Engineering Structures 224 (2020) 111224

Contents lists available at ScienceDirect

Engineering Structures

journal homepage: www.elsevier.com/locate/engstruct

Investigation of vibration serviceability of a footbridge using computer vision-based methods

Chuan-Zhi Dong[a], Selcuk Bas[a,b], F. Necati Catbas[a,*]

[a] Department of Civil, Environmental, and Construction Engineering, University of Central Florida, Orlando, FL 32816, USA
[b] Department of Civil Engineering, Bartin University, Bartin, Turkey

ARTICLE INFO

Keywords:
Computer vision
Footbridge
Vibration serviceability
Human comfort
Structural identification

ABSTRACT

In current standards and codes for vibration serviceability assessment, the serviceability limits are generally defined by acceleration related criteria or indices such as root-mean-square (RMS) or peak value of acceleration time histories. The accelerations are collected by using conventional accelerometers, which have several drawbacks such as traffic closure, setup time, and labor force to deal with the cable wiring work. It may not be convenient or practical to conduct such monitoring, especially for certain field applications. This study proposes to assess the vibration serviceability of a footbridge using computer vision-based methods. The proposed vision-based approach shows great advantages such as non-contact, long distance, low cost, time saving, and ease of use. The proposed approach is validated by a series of experiments on a footbridge under various types of pedestrian loading including walking, jumping and running with different paces (beat per minute). Displacement and velocity are first estimated from the image sequence and then converted to acceleration for the assessment of vibration serviceability. The feasibility of the proposed approach is verified by the comparative analysis between the results of serviceability assessments using the proposed approach and the conventional accelerometers. Consideration and recommendations of the process of converting displacement/velocity to acceleration are also discussed. The proposed approach provides a promising and efficient alternative for the vibration serviceability assessment of footbridges combined with the current standards and codes. As a result, it provides another approach for the serviceability assessment using different data type such as displacement and velocity.

1. Introduction

With the development of high performance structural materials and aesthetic requirements for structures, longer and slender footbridges have attracted great public attention and a large amount of modern footbridges with lightweight and lively structures have been constructed in the last several decades [1]. This trend makes a large achievement of infrastructures in the progress of smart cities. However, it also causes critical issue: excessive vibrations of slender footbridges caused by pedestrian live load. These excessive vibrations may cause another problem, i.e., human comfort, since the main function of footbridges are to convey pedestrians. In this respect, it means that the estimated dynamic response of the footbridges has to be evaluated against human comfort level [2]. In general vibration produced by human-induced loads is a structural vibration serviceability problem rather than a structural safety (structural damage) problem. A famous example related to vibration serviceability problems of the footbridges is the Millennium Bridge over Thames River in England. In 2000, at the opening of the newly built footbridge, excessive vibrations which were descripted as “swaying violently” was reported when a group of pedestrians crossed the footbridge [3]. In a local news article, BBC News descripted this bridge as “Wobbly” Millennium Bridge [4]. It took £5 million and about eight months to solve the problem, while the original cost of the “Wobbly” bridge was £18.2 million [4]. Examples about footbridge collapse incidents due to soldiers marching movement in unison over the bridge such as Broughton Footbridge in England and Angers Bridge in France can date back to nineteenth century [5].

The vibration serviceability has become a hot topic of research and practice in the community of structural engineering and large amount of work has been done to assess the footbridge vibration level and to mitigate the excessive levels of vibrations in slender footbridges. Živanović et al. (2005) [1] summarized the main research focus of footbridge vibration serviceability which generally includes: (1) vibration source related work such as human force estimation and modelling, (2) vibration path related work such as stiffness, mass and damping, (3) human perception related work such as estimation of

* Corresponding author.
E-mail address: catbas@ucf.edu (F.N. Catbas).

https://doi.org/10.1016/j.engstruct.2020.111224
Received 3 January 2020; Received in revised form 29 June 2020; Accepted 13 August 2020
Available online 29 August 2020

human comfort level, (4) human-structure interaction, (5) design guidelines and (6) vibration reduction measures. Among these topics, estimation of human comfort level is one of the most direct and intuitive ways to assess the vibration serviceability considering that the main function of footbridge is to convey pedestrian. Organizations and agencies established standards and codes such as ISO 10137 by International Organization for Standardization [6], Euro code 5 by European Committee for Standardization (ECS) [7], BS 5400 by British Standards Institution [8] and Setra code by French Technical Department for Transport, Roads And Bridges Engineering and Road Safety [9]. In most of the standards and codes, the acceleration related indices combined with vibration frequencies are used to define the serviceability limit. For example, ISO 10137 uses the maximum 1-second running RMS value of the frequency-weighted acceleration time histories. Euro code 5, BS5400 and Setra selected the peak value of acceleration at the fundamental frequencies of structures. In addition to the standards and codes established by the organizations and agencies above, researchers also proposed different serviceability assessment criteria. Mackenzie et al. (2005) [10] proposed the serviceability assessment by defining the acceleration limits as a function of footbridge height, parapet height and route redundancy etc. Kasperski (2006) [11] conducted a series of experiments on a footbridge, and based on the ratio of pedestrians alarmed by vibration, he presented that the RMS limit recommended by ISO 10137 is excessive and should be reduced to 60% for footbridge serviceability. Barker (2007) [12] recommended using the root-mean-quad (RMQ) of acceleration to assess the vibration serviceability instead. Živanović and Pavia (2009) [2] proposed a probabilistic approach for the assessment of vibration serviceability based on the acceleration measurements. Setareh (2016) [5] investigated the relationships between various evaluation parameters such peak value of acceleration, peak value of weighted acceleration, RMS and vibration dose value (VDV) and found that based on the relationships, VDV can also be a good index for the definition of serviceability limit. Dey et al. (2017, 2018) [13,14] conducted a series of experiment on a large scale aluminum pedestrian bridge in laboratory and evaluated and calibrated various guidelines for the serviceability based on the experimental data. Feng et al. (2019) [15] analyzed the correlation between the peak acceleration of footbridge vibration and the pedestrian comfort level collected by doing pedestrian questionnaire and based on the correlation they proposed a procedure of using acceleration data to assess footbridge vibration serviceability.

The current research and practice for vibration serviceability is based on the estimation of vibration level collected from conventional sensors such as the accelerometers [16]. The drawbacks of using conventional sensors are traffic closure, setup time, cost and labor force to deal with the cable wiring work. It is not convenient to conduct such experiments, especially for field application. Placing the accelerometers in the wearable equipment of a human subject on the structure might be a smart way to quickly measure the vibrations, but it is not easy to separate the vibration components of the structures and the human subject. With the development of imaging devices and computer vision technology, computer vision-based applications such as structural vibration/displacement/deformation monitoring [17–19], human load estimation [20], vehicle load estimation [21], local structural damage detection [22] and three-dimensional reconstruction of structures [23] is gathering increasing attention in the field of structural health monitoring (SHM). Due to the advantages such as non-contact, long distance, low cost, time saving, and ease of use, computer vision-based structural vibration monitoring methods has become the viable alternatives to the conventional sensors in current practice of structural monitoring [24–26]. By using cameras to record the structures and tracking the motion of targets on them, the displacement of structures can be easily obtained. To fit the standards and codes of vibration serviceability assessment, acceleration data can be calculated from displacement data with numerical differentiation.

This study proposes a non-contact approach for footbridge serviceability assessment by using computer vision-based methods and combining with current standards and codes. With a non-contact, cost-effective and time-saving way, the vision-based approach can overcome the drawbacks of using conventional sensors for the experimental assessment of vibration serviceability of footbridges. Živanović et al. (2005) [1] in their review paper summarized that in the early stage of research work on vibration serviceability, using displacement and velocity to assess human perception is a more direct and intuitive way and was recommended by lots of researchers. By using vision-based methods, the displacement and velocity are easier to obtain compared with conventional displacement sensor and velocimeter. With the displacement data from vision-based methods, it may supply a possible alternative to define the vibration serviceability criteria based on displacement and velocity.

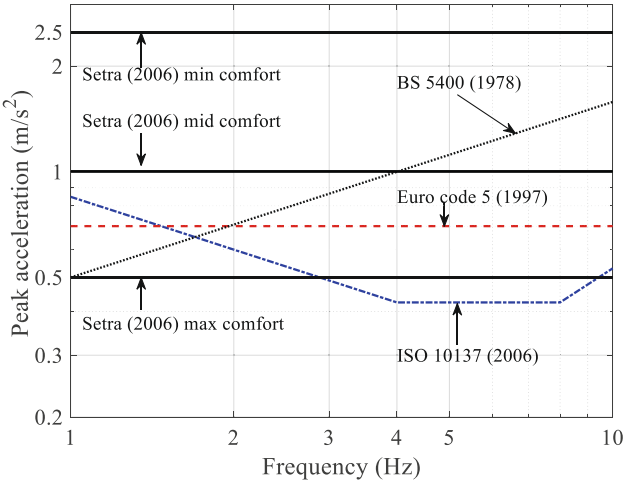

Fig. 1. Acceleration limits for footbridge vibration serviceability in different standards and codes.

