A non-contact vision-based system for multi-point displacement monitoring in a cable-stayed footbridge

Yan Xu[1], James Brownjohn[1], Dali Kong[2].

1Vibration Engineering Section, University of Exeter, Exeter, EX4 4QF, UK
2 Centre for Geophysical and Astrophysical Fluid Dynamic, University of Exeter, Exeter, EX4 4QE, UK

Abstract:

Vision-based monitoring receives increased attention for measuring displacements of civil infrastructure such as towers and bridges. Currently, most field applications rely on artificial targets for video processing convenience, leading to high installation effort and focus on only single-point displacement measurement e.g. at mid-span of a bridge. This study proposes a low-cost and non-contact vision-based system for multi-point displacement measurement based on a consumer-grade camera for video acquisition and a custom-developed package for video processing. The system has been validated on a cable-stayed footbridge for deck deformation and cable vibration measurement under pedestrian loading. The analysis results indicate that the system provides valuable information about bridge deformation of the order of a few cm induced, in this application, by pedestrian passing. The measured data enables accurate estimation of modal frequencies of either the bridge deck or the bridge cables and could be used to investigate variations of modal frequencies under varying pedestrian loads.

Keywords: vision-based system; bridge displacement; cable vibration; pedestrian loads; cable-stayed bridge.

1 INTRODUCTION

Structural health monitoring (SHM) is aimed at providing valuable information about structural performance and characterisation of structural defects to the asset owners, especially for those civil infrastructures beyond the design life. Vibration-based modal tests are a common way for structural condition and serviceability assessment, providing a direct view about the structural stiffness, mass properties and their distributions [1]. As well as for validating designs of civil structures, modal parameters extracted from vibration data obtained in short-term or long-term measurements are widely believed to have potential for identifying changes in structural condition or 'damage' [2]. The sensitivity of these parameters to 'damage' depends on the nature of the 'damage', for example local deterioration of material e.g. due to corrosion may not be detectable against background effects of environmental variability, whereas boundary conditions are known to have a relatively strong effect, for example fixity of bridge supports [3].

Deformation is another important metric for bridge condition and performance assessment. For example, measurement of deformation during controlled vehicle load testing helps to estimate bridge load carrying capacity [4] [5]. Displacement is related to the structural stiffness, and extreme values might

indicate either an extreme load or a deficiency in the structure. When recorded at high sample rates, displacement data provide valuable information about dynamic characteristics, and hence changes in structural condition.

For conventional displacement sensors such as linear variable differential transformers (LVDTs) and dial gauges, a stationary reference point is required that could be challenging in field tests. Global positioning system is the proper choice for monitoring only flexible large-scale structures due to the limitation of the measurement accuracy (i.e. sub-centimetre [6] or centimetre level [7]). The indirect method through integrating the acceleration measurement is usually applied for short duration signals (e.g. a few seconds) and might fail to estimate static or quasi-static displacement components. Limitations of more traditional displacement sensing technologies have driven research in non-contact optical sensing.

1.1 Review of vision-based approaches

Vision-based systems have advantages over other sensors, e.g. easy installation, remote non-contact operation and distributed sensing that promotes use of a single camera for multi-point simultaneous measurement. Efforts have been spent on developing advanced vision-based systems to provide accurate and robust displacement measurement primarily of high-rise buildings [8], short-span bridges [9], [10], [11], [12] and long-span bridges [13], [14], [15], [16], [10]. Previous studies have indicated the significant potential of vision-based systems for structural condition evaluation, especially for system identification [18–20]. Other applications based on the camera measured displacement include finite element model calibration [21], damage detection [22] and bridge weigh-in-motion where another camera is used for the traffic monitoring [23]. However, vision-based systems still face several field challenges, such as the requirement for stable camera mounting [11], measurement error caused by lighting changes [24], and atmospheric effects affecting light refraction –particularly for long-range measurements.

Most of the existing applications have relied on artificial targets for video processing convenience leading to necessity of direct access to the structure as well as increased installation effort. Moreover, the focus is commonly only for single point displacement measurement e.g. at the bridge mid-span, although multi-point simultaneous sensing is supported by the camera sensors.

Non-contact sensing:

There have been relatively few field applications using completely non-contact vision-based systems. In most examples, an artificial target or a set of targets with salient features and some known dimensions [13], [17], [20] were attached to a structure for the convenience of stable target tracking, and more importantly for providing point or line correspondences to determine the projection transformation relating the image coordinate system and the structural coordinate system. Recent non-contact field applications [18,25–27] have eliminated the dependency on artificial targets by using a scaling factor for camera projection transformation. The scaling factor is the simplest method to obtain the projection transformation provided that either the camera-to-target distance or a feature dimension near the region

of interest is known. The scaling factor estimated by the camera-to-target distance is sensitive to the tilt angle of the camera optical axis that is suggested to be less than 10° through laboratory validation tests in short distance (≤ 3.7 m) [28]. Camera positioning is less critical for the scaling factor estimated by a known dimension [9] but the estimated scaling factor is only reliable for displacement measurement along the same direction as the provided dimension.

Distributed sensing:

Vision-based systems allow a single camera to measure structural displacements of multiple points in a structure. The feature of distributed sensing has been used in laboratory structures [18,20,29,30] for multi-storey displacement measurement and system identification, as well as for cable vibration monitoring [31–34] aimed at the estimation of modal frequencies or cable tensions, but applications in bridge deformation measurement are limited, with only a few examples [27,35].

1.2 Purpose of this study

The purpose of this study is to investigate the potential of non-contact vision-based systems for multi-point measurement in field applications. Realisation of the two features, completely non-contacting and multi-point simultaneous sensing are the focus of this study. In most applications to date, the hardware used is a professional high-resolution camera with long-focus lens, thus only a local region over the whole structure, e.g. mid-span of bridge is covered in the field of view. In this study, a low-cost consumer-grade camera with a wide angle lens is used for video acquisition with a wide area of the bridge included in the field of view. A custom-developed package is used for the video processing which supports non-contact sensing for both deck deformation and cable vibration measurement. The developed system enables quick installation/removal, requires no access to the bridge structure and provides simultaneous multi-point displacement measurement.

