1}{T_1^4} - \frac{1}{T_2^4})$ is a positive constant. Thus, we can maximize the objective function S by maximizing $\sum_{i=1}^{n} g_i^4$, which provides the optimal solution

$$g_i = (\frac{T_1}{T_2})^{n-i} \cdot \frac{\sqrt{R} \cdot T_1}{k} \quad \forall i \in \{1, \cdots, n\}. \tag{10}$$

The variable k is a structural characteristics determined value affected by the damping factor of the structure. This result is used to select the optimal amplification gain values in our implementation introduced in Section 4.5.

4.5. Hardware Design using Optimal Solution

To allow the system to obtain signals with different amplification gains, we design the sensing configuration board with multiple amplification settings. As shown in Figure 3, in a situation with n different amplification configurations, the raw signal will go through the sensing unit with each one. Instead of collecting signals from all different configurations, the system selects the optimal one to obtain the signal. Collecting from n configurations limits the sampling rate to $1/n$ due to the system sampling rate limitation as well as radio band width limitation. Then the signal from the selected configuration is digitized and stored.

To obtain the structural variable k for the model, we generate a modeling impulse (for example, a ball-drop with a designated strength) at the edge of the targeting sensing area (a designated R that is determined by the structural noise level), and the system tunes amplification gain g_n to allow the impulse to achieve the highest resolution possible. Then we calculate the value k based on the tuned g_n and the equation $g_n = \sqrt{R} \cdot T_1/k$. After that, we calculate the rest of the gain g_i, $i = 0, ..., n-1$ based on the defined T_1 and T_2, as well as the structural factor k.

5. Algorithm: Collaborative Adaptive Prediction

In order to adapt to signal strength variation caused by pedestrian locations and structural factors, our system operates on two interconnected levels of feedback control as shown in Figure 1: local profiling prediction and global profiling prediction. Local profile prediction refers to the process by which an individual sensing unit uses the data it collects to predict the optimal amplification settings for the next footstep-induced signal. Global profile prediction refers to the collaborative prediction performed by multiple sensing units operating with one another. Together, they serve to provide feedback using known signals to infer and predict optimal amplification selections for future signals on both local and global levels.

5.1. Local Profile Prediction (LPP)

The goal of the LPP is to achieve high resolution for the low signal-to-noise ratio step signals by changing the amplification setting during a pedestrian approaching/leaving the sensor. It predicts the optimal configuration for the next footstep signal that the sensing nodes will detect. To achieve this, the system first detects footstep-induced signals (Step Event Detection). Then, it analyzes the detected signals' resolution condition (Signal Resolution Analysis). Finally, based on the analysis, it makes a prediction on the next step's amplitude (Optimal Configuration Prediction).

5.1.1. Step Event Detection

The system detects distinctive signal segments induced by footstep impulses, which we refer to as *Step Events* in the rest of the paper. They are extracted from the vibration signals through anomaly detection based on a Gaussian model of the background noise (i.e., the signal detected when there is no impulse on the structure) [11]. We utilize a sliding window to collect the background noise signal. The system calculates the signal energy for each windowed signal, with noise modeled by a Gaussian distribution $\mathcal{N}(\mu, \sigma)$. If the signal energy in the window falls outside 3σ range of the Gaussian model, we consider the window to contain a detected step event since it is an abnormal segment.

5.1.2. Signal Resolution Analysis

Understanding the current Step Event's resolution condition allows the system to predict the optimal configuration for the next Step Event. The Step Event resolution is deduced from the relation between the analog signal amplitude and resolution shown in Table 2. For an N-bit analog-to-digital converter configuration, the $T_1(v)$ and $T_2(v)$ are converted to a function of N as $DT_1(N)$ and $DT_2(N)$. These thresholds are applied on the detected Step Event range to determine the signal's resolution class based on the relation demonstrated in Table 2.

5.1.3. Optimal Configuration Prediction

The optimal configuration for the next Step Event is obtained using Algorithm shown in Figure 4 with two main steps: 1) predict the amplitude of the next Step Event and 2) calculate the amplification gain that allows maximum resolution without clipping.

To predict the amplitude of the next step signal, the system looks into $Th_{history}$ number of prior step signals' condition. When there are less than $Th_{history}$ number of steps detected in history, the decision is made by prior step signal. If the step history is almost linear, which is the most common step energy change behavior when the steps are far away due to

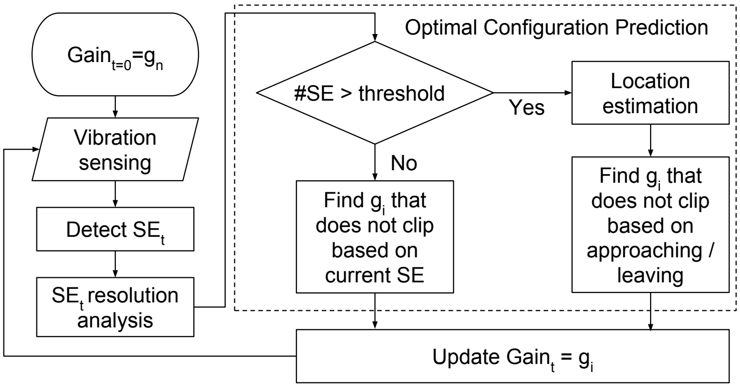

Figure 4: Local profile prediction algorithm.

the noise, the system predicts the next step amplitude Amp_{t+1} with linear model estimated from the step history. On the other hand, if the step history is not linear, which occurs when steps are near the sensor, the system predicts the next step amplitude Amp_{t+1} with the $1/d$ model [11] estimated from the step history.

To calculate the amplification gain, we separate the cases where the pedestrian approaches and leaves the sensing area. When the pedestrian approaches, the system chooses to overestimate the predicted amplification by the $Noise_{Amp}$ in order to find the maximum level of amplification gain that will keep $Amp_{t+1} + Noise_{Amp}$ from getting clipped. On the other hand, when the pedestrian leaves, the system chooses to underestimate the predicted amplification by the $Noise_{Amp}$ to find the maximum level of amplification gain that will keep $Amp_{t+1} - Noise_{Amp}$ from getting clipped. Then the system adjusts the amplification gain based on this calculation.