2. Vibration limit for serviceability assessment in current standards and codes

In this section, the current standards and codes employed by different countries and areas for vibration serviceability assessment of footbridges are discussed. Fig. 1 illustrates the acceleration limits for footbridge vibration serviceability in different standards and codes. The footbridge vibration serviceability assessment is based on the human comfort level of structures. Except for ISO 10137, the others employed the peak value of accelerations as the limits for serviceability. ISO 10137 use RMS value of the frequency weighted accelerations. It provides a base curve using RMS value for serviceability limit and recommends that for footbridges the serviceability limit is calculated by multiplying the base curve with the factor 60 for active pedestrians. To give a comparative view with other standards and codes, in this study the RMS value provided by ISO 10137 is converted to the equivalent acceleration peak value by multiplying by the factor $\sqrt{2}$. The peak acceleration limit of the standards and codes mentioned above are shown in Fig. 1. Detailed formulas to determine the curves in Fig. 1 are listed in Table 1. For ISO 10137, BS 5400 and Euro code 5, a unique curve of serviceability limit of peak acceleration against fundamental frequencies of the structures is provided. The region under the curve is acceptable for serviceability level and the region above the curve is unacceptable. For Setra standard as shown in Fig. 1, it gives three curves to indicate the vibration serviceability level such as min comfort level, mid (medium) comfort level and max (extreme) comfort level.

Table 1
Detailed acceleration limits for footbridge vibration serviceability in different standards and codes.

Code	Vertical peak acceleration (m/s^2)
ISO 10137 (2007)	$\begin{cases} a_{limit} = 0.1414f + 0.9899, & 1 < f \leqslant 4 \\ a_{limit} = 0.4243, & 4 < f \leqslant 8 \\ a_{limit} = 0.053f, & 8 < f \leqslant 80 \end{cases}$
Euro code 5 (European committee for Standardization, 1997)	$a_{limit} = 0.7$
BS5400 (British Standards Institution, 1978)	$a_{limit} = 0.5f^{0.5}$
Setra (French Technical Department for Transport, Roads and Bridges Engineering and Road Safety, 2006)	$\begin{cases} a_{limit} < 0.5, & \text{Extreme} \\ 0.5 < a_{limit} < 0.1, & \text{Medium} \\ 1.0 < a_{limit} < 2.5, & \text{Low} \end{cases}$

3. Vision-based displacement/velocity measurement using feature matching

In general, there are four steps to extract structural displacement/velocity from video or image sequence. In this study, a vision-based displacement/velocity measurement method using feature matching is employed. Fig. 2 shows the procedure of the proposed method.

Firstly, the camera is calibrated to estimate the relationship between the image coordinates and the real-world coordinates. Here the scale ratio is used which calculate the ratio between the actual dimension in physical unit (e.g. millimeter) and the image dimension in pixel [17]. For example, if the actual height of an object in real world is D mm, and the height of the object in image is d pixel, the scale ratio, SR, will be

$$SR = \frac{D}{d} \tag{1}$$

The scale ratio expressed in Eq. (1) is only suitable for the case when the axis of the camera and lens is perpendicular to the motion plane of the measurement target. For the cases that there is an inclination between them, Ref. [17] gives a detailed discussion. The reader is addressed to that reference for more information.

Secondly, the camera records the video or image sequence of the structural motion. Feature points (also called key points, or kps) are extracted from the region of interest (ROI) of each image. A ROI is generally a sub region of an image that represent the measurement target of a structure. As shown in Fig. 3, ROI 1 and ROI 2 are parts of a beam of a bridge. The feature point means a small image patch of the ROI with distinction such as corner, texture and gradient. On the right-top of Fig. 3, extracted feature points are marked with circles in different color. In general, two components are required to define a feature point: feature detector and descriptor. Detector is to locate the region of the feature in an image and descriptor is a vector to descript the feature in mathematical language. In this study, SIFT (Scale-invariant feature transform) detector is applied to locate the feature point and VGG (Visual Geometry Group) descriptor is employed to descript the SIFT feature [27]. It is noted in Ref. [27], that using SIFT detector and VGG descriptor performs better than using the original SIFT feature method (SIFT detector and SIFT descriptor). The ROI selected in Fig. 3 is with the size of 209 × 210 pixels and 120 feature points (kps) are extracted.

Thirdly, the feature points extracted in two images are matched based on the similarity of the feature points. The similarity of two feature points can be calculated by the distance of the descriptors of them. The feature match pairs are selected by the one with the best similarity, i.e., with smallest distance. After the initial feature matching, there might be some wrong matches as shown in Fig. 3. The RANSAC (RANdom SAmple Consensus) method is implemented to remove the outliers [27]. On the right-bottom of Fig. 3, it shows that three wrong matches are removed from the initial match. Depending on whether displacement or velocity is required in the end, the feature matching is performed between different image pairs. Fig. 4 shows two different feature matching strategies. The top of the Fig. 4 illustrates the feature matching between the ROI of the first frame in the image sequence and the ROI of Frame i (i = 2, 3, 4, …). In other words, the first frame is not updated during the feature matching. The location change between the two ROIs is the relative displacement at the time of Frame i to the initial frame (Frame 1) and this is exactly the concept of displacement. While on the bottom of Fig. 4, the feature matching is performed between the consecutive frames such as Frame 1 and Frame 2, Frame 2 and Frame 3, and Frame i-1 and Frame i, etc. The frame is always updated during the feature matching. The location change between the two ROIs of consecutive frames is the incremental of the displacement at the time of the Frame i-1 to Frame i and this refers to the concept of velocity. Using update frame strategy might be better to get more good matches when the structure moves over time and the light condition or surface feature changes along with the motion. The number of matched feature pairs shows the performance of the strategy. A higher number means the high matching quality. Fig. 5 gives an example to show the comparison of the performance of the feature matching using two different strategies. In this case, 120 feature points are extracted in the ROI of each image, and the number of matched pairs using frame update strategy within an image sequence is more than the one without using frame update. It means that feature matching using frame update give better performance. Dong et al. (2019) [28] summarized the pros and cons of using the two different strategies. In this study, whether the former or the latter is used depends on the requirement of whether displacement or velocity is necessary. In the experimental section, more details will be discussed according to the results.

At last, the displacement/velocity can be calculated by taking the average of the location change of the matched feature points in two images. The displacements in x and y direction, X and Y, in physical unit can be calculated by

$$\begin{cases} X_i = SR_x \dfrac{\sum_{j=1}^{n} (x_i^j - x_1^j)}{n} \\ Y_i = SR_y \dfrac{\sum_{j=1}^{n} (y_i^j - y_1^j)}{n} \end{cases} \tag{2}$$

where (x_i^j, y_i^j) and (x_1^j, y_1^j) are the image coordinates of the jth matched feature point of between the ROIs of Frame i and Frame 1, n is the total number of the matched feature point between the ROIs of Frame i and Frame 1, and SR_x and SR_y are the scale ratio in x and y direction. The velocity in x and y direction, V_X and V_Y, in physical unit can be calculated by

$$\begin{cases} V_X = SR_x \dfrac{\sum_{j=1}^{m} (x_i^j - x_{i-1}^j)}{n \cdot \Delta t} \\ V_Y = SR_y \dfrac{\sum_{j=1}^{m} (y_i^j - y_{i-1}^j)}{n \cdot \Delta t} \end{cases} \tag{3}$$

where x_i^j, y_i^j and x_{i-1}^j, y_{i-1}^j are the image coordinates of the jth matched feature point of between the ROIs of Frame i and Frame $i-1$, and Δt is time interval of the image sampling which is the reciprocal of the sampling rate.

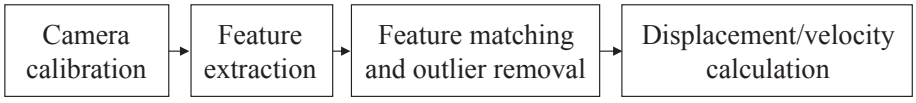

Fig. 2. Procedure for proposed vision-based displacement/velocity measurement method using feature matching.

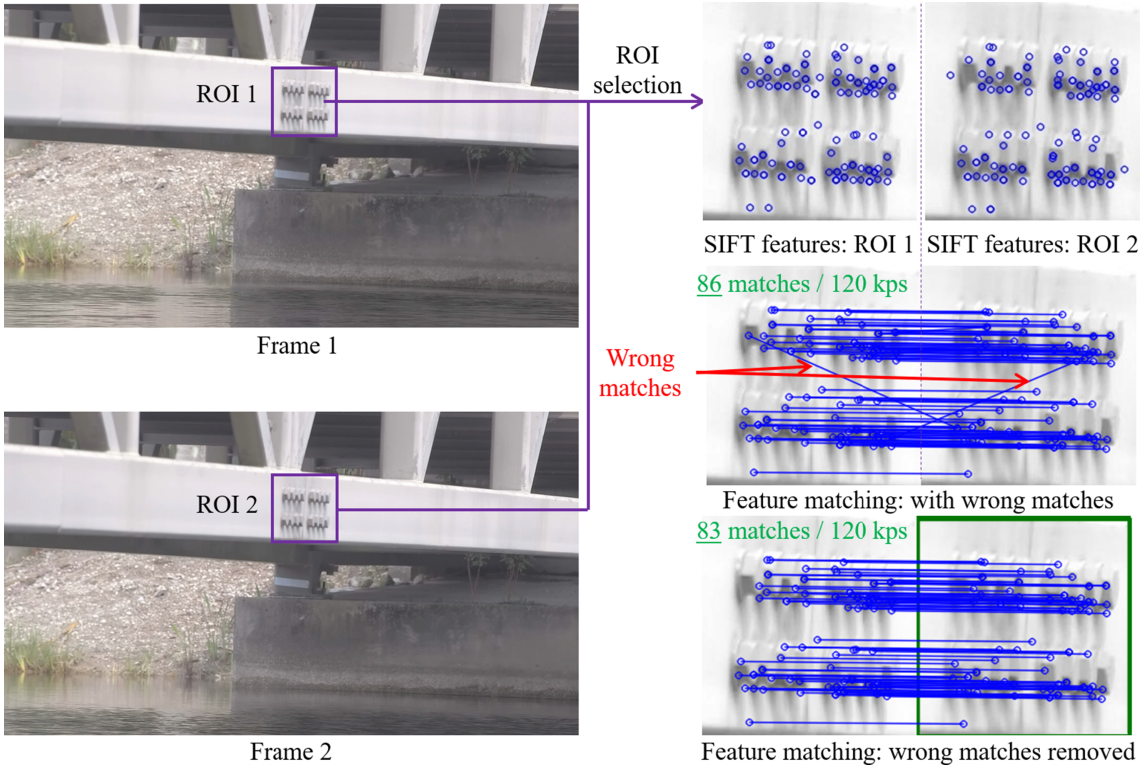

Fig. 3. Feature extraction and outlier removal.

