A crowdsourcing-based methodology using smartphones for bridge health monitoring

Article *in* Structural Health Monitoring · December 2018

DOI: 10.1177/1475921718815457

A Crowdsourcing Based Methodology using Smartphones for Bridge Health Monitoring

[1]Qipei Mei, [2]Mustafa Gül
[1] PhD Student
[2]Associate Professor (Corresponding Author: mustafa.gul@ualberta.ca)
Department of Civil and Environmental Engineering, University of Alberta
Edmonton, Alberta, T6G 2W2, Canada

Abstract

This paper presents a novel framework for monitoring and evaluation of population of bridges using smartphones in a large number of moving vehicles as mobile sensors. Within this framework, a damage detection methodology based on Mel-frequency cepstral coefficients and Kullback-Leibler divergence is developed. For this method, Mel-frequency cepstral coefficients of the vibration data collected from smartphones in vehicles crossing bridges are first extracted as features. Then, Kullback-Leibler divergence is used to compare the distributions of features. The damage in a bridge can be identified by quantifying the difference of the distributions obtained for the same bridge. Both numerical and lab experiments are conducted to verify the proposed framework and methodology. In lab experiments, a smartphone and two wireless accelerometers are used for data collection. From our results, it is concluded that the damage existence can be successfully identified using smartphones in a large number of vehicles. Also, it is observed that there is a significant correlation between magnitude of the damage features and the severity of damage. The results show that the method has the potential to monitor a population of bridges simultaneously and in almost real time.

Keywords
Structural Health Monitoring; Crowdsourcing; Smartphone; Mel-frequency Cepstral Coefficients; Kullback-Leibler Divergence

1. Introduction

Civil infrastructure systems, especially bridges, are a critical component of public transportation system. However, as our infrastructure systems age, they suffer from different types of damage which significantly affects their performance and service life. In this context, bridge health monitoring comes into play to reduce the cost for maintenance and extend the service life. In recent years, vibration-based health monitoring techniques have achieved promising progress thanks to the breakthrough in data analysis techniques and rapid development of wireless sensor networks.[1-7]

Although researchers have cut the cost significantly by introducing new data acquisition techniques like wireless sensors, it should still be acknowledged that those techniques have their limitations. First, wireless sensors are active monitoring systems that require labor to install equipment onsite, in spite of the fact they are already much more convenient than traditional wired sensors. Second, consistent and stable power supply is still a bottleneck for long term monitoring.[8] Third, current monitoring techniques are usually conducted by researchers or engineers, which requires advanced knowledge about structural design and sensing technologies.

To overcome these issues, a new path that utilizes crowdsourcing data have attracted increasing interests from researchers in recent years.[9-11] Owing to the development of wireless technologies and internet of things, almost everyone has at least one smart device which is able to connect to the internet. These smart devices are usually equipped with various sensors, such as accelerometer, camera, gyroscope and GPS. In recent years, these sensors have been successfully used for different applications, such as step counting and sleep tracking. It has been reported by several researchers that smartphones can also be used for structural health monitoring purpose,[10-13] but to the best of authors' knowledge, this paper is the first study that proposes a framework for using large number of smartphones on vehicles for bridge health monitoring and presenting analytical and experimental verifications.

This paper presents a novel crowdsourcing based framework for indirect bridge health monitoring by utilizing data from smartphones in a large number of vehicles. The proposed methodology is developed based on Mel-frequency cepstral coefficients (MFCCs) and Kullback-Leibler (KL) divergence as discussed in the next sections. Afterwards, numerical analysis on a bridge with parameters similar to Yang et al [14] is conducted to verify the proposed methodology. Finally, the smartphone app developed for data collection is introduced and the lab experiments are discussed, which are followed by discussions and conclusions.

It should be noted that the framework can easily be extended to other smart devices including smart vehicles. By definition, indirect health monitoring is to place one or more sensors on a moving vehicle and assess the condition of bridge by analyzing the data collected when vehicle is passing across the bridge. Researchers in this area have been able to identify frequency, mode shapes, damping, or the damage on bridges successfully using a single vehicle.[14-25] However, those methods are still considered as active monitoring since specific vehicles must be dispatched for monitoring purpose. Usually, the same vehicle has to be used for all the measurement in order to make the results valid.[26] The conclusions can vary when using different vehicles or at different speeds.

In the proposed framework, the smartphones in vehicles act as sensors. The data collected from smartphones in a large number of vehicles are used to detect damage. The variation of vehicle configurations is taken into consideration in our research. Figure 1 illustrates the overview of the framework. When people drive across the bridges, their smartphones would automatically collect the vibration data and process the vibration data to extract features. Then, the extracted features will be transmitted to the remote database termed as cloud. Finally, the gathered data are analyzed systematically from the remote computers to assess the condition of bridges.

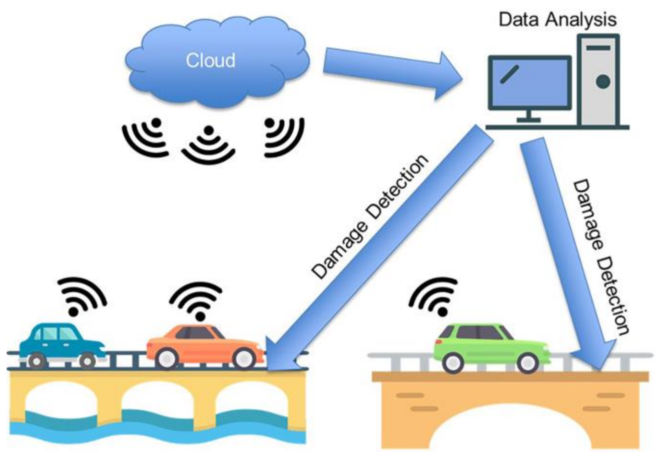


Figure 1 - Overview of crowdsourcing based damage detection technique

There are several advantages when smartphones in a large number of vehicles (i.e., crowdsourcing) that connect to the internet working as mobile sensors instead of a single one is used. First, it can significantly reduce the cost for maintenance since no specific arrangements have to be made for monitoring. The vibration data of bridges can be collected by the smartphones automatically when vehicles cross the bridge. Second, this system has the potential to monitor a population of bridges regularly as long as there are vehicles passing through the bridges. Third, owing to crowdsourcing, this technique is more robust to operational effects and human errors.

2. Methodology

2.1. Overview

In this section, the data analysis methodology within the proposed framework of utilizing smartphones in a large number of vehicles for indirect health monitoring is presented. In our method, it is assumed that there is one smartphone in each vehicle crossing a bridge. As different vehicles pass through the same bridge at different times. MFCCs are extracted for each vehicle (discussed in detail in section 2.4), and the extracted features are transmitted to a feature pool. This feature pool accumulates the features from a large number of vehicles, and an estimate of multi-variate distribution of the features can be made from the feature pool. Considering two different periods, one for baseline state of the bridge and the other for unknown state, it is assumed that there are s_1 vehicles passing across bridge in baseline period and s_2 vehicles for unknown period. Estimates of two distributions (baseline and unknown) can be compared using KL divergence (described in section 2.5). The damage detection procedure described in this paragraph is presented in Figure 2.

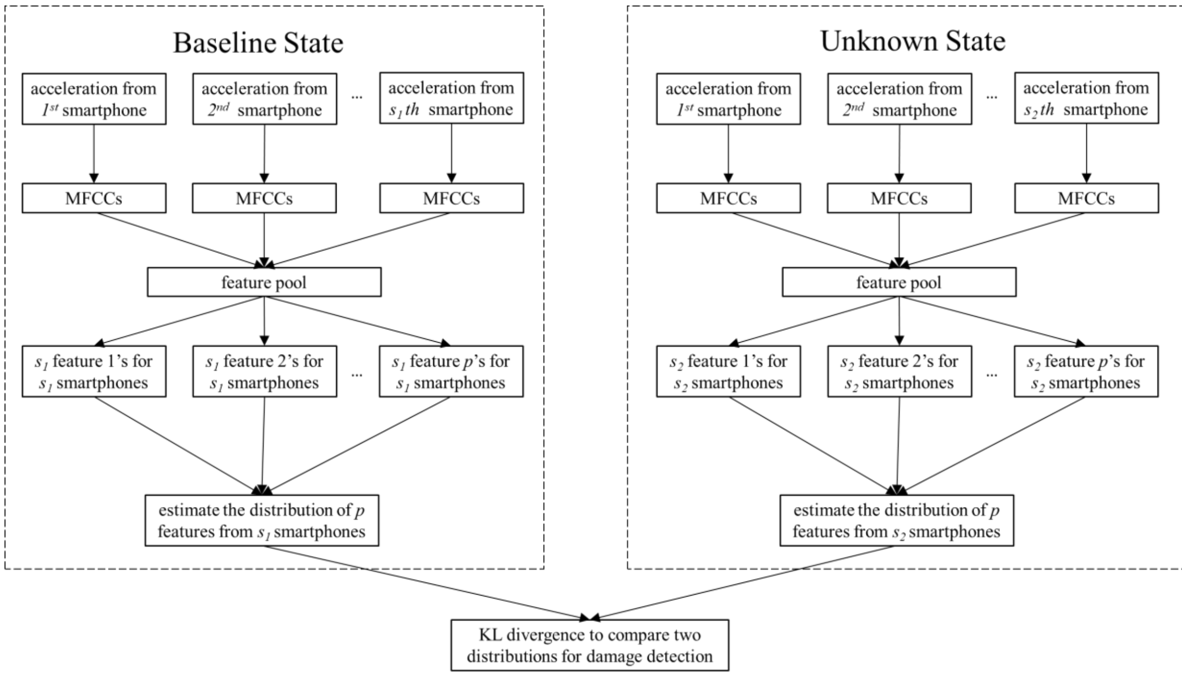


Figure 2 - Damage detection procedure

2.2. Vehicle Bridge Interaction

Vehicle bridge interaction is an important concept for bridge design and maintenance. Due to the dynamic interaction, vibration data collected from either vehicle or bridge would have each other's information in a coupled form.[27] The concept of indirect health monitoring is to indirectly monitor the condition of the bridge utilizing the response of the vehicles. The very early work that extracts the frequencies of the bridge from vibrational data of a passing-by vehicle is from Yang et al.[14, 28] They provide a theoretical solution to represent vehicle's acceleration as function of the dynamic properties of both bridge and vehicle. Following that, researchers have successfully identified the damping[18, 29, 30] and mode shapes[31-33] of the bridge indirectly.