5.2. Global Profile Prediction (GPP)

The goal of the GPP is to achieve low distortion (e.g., clipping) for the high amplitude step signals by utilizing historical information from neighboring sensors. In practical deployments, structural factors such as building beams and partitions, increase footstep strength variance. Such complications may cause different sensors to observe different local sensing behavior, e.g., if a sensor is deployed near a beam, the detected footstep amplification is lower than that of a sensor located between two beams. This type of structural variation between different sensors/locations can be propagated through the sensor network based on the pedestrian moving direction detection and allow sensors to improve their sensing resolution with the historical information from other sensors.

GPP can either perform alone or be used with LPP to improve signal fidelity by taking structural variation into account. In this section, we introduce how the GPP works alone to achieve high resolution signal acquisition for high signal-to-noise ratio step signals. Instead of processing on the Step Event level, GPP works on the Trace Event level (the vibration signal induced by a person passing by the sensor, containing contiguous detected Step Events). First, it obtains the direction of the target trace (Trace Event Direction). Then, it predicts the pedestrian's trace (Trace Prediction), i.e., to specific neighbor sensing node, based on walking direction. GPP propagates the pedestrian walking information towards these neighboring sensing nodes that the pedestrian might pass based on their walking direction. These nodes rely on their location specifications (Location Specification) and the pedestrian walking direction to make predictions..

5.2.1. Trace Event Direction Estimation

The Trace Event direction allows our system to determine which neighboring sensing nodes a pedestrian approaches and which node they are heading away from. So that the system can inform these neighbor nodes of possible structural anomalies causing signal changes, which we will detail in Location Specification. At least two sensing nodes are required to determine the stride direction based on the relative timing of approaching and leaving different sensors [11]. Each sensing node detects the footstep when a pedestrian passes by. When the pedestrian approaches then leaves the sensor, their footstep signal strength will increase then decrease. The spatio temporal information of the footstep signal with the highest energy within a consecutive footstep sequence detected by different sensors indicates the order in which the pedestrian passes sensors. Therefore the system can determine which direction (i.e., from/to which sensor) the pedestrian walks.

5.2.2. Trace Prediction

Propagating the information to the neighboring nodes that need it makes the system robust for ambiguity when people continuously walk by a sensor. To predict which sensor the pedestrian is walking to, the system models all the deployed nodes as vertexes in a graph. If there is a physical route that a pedestrian can walk between two vertexes without passing a third vertex, there is an edge between these two vertexes. We create this graph heuristically at deployment time as a $k \times k$ binary table, where k is the node number, and the table entry value indicates if there is connectivity between two nodes. We choose the binary table for computational search efficiency. When a pedestrian walks in the building and their *stride direction* is detected, the system will notify all the other sensing nodes that share an edge with this node in the graph except the one that the person walked from.

5.2.3. Location Specification

Due to various structural factors such as beams and partitions, sensors may have different sensitivity to the same impulse (i.e., same strength and traveling distance). The goal for the GPP is to achieve high resolution for the high signal-to-noise ratio step signals by utilizing the historical information from neighboring sensors. When multiple pedestrians walk by different sensors/locations, the system learns the different impulse response strength between sensors/locations.

When a pedestrian walks by one sensor and is detected, the system models their step energy change and sends it to the neighboring nodes that the pedestrian will pass by next. The neighboring node then adjusts its own amplification setting based on the historical data, which indicates the impulse response strength variation at these different locations. Then when the pedestrian approaches the neighboring node, the system detects the step signal with highest energy through the structural variation profile as well as detected step signal strength from the last sensor.

5.3. LPP + GPP

To achieve high resolution for both low and high amplitude step signals, we combine LPP and GPP. LPP performs better with low amplitude step signals because the local adjustment mechanism allows these signals to have higher resolution. However, for high amplitude step signals, the prediction is highly affected by the variation/noise in the human step strength, which could lead to over compensation for estimation. On the other hand, GPP performs better with high amplitude step signals because for those low amplitude step signals within one trace, there is no adjusting mechanism. However, the fixed amplification means low amplitude step signals will have low resolution. Therefore, by combining the LPP and the GPP, the system can achieve better performance in step signal resolution.

By combining the LPP and GPP, the system utilizes the LPP to handle step signals with low amplitude when they are far from the sensor. When the amplitude increases and the step history is not linear, instead of using the 1/d model as described in Section 5.1.3, the system relies on the GPP to make the decisions. Instead of using the detected highest step signal energy, the GPP utilizes the step signal energy changing rate detected by the prior sensors and matches the current step history changing rate. The system searches the entire step history of the neighbor nodes and matches the changing rate between continuously detected $Th_{history}$ number of steps that has the least square error to that on record. It then predicts the next step strength.

6. Implementation

To validate our design, we develop a prototype sensing node with $n = 3$ amplification settings. We install three operational amplifiers (LMV385) with customized amplification gains on the sensing configuration board. The processor board is connected to the amplifiers through three analog-to-digital converter pins. Based on Eq. 10 and the sufficient resolution range we defined in Section 4, we have $T_1/T_2 = 1/2$, which leads to the ratio of the optimal gains as $(1/2)^2 : (1/2)^1 : (1/2)^0 = 1 : 2 : 4$. Through empirical measurements of the other constants ($T_1 = 1.5, k = 3 \times 10^{-4}, R = 9$) we obtain optimal gains of 2000X, 4000X, and 8000X.