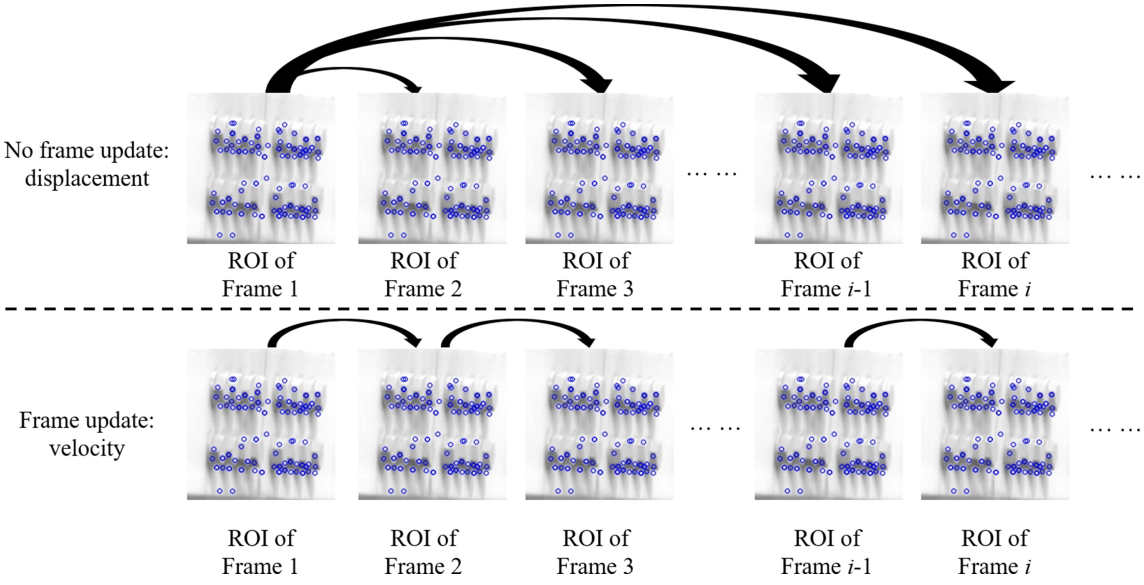

Fig. 4. Feature matching for displacement and velocity.

The procedure for proposed vision-based displacement/velocity measurement method using feature matching is implemented by using Python programming language and OpenCV (Open Source Computer Vision Library). The reason of the programming language selection is that they are open source and free for users, and it is easy to employ Python and OpenCV to develop a user-friendly software. The image data acquisition device used in this study is a customer-grade portable camera and the cost of the customer-grade camera is within hundreds of US dollars. And a camera can be employed to achieve multiple points measurement. The total cost of the proposed monitoring system is order of magnitude lower than the conventional sensor-based monitoring which would cost over thousands of US dollars including the sensors, cable, data acquisition module and commercial software.

4. Experimental verification and field application

4.1. Experimental setup

A series of experiments were conducted on a footbridge on a campus in the southeast of the United States. As shown in Fig. 6, the footbridge comprises of 19.5 m long vertical truss frames which are connected via splice connection in the middle and spans an entire length of 39 m over a pond. The width of the bridge is 4.17 m. The vertical truss members on the left and the right side have HSS 10 × 10 × 3/8 top and bottom

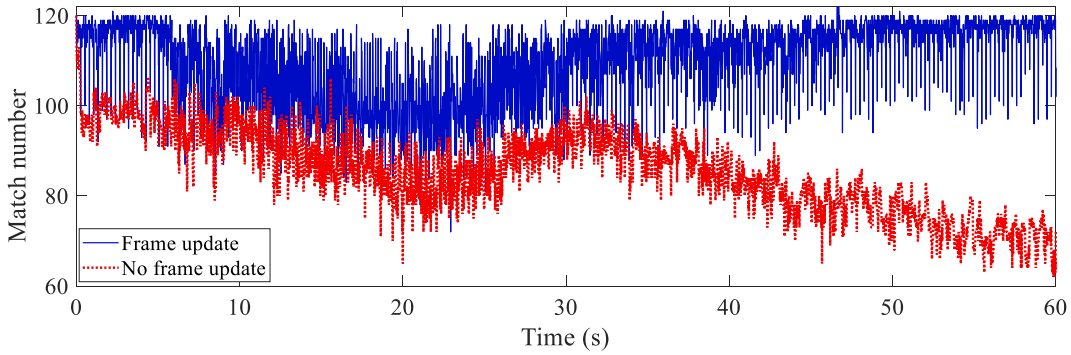

Fig. 5. Matched pairs of an image sequences using different feature matching strategies.

Fig. 6. Experimental setup.

chords and are stabilized with HSS 6 × 4 × 3/8 type vertical and HSS 4 × 4 × 1/4 type diagonal steel members. The lateral stability is provided by another truss frame that is 3.65 m wide which is constructed with HSS 3 × 3 × 1/4 type diagonal cross braces, W12 × 22 type lateral members. Two separate spans are spliced in the middle and the entire frame holds a thin layered aluminum-concrete composite deck [29]. In general, the bridge is under light pedestrian traffic loads and small vehicles such as golf carts. The fundamental frequency of the footbridge is 2.54 Hz as presented in the authors previous publication [30]. In this experiment, a camera with the resolution of 1920 × 1080 pixels and the speed of 60 frame per second was employed to monitor the vibration of the mid-span, marked as P1. An accelerometer was also installed at the mid-span to record the vibration of the footbridge. The sampling rate of the accelerometer was 200 Hz.

Eight people were employed to conduct this experiment. One person stood on the mid span as the passive subject (footbridge bystander) and the other seven people moved as active subjects (pedestrians) by walking, running and jumping with different paces in different loading cases respectively. The reason why eight people was employed is that ISO 10137 states that a group size of 8 to 15 people are the average pedestrian flow based on the daily occurrence rate. For this footbridge, the daily occurrence rate as observed is smaller than the range stated in ISO 10137. Here the low bound of the group size of 8 to 15 people was selected. Table 2 lists the loading cases conducted in this experiment. In Case 1 to 3, the seven people walked on the footbridge in a group by following the beats played by a metronome which can guide the people to walk with the predesignated paces. The paces produced by the metronome are 101, 120 and 201 beats per minute (bpm), respectively, which are equivalent to 1.68, 2.0 and 3.35 Hz. Here 1.68 Hz and 3.35 Hz are close to the value calculated by subtracting and adding the fundamental frequency of the footbridge with one third of it. In Refs. [1,31], it is presented that the frequency of people walking on footbridges follows a normal distribution with a mean pacing rate of 2.0 Hz and standard deviation of 0.173 Hz. That is the reason that 2.0 Hz was chosen in this experiment. In Case 4, the seven people randomly walked on the footbridge with their normal paces. In Case 5, the seven people ran by following the metronome with the pace of 180 bpm (3 Hz) and the frequency 3 Hz for running is in the range 2.0–3.5 Hz defined by [32]. In Case 6, the seven people randomly ran on the footbridge. In Case 7, the seven people jumped on the footbridge by following the metronome with the pace of 150 bpm (2.5 Hz) which is close to the fundamental frequency of the structure (2.54 Hz). In Case 8, the seven people jumped on the footbridge with random paces. During the experiment, the camera and accelerometer both recorded the vibrations of all the loading cases.

4.2. Result analysis

4.2.1. Comparison of displacement results from two different feature matching strategies

In this section the measurement results from Case 5 is selected to compare the performance of displacement measurement using the vision-based methods with frame update and no frame update strategies. With no frame update, the displacement can be calculated directly from Eq. (2). With frame update, the displacement is calculated by accumulating the displacement change between the two consecutive images since the location change in two consecutive images actually is the incremental value of displacement at the time of current frame.

Fig. 7 shows the displacements from two different matching strategies. The correlation coefficient between the two time histories is 91.68%. As stated in [28], the displacement by using the strategy of frame update has the accumulating errors when converting the velocity/acceleration to displacement and this phenomena also occurs in Fig. 7a. While in frequency domain, both of them give the consistent frequencies: (1) 2.534 Hz (close to fundamental frequency, 2.54 Hz) and (2) 3.001 Hz (frequency of people running beat, 3.0 Hz). From this comparison, it is recommended that if displacement is required for serviceability assessment, the vision-based method with no frame update should be used.

4.2.2. Comparison of velocity results from two different feature matching strategies

Here Case 5 is also selected to do the comparison: to compare the performance of velocity measurement using the vision-based methods with frame update and no frame update strategies. With frame update, the velocity can be calculated directly from Eq. (3). With no frame update, the velocity is obtained by calculating the numerical differentiation (derivative) of the displacement which is calculated by using Eq. (2). The procedure of calculating the velocity (V) and acceleration (a) from displacement by using numerical differentiation are expressed as:

$$V_i = \frac{Y_{i+1} - Y_i}{\Delta t} \tag{4}$$

Table 2
Experimental cases.