To that end, Section 2 provides descriptions of the proposed vision-based system including the hardware and video processing methods used. Section 3 describes the validation test on a simple laboratory beam structure and evaluates the measurement accuracy of vision-based system. Section 4 introduces a field test in a cable-stayed footbridge under pedestrian crowd loading. Section 5 provides the results of the field test including the bridge deck deformation and cable vibration in time and frequency domain. The analysis results illustrate the changing bridge modal frequencies under varying pedestrian loads.

2 PROPOSED VISION-BASED SYSTEM

Applying a vision-based system for structural displacement measurement requires setting up one or more cameras in a stable location aimed at the ‘targets’ of interest and deriving target motions through video processing techniques. ‘Targets’ could be either artificial (i.e. LED markers or planar targets with chessboard pattern) or natural structure features (i.e. bolts or holes). However, natural targets are preferred on site with reduced installation efforts for the monitoring system.

In the proposed system, the hardware comprises one consumer-grade camera (i.e. GoPro Hero 4 Black) and a tripod shown in Figure 1. The recorded video files are post-processed in a custom-developed video-processing package to extract the displacement information of structure. The programming environment is Visual Studio 2015 using C++ language and partly referring to OpenCV library.

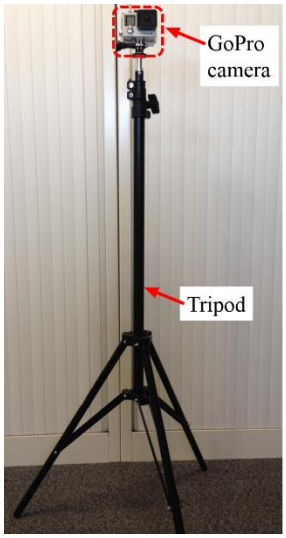


Figure 1 Hardware of vision-based system consisting a GoPro Camera and a tripod

The role of the video-processing package is tracking the target locations in image sequences and transforming the target location information in images to a time history of structural displacements. The procedures could be fitted into a three-component framework shown in Figure 2, namely camera calibration, target tracking and structural displacement calculation. The measured displacement data could be interpreted for the evaluation of structural condition e.g. system identification.

When the monitoring campaign is only for system identification and precise spatial measurements are not necessary [36], [37] e.g. cable vibration measurement, target tracking may be the only part of the whole video processing procedure needed. The prerequisite is that the cable's depth change is much smaller than the camera-to-cable distance and the cable location is close to the camera's principal axis so that the mapping between the world coordinates and the image plane becomes approximately linear [36].

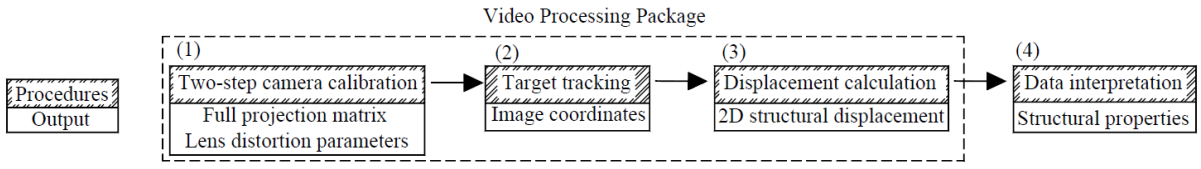


Figure 2 Procedures of video processing package and the corresponding output at each step

This section mainly introduces the video processing package and data interpretation methods used in bridge monitoring campaigns. Section 2.1 and 2.2 provide details of the two components, camera calibration and target tracking while section 2.3 demonstrates the system identification methods to analyse the monitoring data.

2.1 Two-step camera calibration

Camera calibration is aimed at determining the transformation metric between the image natural units (pixels) and the real world units (e.g. mm), and the structural displacement could be easily derived from the change of structural coordinates given the image location of a target (output of target tracking) and a transformation metric (output of camera calibration).

Mathematically, the projection process from a 3D spatial domain to a 2D image plane loses some geometric information of the target. Thus, in a single-camera system, the calibration is realised by reducing the dimensions of target motion, i.e. assuming that the target moves within a plane in 3D space. The projection is then simplified as a 2D-to-2D transformation, enabling the recovery of the 2D structural displacement. For bridge applications, the dominant motions under traffic or pedestrian loads are in the vertical direction, making it feasible to neglect the lateral or the longitude motions in a short-time monitoring campaign.

Several methods are available to determine the transformation metric.

- Scaling factor is the simplest method based on one dimensional feature correspondence or camera-to-target distance, thus is very popular in civil applications e.g. [8], [9], [13], [16], [18]. This method is based on an assumption that the camera principal axis is perpendicular to the structural surface plane or the two motion directions of interest, which sets constraints on camera position and orientation on site. Although the scaling factor estimated by a known dimension is less sensitive to camera positioning, the calibration should be applied separately to each target using one adjacent dimensional feature along the same direction as the movement of interest.
- Planar homography matrix is a transformation metric that links the 2D image plane with the 2D structural surface plane and is applied for the 2D motion estimation [38], [39]. The calibration is based on at least four sets of 2D-to-2D point correspondences [40], i.e. structural coordinates of points in 2D structural surface plane and image coordinates of their projections in 2D image plane.
- Full projection matrix is the general form of transformation metric between the 2D image plane and the 3D structural coordinate system with no assumption, with example applications in [17], [20], [41], [42]. The calibration process usually comprises two steps, i) offline calibration in the laboratory to determine camera intrinsic parameters [43], and ii) site calibration to estimate the camera extrinsic matrix (i.e. camera position and orientation relative to the structural coordinate system) based on at least four sets of point correspondences. The full projection matrix is the multiplication of the camera intrinsic matrix and camera extrinsic matrix.