To briefly describe the vehicle bridge interaction process, a simply supported bridge is presented in Figure 3 and a spring mass model is used to simulate vehicle, the VBI equations can be written in equation (1).[28] The first equation is the dynamic equilibrium of the bridge, and the second is for the vehicle.

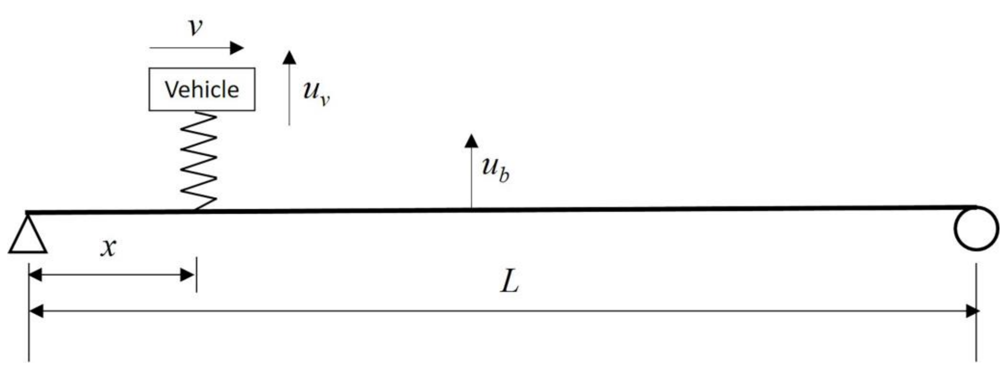


Figure 3 - Spring mass vehicle bridge interaction model

$$\begin{cases} \bar{m}(x)u_b''(x,t) + E(x)I(x)u_b''''(x,t) = c(x,t) \\ m_v u_v''(t) + k_v[u_v(t) - u_b(vt,t)] = 0 \end{cases} \tag{1}$$

where x is the distance to an end of the bridge, v is the speed of the vehicle, t is the current time, $\bar{m}(x)$ is the mass of bridge per unit length, $E(x)$ and $I(x)$ are elastic modulus and moment of inertia of the bridge section, m_v is the mass of the vehicle, k_v is the stiffness of the vehicle, and $c(x, t)$ is the contact force. In addition, u_b and u_v are the displacements of bridge and vehicle relative to the equilibrium position. It should be noted that u_b is related to the location x and time t, and u_v is only related to time t. Also note that equation (1) is a general form which suits all beam-like structures, and $\bar{m}$, E and I will be constants if cross-sectional and material properties are uniform.

The detailed derivation and solution of the equations can be found in Yang and Lin.[28] It could be seen in their paper that the displacement, velocity and acceleration from vehicle include dynamic characteristics of both bridge and vehicle. Standard signal techniques such as Fourier transform is a reasonable way to separate the dynamic information. However, with the existence of road surface roughness, operational effects and measurement noise, the direct use of Fourier transform usually does not work.

2.3. Crowdsourcing with Smartphones in Passing by Vehicles

In traditional indirect health monitoring methods, damage can only be detected when the properties of the vehicle (e.g. natural frequency, speed, etc.) are known or the same vehicle is used for both baseline and unknown cases. In reality, it is impractical to measure all the properties of the vehicle. Also, the damage could propagate slowly, and it is hard to keep the vehicle's properties the same for every test. Any inconsistency in the vehicle configurations would lead to increased uncertainties in the results.

In recent years, thanks to the fast development of the Internet of Things technologies, we can easily access a variety of sensors that are embedded in smartphones or smart connected vehicles, which have access to internet. In our method, instead of using a single vehicle, crowdsourcing data collected from smartphones in a large number of vehicles are processed to extract features. Unlike traditional indirect health monitoring where features are directly compared to identify damage, our proposed method compares the distributions of features estimated from data collected from smartphones in a large number of vehicles. In this way, although the configurations of vehicles may vary for different tests, the statistical characteristics of all the vehicles are expected to be stable and thus changes in the distributions of the features can be used for damage detection. Using crowdsourcing data from smartphones, a population of bridges could continuously be monitored in near real time with the help of several thousands or tens of thousands of vehicles, if not more.

2.4. Mel-frequency cepstral coefficients (MFCCs)

Cepstral analysis is the inverse Fourier transform of the logarithm of the spectrum for a signal, which is a powerful tool for speech recognition and natural language processing. There are a number of varied cepstral analysis techniques, among which Mel-frequency cepstral analysis is one of the most popular ones. The features extracted from Mel-frequency cepstral analysis is called MFCCs, i.e., Mel-frequency cepstral coefficients. The major advantages of MFCCs compared to other feature extraction methods include: 1) Mel-cepstral analysis looks for patterns in a frequency range instead of searching peaks. 2) It gives higher weights for lower frequencies and lower weights for higher frequencies. MFCC has gained great success in the area of voice recognition and natural language processing,[34-36] but has only started to attract a little attention from structural health monitoring researchers in recent years.[37, 38]

The procedure of calculating MFCCs from acceleration data collected from the smartphones can be summarized in four main steps: 1) the Fourier transform of the acceleration signal is computed; 2) the power spectrum in Hertz scale is mapped to Mel scale using Mel filter bank; 3) the logarithms of the mapped powers at each Mel frequency is calculated; 4) the discrete cosine transforms of the log powers are calculated. After implementing these four steps, the amplitudes of the resulting spectrum will be MFCCs. In this way, all the information in the power spectrum would be integrated in MFCCs with higher weights for lower frequencies.

Mel-frequency cepstral analysis is originally designed to simulate human being's auditory system. One of its key steps is to process the data with a set of triangular filters with unevenly spaced intervals, which considers the fact that difference in lower frequency is more significant than in higher frequency. However, MFCCs are built for hearing system which is most sensitive to frequencies from 2000 to 5000 Hz.[39] This frequency range is far larger than the normal frequency range for bridges, which indicates that the algorithm to calculate MFCCs should be adapted before being applied to bridge health monitoring.[40] Considering the characteristics of bridges, two equations are proposed as described in equation (2) and equation (3) for conversion between Mel-scale and Hertz-scale frequencies.

$$m = M(f) = 5\ln(1+\frac{f}{5}) \tag{2}$$

$$f = M^{-1}(m) = 5(e^{m/5}-1) \tag{3}$$

where f is the Hertz-scale frequency and m is the Mel-scale frequency.

The relationship between frequencies in Hertz-scale and Mel-scale are shown in Figure 4. The main function of this conversion is to find points in Hertz-scale with exponentially increasing intervals. In Figure 4, it is seen that evenly spaced points on adapted Mel-scale are mapped to increasingly spaced points in Hertz-scale. The equation for locating the unevenly spaced points are as below,

$$f_i = M^{-1}(M(f_1)+(i-1)\times\frac{M(f_n)-M(f_1)}{n-1}) \tag{4}$$

in which f_i is the i^{th} Hertz-scale frequency corresponding to the i^{th} evenly spaced Mel-scale frequency, and n is the number of key frequencies within the range from f_1 to f_n.

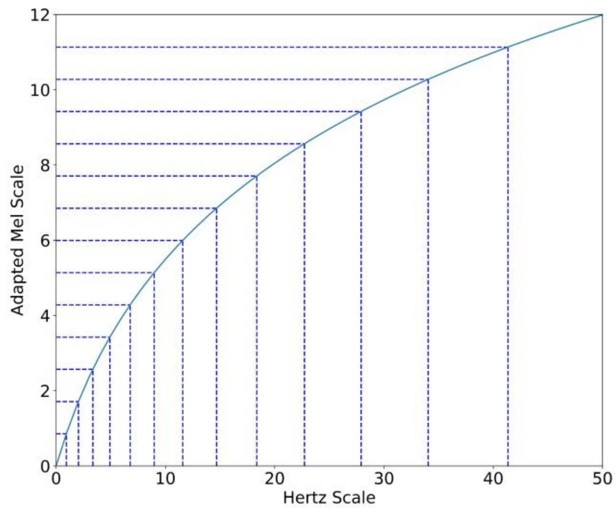


Figure 4 - Conversion between Hertz Scale and Adapted Mel Scale

The next step is to build a series of triangular filters based on key points calculated from equation (4). A single triangular filter is defined according to equation (5).

$$H_i(k) = \begin{cases} 0 & k < f_{i-1} \\ \dfrac{k - f_{i-1}}{f_i - f_{i-1}} & f_{i-1} \le k < f_i \\ \dfrac{f_{i+1} - k}{f_{i+1} - f_i} & f_i \le k < f_{i+1} \\ 0 & k \ge f_{i+1} \end{cases} \tag{5}$$

in which $H_i(k)$ is the i^{th} filter bank that will be applied to the power spectrum and i could be 2,3,4,…,n-1 , and k is the frequency. Figure 5 plots these triangular filters in frequency domain.