The geophone we used is SM-24 [24], with the sensitivity of 28.8V/m/s. The theoretical sensing range of the sensor is limited by its max coil excursion, which is 2mm. However, in practical scenarios, the sensing range is limited by the amplifier voltage, which in our system is 3.3V. Therefore, when an amplifier with $g_0 = 1$ is applied, the sensing range of the sensor is 0.1146m/s. When a 10-bits analog-to-digital converter is used, the resolution of the system is 1.12×10^{-4}m/s, which is not enough to observe signals with peak values fall in the range of 10^{-6}m/s and 10^{-4}m/s. Therefore, when an amplifier with a gain of 2000X is applied, the sensing range of the sensor is 5.73×10^{-5}m/s, with a resolution of 5.6×10^{-8}m/s. Compare to the setting of $g_0 = 1$, this setting has less sensing range but higher resolution. Similarly, the gain of 4000X and 8000X enables even higher resolution (respectively 2.8×10^{-8}m/s and 1.4×10^{-8}m/s) with less sensing range (respectively 2.865×10^{-5}m/s and 1.43×10^{-5}m/s). Therefore, by combining multiple settings, the system achieves

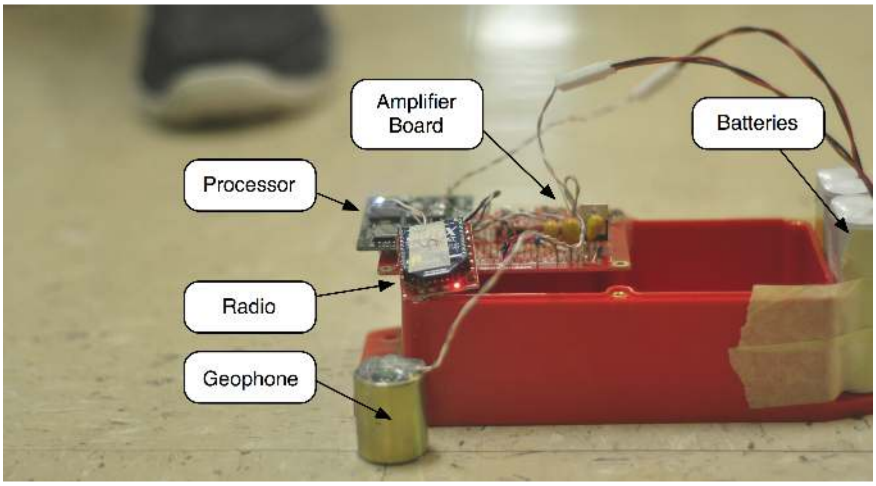

Figure 5: Sensing node.

high resolution (1.4×10^{-8}m/s) as well as high sensing range (5.73×10^{-5}m/s) at the same time.

Implemented amplification gains, however, differ slightly due to practical constraints. We use two-stage amplifiers in the implementation for better signal filtering, because each stage has a differential amplifier serving as a band pass filter. For the first-stage amplifier, we selected the resistor value of $470K\Omega$ over $10K\Omega$ for the amplification gain $470/10 = 47$. When selecting the first-stage gain, the corresponding resistor should be available and the gain should not cause clipping under most circumstances; otherwise, the clipped signal is smoothed by the second-stage's filter. If that happens, the output signal of the second stage will not show evidence of clipping, even though it is distorted. For the second-stage amplifier, we selected the resistor values of $470K\Omega$, $1M\Omega$, and $2M\Omega$ to achieve the designated gain. The calculated gains from this combination were $2200 \approx 47 \times 47$, $4700 = 47 \times 100$, and $9400 = 47 \times 200$, respectively. However, due to the limited open loop gain and filtering effects of the two-level op-amp circuit, the actual gains of the configuration were approximately $g_1 = 2200$, $g_2 = 4400$, and $g_3 = 6400$ [25]. With chosen configurations, over 90% of the impulses induced by detected footsteps are not clipped with g_1, and the background structural vibration noise after amplification is still less than 1/10 of the entire resolution range with g_3.

We placed a prototype sensing node, which is shown in Figure 5, in a hallway and collected data from all configurations when a pedestrian passed by, and the signals are shown in Figure 6. The blue, red and black lines mark signals collected with configuration of g_3, g_2, and g_1 respectively. Figures 6 (a, b, and c) are signals collected with fixed configuration, from which we can see footstep signals of different amplitude. Figure 6 (d) demonstrates the footstep signals of highest resolution without clipping, i.e., the first six footsteps of g_1 configuration, and the rest signals of g_2 and g_3 configurations. To automatically adapt these configurations during sensing, the signal condition prediction is needed, which we will explain in the next section.

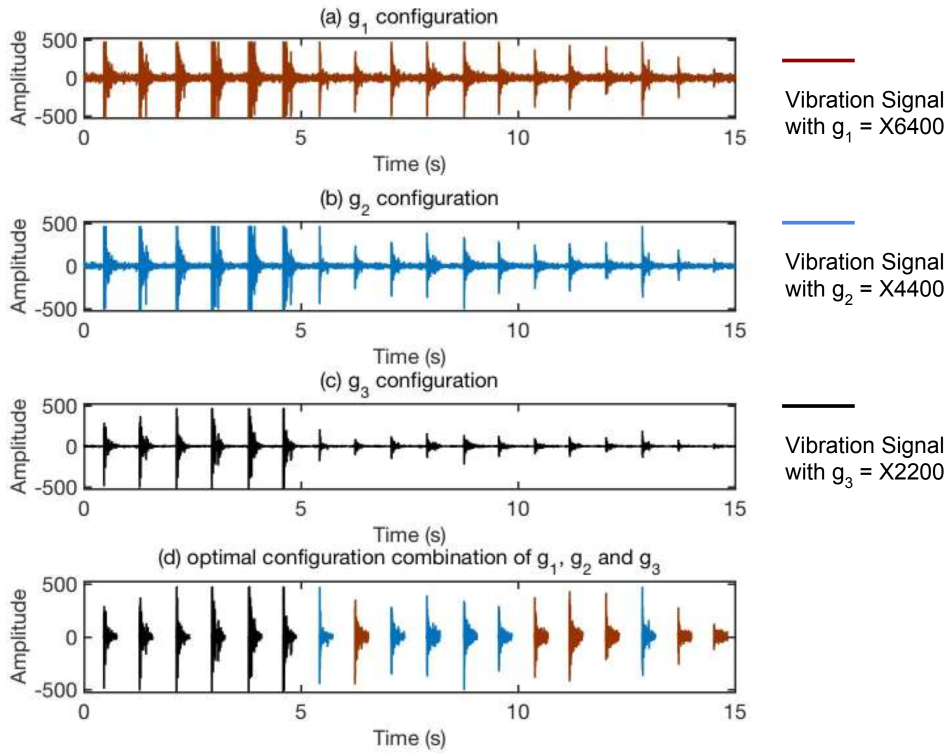

Figure 6: Example of multiple amplification gain configuration. (a, b, and c) are signals collected with amplification gain of 2200, 4400, and 6400. (d) is the signal of optimal resolution selection from events detected in (a, b, and c). The impulses shown as black lines are of gain 2200, and those shown as red lines are of gain 4400. The blue lines are the original line from (a) which is of the starting amplification gain 6400.