Scaling factor is inappropriate for site applications due to the prerequisite of camera perpendicular configuration or the required dimensional features. Estimation of either planar homography matrix or full projection matrix requires some known geometric information in the structure which is commonly acquired with the assistance of some artificial targets e.g. planar targets [42], [44] and a 3D calibration object [17]. Since the offline camera calibration step in the full projection matrix method could consider

lens distortion that is common in consumer-grade cameras, the full projection matrix method is used in the video processing package.

In the package, offline camera calibration is performed in the laboratory using the camera to observe a chessboard target in different views in order to obtain the camera intrinsic matrix and the lens distortion parameters. For the site calibration, the camera extrinsic matrix is derived based on at least four sets of 2D-to-3D point correspondences. Since completely non-contact sensing is preferred, the required geometrical information is acquired from the as-built drawings e.g. the bridge span length and the pylon height. To consider the lens distortion, instead of correcting the full frame before the target tracking step, the correction occurs after the target tracking step only to the image coordinates observed from the raw frame in order to save computation efforts. Finally, the 2D structural displacement along the vertical and longitudinal directions is derived based on the corrected image coordinates and the full projection matrix.

2.2 Target tracking techniques

Target tracking is aimed at determining target locations in frame sequences of a video record with several techniques available:

- Correlation-based template matching is a classic and widely-used technique [8], [9], [13], [14], [16] which is realised by searching for an area in a new frame most closely resembling the reference (or template) that is pre-defined as a rectangular subset in the initial frame.
- Optical flow estimation is an established method which detects motions or flows of each pixel within the pre-determined target region based on one temporal and one spatial constraints [45]. Lucas and Kanade optical flow estimation [46] has been validated in a laboratory test of a multi-storey metal tower for system identification [18] and applied in field monitoring of bridge stay cables during normal operation [19].
- Feature point matching is an efficient approach that detects key-points in two images independently and then finds point correspondences based on their local appearance. Currently the applications in structural monitoring are limited, two examples being displacement monitoring test in a stadium structure using FREAK matching [26] and in a viaduct system using SIFT matching [25].
- Shape-based tracking is used to match special target shapes and patterns between two images, i.e. line-type target [34] or custom-made targets with white and black squares [47], [48]. They do not have generality for all target patterns.

Target tracking methods for the deck and cable targets are chosen separately considering their pattern features.

2.2.1 Tracking deck targets

Correlation-based template matching and feature point matching are the two potential approaches for tracking the deck target regions. Correlation-based template matching has been applied for structural displacement monitoring on a railway bridge [9], a long-span bridge [16] and a high-rise building for tracking specific patterns [13,14,44], LED lamp targets [16] and feature targets on structural surfaces [9]. One critical advantage of this method is the minimal user intervention, limited to specifying the

template region in the reference frame. However, the method is sensitive to lighting changes [24,49] and changes of background conditions [50]. Also, the method is not the ideal choice for tracking slender structural components, since a rectangular template might include background pixels that move differently from the structural elements.

Feature point matching is an alternative for the target tracking based on the key-point detection and matching. Key-points in computer vision are those that are stable, distinctive and invariant to image transformation like building corners, connection bolts or other patches with interesting shapes [51]. Instead of the raw image intensities, a feature descriptor is used for matching that is a complex representation based on the shape and appearance of a small window around the key-point. Thus this technique is less sensitive to illumination change, shape change and scale variation. However, feature point matching requires the target region to have rich textures for saliency during the whole recording period. Also several threshold parameters need to be specified according to users' experience or judgement, e.g. contrast threshold for the feature detector and distance threshold in the matching criteria. These threshold values might depend on environmental conditions, e.g. the threshold for outlier removal varies with the illumination condition [26]. The existing applications are mainly focused on the short-range measurement [25,26,50,52] while the feasibility for long-range monitoring and the stability over several hours are not validated yet.

For the studied footbridge, natural features near the bridge deck along the length direction are available for tracking, but these features are not very distinctive. The monitoring was continued over several hours, recording the different occupation states of the bridge. Thus the automatic tracking with little user adjustment under varied environmental conditions was preferred. With these considerations, correlation-based template matching is used for deck target tracking in the video processing package.

The tracking process is given in Figure 3. A target region is selected as the template that is a subset image in the reference frame. A correlation criterion is defined to evaluate the similarity degree between the template and the new frame. Zero-mean normalised cross correlation coefficient (ZNCC) is used as the correlation criterion which offers robust noise-proof performance and is insensitive to offset and linear scale in illumination [53]. The target location in the new frame corresponds to the peak location in the ZNCC matrix that has the resolution at pixel level. Subpixel interpolation schemes [9] are required to refine the tracking results. The interpolation method used in this study is zero-padding in frequency domain using matrix multiplication form of discrete Fourier transform [54].