Case	Loading form	Pace (bpm, beat per minute)	Frequency (Hz)
1	Seven people, walking with a metronome	101	1.68
2	Seven people, walking with a metronome	120	2.0
3	Seven people, walking with a metronome	201	3.35
4	Seven people, random walking	Random pace	–
5	Seven people, running with a metronome	180	3.0
6	Seven people, random running	Random pace	–
7	Seven people, random jumping	150	2.5
8	Seven people, random jumping	Random pace	–

$$a_i = \frac{V_{i+1} - V_i}{\Delta t} \tag{5}$$

where i is number of data sample, V_i is the velocity of the ith data sample, a_i is the acceleration of the ith data sample, and Δt is sampling time of each data sample. In this study, Δt is equal to $1/f_{cam}$, where f_{cam} is the frame rate of the camera. Fig. 8 shows the velocities from two different matching strategies. The correlation coefficient between the two time histories is 97.03%. Comparing to the results of displacement (correlation coefficient is 91.68%), the velocity results show higher consistency, which means calculating velocity from displacement using numerical differentiation gives more reliable results. In frequency domain, the two frequencies are also 2.534 Hz and 3.001 Hz which are consistent with those in Section 4.2.1. From this comparison, it is suggested that if velocity is required for serviceability assessment, the vision-based method with either frame update or no frame update is good.

4.2.3. Comparison of acceleration results from two different feature matching strategies and accelerometer

The acceleration can be calculated by taking the first derivative of the velocity or by taking the second derivative of the displacement using numerical methods as expressed in Eqs. (4) and (5). Fig. 9 and Fig. 10 shows the acceleration data (raw data, without filtering) directly calculated from displacement (no frame update) and velocity (frame update). Fig. 11 shows the acceleration data collected by accelerometer. Comparing with Fig. 9a, Fig. 10a and Fig. 11a, it can be seen that the ranges of the raw acceleration data obtained from vision-based methods are both from -2 m/s^2 to 2 m/s^2, while the range of the acceleration collected by accelerometer is from -0.5 m/s^2 to 0.5 m/s^2. Comparing Fig. 7b, Fig. 8b, Fig. 9b, Fig. 10b, and Fig. 11b, it is indicated that high order modes are mixed into the acceleration time histories when converting displacement to acceleration and velocity to acceleration. The reason is thought due to the numerical differentiation. It might also induce the large spikes in acceleration data and a large range, e.g. from -2 m/s^2 to 2 m/s^2.

To eliminate the high order modes mixed in the acceleration data obtained from vision-based methods in Fig. 9 and Fig. 10, a low-pass filtering is required. Fig. 12 shows the comparison of filtered acceleration obtained from vision-based methods with the data from accelerometer. From Fig. 12a, it is suggested that the three acceleration time histories are consistent with each other very well. Fig. 12b shows the zoomed window of Fig. 12a within the range of 18 s to 22 s. Large

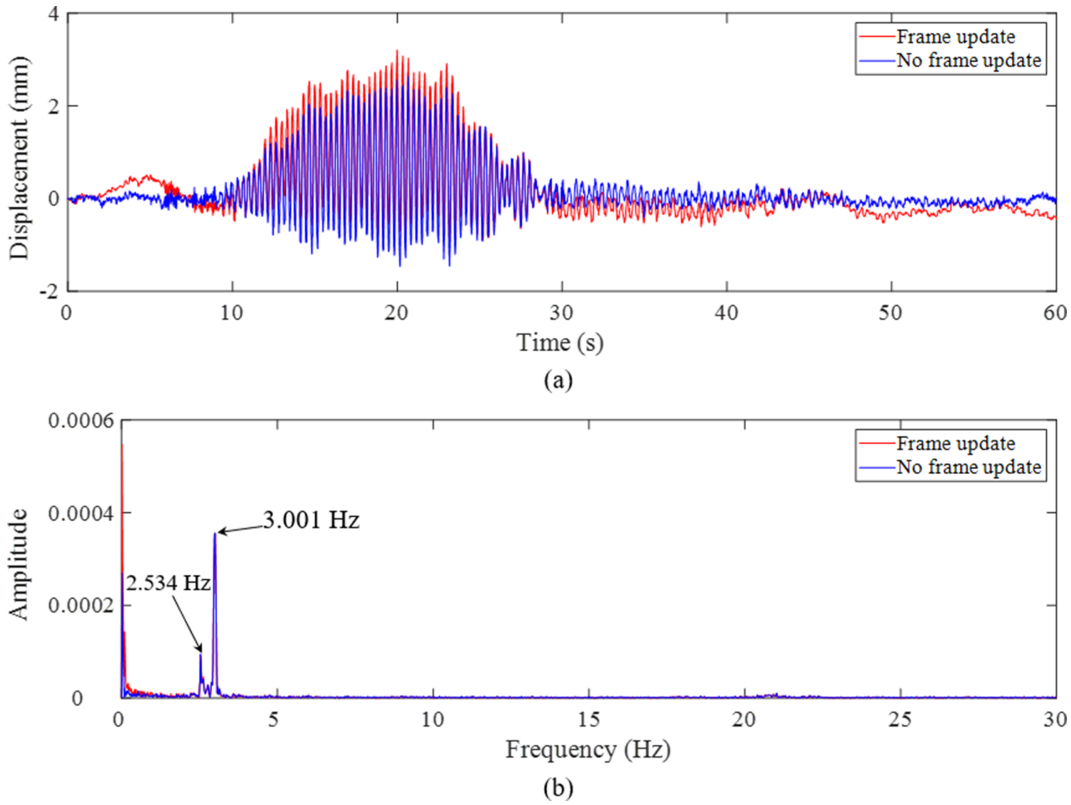

Fig. 7. Comparison of displacement results from two different feature matching strategies of Case 5: (a) comparison in time domain, and (b) comparison in frequency domain.

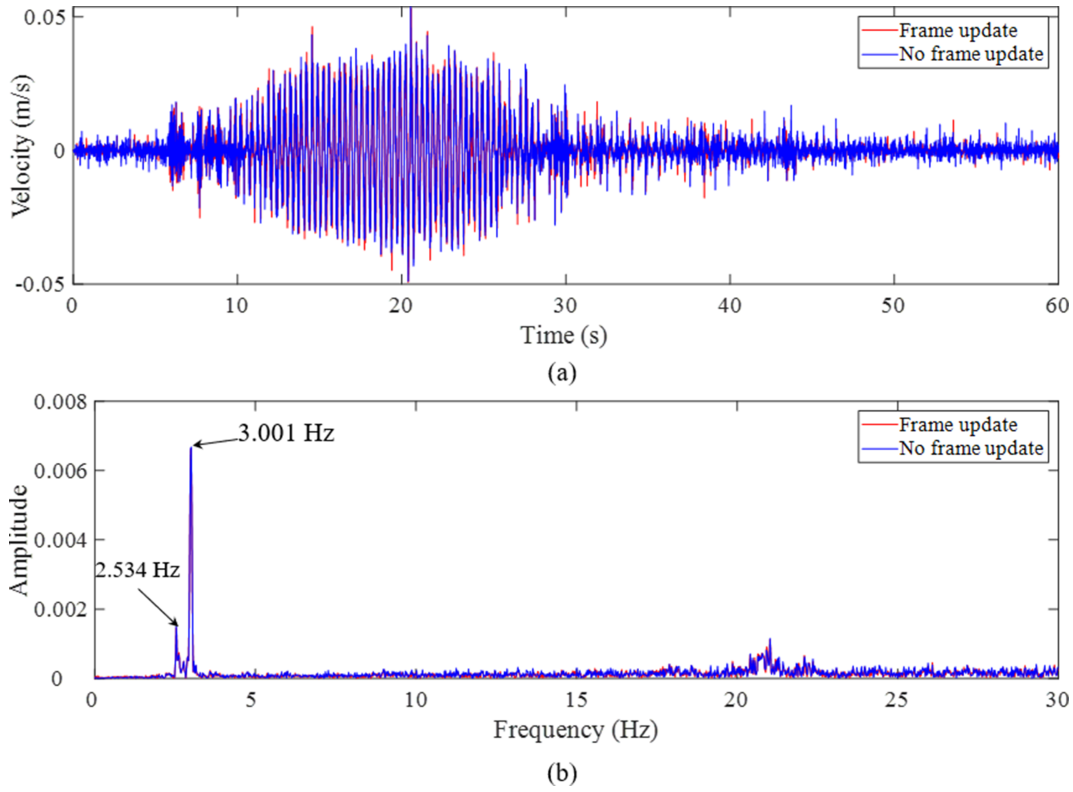

Fig. 8. Comparison of velocity results from two different feature matching strategies of Case 5: (a) comparison in time domain, and (b) comparison in frequency domain.