7. Evaluation

To understand the system's performance on high fidelity signal acquisition, we conduct pedestrian monitoring experiments to evaluate the system. First of all, we introduce the metrics used to define the 'high fidelity signal', which is used to measure the performance of the system. Next, we present the experiments. Finally, we analyze the results of experiments to verify our system design.

7.1. Evaluation Metrics

Signals which exhibit high distortion or low resolution make it difficult if not impossible to acquire accurate information on vibrations induced by different impulses. Thus, we define 'high fidelity signals' to be signals that minimize signal distortion and noise while maximizing signal resolution. In this subsection, we present the metrics we use to measure and evaluate high fidelity signals quantitatively.

7.1.1. Signal Resolution

Signal resolution in the context of this paper refers to the number of bits used to represent a signal. We defined the sufficient resolution range in Section 4. To determine if an Step Event is of sufficient resolution, its *magnitude* is calculated as the maximum absolute value of the zero mean Step Event signal, and if the magnitude falls into the defined sufficient resolution range, we consider this Step Event is of sufficient resolution. Therefore, the rate of Step Events that of sufficient resolution over all the detected Step Events measures the general signal resolution level. Based on such definition, we define sufficient resolution rate (SRR) as

$$SRR = \frac{\#sufficient\ resolution\ StepEvents}{\#detected\ StepEvents} \tag{11}$$

The higher the SRR value, the more signals of high resolution, and the higher the general signal resolution. In the analysis, we normalize the SRR by the maximum possible SRR value the given system hardware configurations can achieve. This normalized SRR evaluates the performance of LPP and GPP.

7.1.2. Signal Distortion

Signal distortion refers to the degree a measured signal shape differs from the defined baseline. In this work, we focus on the distortion caused by clipping. Therefore, to measure the proportion of Step Events that suffers from such distortion, we calculate the clipping rate of the detected Step Events. The lower the clipping rate, the less signal distortion the system experiences.

7.1.3. Signal Magnitude

Signal magnitude is defined as the maximum absolute value of a zero-mean step event signal. It indicates how many digits are actually used to represent the signal. In the ideal scenario, the system should achieve maximum signal magnitude for each predicted step event signal. However, due to the variation and randomness in human activities as well as the monitored structure, the prediction result can vary, i.e., even an Step Event is count as of sufficient resolution, it might not have maximum magnitude. On the other hand, for different definitions of sufficient resolution, the same magnitude may be of sufficient or insufficient resolution. Therefore, we used magnitude to reveal detailed information about each Step Event.

7.2. Experiment

We conducted experiments to evaluate the system from three different perspectives. First of all, to understand the variables of the proposed system, we evaluated the calculated configuration setting, LPP, and GPP respectively through a simulation with different numbers of amplification levels ($l < n$) implemented (Section 7.3). Then to evaluate the signal quality with the implemented hardware, we placed five sensing nodes in a busy hallway and measured the signal condition with and without our system (Section 7.4). Finally, we evaluated the system's localization performance by comparing the localization accuracy with and without the adaptive amplification design (Section 7.5).

7.3. Evaluation I: System Variables

The system design is determined by two factors as discussed in Section 4: 1) the definition of sufficient resolution and 2) the implemented number of amplification gains. In this section, we specifically evaluate the system behavior in these two factors under perfect amplification settings by generating an amplified 10-bit signal through a high resolution oscilloscope signal of people walking by one sensor.

In total, 15 traces are collected as the seeds for the 10-bit signal generation. Each seed generates N traces of different amplification settings. The minimum amplification gain does not have any signal beyond the sufficient resolution, while the maximum amplification gain have maximum 0.5% clipped signal among the entire trace of signals. This discrepancy means the starting and ending steps are not clipped while most of the close-to-sensor step signals are clipped. In total 5 sensors with different structural impulse response strength rates are simulated for each collected trace. For the first sensor, the step strength for each trace is derived from the seed, and for the rest of the sensors, the step strength for each step is calculated with a ratio of $structural\_rate \times (1 + human\_noise)$ to simulate the human behavior noise as well as structural variation.

We compare five cases in general: 1) only the LPP algorithm; 2) the baseline, which is defined as the median amplification level available; 3) the ground truth, which is the upper bound performance the system can achieve with the implemented hardware, i.e., the system rejects the settings that result in clipping signal and keeps the highest resolution signal that is not clipped; 4) only the GPP algorithm; and 5) both the LPP and GPP conducted collaborative sensing as discussed in Section 5.3. The acronyms used in the evaluation section are summarized in Table 3.

7.3.1. Sufficient resolution definition

To understand the effects of different sufficient resolution definition, we define the sufficient resolution parameter as $T_2 = 1024$ and $T_1 = i/16\ T_2$, with $i = 1...15$. For each definition case, we generate N level of amplified traces as described earlier and run the LPP algorithm through the N level amplifications. Figure 7 demonstrates the SRR, clipping rate, and signal magnitude of the results: 1) the blue line with + markers demonstrates the LPP algorithm, 2) the red line demonstrates the baseline, 3) the yellow line demonstrates the ground truth result, 4) the purple line with circle markers shows the GPP algorithm, and 5) the green line with cross markers demonstrates results with both LPP and GPP.