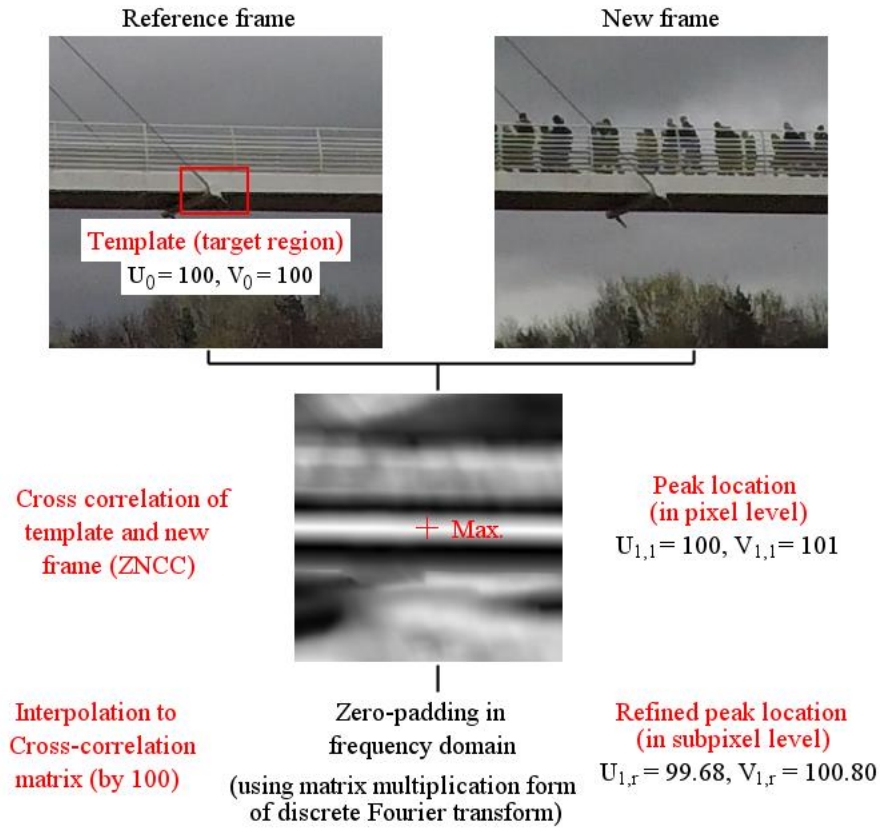


Figure 3 Demonstration of correlation-based template matching for target tracking

2.2.2 Tracking cable targets

To enable a wider field of view covering the majority of the bridge, bridge cables are projected to be slender lines in a camera image e.g. with less than four pixels along the width direction. Correlation-based template matching is inappropriate in this case since pixels within a selected template (a rectangular subset from the reference frame) might cover cable segments as well as some background (e.g. clouds and tree branches) with inconsistent motions. Optical flow estimation method faces challenges due to the limited numbers of salient feature points. The cable tracking method based on edge detection is more robust to the variations of local features and is used in the video processing package.

The cable tracking consists of two steps; edge detection and motion estimation. Edge detection is aimed at determining the cable location in a small subset window while the cable motion is estimated from the distance between two extracted edges.

In the edge detection step, a region of interest including a small cable segment is selected for tracking shown in Figure 4(a). Since edge points have significant local changes in image intensity that lead to a local peak in the first derivative, image gradient is a common measure for edge detection. One of the gradient-based edge detectors, Sobel operator [55] is used to detect the probable edge points (in pixel level) through calculating the image gradients among 3 by 3 neighbourhood and thresholding the magnitude of gradients. Zernike moment operator [56], [57] is then applied to re-located the edge

precisely from the points detected by Sobel operator in Figure 4(b). Zernike moments are constructed by mapping the image onto a set of complex polynomials through convolving the image intensity matrix with three pre-determined masks. Three edge parameters (i.e. step height, perpendicular distance from the mask centre and the edge direction with respect to one image axis) are estimated from Zernike moments for the probable edge points and the edge parameters are then used as criteria to remove outliers and to refine image coordinates for the remaining edge points (in subpixel level). The direction of cable segment within the selected image region is determined by line fitting among the remaining edge points in Figure 4(d).

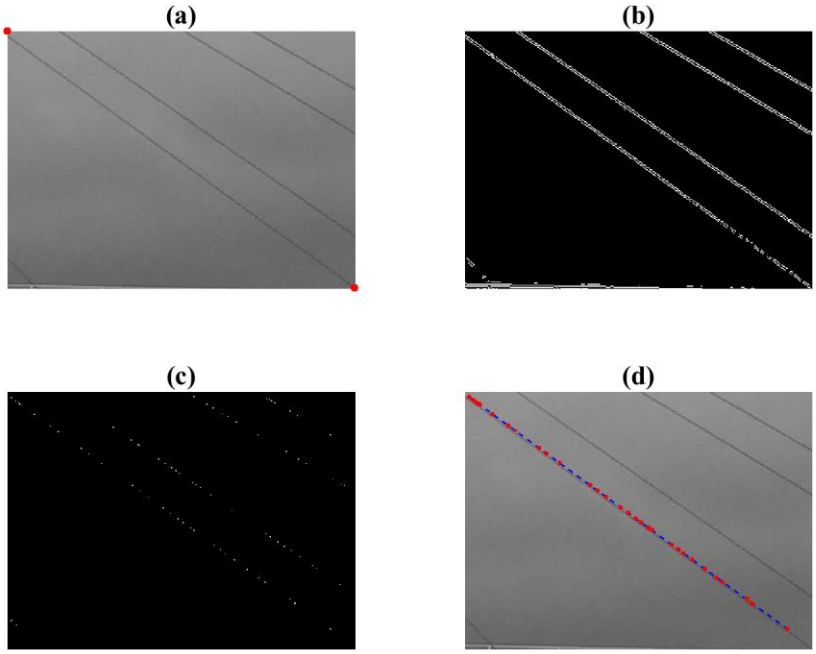


Figure 4 Edge detection procedures for cable targets in the video processing package: (a) the selected cable element located at the diagonal of the cropped frame; (b) the edge detection result by Sobel operator; (c) the refined edge points by Zernike moment operator; and (c) line fitting results of one cable edge.

The cable motion is then estimated from the distance between two edge lines along one assumed motion direction (see Figure 5). Even if the assumed motion direction deviates from the true direction, the motion estimate is proportional, so not affecting identification of cable modal frequencies.