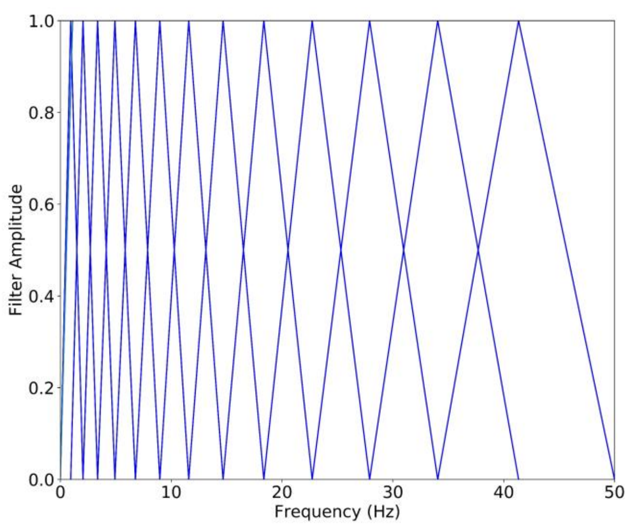


Figure 5 - Triangular filter bank

As described in equation (6), since both filter bank and spectrum are in frequency domain, the total energy for each filter is calculated as dot product of filter and spectrum.

$$E(i)=\sum_k H_i(k)\times F(k) \tag{6}$$

where $F(k)$ is the Discrete Fourier Transform of the original acceleration signal, and $H_i(k)$ is the i^{th} filter in the filter bank.

Then, the logarithms of the total energy for each filter can be taken at each of the Mel frequencies.

$$\log E(i)=\log(\sum_k H_i(k)\times F(k)) \tag{7}$$

The last step to calculate MFCCs is to take the Discrete Cosine Transform (DCT) of the logged energy. Since there are in total n-2 filters, the DCT can be written as in equation (8).

$$X_j=\sum_{i=2}^{n-1}\log E(i)\cos\left[\frac{\pi}{n\text{-}2}\left(i-\frac{3}{2}\right)j\right] \tag{8}$$

where X_j is denoted as MFCC. The number of MFCCs is equal to the number of triangular filters, but practically only some of them are selected for further analysis.

2.5. Kullback–Leibler (KL) divergence

For baseline case, extracting MFCCs from s_1 vehicles, an $s_1\times p$ matrix can be formed, in which p is the number of coefficients kept from MFCCs. It should be acknowledged that this $s_1\times p$ matrix should follow a multi-variate distribution. For an unknown state, if there are s_2 vehicles passing through the bridge, the resulting $s_2\times p$ matrix should follow another multi-variate distribution. The damage can be identified by comparing these two distributions. In this paper, KL divergence, which is deemed as one of the few powerful methods for comparing multi-variate distribution, is utilized to compare two probability distributions, i.e., distributions of features from baseline and unknown cases.[41] The general form of KL divergence is presented in equation (9).

$$D_{KL}(N_0\,||\,N_1)=\int_X \log\frac{dN_0}{dN_1}dN_0 \tag{9}$$

where N_0 and N_1 are the probabilities over a set X, and $D_{KL}(N_0||N_1)$ stands for the KL divergence from N_1 to N_0. The KL divergence comes from information theory, and measures the information loss when N_1 is used to approximate N_0. KL divergence is always non-negative, and is well studied for multi-variate distribution comparison, which makes it a suitable mathematical tool for designing damage features (DFs).

In this paper, the features collected from vehicles within a certain period are assumed to follow multi-dimensional Gaussian distribution,[42, 43] the KL divergence can be written as below.[44]

$$D_{KL}(N_0\,||\,N_1)=\frac{1}{2}\left[\ln\frac{|\Sigma_1|}{|\Sigma_0|}+trace(\Sigma_1^{-1}\Sigma_0)+(\mu_1-\mu_0)^T\Sigma_1^{-1}(\mu_1-\mu_0)-k\right] \tag{10}$$

where μ_0 and μ_1 are the mean matrices and Σ_0 and Σ_1 are the covariance matrices for baseline and unknown cases, *trace*() stands for the trace of matrix, and k is the number of features.

The range of KL divergence is [0, ∞). Due to its mathematical form, the KL divergence is exponentially related to the distance of distributions. In order to detect the damage more robustly, it is mapped to a linear relationship with range of [0, ∞) as well. The following transformation is applied and the output is defined as damage feature (DF).

$$DF(N_0, N_1) = \ln\left[D_{KL}(N_0 \| N_1) + e\right] - 1 \tag{11}$$

where e is the Euler's number.

3. Numerical Analysis

In this section, a series of numerical simulations conducted in Abaqus on a simply supported bridge with the same dimensions as in Yang et al. is presented.[14] The length of the bridge is 25 m. The bridge is made of reinforced concrete, which has density of 2400 kg/m^3 and elastic modulus of 27.5 GPa. The cross section of the bridge has an area of 2.0 m^2 and moment of inertia of 0.12 m^4. The first three frequencies of the bridge are 2.08 Hz, 8.33 Hz and 18.75 Hz. The bridge is modeled by beam elements and is divided into 16 elements as numbered in Figure 6.

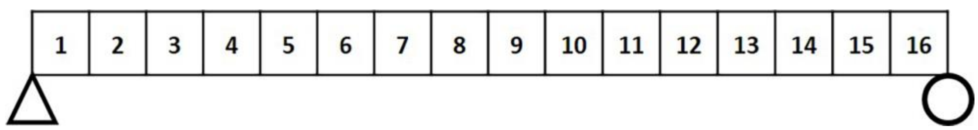


Figure 6 - Mesh grid of the bridge

The vehicle is simply modeled as a spring mass moving over the bridge. The spring constant, mass and speed are defined according to Yang et al.[14] and are varied to model different vehicles. The configurations are summarized in Table 1.

Table 1 - Variation of vehicle configurations

	Vehicle configurations
Mass (kg)	960, 1200, 1440, 1680, 1920, 2160, 2400
Spring constant (kN/m)	200, 250, 300, 350, 400, 450, 500
Speed (m/s)	8, 10, 12, 14, 16, 18, 20

Combing all the possible values for three parameters, there are 7×7×7=343 simulations with different vehicles. If each vehicle is repeated 5 times (artificial random noise is added to each simulation), the vehicle test pool is extended to 343×5=1715 vehicles. In Abaqus, all the vehicles passing through the bridge are simulated. The sampling frequency of the acceleration data is 100 Hz. To consider the measurement noise, all the data are corrupted with 5% artificial noise.

In order to consider the fact that the vehicles passing across the bridge during different periods are not the same, a randomly sampling process is taken for each state of the bridge. For the intact bridge, 50% of 1715 data records are sampled randomly and reserved as baseline case, and the remaining 50% data records are taken for validation (called DC0_N).

In addition to baseline and validation cases, 5 damage cases that are grouped into 3 categories are introduced to verify the proposed method. Similarly, 50% data records for each damage case are randomly sampled as well. To verify the robustness of the method, the random sampling process is repeated 30 times, and at each time the data for the baseline, validation and damage cases are selected differently. The list of numerically introduced damage cases is shown as below:

1) DC1a_N: 15% reduction of stiffness at elements 8 and 9
2) DC1b_N: 30% reduction of stiffness at elements 8 and 9
3) DC2a_N: 15% reduction of stiffness at elements 4 and 5
4) DC2b_N: 30% reduction of stiffness at elements 4 and 5
5) DC3_N: Both ends are changed to fixed supports.

Among the above damage cases, DC1_N and DC2_N are related to stiffness reduction of elements, while DC3_N is the boundary condition change at both ends. DC1_N and DC2_N introduce damage at different locations, and *a* and *b* stand for 15% and 30% stiffness reduction. The stiffness reduction cases are simulated by reducing elasticity modulus of elements.

3.1. DC1a_N and DC1b_N:15% and 30% reduction of stiffness at the mid-span

The DFs for the cases of stiffness reduction at the mid-span are presented in Figure 7. Each point on the figure represent one complete implementation of the algorithm for a single damage case. As shown in the figure, the DFs for DC1a_N and DC1b_N are about 0.3 and 0.45, which are both higher than DC0_N, i.e., validation case. Also, as the damage becomes more severe, the DFs become higher. This indicates that the severity of the damage is also successfully identified. Due to the randomness of the sampling process, the results for 30 trials are different. However, as can be seen in the figure, the DFs for each damage case fluctuate about the average within a reasonable range. The coefficient of variation of DFs for DC0_N, DC1a_N and DC1b_N are 0.128, 0.034 and 0.038, respectively. The difference of DFs among different damage cases are larger than fluctuation for a single damage case.