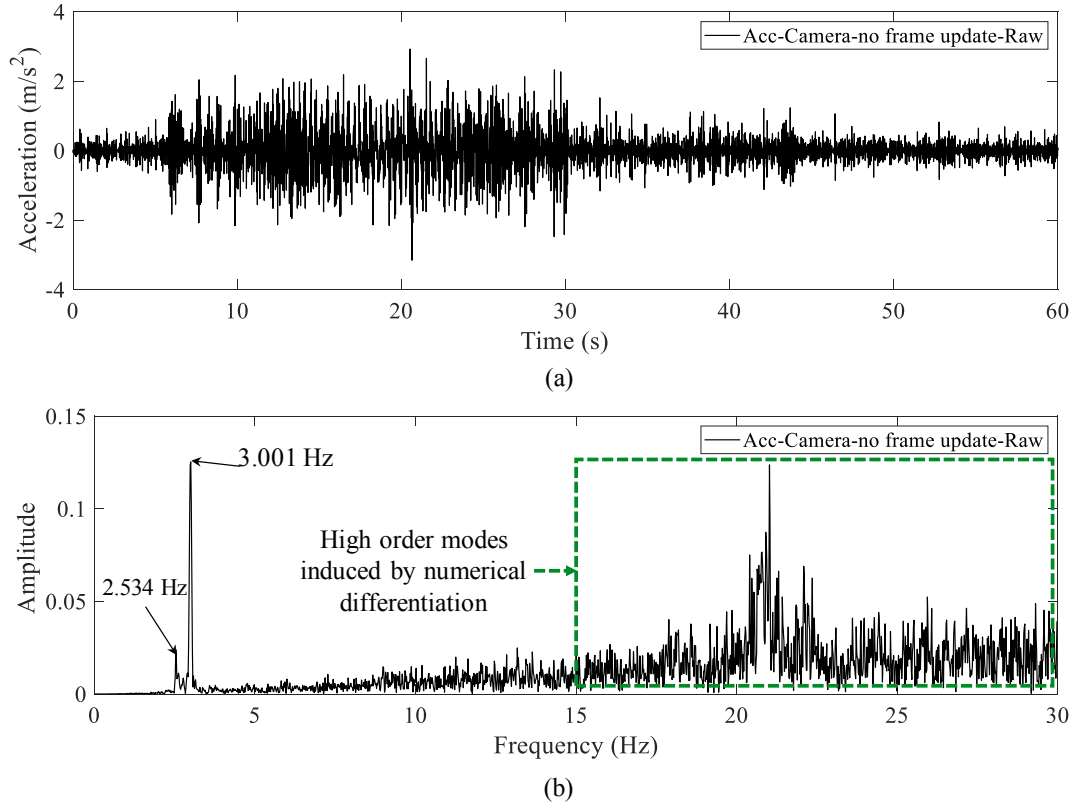

Fig. 9. Acceleration (raw data) directly calculated from the displacement data by using vision-based method with no frame update of Case 5: (a) in time domain, and (b) in frequency domain.

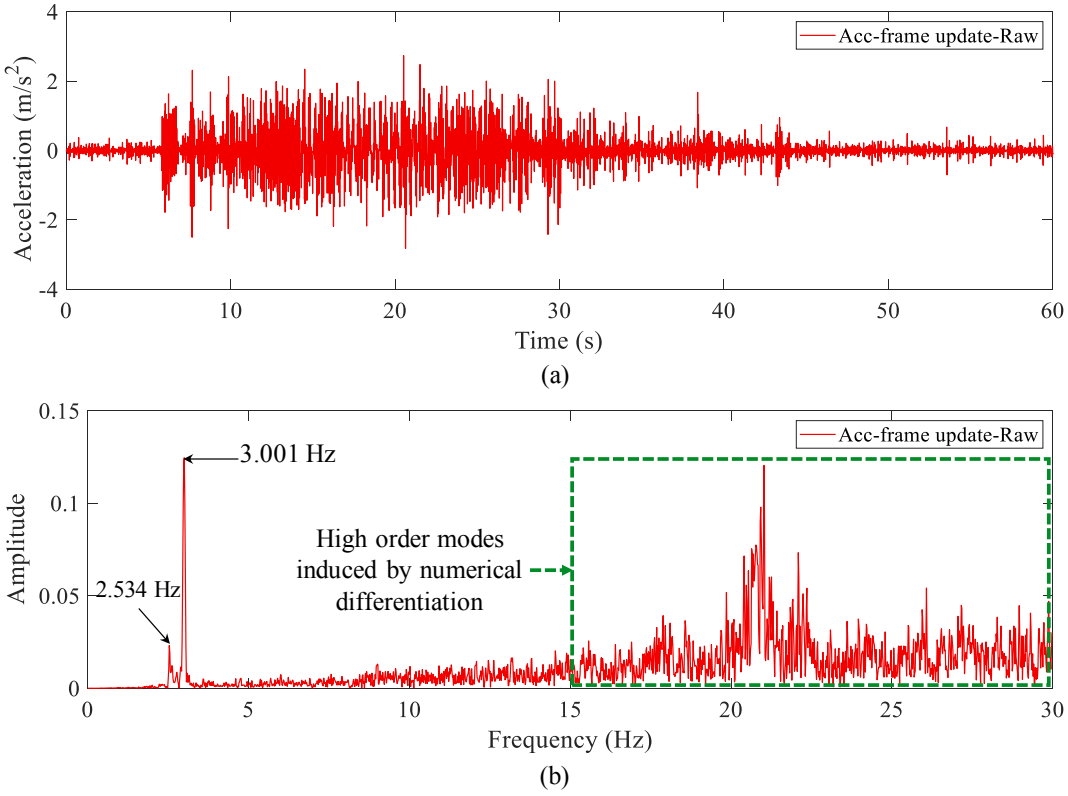

Fig. 10. Acceleration (raw data) directly calculated from the velocity data by using vision-based method with frame update of Case 5: (a) in time domain, and (b) in frequency domain.

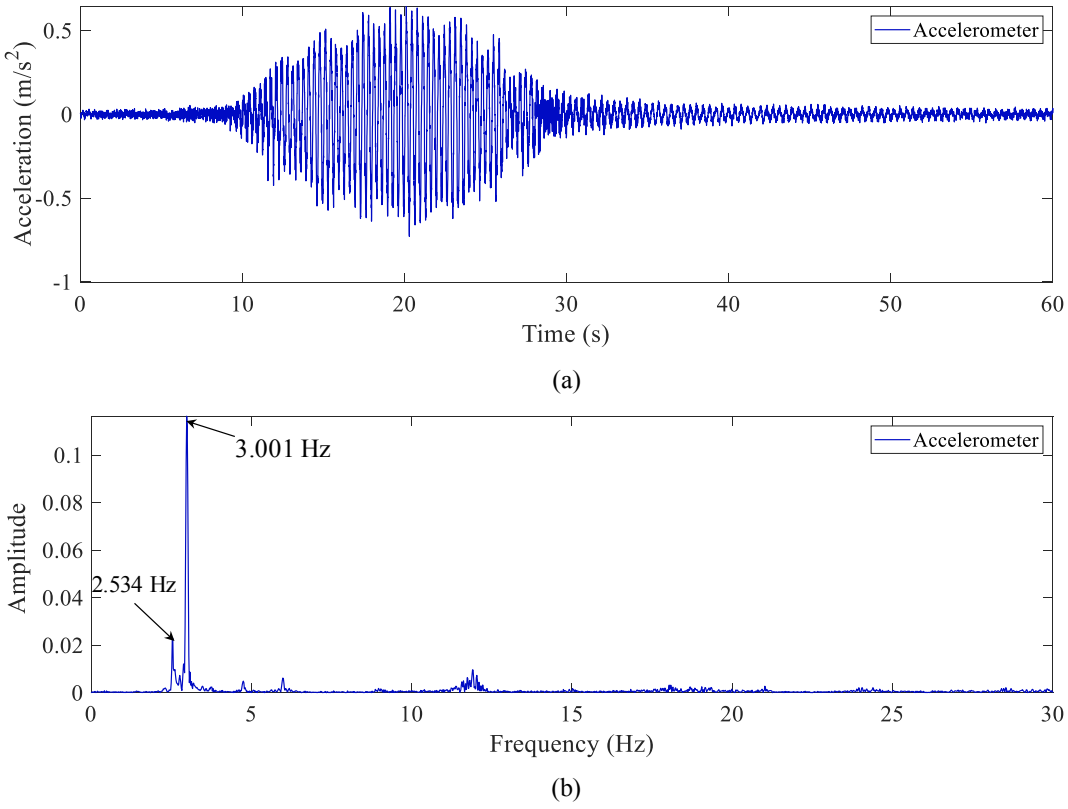

Fig. 11. Acceleration data collected by accelerometer of Case 5: (a) in time domain, and (b) in frequency domain.