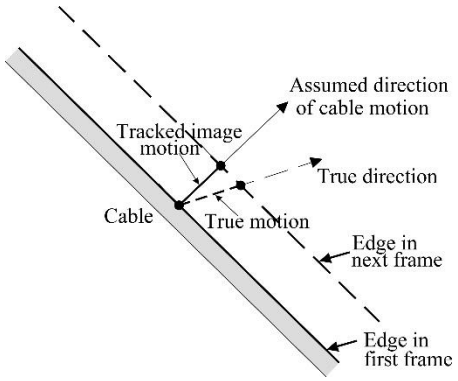


Figure 5 Cable motion estimated from edge shift

2.3 Monitoring data interpretation

The acquired displacement data from the video processing package could be used for data interpretation e.g. extracting the structural dynamic properties. Since the monitored structure in this study is a cable stayed footbridge loaded by crowds of passing pedestrians, the modal properties might vary with the loading.

For system identification, Welch's method [58] was used to estimate the power spectral densities of monitoring data by computing the average of periodograms, and the modal frequencies were estimated through peak picking. The data-driven stochastic subspace identification (SSI) method [59] was also used to extract the modal frequencies and mode shapes through estimating a state-space model from measurement data and performing eigenvalue decomposition to the state-space model.

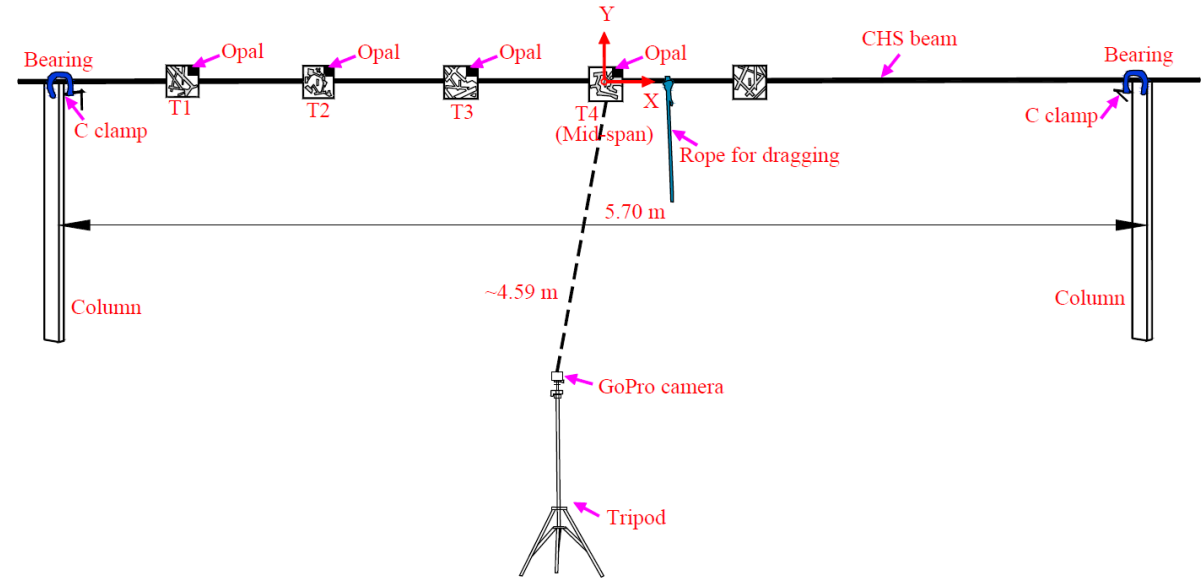
In time-frequency analysis, the continuous wavelet transform (CWT) was used to acquire the time-frequency distribution of measured signals including the displacement and acceleration responses. The complex Morlet wavelet was set as the mother wavelet with the relevant parameters (frequency bandwidth and central frequency) tuned according to the minimisation of the Shannon wavelet entropy [60–62]. The instantaneous frequencies were extracted from the ridges of wavelet transform modulus using the modulus maxima method [63].

3 VALIDATION TEST IN LABORATORY

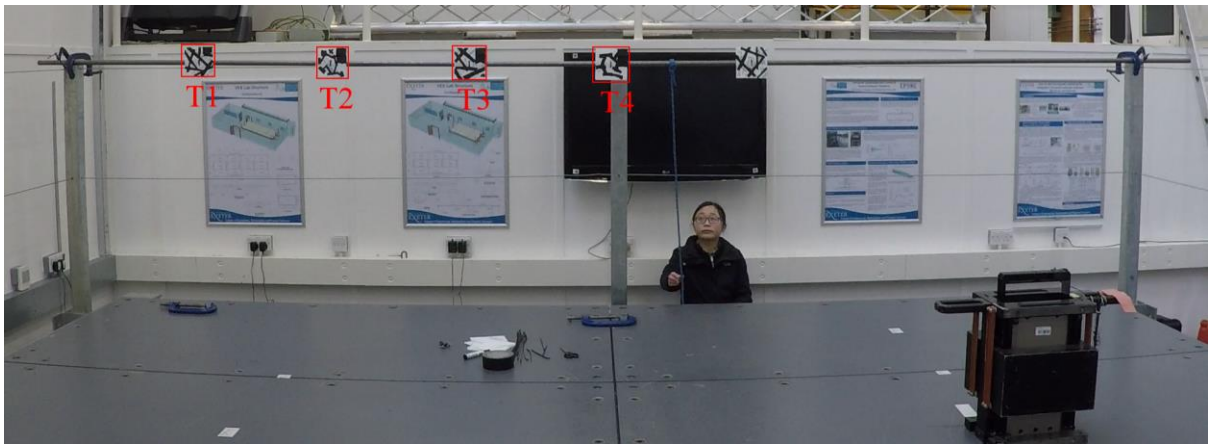
The proposed vision-based system was first validated on a beam structure in controlled laboratory conditions. Since the cable tracking method based on the edge detection has been validated in the previous work [24], the focus of this laboratory test is to investigate the working performance of correlation-based template matching method for structural displacement measurement. The system was applied to measure the displacement responses of several points in the beam when repeatedly set into free vibration. The measured data during the stationary periods were used to evaluate the measurement accuracy, while the measured data under the excitation periods were evaluated by comparison with the accelerometer measurement. Section 3.1 describes the tested beam structure and sensors used and section 3.2 evaluates the accuracy of the displacement data measured by vision-based system.