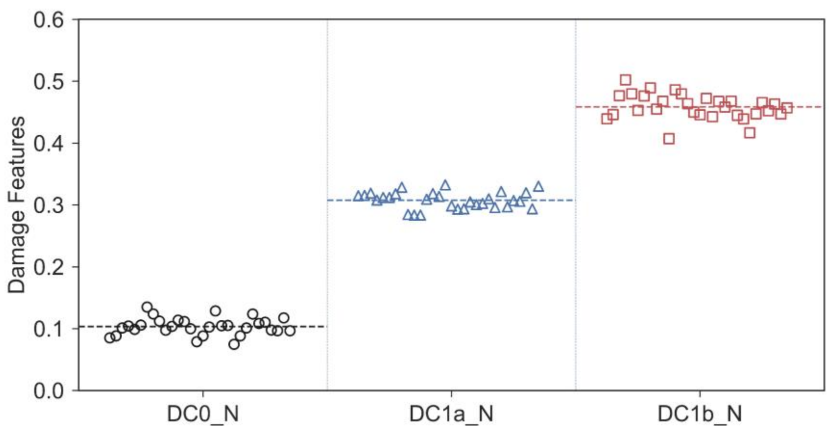


Figure 7 - DFs for DC1a_N and DC1b

3.2. DC2a_N and DC2b_N:15% and 30% reduction of stiffness at the ¼ span

DC2a_N and DC2b_N are also related to stiffness reduction, but at different location than DC1a_N and DC1b_N. In Figure 8, it is seen that both two damage cases are successfully identified as well. The DFs also increases as the damage become more severe. The average DFs for DC2a_N and DC2b_N are about 0.3 and 0.4, and the coefficient of variation for them are 0.050 and 0.037. Comparing Figure 7 and Figure 8, the DFs for the cases with the same extent of stiffness reduction, i.e., DC1a_N and DC2a_N, DC1b_N and DC2b_N, are at about the same level, which in a way shows that the values of DFs are mainly related to the severity of damage.

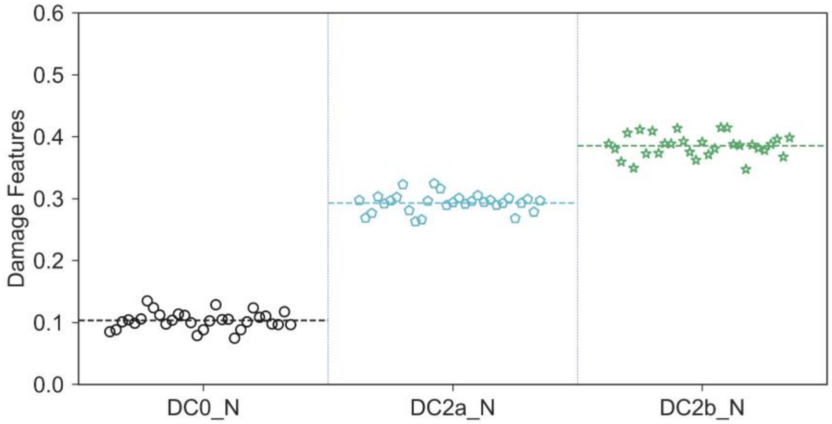


Figure 8 - DFs for DC2a_N to DC2b_N

3.3. DC3_N: Boundary condition change at both ends

Comparing with the local damage cases, DC3_N, boundary condition change at both ends, is the most severe damage. Figure 9 shows the DFs for both DC0_N and DC3_N. It is seen that the average of DFs for DC3_N are around 2.8, which is significantly larger than validation case and the local damage cases. This again shows that the level of DFs is a very useful indicator for assessing the severity of the damage. The coefficient of variation of DFs for DC3_N is 0.011.

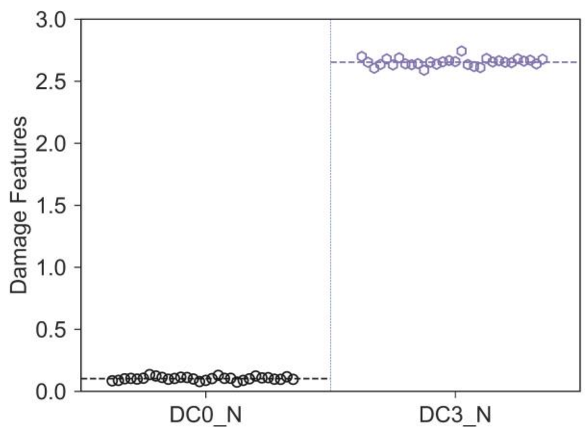


Figure 9 - DFs for DC3_N

3.4. Influence of number of vehicles

As described in previous sections, the DFs fluctuate due to the fact that data from different sets of vehicles are used for baseline and unknown states as well as the influence of operational effect. In this section, the relationship between number of vehicles and the dispersion of DFs is discussed. As shown in Figure 10, the coefficient of variation of DFs is plotted against different number of vehicles. It is seen that the coefficient of variation of DFs are higher for all 5 damaged cases when small number of vehicles are used. As the number of vehicles increases, they all decrease gradually. Therefore, it is safe to say that the approach is more robust when larger number of vehicles are used which is very promising for real-life applications with big-data extracted from large number of vehicles. It should also be noted that curves for coefficient of variation of DC1a_N, DC1b_N, DC2a_N and DC2b_N are very close to each other, which are all far from the curve for DC3_N. This is because the first 4 damage cases are all related to stiffness reduction but DC3_N is about boundary condition change, which shows that the dispersion of DFs is affected by the type of damage.

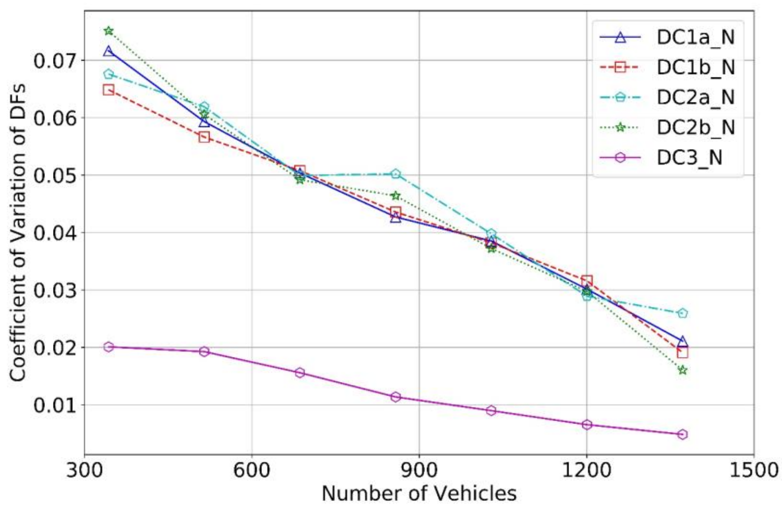
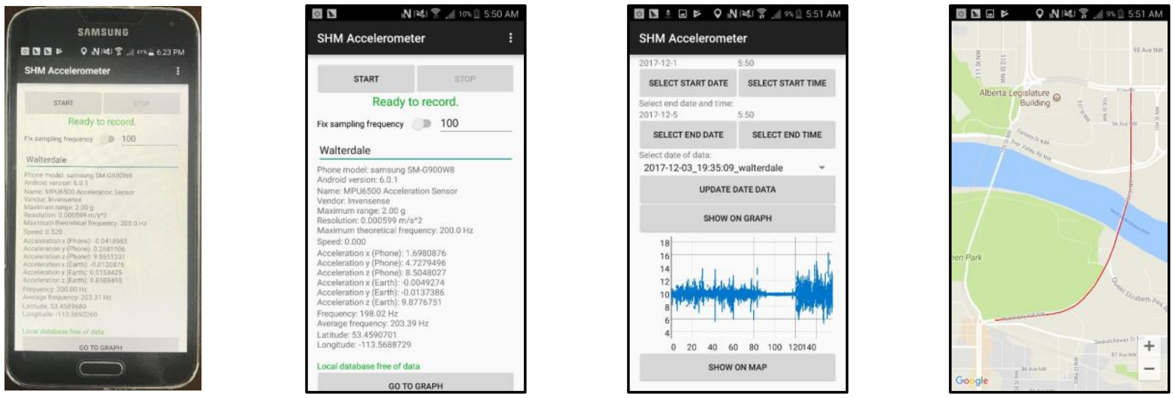


Figure 10 - Relationship between coefficient of variation of DFs and number of vehicles

4. Experimental Studies

4.1. Data collection app

In order to collect the data smoothly, our research team developed an android app that can collect acceleration data and GPS information simultaneously. The screens of the app are presented in Figure 11. The data is collected from the sensors on smartphone and stored remotely on Amazon web service remote database. A Matlab based post-processing program compatible to this app is also developed. This app has the flexibility to set the sampling frequency, collect acceleration data along different axes and record the GPS information along with the vibration information. It also has the ability to visualize the vibration data and GPS information. This app is our first step to apply our techniques onto experiments or real bridges.



(a) Android app | (b) setup screen | (c) acceleration data visualization | (d) GPS data visualization

Figure 11 - Data collection app

4.1.1. Global coordinates

In practice, the smartphone is likely to be placed in arbitrary orientation in the vehicle. The acceleration data collected from the built-in triaxial accelerometer is relative acceleration regarding to smartphone's local coordinate system. However, in order to identify damage in a bridge, the global vertical acceleration is used. To resolve this issue, the orientation of the smartphone is measured through the gyroscope and magnetometer in combination with accelerometer, and a rotation matrix, R, can be extracted from Android API. The global acceleration can be simply calculated by multiplying the inverse of R to the relative acceleration vector.[45]

4.1.2. Sampling frequency correction

It is observed that Android smartphones usually have imperfect sampling frequency, i.e. the sampling intervals would deviate from the preset values. This phenomenon is also reported by other researchers.[46] This not only affects performance but also causes problems to signal processing techniques such as Fourier transform.

In order to study this effect, a simple impact test is conducted on an intact simply supported bridge in the lab, and 10 s acceleration data is recorded using our app. The plot that shows the relationship between time intervals and time is presented in Figure 12. It is seen that the time interval fluctuates around the given value, 0.01 s, except for a couple of samples with very large or very small intervals. The average of all the intervals in this period is 0.00994 s and the standard deviation is 0.00055 s.