When the value of T_1/T_2 is low, meaning a large portion of the signal between -512 and 512 is considered as sufficient resolution, the change between different amplification gain

Table 3: Acronyms

Acronym	Meaning
LPP	Local Profile Prediction
GPP	Global Profile Prediction
SRR	Sufficient Resolution Rate

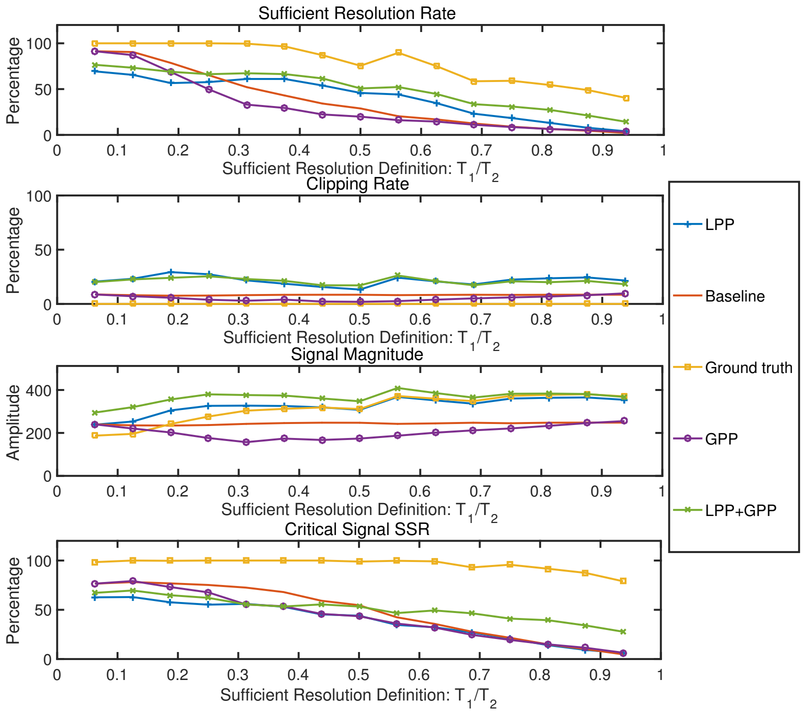

Figure 7: System parameter change: sufficient resolution definition. When the sufficient resolution definition becomes restrict (range $[T_1, T_2]$ reduces), the SRR reduced for all cases, including LPP, baseline, ground truth, GPP, and LPP+GPP.

is large ($g_{i+1}/g_i = T_2/T_1$). Therefore a lower number of amplifiers (N) is needed to cover the variation of the footstep signals. This also means that more low magnitude step signals are considered sufficient resolution, and have a high SRR value and low signal magnitude value. With the increase of the value of T_1/T_2, the clipping rate remains stable, while the signal magnitude increases. This means that the signal quality increases, but due to the increment of the sufficient resolution definition, the SRR decreases. In addition, since the GPP is focused on decreasing the clipping rate and hence increasing the sufficient resolution rate, we further explore a fourth metric, the critical signal SRR, which includes only 5 steps with the highest signal-to-noise ratio in a trace.

LPP in general outperforms the baseline when the definition of the sufficient resolution is over $1/4$ of the entire resolution range in terms of SRR and signal magnitude by an average of 5% and 34% respectively. GPP reduces the clipping rate when compared to the baseline when the sufficient resolution is between $1/4$ and $3/4$ of the entire resolution range, therefore causing a clipping rate 1.6X lower and lowering the signal magnitude as well. When LPP and GPP are combined, the SRR is higher than either algorithm performing alone by 10% on average and raises the signal magnitude by 12% on average. In general, for all the metrics, the LPP and GPP combination follows the trend of LPP and outperforms the LPP most in the critical step signals with high signal-to-noise ratio. This advantage shows an average increase of 10% and up to 4X increase for the highest T_1/T_2 value when the definition of the

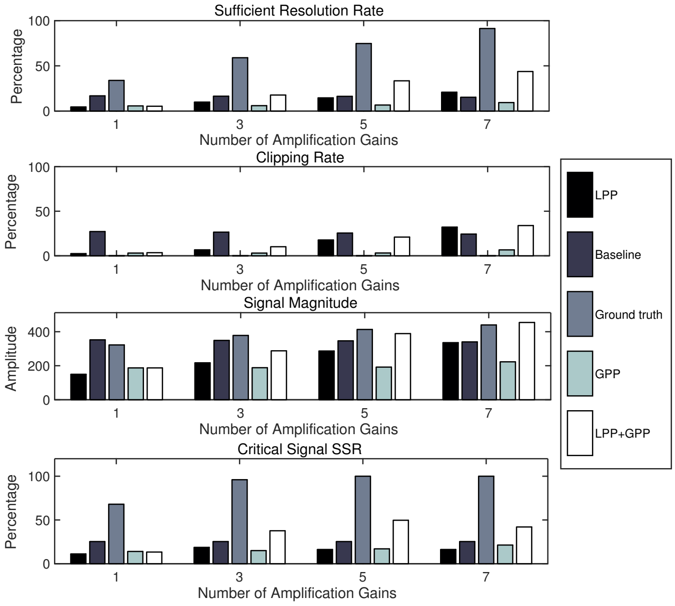

Figure 8: System parameter change: number of amplification gains. When the number of amplification gain implemented is increased, the SRR of the ground truth and the LPP+GPP increases. The clipping rate of the ground truth remains zero since the system can always reject the clipped signal, while that of the LPP+GPP increase due to the prediction error.

sufficient resolution is of a high standard (T_1/T_2 value high) for the critical step signal SRR.

7.3.2. Number of amplifications

In an ideal scenario, the system could have an infinite number of amplification levels to cover an infinite range of amplification needs. However, in reality, only a limited number of amplification levels can be implemented. Because of this, the number of amplifications actually implemented affects the amplification range the system can achieve and therefore affects the system performance. Based on the results from the experiment results in Section 7.3.1, we selected the definition of $T_1/T_2 = 12/16$, which introduces seven levels of amplification gains. The number is selected so that there are large enough available amplification gains involved to demonstrate the system performance when different numbers of amplification gains are implemented.