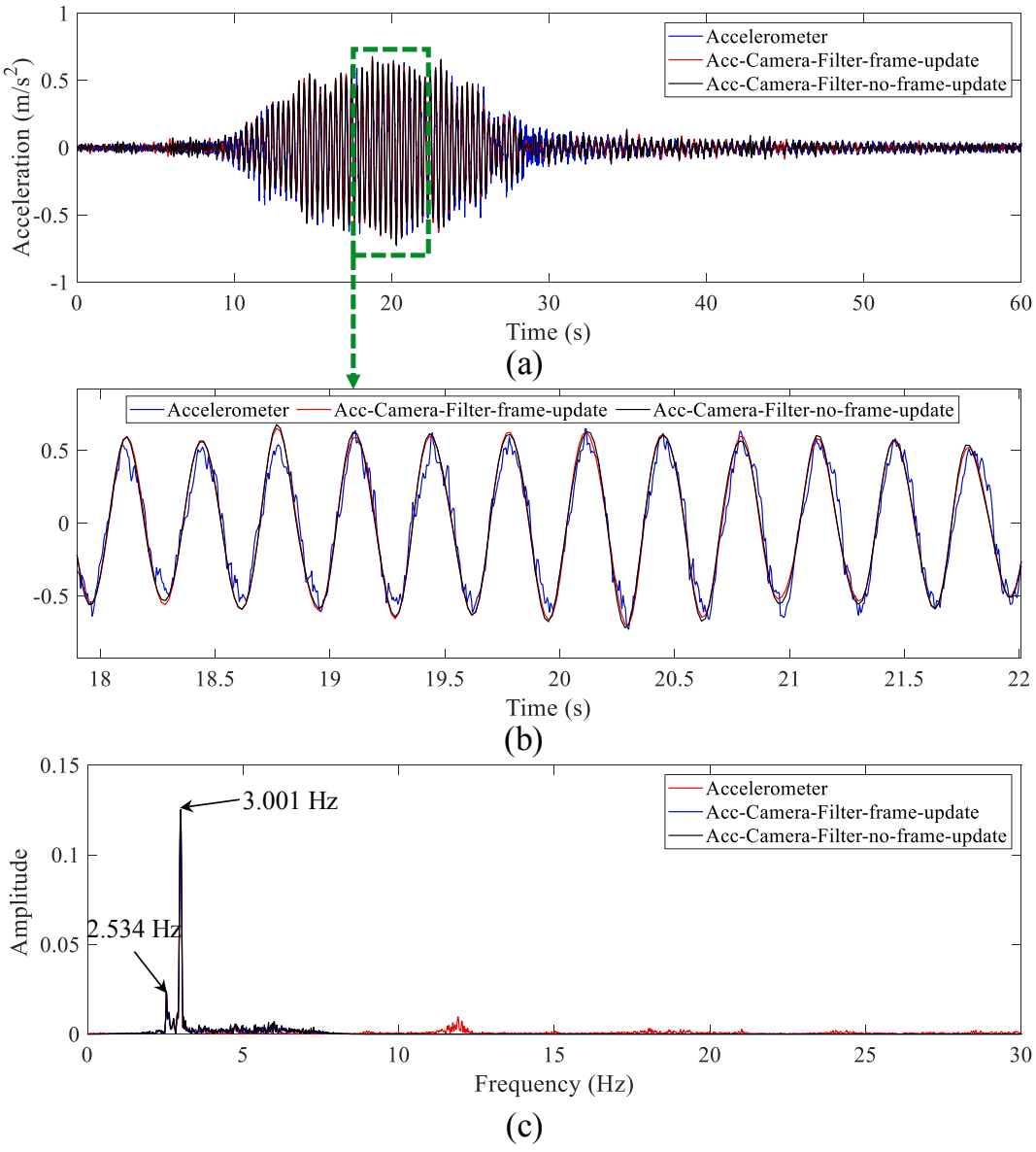

Fig. 12. Comparison of filtered acceleration obtained from vision-based methods with the data from accelerometer of Case 5: (a) in time domain, and (b) in frequency domain.

Table 3
Correlation matrix of time acceleration histories.

	Accelerometer	Camera-frame update	Camera-no frame update
Accelerometer	1	96.07%	95.72%
Camera-frame update	96.07%	1	99.38%
Camera-no frame update	95.72%	99.38%	1

spikes in the raw acceleration data obtained from vision-based methods shown in Fig. 9a and Fig. 10a are removed and the amplitude range is within -0.5 m/s^2 to 0.5 m/s^2.The high order modes in Fig. 9b and Fig. 10b are eliminated as shown in Fig. 12c. This is important because in current standards and codes for vibration serviceability assessment, the peak acceleration value or RMS is used to define the serviceability limits. Spikes in the raw acceleration time histories due to the numerical differentiation would cause higher peak value and RMS, which result in underestimated serviceability assessment results. Table 3 lists the correlation matrix of time acceleration histories obtained from vision-based methods (camera) with low-pass filtering and accelerometer. It can be seen that the correlation coefficients between the data extracted from camera and accelerometer are 96.07% (frame update) and 95.72% (no frame update), which gives a quite consistent verification. Also, the correlation coefficient between the data obtained from two feature matching strategies is very high, 99.38%. If the acceleration collected by the conventional sensor, i.e., accelerometer is regarded as the ground truth, considering the high consistencies between the three acceleration time histories, the acceleration obtained from both the two feature matching strategies are suitable for serviceability assessment.

4.2.4. *Vibration serviceability assessment based on human comfort level*

From the comparisons in Section 4.2.2 and 4.2.3, it is noted that the measurement results of velocity and acceleration obtained by using two different feature matching strategies are very consistent with each other and the acceleration results derived by using the two strategies are also very close to the results obtained from the conventional accelerometers. However, as indicated in Section 4.2.1, the displacement data extracted from the feature matching strategy of frame update could produce

Table 4
RMS and peak value of the acceleration data.

Case	RMS (m/s^2)		$a_{peak\text{-}RMS}$ (m/s^2)		a_{peak} (m/s^2)	
	Cam	Accl	Cam	Accl	Cam	Accl
1	0.0324	0.0341	0.0458	0.0482	0.099	0.124
2	0.0564	0.0389	0.0798	0.055	0.176	0.118
3	0.147	0.110	0.207	0.156	0.341	0.467
4	0.0373	0.0351	0.0527	0.0497	0.0857	0.109
5	0.369	0.330	0.522	0.468	0.659	0.650
6	0.474	0.495	0.671	0.701	1.051	1.178
7	0.904	0.892	1.279	1.262	1.919	1.836
8	0.6151	0.584	0.870	0.826	1.500	1.426

accumulated errors, which make the displacement measurement accuracy be lower than the strategy with no frame update. Based on the consideration for more accurate data sources including displacement, velocity and acceleration obtained from vision-based methods, the feature matching strategy with no frame update is implemented in this study to first extract the displacement data from the images. Then the displacement data are converted to velocity and acceleration by using numerical differentiation and low-pass filtering for the purposes of structural vibration serviceability assessment as expressed in Eqs. (4) and (5). All the loading cases in this study are processed by following this procedure to get the final acceleration time histories from image sequences. The computation times for acceleration data sample from image sequence collected in each case are very close and the average time is around 0.06 s including the image feature matching and the derivation operation. The data processing was carried out on a personal computer with the AMD Ryzen 5 2600X CPU, 16 Gb RAM and Windows 10 system. To keep the paper reasonably concise, the displacement and acceleration time histories of all the loading cases are not listed here. The RMS and peak value calculated from the acceleration data by using vision-based methods and conventional accelerometers are shown directly. Table 4 lists the RMS and peak value of the acceleration data (a_{peak}). Here the RMS is also converted to the equivalent peak acceleration value ($a_{peak\text{-}RMS}$) as indicated in Fig. 1. $a_{peak\text{-}RMS}$ and a_{peak} are applied to assess the vibration serviceability comparing with the serviceability limits based on human comfort shown in Fig. 1. In Table 4, the column titled by “Cam” (Camera) is the value calculated by the displacement data obtained from the vision-based method with no frame update and the column titled by “Accl” (Acceleration) is the value calculated by the acceleration obtained from accelerometer.

Table 5 lists the results of the vibration serviceability assessment of different cases using vision-based methods and accelerometer. Since the assessment of the vibration serviceability is also related to the vibration frequencies as presented in the standards such as BS 5400 and ISO 10137 (Fig. 1), in Fig. 13 the vibration results of all the loading cases in frequency domain are also listed. The results of two different approaches are very consistent. It can be seen that the level of vibration serviceability of Cases 1 to 4 is within the limits of ISO 10137, Euro code 5 and BS 5400. For the Setra code, it is within the maximum comfort level. Cases 1 to 4 are the cases of human walking with different paces (1.68 Hz, 2.0 Hz, 3.35 Hz and random pace). Among the four cases, walking with the frequency of 3.35 Hz (Case 4) causes the largest RMS and a_{peak}.

Fig. 13a to 13c show the vibration results of the walking cases with different paces (loading frequencies). It can be seen that the walking loads excite both the fundamental frequencies and loading frequencies. And the vibration modes related to the two frequencies have higher energy, which can be seen from the amplitudes of related frequencies. However, there are still some small differences of the contributions of the two frequencies in the vibration signals of the three cases. In Case 1 as shown in Fig. 13a, the amplitude of the loading frequency (1.68 Hz) is just around 25% of the amplitude of the fundamental frequency (2.54 Hz), while in Case 2 as shown in Fig. 13b, the amplitude of the loading frequency (2.0 Hz) is around 3.8 times of the amplitude of the fundamental frequency (2.54 Hz). In Case 3 as shown in Fig. 13c, the amplitudes of the loading frequency (3.35 Hz) and the fundamental frequency (2.54 Hz) are very close. From Fig. 13d, it is noted that more vibration frequencies are excited when the walking pace (or loading frequency) is random. Especially the four frequencies around the fundamental frequency (2.54 Hz) exhibit the highest energy in the vibration signal and the amplitude of the fundamental frequency is around twice of the other frequencies such as 1.955 Hz, 3.787 Hz and 4.773 Hz. In Fig. 13a to 13d, it is also observed that there are some discrepancies in the comparison between the data from camera and accelerometer and the data from camera shows some noise around 5 Hz. The reason might be that during the walking load cases, the structural displacement components around 5 Hz, which is higher than the fundamental frequency of the structure and the loading frequency, are too small and difficult to capture. Noise inevitably exists in the displacement measurements when using the vision-based methods. In this case, noise would not affect the serviceability assessment since the vibration components of fundamental frequency and loading frequencies are not contaminated by the noise around 5 Hz. The noise problem can be mitigated by applying some possible solutions such as improving the camera resolution or attaching the manual markers with rich features on the surface as tracking targets instead of using natural features on the bridge surface.

In Case 5, seven people running at the frequency of 3 Hz caused larger RMS and a_{peak} than walking cases, but the vibration serviceability level is still within the limits of ISO 10137, Euro code 5 and BS 5400. For the Setra code, it is in the range between mid and maximum comfort level. However, when the seven people ran with random paces (Case 6), the vibration serviceability level exceeds the limits defined by ISO 10137, Euro code 5 and BS 5400. Also, for Setra code, it moves to the level between minimum and mid comfort level. The reason for the increase of RMS and a_{peak} might be that during the random running case, as shown in Fig. 13f the seven people’s random running induced a

Table 5
Serviceability assessment of different cases.