Since the time interval does not deviate from preset value significantly, to resolve this issue, a sampling frequency correction technique similar to the one proposed in Ozer et al.[46] is used. The corrected sampling frequency is calculated as the multiplication of original sampling frequency and correction coefficient, where correction coefficient is the ratio of the number of data points for target sampling frequency and original sampling frequency. After the correction, the acceleration data is illustrated in Figure 13. It is seen that the pattern of acceleration data does not change considerably, but now a variety of signal processing techniques can be used.

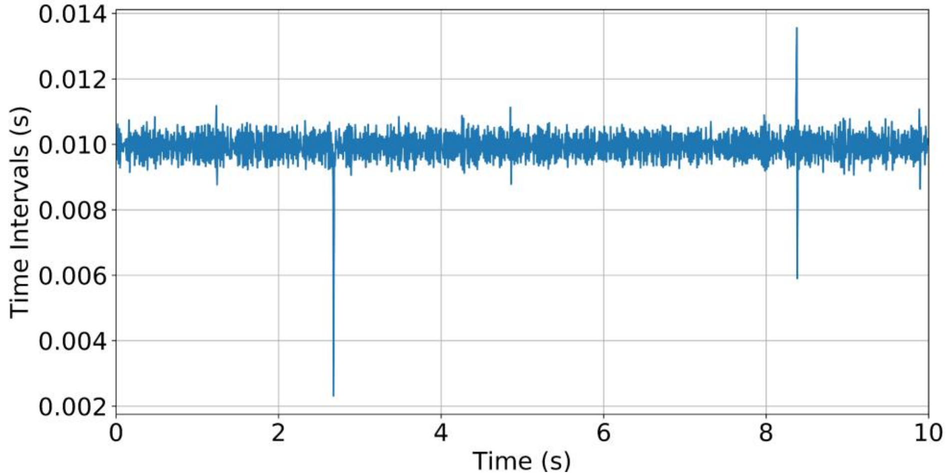


Figure 12 - Time intervals versus time

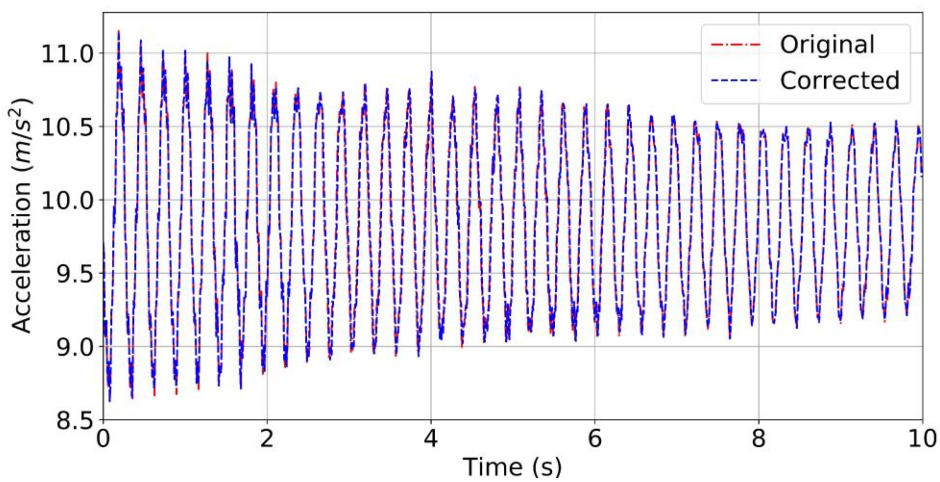


Figure 13 - The original and corrected acceleration data

4.2. Experimental setup

In this section, a simply supported bridge as shown in Figure 14 is setup in our lab to verify the proposed method experimentally. The bridge deck is made of hot rolled steel W44, which has a yield strength of 250 MPa, and ultimate strength of 310 MPa. The modulus of elasticity of the steel is 200 GPa. The dimensions of the bridge are as follows: length 2 m, width 330 mm and thickness 6.35 mm. Similar to numerical analysis, 5 artificial damage cases grouped in 3 categories are applied to the experimental bridge model:

1. DC1a_E: 15% stiffness reduction at the mid-span; (See Figure 15)
2. DC1b_E: 30% stiffness reduction at the mid-span; (See Figure 16)
3. DC2a_E: 15% stiffness reduction at the ¼-span. (See Figure 15)
4. DC2b_E: 30% stiffness reduction at the ¼-span. (See Figure 16)
5. DC3_E: Boundary condition change at both ends (See Figure 17)

Figure 14 - Setup of the lab experiment

In above damage cases, DC1a_E has a 24.8 mm by 250 mm cut centered at mid-span at each side, while in DC2a_E the cut with the same size is centered at 0.5 m from an end. Similarly, a 49.5 mm by 250 mm cut is made at each side of steel bridge as well to simulate 30% stiffness reduction, where the damage at the mid-span is called DC1b_E and at the ¼ span is called DC2b_E. In order to eliminate the effect of mass reduction when area is cut off the bridge, for all stiffness reduction damage cases, steel flat bars with the same size as cut area (see Figure 15 and Figure 16) are loosely attached to the bridge using hot glue. DC3 is applied by mounting each end to a short I-beam through four bolts.

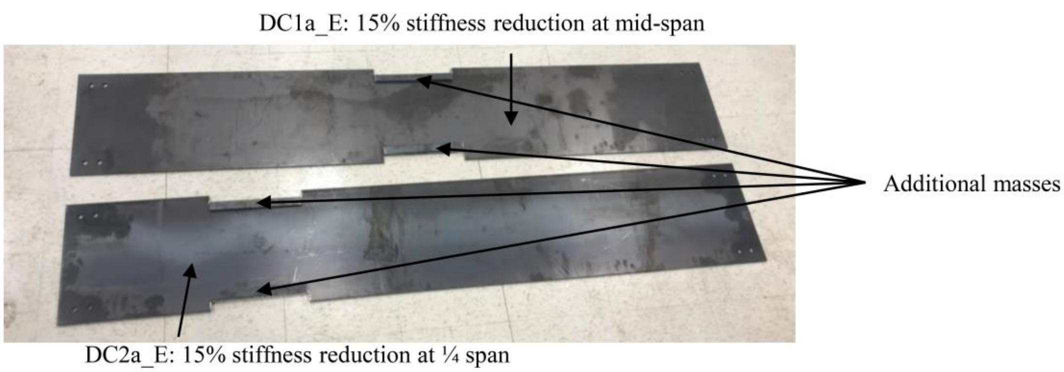


Figure 15 - Artificial damage applied for DC1a_E and DC2a_E

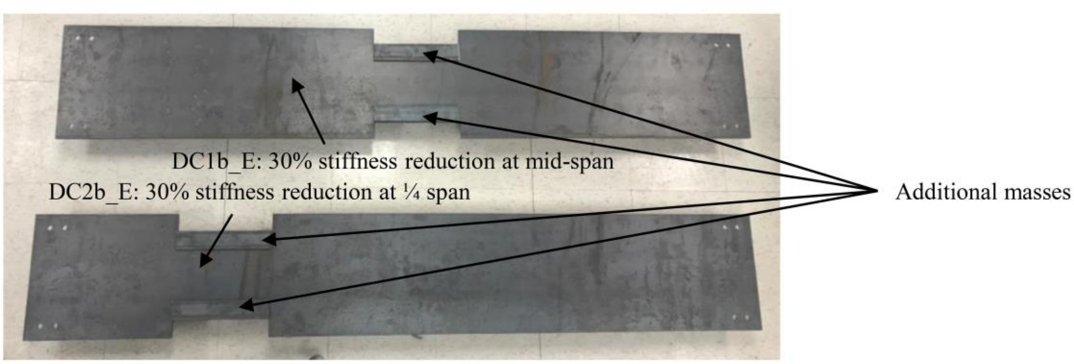


Figure 16 - Artificial damage applied for DC1b_E and DC2b_E

Figure 17 - Boundary condition change for DC3_E

A model vehicle is developed to mimic the behavior of real vehicles. This model vehicle can be used to simulate different speed, suspension, weight etc. As illustrated in Figure 18, the vehicle is mainly composed of two aluminum plates. The motors and wheels are connected to the bottom plate and controlled by an Arduino programmable board. The whole vehicle is powered by 5 AA batteries. The top plate is connected to the bottom plate through rods, linear bearings and springs to model the suspension system. Two G-Link®-200 wireless accelerometers are mounted at sides of the top plate. Galaxy S5 smartphone is mounted at the center of the top plate. The wireless accelerometers have a sampling frequency of 128 Hz, while smartphone has a sampling frequency of 100 Hz. The average is taken when data are collected from two wireless accelerometers.