3.1 Description of a beam structure and the monitoring test

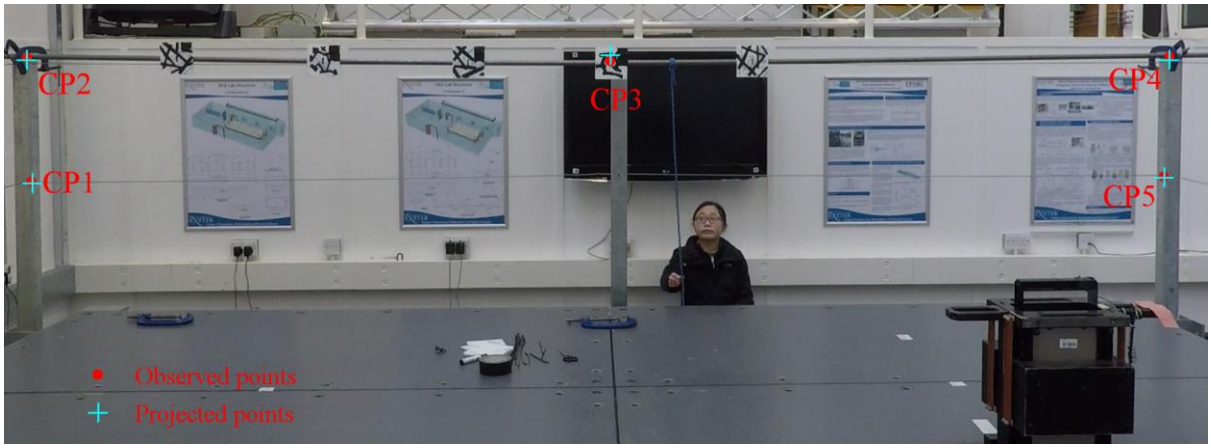
A simply supported beam structure was created in the Structures laboratory of University of Exeter by mounting a steel circular hollow section (tube) on the top of two columns and holding in place with C-clamps. The beam structure, with a span of 5.70 m is shown in Figure 6.



(a)



(b)



(c)

Figure 6 A beam structure mounted on the top of two columns in the laboratory with sensors installed for vibration monitoring: (a) configuration of the beam structure including sensor locations; (b) sample frame (with the lens distortion influence) from recorded videos by GoPro camera with the marked locations of tracking targets; and (c) corrected image of the sample frame in (b) (after removing lens distortion influence) with the marked locations of control points for camera calibration.

To generate vertical free vibration the beam was repeatedly pulled down with a rope and released, and the free vibration response was monitored by a combination of vision-based system and four wireless accelerometers.

In the vision-based system, a GoPro Hero 4 Black camera was mounted on the top of a tripod approximately 4.59 m from beam mid-span. The entire beam was in the field of view with one sample frame indicated in Figure 6(b) showing an obvious lens distortion effect in the four corner regions of the captured frame. The nominal sample rate was set as 60 Hz while the actual rate was 59.94 Hz. Narrow field of view setting was selected with the corresponding focal length equivalent to 30-34 mm. The image dimensions was 1920 pixels × 1080 pixels.

The video processing consists of the three main steps to extract the time histories of beam displacement, camera calibration, target tracking and displacement calculation. In terms of camera calibration, camera intrinsic matrix and lens distortion parameters were pre-determined by analysing the chessboard images taken from different views in the laboratory. The lens distortion parameters were used to correct the lens distortion influence with the corrected frame shown in Figure 6(c). Camera extrinsic matrix was determined based on five pairs of point correspondences between the structural coordinate system and the image plane. The structural coordinate system was defined as the origin at the mid-span of the beam with the X axis along the beam span direction and the Y axis in the vertical direction (See Figure 6(a)). The control points (CP) with known structural coordinates used for calibration are marked in Figure 6(c) as red dots. The camera extrinsic matrix was derived by minimising the total re-projection error between the observed image points and the calculated projected image points given the estimated projection relation. The re-projected image points according to the estimated camera extrinsic matrix and the structural coordinates are indicated in Figure 6(c) as the '+' markers in light colour. The coordinate information and re-projection errors are given in Table 1. An obvious deviation between the observed and the projected image points (re-projection error = 10.4 pixel) occurs at control point CP3 that is the mid-span point of the beam. This deviation might be caused by the error in the provided structural coordinates of CP3 since the initial deformation of the beam induced by self-weight was not considered.

In the second step of target tracking, four targets (T1~4) were chosen for tracking along the span direction located at 1/8, 1/4, 3/8 and 1/2 span points of the beam, respectively. When using the correlation-based template matching method, a planar area with a proper projection size in the video frame (e.g. 40 pixels) is required as the target region. This could be easily satisfied in the bridge

monitoring test, for example using the deck area in Baker Bridge. However, the required planar area was not available in the structural surface of this simply-supported beam due to the long, thin shape. Thus several planar boards with black random patterns were attached to the beam for selection as regions of interest, shown in Figure 6(a).

In the last step, the structural displacement along the vertical and longitude directions was estimated based on the camera calibration and target tracking results.

Reference sensors were required for evaluating the measurement by the vision-based system. Conventional sensors for displacement measurement like LVDT and dial gauges would not work because the target regions on the beam structure were over 2 m higher than the stationary base (the ground). Integration schemes from accelerometer data have some limitations due to low frequency noise and attempts to filter it out that result in loss of quasi-static components. However, for a short signal, e.g. a few seconds, it is possible to derive the reliable displacement information from accelerometer measurement [64]. With the consideration of feasibility and installation effort, wireless accelerometers were chosen as the reference in this laboratory test. Four tri-axial APDM Opal^{TM} wireless sensors were attached to the beam to record the acceleration responses. The sensors were fixed at the top right of target plate using black tape, as indicated in Figure 6(a). The Opal sample rate was set to 128 Hz.