Case	Serviceability Assessment							
	ISO 10137		Euro code 5		BS 5400		Setra	
	Cam	Accl	Cam	Accl	Cam	Accl	Cam	Accl
1: 1.68 Hz walking	Under limit	Under limit	Under limit	Under limit	Under limit	Under limit	Max comfort	Max comfort
2: 2.0 Hz walking	Under limit	Under limit	Under limit	Under limit	Under limit	Under limit	Max comfort	Max comfort
3: 3.35 Hz walking	Under limit	Under limit	Under limit	Under limit	Under limit	Under limit	Max comfort	Max comfort
4: random walking	Under limit	Under limit	Under limit	Under limit	Under limit	Under limit	Max comfort	Max comfort
5: 3.0 Hz running	Under limit	Under limit	Under limit	Under limit	Under limit	Under limit	Mid-Max comfort	Mid-Max comfort
6: random running	Exceed limit	Exceed limit	Exceed limit	Exceed limit	Exceed limit	Exceed limit	Min-Mid comfort	Min-Mid comfort
7: 2.5 Hz jumping	Exceed limit	Exceed limit	Exceed limit	Exceed limit	Exceed limit	Exceed limit	Min-Mid comfort	Min-Mid comfort
8: random jumping	Exceed limit	Exceed limit	Exceed limit	Exceed limit	Exceed limit	Exceed limit	Min-Mid comfort	Min-Mid comfort

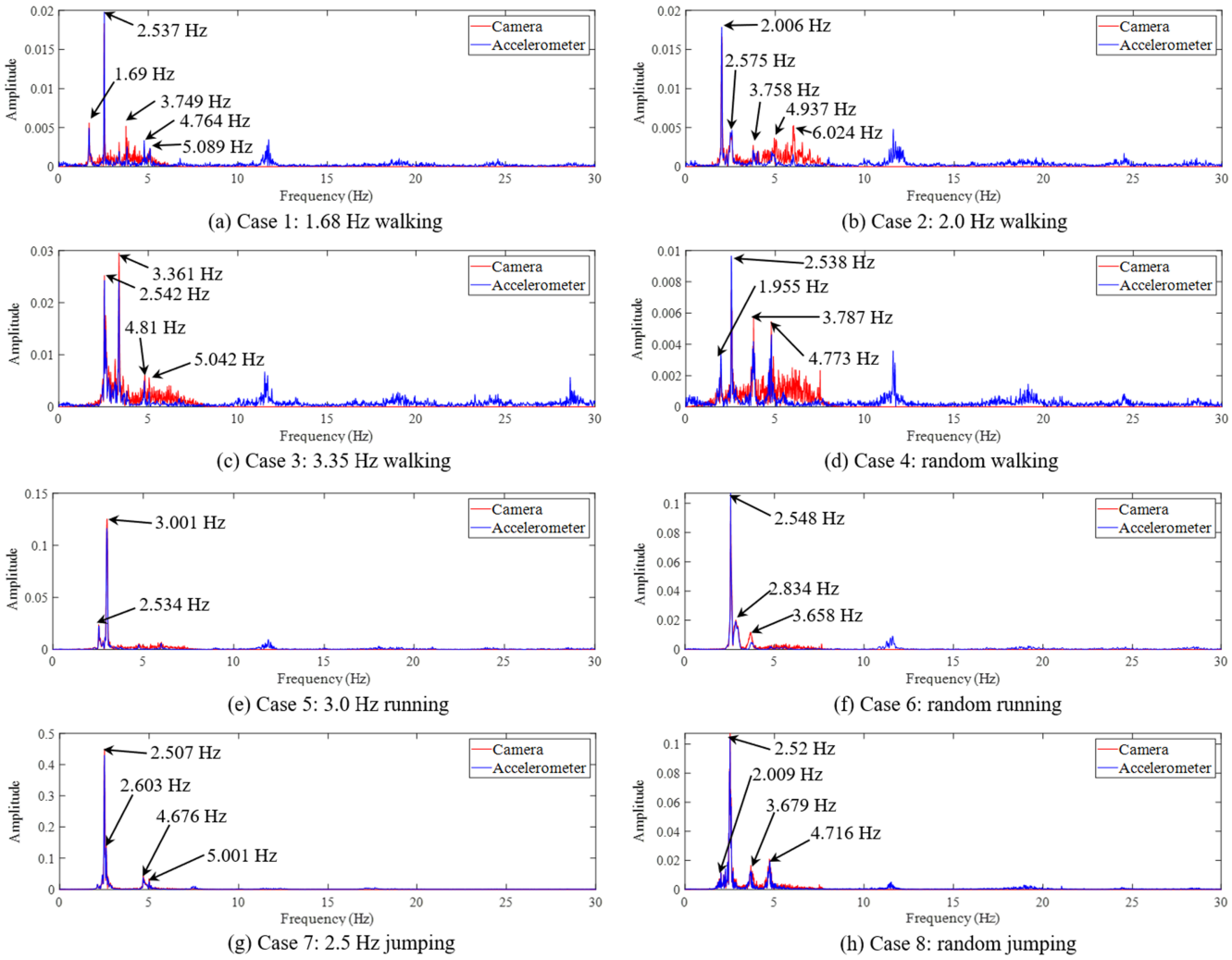

Fig. 13. Frequency spectra of the acceleration time histories obtained from vision-based method and accelerometer of all cases.

dominant frequency, 2.548 Hz which is the resonant frequency of the footbridge and this component make up a large portion in the time history. The resonant effect induced larger amplitude of acceleration responses and larger RMS and a_{peak}. While in Case 5 shown in Fig. 13e, the dominant frequency of running with the frequency of 3.0 Hz is 3.001 Hz which is not the resonant frequency.

In Case 7 as shown in Table 4, when the seven people kept jumping on the bridge with the frequency (2.54 Hz) which is identical with the fundamental frequency of the structure, the RMS and the peak value are the largest among all the eight loading cases. In Fig. 13g, it can be seen that the loading frequency (2.5 Hz) in the Case 7 show the highest amplitude of the structural vibration. In addition, the vibration serviceability assessment result of Case 7 also exceeds the limits defined by ISO 10137, Euro code 5 and BS5400, and it reaches the Min-Mid comfort level defined by Setra. When seven people jumping with random frequencies on the bridge, the vibration serviceability assessment result of Case 8 is similar to the results of Case 6 and Case 7. In Fig. 13h, the dominant frequency is also close the fundamental frequency of the structure. It is also observed from Fig. 13d, 13f and 13 h, when the people walk, run and jump with random frequency, the dominant frequency of the vibration responses of the structure is always around the fundamental frequency of the structure. The fundamental frequency of the structure might be slightly affected by the loading types and paces of the pedestrian. In addition to support the vibration serviceability assessment, these observations can also benefit the engineering practices in structural design for the purposes of vibration serviceability.

From the results shown in Table 4, Table 5 and Fig. 13, it can be seen that the proposed method by using computer vision techniques to assess the vibration serviceability can give very consistent results compared to that obtained from conventional accelerometers. It is promising to apply the proposed method in the future engineering practices for vibration serviceability assessment. However, as a non-contact and optical-based method, the proposed method might be influenced by the adverse environmental factors such as illumination change and camera shaking induced by the ground vibration or wind effects. The SIFT feature detector and VGG descriptor implemented in this study were presented in Refs. [33,34] to be robust to the illumination change within selected datasets, the robustness to the illumination change in field application remains to be validated in the future. The measurement errors due to the camera shaking in real environment can be mitigated by using the background reference subtraction approach as stated in the previous work [28]. The reader is addressed to Ref. [28] for more details.

5. Conclusions

In this study, the vibration serviceability assessment based on human comfort of a footbridge using computer vision techniques is investigated. A series of experiments of footbridge under different types of human loading including walking, running and jumping with different paces (frequencies) are conducted to verify the proposed

approach. The main approaches, findings, and conclusions are as follows:

(1) The vibration serviceability assessment criteria in current standards and codes are discussed and ISO 10137, Euro code 5, BS 5400 and Setra are taken as the reference for serviceability assessment;
(2) A vision-based displacement/velocity monitoring approach is proposed by using feature matching. Two different feature matching strategies such as matching between first frame and current frame (no frame update) and matching between consecutive frames (frame update) are compared and it is suggested that the feature matching with no frame update gives good displacement, velocity and acceleration, while the feature matching with no frame update does not perform good in terms of displacement measurement.
(3) The feature matching with no frame update is applied to estimate the displacement data from image sequence and the displacement is first converted to acceleration and then indicators such RMS and a_{peak} are calculated for the serviceability assessment compared with the conventional accelerometer.
(4) The random running, random jumping and jumping with the frequency that is close to the fundamental frequency of the footbridge can induce the serviceability level to exceed the limits defined by the current standards and codes.
(5) The human loads of different types and paces can induce a slight change (within 0.1 Hz in this experiment) of the fundamental frequencies of the footbridge and this would affect the serviceability assessment since the serviceability limit defined in standards and codes is related the fundamental frequency.