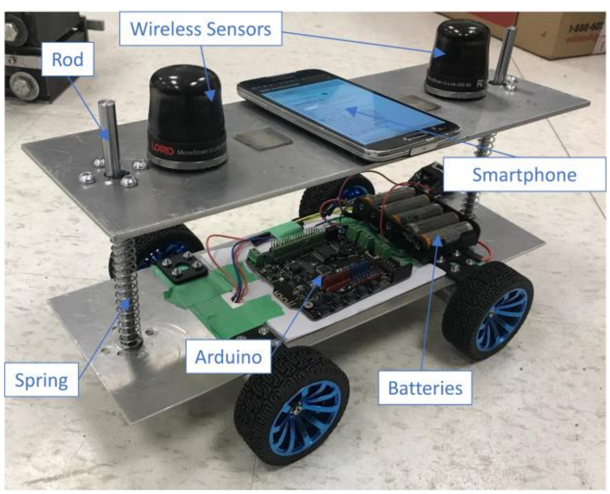


Figure 18 - Setup of the model vehicle

As presented in Figure 19, to verify the capacity of the smartphone to collect acceleration data, the same impact test as described in section 4.1.2 is conducted on the intact simply supported bridge. The smartphone as well as wireless accelerometer are mounted at the mid-span of the bridge for data collection. The acceleration data measured from both devices are presented in Figure 20(a), and Fourier transform are conducted and shown in Figure 20(b). It is seen that smartphone can measure the acceleration very accurately and can identify the modal frequencies of the bridge in the given range successfully, which are 3.71 Hz, 14.9 Hz, 33.4 Hz.

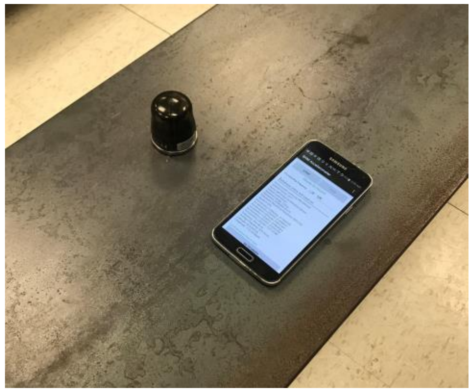
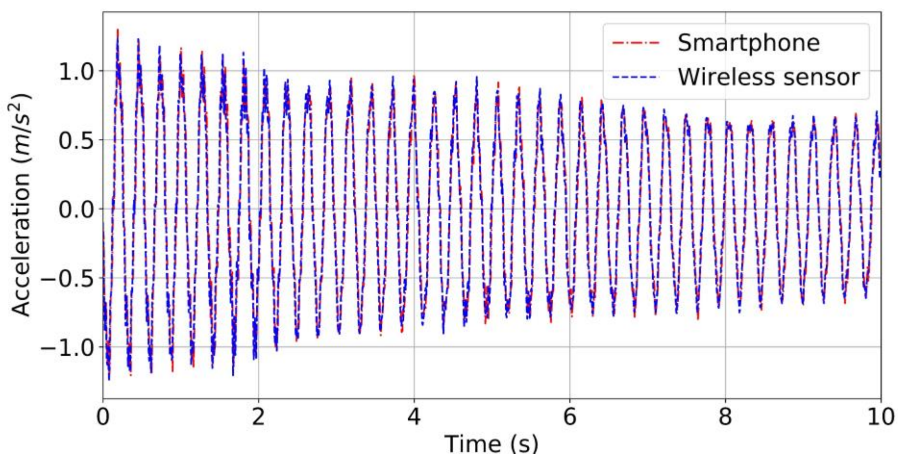
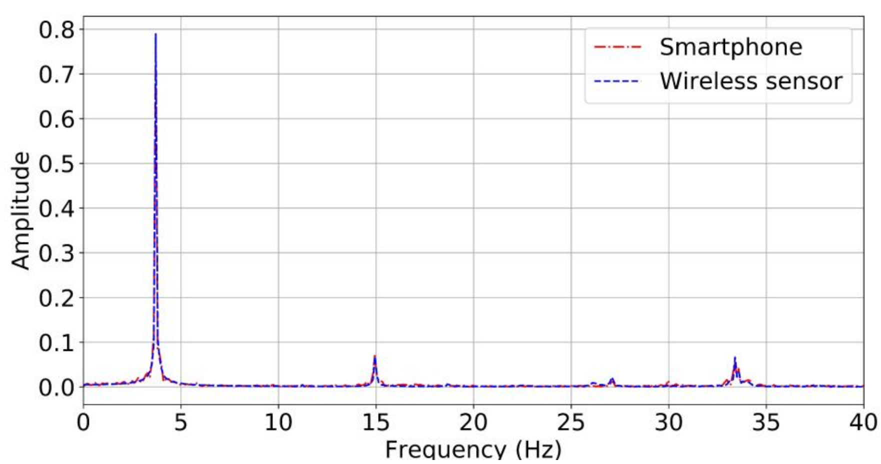


Figure 19 - Placement of smartphone and wireless accelerometer for impact test



(a) time domain



(b) frequency domain

Figure 20 - Comparison of smartphone and wireless accelerometer

In the model vehicle, the springs are replaceable, and the spring constant is changed among 5 values: 155 N/m, 288 N/m, 425 N/m, 615 N/m, and 726 N/m. The weight of the model vehicle can be changed by placing additional masses on the top plate, and it is changed among five levels: 0.898 kg, 0.988 kg, 1.084 kg, 1.170 kg, 1.270 kg. The speed can be changed by programming the Arduino board, and it is varied among 3 values: 0.25 m/s, 0.33 m/s, 0.40 m/s. In addition, each test is repeated 3 times for each model vehicle configuration.

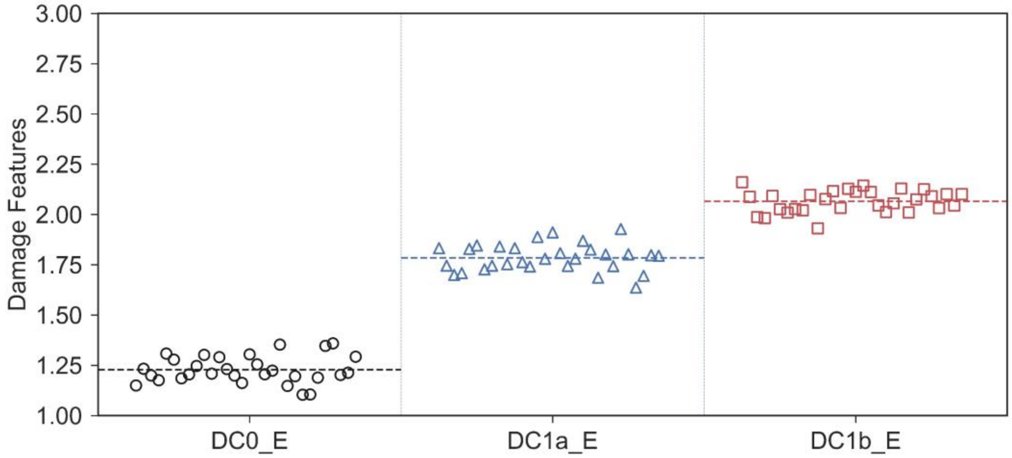
Combining all the parameter changes and repeated tests, there are in total $5 \times 5 \times 3 \times 3 = 225$ tests for each bridge state. To avoid using exactly the same set of vehicles for different damage cases, 50% data entries of the vehicle pool are randomly sampled. Similar to the numerical analysis, 50% of the tests on intact bridge are reserved as baseline case and the remaining 50% are for validation. For other damage cases, only 50% of the tests are selected for damage detection. In this way, the sets of vehicles would be different for different damage cases, which is close to real situation because it is impossible to have exactly the same set of vehicles passing across bridge during different periods. However, since they are sampled from the same set of vehicle configurations, they should follow similar distributions. The sampling process has randomness. Therefore, in order to verify the robustness, the method is implemented 30 times with different samples included.

4.3. Experimental results

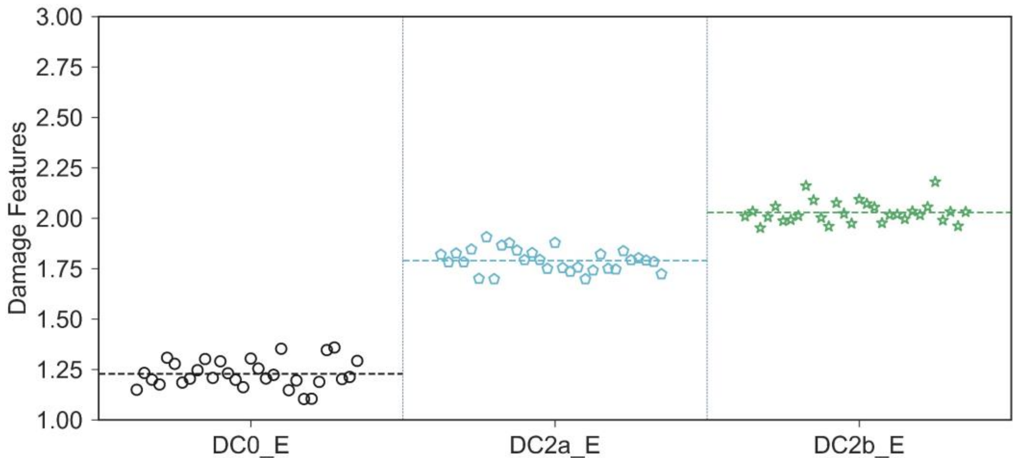
The results obtained using the wireless sensors and the smartphone are presented in Figure 21 and Figure 22, respectively. Each point in the figures represents the DF for a complete implementation of the algorithm from one sampling process, and therefore there are in total 30 points for each damage cases.