The GoPro camera and the wireless sensors had independent clocks but were separately synchronised with an online reference time. Before comparing the signals, the small time shift (i.e. 0.004 s) between the two sensing systems clocks was corrected by finding the maximum of cross correlation of two velocity signals respectively derived from acceleration data and displacement, both resampled at 256 Hz.

3.2 Measurement and analysis results

Four points (T1~4) along the half span of the beam structure were tracked by the vision-based system with the measured displacement in vertical direction shown in Figure 7. Free vibration was induced three times by pulling and releasing at mid-span and in each case vibrations decayed within 5 seconds. The third free vibration response is the strongest and clearest and is replotted in Figure 7(b) with expanded timescale for a clearer visualisation. The maximum deformation at the mid-span (T4) reaches 5.93 mm at the time T = 40.5 s with the corresponding deflection at 1/8, 1/4 and 3/8 points (T1/2/3) at 1.20 mm, 3.09 mm, and 4.85 mm, respectively. Vibrations decayed to less than 0.3 mm within 2 s and showed a modal frequency of approximately 4.5 Hz.

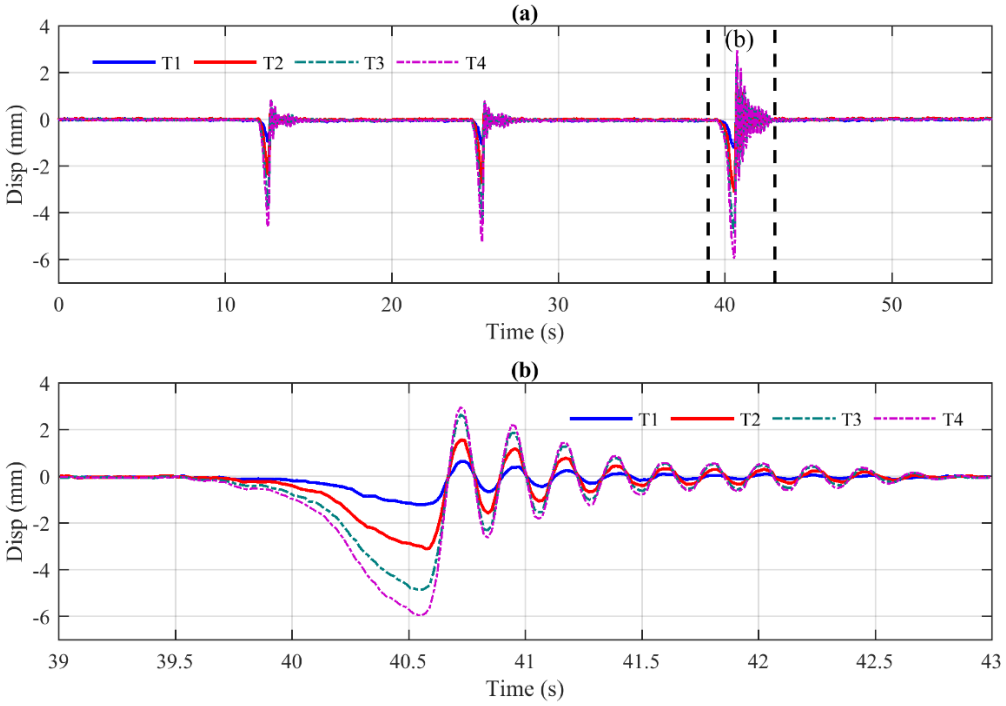


Figure 7 Vertical displacement of four targets along the span direction measured by vision-based system: (a) time history signals within a 55-s duration recording three excitations; and (b) zoom-in view of time history signals during the third excitation.

The displacement measurement shown in Figure 7(a) include data during several stationary periods (i.e. the time intervals of [0, 12] s, [17, 24] s, [31, 38] s and [46, 55] s). The data samples collected during these periods were used to evaluate the measurement accuracy of the vision-based system. The non-zero measured data is regarded as the measurement error since the true value of displacement is zero. The estimated distribution of measurement error is shown in Figure 8 indicating standard deviations of measurement error at the four targets (T1~4) to be almost identical, varying from 0.018 mm to 0.019 mm. The measurement accuracy with 95% confidence interval was estimated by the standard deviation multiplying a critical value determined from the T distribution (i.e. ±1.96). Thus the measurement accuracy during the stationary period was ±0.037 mm. For dynamic displacement data, the measurement accuracy might be decreased because the possible deviation in estimated projection transformation influences more on the measured displacement with larger amplitudes.

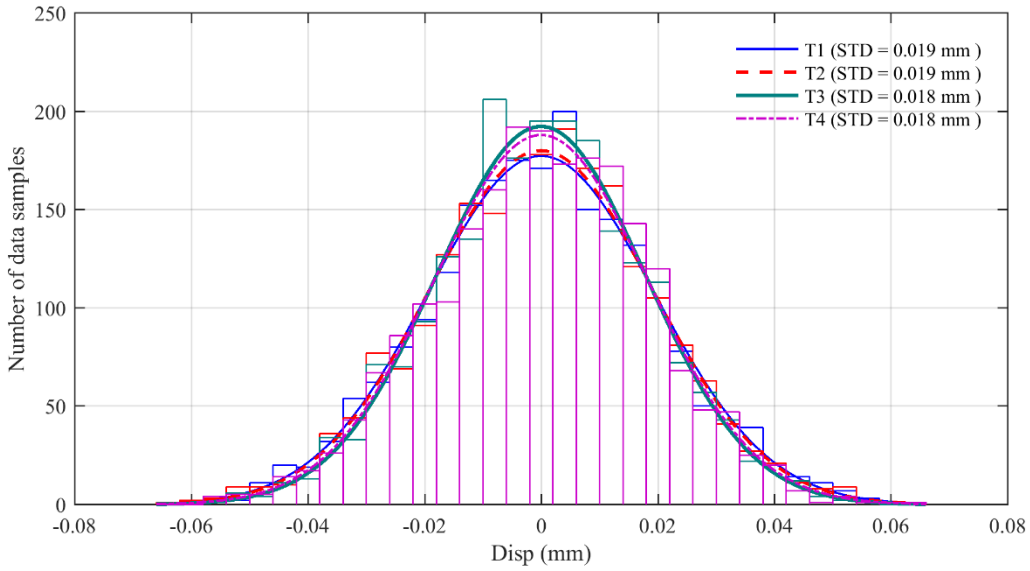


Figure 8 Distribution of error in measured displacement in vertical direction collected during the stationary periods corresponding to four time intervals in Figure 7(a) (i.e. [0, 12] s, [17, 24] s, [31, 38] s and [46, 55] s)