To understand the number of implementation of amplifications, we selected the median level of amplification, then increase the number of levels by adding one smaller and one larger amplification gain each time, and explore the system performance with these different number of gains. Figure 8 shows the evaluation results of SRR, clipping rate and the signal magnitude when these different numbers of amplification gains are used. Each metric shows an increase trend for all evaluated scenarios except baseline, since baseline is a fixed amplification setting only affected by the definition of the sufficient resolution rate. The more amplification gain levels are implemented, the more adaptable levels can be used for selection, therefore increasing the sufficient resolution rate and signal magnitude. On the

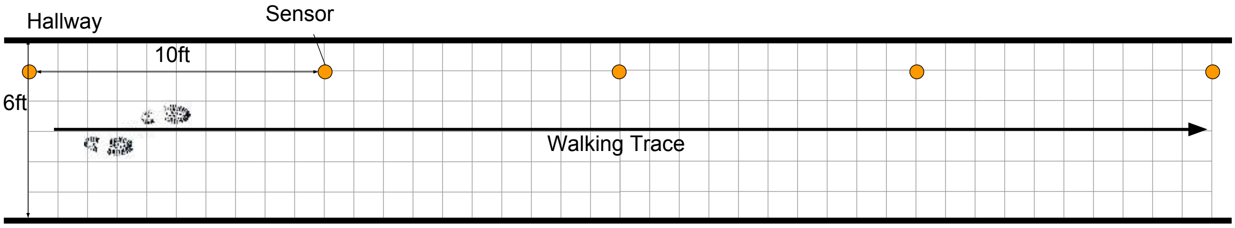

Figure 9: Deployment floor-plan of experimental setups in a school building. Five sensing nodes are deployed in a straight line, approximately three meters apart. Sensors are directly attached to the floor.

other hand, the more choices on the high amplification gains the system is allowed to have, the higher chance the system may select a high amplification gain that causes clipping, hence the increasing clipping rate as well.

7.4. Evaluation II: Adaptive Amplification

To evaluate the system performance in the real-world scenario, we conducted the experiment with a small-scale deployment of five sensing nodes in a school building. We mounted these sensing nodes in a hallway (approximately $20m \times 2m$ area, tile floor) inside the school building as shown in Figure 9. The system sampled the vibration data at 1000 Hz in three amplification configurations. The Real-Time-Clock module on each sensing node provided timestamps for each sensing node's data collection. 10 subjects were asked to walk naturally down a hallway with no restriction on activities (e.g. cell phones, conversing), with the

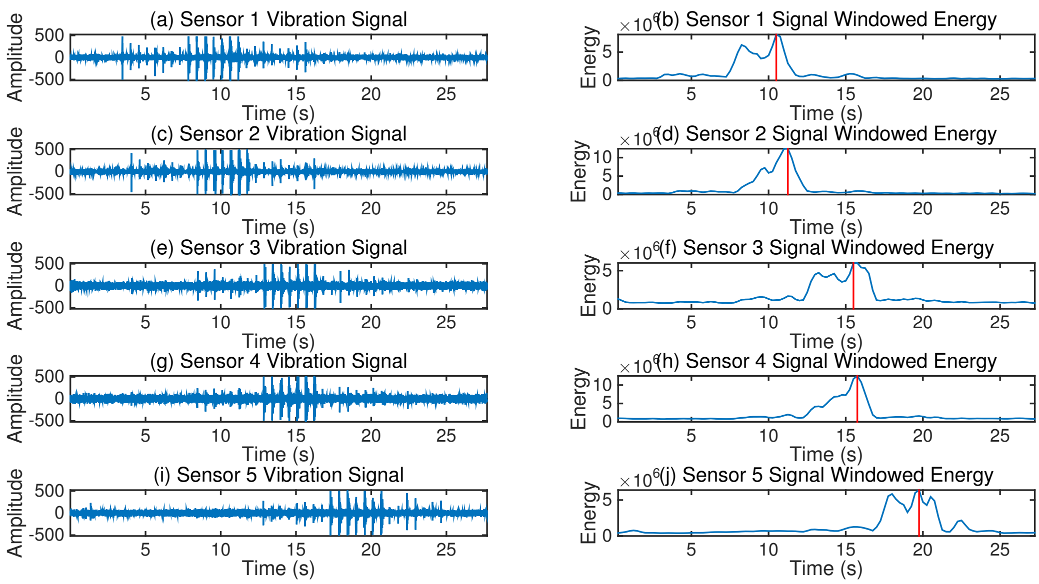

Figure 10: Structural vibration signal detected by sensors when a pedestrian walks by.

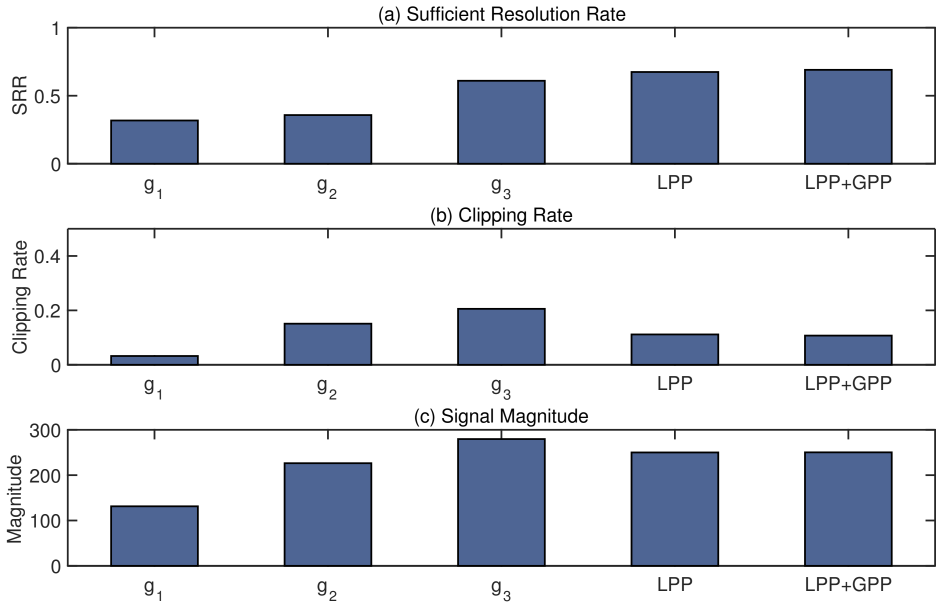

Figure 11: Evaluation of the system with the system performance experiment. Approximate amplification ratios for fixed sensing configurations are $g_1 = 2200$, $g_2 = 4400$, and $g_3 = 6400$. The performance of LPP, as well as LPP+GPP has higher SRR than the fixed configurations. g_3 has highest average signal magnitude resulting from its high clipping rate. Therefore, LPP+GPP's over all performance is improved compared to fixed amplification gains. GPP results are similar to LPP results due to the lack of structural effects in this experiment.

footstep data being picked up by the system. Figure 10 demonstrates an example of one of the subjects walks along the hallway passing five sensors deployed.