Similar to numerical analysis, the horizontal dashed lines stand for the average DF over all 30 trials. In Figure 21, it is shown that DFs for all damage cases are higher than DC0_E, which indicates that the damage is successfully identified. Among these damaged cases, DFs for DC1a_E and DC1b are about the same level as DC2a_E and DC2b_E. DFs for DC1b_E and DC2b_E are higher than DC1a_E and DC2a_E, because 30% stiffness reduction is more severe than 15% reduction. DC3_E has the highest DFs, which makes sense because boundary condition change is the most severe damage. It should be noted that the difference of DFs among different damage cases is not as significant as in numerical analysis. The reason is that the vibration of the vehicle and bridge in experiments is not as ideal as in numerical models. The existence of noise in frequency spectrums makes the extracted features harder to distinguish. In spite of that, the variation of DFs in a single damage case is still within an acceptable range so that different damage cases are separable. The observations above show that the approach is robust when different sets of vehicles following the same distribution are used for damage detection.

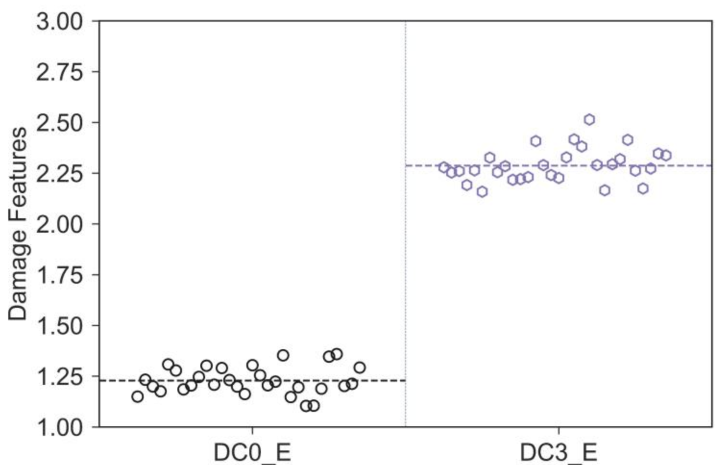
Comparing DFs obtained using wireless accelerometers and the smartphone, it is seen that the DFs from both devices are very similar, but the fluctuation of DFs from smartphone data is generally larger than from sensors. The reason for this is that the location of wireless accelerometers is different than the smartphone. It can be seen in Figure 18 that smartphone is placed at the center of the top plate, and therefore the vibration of the top plate is likely to be included in smartphone data.



(a) DC1a_E and DC1b_E

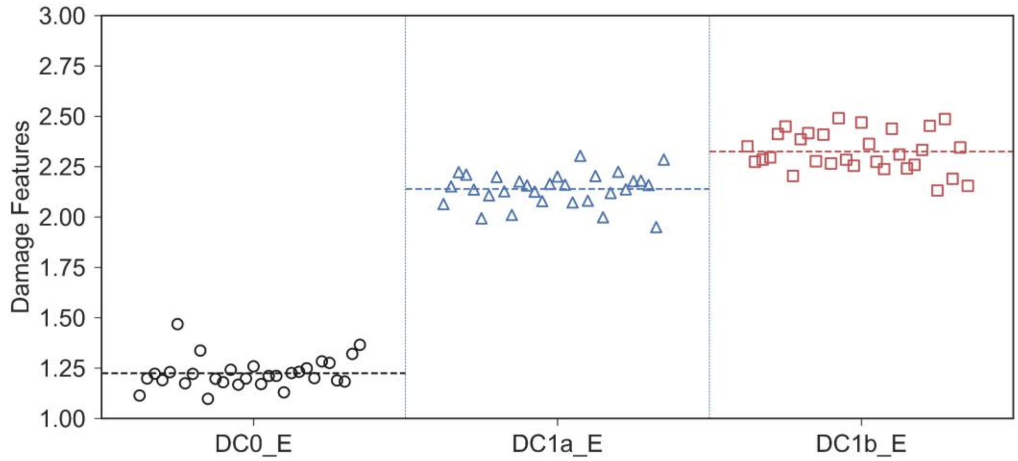


(b) DC2a_E and DC2b_E

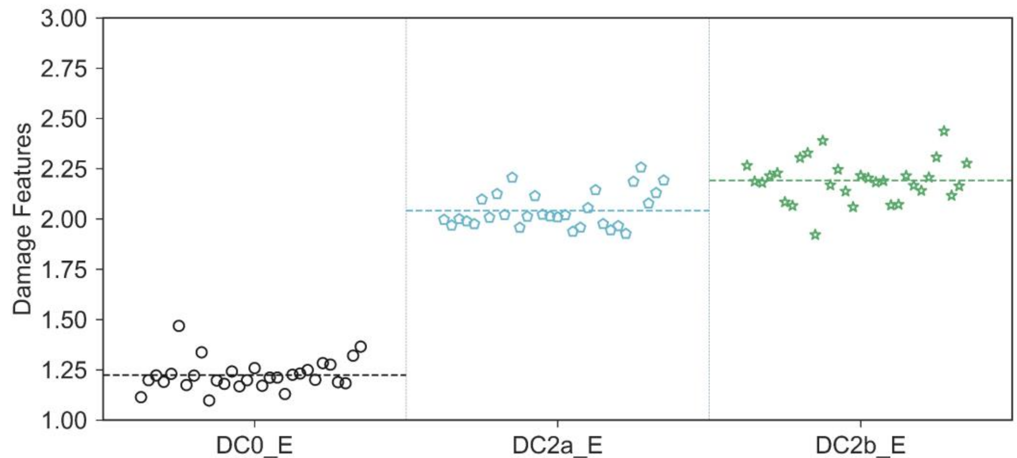


(c) DC3_E

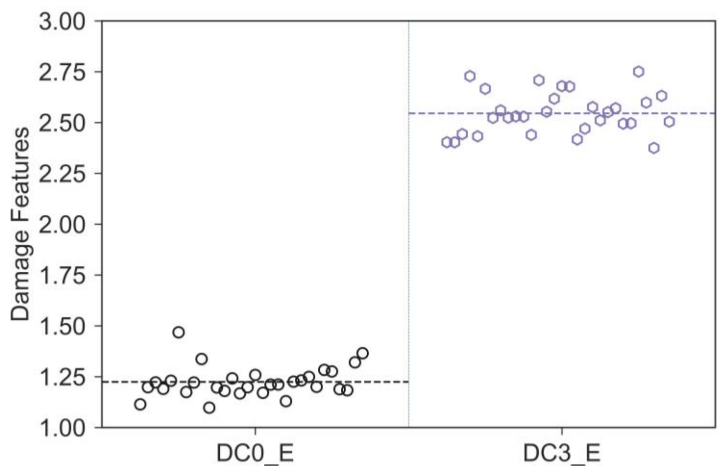
Figure 21 - DFs obtained using the wireless accelerometers



(a) DC1a_E and DC1b_E



(b) DC2a_E and DC2b_E



(c) DC3_E

Figure 22 - DFs obtained using the smartphone

5. Discussion

To the best of the authors' knowledge, the proposed framework is the first of its kind in the world and it is anticipated that the expected outcomes of the proposed research program will be extremely useful for developing smart transportation infrastructure systems using connected vehicles and other mobile sensing technologies. Although the proposed framework and methodology of utilizing crowdsourcing data from smartphones in vehicles has shown promising results, there are still challenges that need to be resolved before routine real-life applications.

First, the increased uncertainties on the collected data introduced by the random placement of the smartphone in the vehicle is an important factor. The influence of location and direction of the smartphone should be investigated and eliminated. One possible solution is to use the built-in gyroscope and magnetometer data to aid the accelerometer and convert the local acceleration data to global vertical acceleration, which is already implemented in our Android application. Also, active filtering techniques could be applied to identify the validity of the data, and unstable data could be discarded automatically.

Besides, this proposed framework requires the smartphone to measure the acceleration and transmit data to remote database. To minimize the battery and data usage in real-life applications, the method presented in this paper only requires the transmission of a small number of features extracted from the data. Also, this issue can be overcome by using GPS data. For example, the data recording process will not be activated until the GPS tells that the vehicle is about to pass a bridge. The data accessibility is another issue that should be resolved for future real life implementations. Currently, we rely on volunteers to install the app and ask their permission to collect and transmit data. In the future, the method can be integrated into commercial apps such as Google map to increase the volume of data or apps that are developed by municipalities and other local authorities for the use of the citizens.

In addition, the challenges from traditional indirect health monitoring problem such as road profile roughness, limited vehicle-bridge interaction time and environmental effects, will affect the framework as well and should be studied. In practice, to reduce the influence of environmental and operational, the framework can also be used with other monitoring data such as weather information or temperature. Also, it should be noted that the framework neither requires the collection window to be continuous nor restrict the length of it as long as a large number of vehicles passing through the bridge within that window. In practice, the collection window can be chosen as large as possible to average out the environmental and operational effects.

6. Conclusions

This paper proposes a novel framework for damage detection of bridges using crowdsourcing data collected from a large number of vehicles passing through a bridge at different times. The method developed within this framework utilizes Mel-frequency cepstral coefficients and KL divergence under this framework to detect damage on the bridge. Both numerical analysis and experiments are conducted to verify this method. The crowdsourced vehicles are simulated by changing the configurations of a single vehicle (i.e. weight, suspension spring and the speed of the vehicle). In the experiments, data collection using both wireless sensors and smartphones are investigated. The following main conclusions are obtained from our studies:

1. The proposed DFs calculated based on crowdsourcing data from smartphones in vehicles successfully identify the existence of damage and extract useful information about severity.
2. One of the main advantage of the proposed method is that the requirement of knowing vehicle configuration or using the same vehicle for damage detection is eliminated.
3. It is shown that smartphones can potentially be used to indirectly detect damage in bridges.