The serviceability assessment in current standards and codes is based on the human comfort and human comfort is a concept of perception. It is stated in literature that both the displacement and velocity can trigger the perception of human [1]. For example, the minimum displacement that human can percept is 0.001 mm [35]. Smith (1969) applied displacement as the indicator for human comfort and serviceability [36]. It is also noted that in Japan by Yoneda (2002), velocity is used as the index to assess footbridge serviceability [37]. Živanović et al (2004) stated that "*Usually, acceleration response was measured because it was established as the best parameter for describing people's reaction to vibrations and, also, it was easy to measure it using widely available accelerometers*" [1]. This explains the main reasons why acceleration is used in most of the standards and codes for serviceability assessment. Although in current standards and codes, acceleration is still the most common data type for serviceability assessment, displacement and velocity can also be employed as another alternative. Considering the advantages of vision-based methods such as non-contact, long distance, low cost, time saving, and ease of use compared to conventional accelerometers [38], it is shown that using vision-based methods to collect displacement, velocity and acceleration would be a practical and efficient choice for vibration serviceability assessment for many field scenarios. As a result, this paper presents a novel approach for such practical serviceability assessment using different data types, i.e. displacement and velocity. It is interesting to observe that such novelties that were discussed in a forward look for sensing and monitoring for structures for serviceability [39] are now realized and many others technologies can be expected to positively impact structural engineering in the near future.

CRediT authorship contribution statement

Chuan-Zhi Dong: Methodology, Software, Formal analysis, Data curation, Writing - original draft. **Selcuk Bas:** Visualization, Resources, Writing - original draft. **F. Necati Catbas:** Conceptualization, Supervision, Project administration, Funding acquisition, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The financial support for this research was provided by U.S. National Science Foundation (NSF) Division of Civil, Mechanical and Manufacturing Innovation [grant number 1463493]. The authors would like to acknowledge members of the Civil Infrastructure Technologies for Resilience and Safety (CITRS) research group at University of Central Florida for their endless support in creation of this work. The second author would like to kindly acknowledge the Scientific and Technological Research Council of Turkey (TUBITAK) through grant number 2219. The authors specially acknowledge Dr. Ninel Alver for the valuable suggestions and editing help during the writing of this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.engstruct.2020.111224.

References

[1] Živanović S, Pavic A, Reynolds P. Vibration serviceability of footbridges under human-induced excitation: A literature review. J Sound Vib 2005;279:1–74. https://doi.org/10.1016/j.jsv.2004.01.019.
[2] Živanović S, Pavia A. Probabilistic assessment of human response to footbridge vibration. J Low Freq Noise Vib Act Control 2009;28:255–68. https://doi.org/10.1260/0263-0923.28.4.255.
[3] BBC, 2000: Swaying Millennium Bridge closed, BBC News. http://news.bbc.co.uk/onthisday/hi/dates/stories/june/10/newsid_2510000/2510839.stm.
[4] BBC. "Wobbly" Millennium Bridge fixed. BBC News 2002.
[5] Setareh M. Vibration Serviceability Issues of Slender Footbridges. J Bridg Eng 2016;21:1–12. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000951.
[6] ISO. Bases for design of structures -serviceability of buildings and walkways against Vibrations, 2nd ed., International Organization for Standardization, Geneva, Switzerland; 2007.
[7] ECS. Eurocode 5: design of timber structures—part 2: bridges (ENV 1995- 2). European Committee for Standardization; 1997.
[8] BSI, Steel, Concrete and Composite Bridges—Part 2: Specification for Loads, British Standards Association, London; 1978.
[9] Setra, Footbridges: Assessment of Vibrational Behavior of Footbridges Under Pedestrian Loading (Technical guide), Paris; 2006.
[10] Mackenzie D, Barker C, Mcfadyen N, Allison B. Footbridge pedestrian vibration limits Part2: Human sensitivity. In: Footbridge 2005 - Second Int. Conf., Venice, Italy; 2005. http://www.bridgeforum.com/bof/meetings/bof18/PAPmackenzie-part2.pdf.
[11] Kasperski M. Vibration serviceability for pedestrian bridges. Proc Inst Civ Eng Struct Build 2006;159:273–82. https://doi.org/10.1680/stbu.2006.159.5.273.
[12] Barker C. Vibration Limits Background to Response Calculation. Int J Sp Struct 2007;22:35–43.
[13] Dey P, Narasimhan S, Walbridge S. Evaluation of design guidelines for the serviceability assessment of aluminum pedestrian bridges. J Bridg Eng 2017;22:1–15. https://doi.org/10.1061/(ASCE)BE.1943-5592.0000983.
[14] Dey P, Narasimhan S, Walbridge S. Calibrating Pedestrian-Bridge Standards for Vibration Serviceability. J Bridg Eng 2018;23:1–17. https://doi.org/10.1061/(ASCE)BE.1943-5592.0001270.
[15] Feng P, Wang Z, Jin F, Zhu S. Vibration serviceability assessment of pedestrian bridges based on comfort level. J Perform Constr Facil 2019;33:1–10. https://doi.org/10.1061/(ASCE)CF.1943-5509.0001316.
[16] Heinemeyer C, Butz C, Keil A, Schlaich M, Goldack A, Trometer S, et al. Design of lightweight footbridges for human induced vibrations. Aachen 2009. https://doi.org/10.2788/33846.
[17] Dong CZ, Celik O, Catbas FN. Marker free monitoring of the grandstand structures and modal identification using computer vision methods. Struct Heal Monit 2019;18:1491–509. https://doi.org/10.1177/1475921718806895.
[18] Zhong F, Quan C. Efficient digital image correlation using gradient orientation. Opt Laser Technol 2018;106:417–26. https://doi.org/10.1016/j.optlastec.2018.04.024.
[19] Chen Y, Joffre D, Avitabile P. Underwater dynamic response at limited points expanded to full-field strain response. J Vib Acoust 2018;140:051016. https://doi.org/10.1115/1.4039800.
[20] Celik O, Dong CZ, Catbas FN. A computer vision approach for the load time history estimation of lively individuals and crowds. Comput Struct 2018;200:32–52.

https://doi.org/10.1016/j.compstruc.2018.02.001.
[21] Jian X, Xia Y, Lozano-Galant JA, Sun L. Traffic sensing methodology combining influence line theory and computer vision techniques for girder bridges. J Sensors 2019;2019:1–15. https://doi.org/10.1155/2019/3409525.
[22] Spencer BF, Hoskere V, Narazaki Y. Advances in computer vision-based civil infrastructure inspection and monitoring. Engineering 2019;5:199–222. https://doi.org/10.1016/j.eng.2018.11.030.
[23] Pan Y, Dong Y, Wang D, Chen A, Ye Z. Three-dimensional reconstruction of structural surface model of heritage bridges using UAV-based photogrammetric point clouds. Remote Sens 2019;11. https://doi.org/10.3390/rs11101204.
[24] Dong CZ, Celik O, Catbas FN, Obrien E, Taylor S. A robust vision-based method for displacement measurement under adverse environmental factors using Spatio-Temporal context learning and Taylor approximation. Sensors 19 (2019) 3197. doi: 10.3390/s19143197.
[25] Feng D, Feng MQ, Ozer E, Fukuda Y. A vision-based sensor for noncontact structural displacement measurement. Sensors (Switzerland) 2015;15:16557–75. https://doi.org/10.3390/s150716557.
[26] Xu Y, Brownjohn J, Kong D. A non-contact vision-based system for multipoint displacement monitoring in a cable-stayed footbridge. Struct Control Heal Monit 2018;25:1–23. https://doi.org/10.1002/stc.2155.
[27] Dong CZ, Catbas FN. A non-target structural displacement measurement method using advanced feature matching strategy. Adv Struct Eng 2019;22:3461–72. https://doi.org/10.1177/1369433219856171.
[28] Dong CZ, Celik O, Catbas FN, O'Brien EJ, Taylor S. Structural displacement monitoring using deep learning-based full field optical flow methods. Struct Infrastruct Eng 2020;16:51–71. https://doi.org/10.1080/15732479.2019.1650078.
[29] Dong CZ, Bas S, Catbas FN. A completely non-contact recognition system for bridge unit influence line using portable cameras and computer vision. Smart Struct Syst 2019;24:617–30.
[30] Celik O, Dong CZ, Catbas FN. Investigation of human induced excitations using hybrid multivariate empirical mode decomposition, J Struct Eng ASCE, II Press. 2019.
[31] Matsumoto Y, Nishioka T, Shiojiri H, Matsuzaki K. Dynamic Design of Footbridges. Proc Int Assoc Bridg Struct Eng 1978;2:1–15. https://doi.org/10.5169/seals-33221.
[32] Bachmann H, Pretlove AJ, Rainer JH. Dynamic forces from rhythmical human body motions. In: Vib Probl Struct Pract Guidel, Birkhauser, Basel; 1995.
[33] Lowe DG. Distinctive image features from scale-invariant keypoints. Int J Comput Vis 2004;60:91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94.
[34] Simonyan K, Vedaldi A, Zisserman A. Learning Local Feature Descriptors Using Convex Optimisation. IEEE Trans Pattern Anal Mach Intell 2014;36:1573–85. https://doi.org/10.1109/TPAMI.2014.2301163.
[35] Pretlove AJ, Rainer JH. Human response to vibrations. In: Vib Probl Struct Pract Guidel, Birkhauser, Basel; 1995.
[36] Smith JW. The vibration of highway bridges and the effects on human comfort. UK: University of Bristol; 1969.
[37] Yoneda M. A simplified method to evaluate pedestrian-induced maximum response of cable-supported pedestrian bridges. In: Proc Int Conf Des Dyn Behav Footbridges, Paris, France; 2002.
[38] Dong C-Z, Catbas FN. A review of computer vision–based structural health monitoring at local and global levels. Struct Health Monit 2020:1–52. https://doi.org/10.1177/1475921720935585.
[39] Catbas FN, Celik O, Avci O, Abdeljaber O, Gul M, Do NT. Sensing and monitoring for stadium structures: A review of recent advances and a forward look. Front Built Environ 2017;3:1–18. https://doi.org/10.3389/fbuil.2017.00038.