With the data from the experiment, we conduct configuration adaptation to compare our system (LPP + GPP) with fixed configurations. Figure 11 (a) shows the normalized SRR from three different fixed amplification configurations ($g_1 = 2200$, $g_2 = 4400$, and $g_3 = 6400$,), an adaptive configuration using only LPP, and an adaptive configuration with LPP+GPP respectively 32%, 36%, 61%, 67%, and 69%. The system improvement comparing to g_1, g_2, and g_3 are at least 1.7X and up to 2X. Note that the algorithm is designed for regular footsteps, i.e., footsteps from the same person are assumed to be same impulses, and uses fixed padding values (P_1 and P_2 as described in Section 5.1.3). However, the randomness in human footsteps introduced prediction errors, leading to an approximately 30% lower SRR value compared to hardware limitation. The LPP achieves higher SRR compared to that of g_1, g_2, and g_3. g_3 and g_2 amplify the near field signal so that many of the signals are clipped, leading to low count on sufficient resolution rate. To validate that, we also demonstrated clipping rate of these configurations in Figure 11 (b), of which values are respectively 3%, 15%, 21%, 11%, and 11%. g_1 obtains most of the near field signals without clipping, but the far field signals are of low resolution due to insufficient amplification, therefore lowering the SRR. In order to understand the low resolution effects, we also present average signal magnitude in Figure 11 (c). As mentioned earlier, the magnitude of a signal is defined as the maximum absolute value of the zero mean signal. The figure shows that fixed gains have an expected effect on magnitude while LPP and GPP sometimes reduce and increase gain as needed. The GPP only made slightly higher SRR comparing to LPP in this experiment due to the relative uniform nature of the structure.

7.5. Evaluation III: Application

We further investigated the system with the application of 1-D localization based on a footstep induced vibration amplitude decay model [26]. Based on the Rayleigh wave propagation model, we used the system to locate where the pedestrian passes the sensor in a hallway. Accurately detecting the passing point allows localization of the person in one dimension. To evaluate that, we fixed the parameters we investigated in Section 7.3 to $T_1/T_2 = 12/16$ and the number of amplification levels as 7. Then we selected the detected step signal with the highest amplitude as the passing point. We compared the step count error of our system to that of the fixed amplification, in this case selecting the middle level (level 4). The average error for our system in detecting the step where the pedestrian is passing the sensor is 0.47m, and the average error for the fixed amplification is 1.13m. Our system shows a 2X less step error when used to locate the pedestrian steps.

8. Discussion

In this section, we discuss the system limitations, the design trade off, the multiple pedestrian sensing condition, and the motivating use-cases for the system.

8.1. System Limitations

The limitations of our system come from mainly two assumptions: 1) the assumption that pedestrian induced structural vibrations have the signal-strength that can be predicted, and 2) the assumption that the algorithm selects from the amplification configurations so that the monitored signal has a sufficient resolution using at least one of the amplifier gains. When the pedestrian induced structural vibration strength is not predictable, e.g., erratic crowd behavior, the system prediction accuracy will decrease, which will reduce the signal fidelity. When the monitored signal is extremely high or low in amplitude, the system configuration may always be clipping or of insufficient resolution, despite the accurate prediction, due to limited number of amplifier configurations.

8.2. Design Trade-offs

Our system implementation considers the trade-offs between a number of analog-to-digital converters and the sampling rate. When the system has access to a large enough number of analog-to-digital converters, which connects to a large enough number of amplification settings, and can sample at a high enough rate, the system, in theory, can obtain highest resolution signal for all monitored structural responses. When the number of analog-to-digital converters is limited, and the sampling rate is high enough, the system can still obtain signals from all available amplification settings. In this case, the system can reject clipped signals, and keep the highest resolution signal without clipping, which is the ground truth scenario in our evaluation. In many practical scenarios, however, it is difficult if not impossible for the system to sample many analog-to-digital converters at the same time, due to limited sampling rates. Then the LPP and GPP are used to predict and select the amplification settings needed, and the prediction errors cause the clipping and insufficient resolution incidences we see in the evaluation.

8.3. Multiple People Sensing

When multiple people passing the sensing area at the same time, the vibration signals induced by their steps mix. When people passing by the sensing area in a different manner (side by side, one after another, towards each other, etc.), their footstep signals may show different energy change patterns, which may not agree with the heuristic rules used in LPP. In this case, our system can utilize the mobility model of the pedestrians and rely on the GPP more than LPP to achieve more stable prediction of the structural response strengths.

8.4. Motivating Use-cases

Monitoring human activity induced excitations enables human information inference. When people walk on the floor, the footstep induced structural vibration can be used to tracking, identify, and count pedestrian in the sensing area [11, 9, 3, 13]. When people lie on the bed, their heartbeats induced vibration can also be detected, hence be used for health status estimation [27]. When people cook in the kitchen, play games in the living room, or cleaning in the house, their interaction with the physical environment induce structural vibration too, which enables activity recognition [28]. Furthermore, this inevitable interaction with the objects in the physical environment makes it possible to turn ambient objects with a flat solid surface into a touch screen [14]. These types of information enable smart home applications such as kid monitoring, kitchen safety monitoring. When deployed in large-scaled scenarios, such as in a nursing home or hospital, the human activity induced excitation monitoring can enable patient/elderly monitoring.

9. Conclusion

In this paper, we introduce a high fidelity structural vibration acquisition sensing system. It is an easy-to-install sparse sensing system that improves the sensing signal fidelity through adapting hardware configurations based on target signal prediction. The prediction is achieved through two key aspects: 1) each individual sensor predict the step strength change based on a pedestrian walking model, 2) networked devices collaboratively predict the step strength through a global profile on a structural variation model. In our pedestrian footstep monitoring application, our system demonstrated up to 2X increase on SRR in our evaluation experiments and up to 2X less error rate when used to locate the pedestrian when they walk along the hallway. We believe that such a signal acquisition system can be applied to various future applications in smart buildings for human activity induced excitation vibration data acquisition.

Acknowledgements

This work is partially supported by Intel, Pennsylvania Infrastructure Technology Alliance (PITA), CMU-SYSU Collaborative Innovation Research Center (CIRC), and Google.