In this paper, the smartphones are considered as sensors for data collection when dealing with a large number of vehicles. It has to be acknowledged that smartphones have several defects compared to wireless sensors, such as lower sampling frequency and lower resolution, but they are widely used and have access to internet. This paper has demonstrated that the smartphone can give just as good results as wireless sensors for indirect bridge health monitoring. With the help of crowdsourcing with smartphones, this method has the potential to monitor a population of bridges simultaneously and in almost real time.

In the future, the localization of damage using a large number of vehicles will be investigated. Factors such as road surface roughness and operational effect will be considered as well. It should be noted that the method can be extended to other smart devices including smart vehicles considering the smart and connected vehicles have several embedded sensors which can be exploited for assessing different infrastructure systems including bridges.

Declaration of conflicting interests

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclose receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the NSERC Discovery Grants (no. NSERC RGPIN 436051).

References

1. Mei Q and Gül M. A fixed-order time series model for damage detection and localization. *Journal of Civil Structural Health Monitoring*. 2016; 6: 763-77.
2. Gul M and Catbas FN. Damage Assessment with Ambient Vibration Data Using a Novel Time Series Analysis Methodology. *Journal of Structural Engineering*. 2011; 137: 1518-26.
3. Zhu D, Guo J, Cho C, Wang Y and Lee K-M. Wireless mobile sensor network for the system identification of a space frame bridge. *Mechatronics, IEEE/ASME Transactions on*. 2012; 17: 499-507.
4. Soh C, Tseng KK, Bhalla S and Gupta A. Performance of smart piezoceramic patches in health monitoring of a RC bridge. *Smart materials and Structures*. 2000; 9: 533.
5. Catbas FN, Gul M and Burkett JL. Conceptual damage-sensitive features for structural health monitoring: laboratory and field demonstrations. *Mechanical Systems and Signal Processing*. 2008; 22: 1650-69.
6. Bao Y, Shi Z, Wang X and Li H. Compressive sensing of wireless sensors based on group sparse optimization for structural health monitoring. *Structural Health Monitoring*. 2017: 1475921717721457.

7. Loh C-H, Hsueh W, Tu Y-C, Lin J-H and Kuo T-J. Vibration-based damage assessment of structures using signal decomposition and two-dimensional visualization techniques. *Structural Health Monitoring*. 2018: 1475921718765915.

8. Spencer BF, Ruiz-Sandoval ME and Kurata N. Smart sensing technology: opportunities and challenges. *Structural Control and Health Monitoring*. 2004; 11: 349-68.

9. Harris DK, Alipour M, Acton ST, Messeri LR, Vaccari A and Barnes LE. The Citizen Engineer: Urban Infrastructure Monitoring via Crowd-Sourced Data Analytics. *Structures Congress 2017*. 2017, p. 495-510.

10. Matarazzo T, Vazifeh M, Pakzad S, Santi P and Ratti C. Smartphone data streams for bridge health monitoring. *Procedia Engineering*. 2017; 199: 966-71.

11. Ozer E, Feng MQ and Feng D. Citizen sensors for SHM: Towards a crowdsourcing platform. *Sensors*. 2015; 15: 14591-614.

12. Chugh G, Bansal D and Sofat S. Road condition detection using smartphone sensors: A survey. *International Journal of Electronic and Electrical Engineering*. 2014; 7: 595-602.

13. Ramana L, Choi W and Cha Y-J. Fully automated vision-based loosened bolt detection using the Viola–Jones algorithm. *Structural Health Monitoring*. 2018: 1475921718757459.

14. Yang YB, Lin CW and Yau JD. Extracting bridge frequencies from the dynamic response of a passing vehicle. *Journal of Sound and Vibration*. 2004; 272: 471-93.

15. Kim J and Lynch JP. Experimental analysis of vehicle–bridge interaction using a wireless monitoring system and a two-stage system identification technique. *Mechanical Systems and Signal Processing*. 2012; 28: 3-19.

16. Zhang Y, Lie ST and Xiang Z. Damage detection method based on operating deflection shape curvature extracted from dynamic response of a passing vehicle. *Mechanical Systems and Signal Processing*. 2013; 35: 238-54.

17. Li Z. Damage identification of bridges from signals measured with a moving vehicle. *HKU Theses Online (HKUTO)*. The University of Hong Kong (Pokfulam, Hong Kong), 2014.

18. Keenahan J, OBrien EJ, McGetrick PJ and Gonzalez A. The use of a dynamic truck–trailer drive-by system to monitor bridge damping. *Structural Health Monitoring*. 2014; 13: 143-57.

19. Cerda F, Chen S, Bielak J, Garrett JH, Rizzo P and Kovacevic J. Indirect structural health monitoring of a simplified laboratory-scale bridge model. *Smart Structures and Systems*. 2014; 13: 849-68.

20. He X, Kawatani M, Hayashikawa T, Kim C-W, Catbas FN and Furuta H. A structural damage detection approach using train-bridge interaction analysis and soft computing methods. *Smart Structures and Systems*. 2014; 13: 869-90.

21. OBrien EJ, Malekjafarian A and González A. Application of empirical mode decomposition to drive-by bridge damage detection. *European Journal of Mechanics-A/Solids*. 2017; 61: 151-63.

22. Kim C-W, Chang K-C, McGetrick PJ, Inoue S and Hasegawa S. Utilizing moving vehicles as sensors for bridge condition screening—a laboratory verification. *Sensors and Materials*. 2017; 29: 153-63.

23. Hester D and González A. A discussion on the merits and limitations of using drive-by monitoring to detect localised damage in a bridge. *Mechanical Systems and Signal Processing*. 2017; 90: 234-53.

24. McGetrick PJ, Kim C-W, González A and Brien EJ. Experimental validation of a drive-by stiffness identification method for bridge monitoring. *Structural Health Monitoring*. 2015; 14: 317-31.

25. Pakrashi V, O'Connor A and Basu B. A bridge-vehicle interaction based experimental investigation of damage evolution. *Structural Health Monitoring*. 2010; 9: 285-96.

26. Malekjafarian A, McGetrick PJ and OBrien EJ. A review of indirect bridge monitoring using passing vehicles. *Shock and Vibration*. 2015; 2015.

27. Yang Y-B and Yau J-D. Vehicle-Bridge Interaction Element for Dynamic Analysis. *Journal of Structural Engineering*. 1997; 123: 1512-8.

28. Yang YB and Lin CW. Vehicle–bridge interaction dynamics and potential applications. *Journal of Sound and Vibration*. 2005; 284: 205-26.

29. McGetrick PJ, Gonzlez A and OBrien EJ. Theoretical investigation of the use of a moving vehicle to identify bridge dynamic parameters. *Insight-Non-Destructive Testing and Condition Monitoring*. 2009; 51: 433-8.

30. González A, Obrien EJ and McGetrick PJ. Identification of damping in a bridge using a moving instrumented vehicle. *Journal of Sound and Vibration*. 2012; 331: 4115-31.

31. Malekjafarian A and OBrien EJ. On the use of a passing vehicle for the estimation of bridge mode shapes. *Journal of Sound and Vibration*. 2017; 397: 77-91.

32. Zhang Y, Wang L and Xiang Z. Damage detection by mode shape squares extracted from a passing vehicle. *Journal of Sound and Vibration*. 2012; 331: 291-307.

33. Yang Y, Li Y and Chang K. Constructing the mode shapes of a bridge from a passing vehicle: a theoretical study. *Smart Structures and Systems*. 2014; 13: 797-819.

34. Muda L, Begam M and Elamvazuthi I. Voice recognition algorithms using mel frequency cepstral coefficient (MFCC) and dynamic time warping (DTW) techniques. *arXiv preprint arXiv:10034083*. 2010.

35. Murty KSR and Yegnanarayana B. Combining evidence from residual phase and MFCC features for speaker recognition. *IEEE signal processing letters*. 2006; 13: 52-5.

36. Zheng F, Zhang G and Song Z. Comparison of different implementations of MFCC. *Journal of Computer Science and Technology*. 2001; 16: 582-9.

37. Dackermann U, Smith WA and Randall RB. Damage identification based on response-only measurements using cepstrum analysis and artificial neural networks. *Structural Health Monitoring*. 2014; 13: 430-44.

38. Zhang G, Harichandran RS and Ramuhalli P. Application of noise cancelling and damage detection algorithms in NDE of concrete bridge decks using impact signals. *Journal of Nondestructive Evaluation*. 2011; 30: 259-72.

39. Gelfand SA. *Essentials of audiology*. Thieme, 2009.

40. Balsamo L, Betti R and Beigi H. A structural health monitoring strategy using cepstral features. *Journal of Sound and Vibration*. 2014; 333: 4526-42.

41. Kullback S and Leibler RA. On information and sufficiency. *The annals of mathematical statistics*. 1951; 22: 79-86.

42. Berry DS and Belmont DM. Distribution of vehicle speeds and travel times. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. The Regents of the University of California, 1951.

43. Dhamaniya A and Chandra S. Speed characteristics of mixed traffic flow on urban arterials. *International Journal of Civil, Architectural Science and Engineering*. 2013; 7: 330-5.

44. Duchi J. Derivations for linear algebra and optimization. *Berkeley, California*. 2007.

45. Jin Y, Toh H-S, Soh W-S and Wong W-C. A robust dead-reckoning pedestrian tracking system with low cost sensors. *Pervasive Computing and Communications (PerCom), 2011 IEEE International Conference on*. IEEE, 2011, p. 222-30.

46. Ozer E, Feng D and Feng MQ. Hybrid motion sensing and experimental modal analysis using collocated smartphone camera and accelerometers. *Measurement Science and Technology*. 2017; 28: 105903.

List of Tables

List of Figures