References

[1] A. Ekimov, J. M. Sabatier, Vibration and sound signatures of human footsteps in buildingsa), The Journal of the Acoustical Society of America 120 (2) (2006) 762–768.

[2] A. Itai, H. Yasukawa, Footstep classification using simple speech recognition technique, in: Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on, IEEE, 2008, pp. 3234–3237.

[3] S. Pan, N. Wang, Y. Qian, I. Velibeyoglu, H. Y. Noh, P. Zhang, Indoor person identification through footstep induced structural vibration, in: Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications, ACM, 2015, pp. 81–86.

[4] D. Bales, P. Tarazaga, M. Kasarda, D. Batra, Gender classification using under floor vibration measurements, in: Dynamics of Coupled Structures, Volume 4, Springer, 2016, pp. 377–383.

[5] D. Bales, P. Tarazaga, M. Kasarda, D. Batra, A. Woolard, J. D. Poston, V. Malladi, Gender classification of walkers via underfloor accelerometer measurements, IEEE Internet of Things Journal (2016) 1259 – 1266.

[6] M. Mirshekari, S. Pan, A. Bannis, Y. P. M. Lam, P. Zhang, H. Y. Noh, Step-level person localization through sparse sensing of structural vibration, in: Proceedings of the 14th International Conference on Information Processing in Sensor Networks, ACM, 2015, pp. 376–377.

[7] J. D. Poston, J. Schloemann, R. M. Buehrer, V. Malladi, A. G. Woolard, P. A. Tarazaga, Towards indoor localization of pedestrians via smart building vibration sensing, in: Localization and GNSS (ICL-GNSS), 2015 International Conference on, IEEE, 2015, pp. 1–6.

[8] J. Schloemann, V. S. Malladi, A. G. Woolard, J. M. Hamilton, R. M. Buehrer, P. A. Tarazaga, Vibration event localization in an instrumented building, in: Experimental Techniques, Rotating Machinery, and Acoustics, Volume 8, Springer, 2015, pp. 265–271.

[9] M. Mirshekari, S. Pan, P. Zhang, H. Y. Noh, Characterizing wave propagation to improve indoor step-level person localization using floor vibration, in: SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring, International Society for Optics and Photonics, 2016, pp. 980305–980305.

[10] K. Dobbler, M. Fišer, M. Fellner, B. Rettenbacher, Vibroacoustic monitoring: Techniques for human gait analysis in smart homes, in: Ambient Assisted Living, Springer, 2014, pp. 47–58.

[11] S. Pan, A. Bonde, J. Jing, L. Zhang, P. Zhang, H. Noh, Boes: building occupancy estimation system using sparse ambient vibration monitoring, in: SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring, International Society for Optics and Photonics, 2014, pp. 90611O–90611O–16.

[12] A. Subramanian, K. G. Mehrotra, C. K. Mohan, P. K. Varshney, T. Damarla, Feature selection and occupancy classification using seismic sensors, in: International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, Springer, 2010, pp. 605–614.

[13] S. Pan, M. Mirshekari, P. Zhang, H. Y. Noh, Occupant traffic estimation through structural vibration sensing, in: SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring, International Society for Optics and Photonics, 2016, pp. 980306–980306.

[14] S. Pan, C. 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: Information Processing in Sensor Networks (IPSN), 2017 15th ACM/IEEE International Conference on, IEEE, 2017, pp. 1–12.

[15] A. Barzilai, Improving a geophone to produce an affordable broadband seisometer, Mechanical Engineering, Stanford University Jan. 25.

[16] A. Janssen, R. Veldhuis, L. Vries, Adaptive interpolation of discrete-time signals that can be modeled as autoregressive processes, Acoustics, Speech and Signal Processing, IEEE Transactions on 34 (2) (1986) 317–330.
[17] S. Kitic, L. Jacques, N. Madhu, M. P. Hopwood, A. Spriet, C. De Vleeschouwer, Consistent iterative hard thresholding for signal declipping, in: Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on, IEEE, 2013, pp. 5939–5943.
[18] S. Miura, H. Nakajima, S. Miyabe, S. Makino, T. Yamada, K. Nakadai, Restoration of clipped audio signal using recursive vector projection, in: TENCON 2011-2011 IEEE Region 10 Conference, IEEE, 2011, pp. 394–397.
[19] W. Loetwassana, R. Punchalard, A. Lorsawatsiri, J. Koseeyaporn, P. Wardkein, Adaptive howling suppressor in an audio amplifier system, in: 2007 Asia-Pacific Conference on Communications, IEEE, 2007, pp. 445–448.
[20] J. Zhang, R. Wang, S. Lu, J. Gong, Z. Zhao, H. Chen, L. Cui, N. Wang, Y. Yu, Easicprs: design and implementation of a portable chinese pulse-wave retrieval system, in: Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems, ACM, 2011, pp. 149–161.
[21] A. C. Krochmal, D. P. Stewart, J. E. Whitecar, Variable distortion limiter using clip detect predictor, uS Patent 7,202,731 (Apr. 10 2007).
[22] T. G. Gutowski, C. L. Dym, Propagation of ground vibration: a review, Journal of Sound and Vibration 49 (2) (1976) 179–193.
[23] H. Verhas, Prediction of the propagation of train-induced ground vibration, Journal of Sound and Vibration 66 (3) (1979) 371–376.
[24] Geophone sm-24, `https://www.sparkfun.com/products/11744`, accessed: 2017-01-24 (2006).
[25] R.-t.-R. Low-Voltage, LMV321, LMV358, LMV324 (2009).
[26] I. A. Viktorov, Rayleigh and Lamb waves: physical theory and applications, Plenum press, 1970.
[27] Z. Jia, M. Alaziz, X. Chi, R. E. Howard, Y. Zhang, P. Zhang, W. Trappe, A. Sivasubramaniam, N. An, Hb-phone: a bed-mounted geophone-based heartbeat monitoring system, in: Information Processing in Sensor Networks (IPSN), 2016 15th ACM/IEEE International Conference on, IEEE, Vienna, Austria, 2016, pp. 1–12.
[28] P. A. Kodeswaran, R. Kokku, S. Sen, M. Srivatsa, Idea: A system for efficient failure management in smart iot environments, in: Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services, ACM, 2016, pp. 43–56.