Accelerometers were used as reference sensors to evaluate the measurement accuracy of vision-based system because their acceleration resolution at 4.5 Hz, limited by noise of 128 μg/√Hz translates to velocity resolution of 0.044 mm/s and displacement resolution of 0.0016 mm in the band 4.5±0.5 Hz. Ideally, the displacement data could be directly recovered from the accelerometer measurement through double integration. However, to mitigate the amplification and accumulation of acceleration error during the integration procedures, the accelerometer measurement was integrated to velocity response that was then compared with velocity derived from vision-based measurement.

Acceleration response at the beam mid-span (T4) during the third period of free vibration was truncated for comparison with displacement data for the corresponding period shown in Figure 7(b). Figure 9(a) indicates the measured acceleration and displacement from Opal accelerometer and vision-based system. The derived velocity results are shown in Figure 9(b) indicating high similarity with 98.86% cross correlation coefficient, while the normalised root mean square deviation between two velocity signals is 3.70 mm/s compared to the maximum amplitude value of velocity response at 95.89 mm/s.

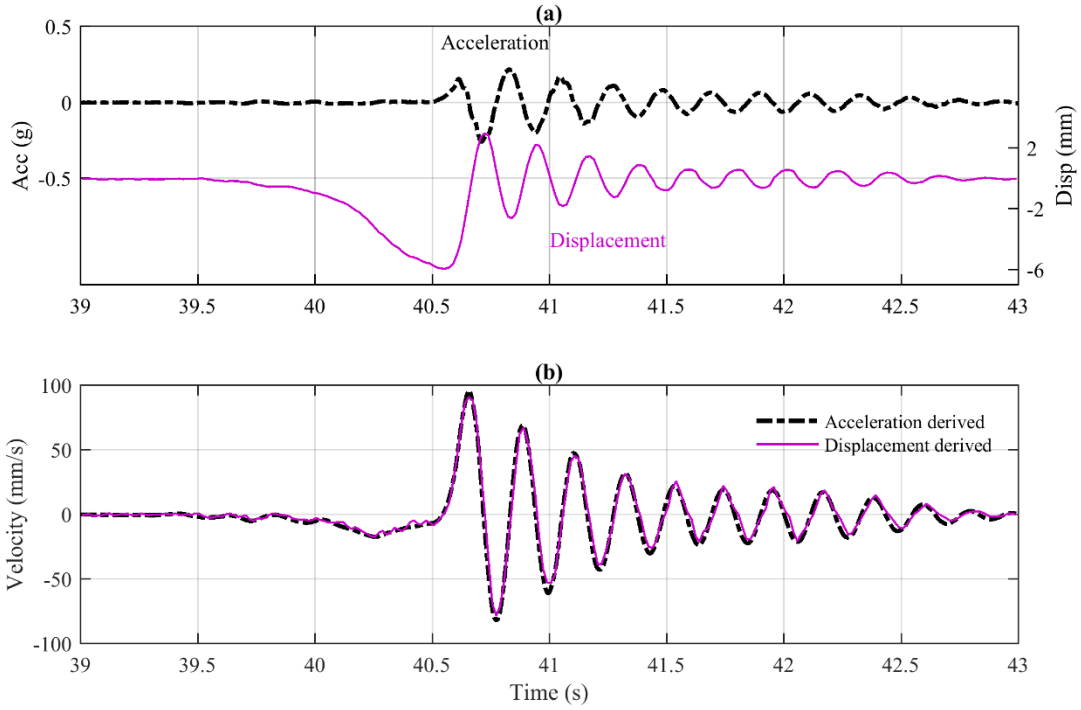


Figure 9 Velocity responses of the beam structure at mid-span (T4) including the third free vibration period derived from accelerometer and vision-based measurement: (a) time history signals of acceleration and displacement by Opal accelerometer and vision-based system; (b) time history signals of velocity derived from the acceleration integration and displacement differentiation.

Having demonstrated the reliability of the vision-based system for displacement measurement in the laboratory, the proposed system was applied in a monitoring test of a full-scale footbridge having dominant vibration modes with frequencies below 2.5 Hz. The low frequencies make it more challenging to recover the deflection information from accelerometer measurement. Therefore, the vision-based system has the advantage in quantifying the quasi-static deflection under heavy loads and also provides the capacity for the evaluation of bridge dynamic performance.

4 FIELD TEST ON A CABLE-STAYED FOOTBRIDGE

The vision-based system was applied in a monitoring test of a cable-stayed footbridge, Baker Bridge in Exeter, UK. This section described both the bridge and the configuration of vision-based system on site.

4.1 Bridge description

Baker Bridge is a 109 m cable-stayed footbridge crossing the A379 dual-carriageway in Exeter, UK (See Figure 10). The bridge provides cyclist and pedestrian access to Sandy Park Stadium (south side of bridge), the home ground of Exeter Chiefs Rugby Club, and thus experiences heavy pedestrian traffic on match days. The bridge comprises a single A-shaped tower that supports the continuous steel deck over a simple support at the pylon cross-beam as well as via seven pairs of stay cables.

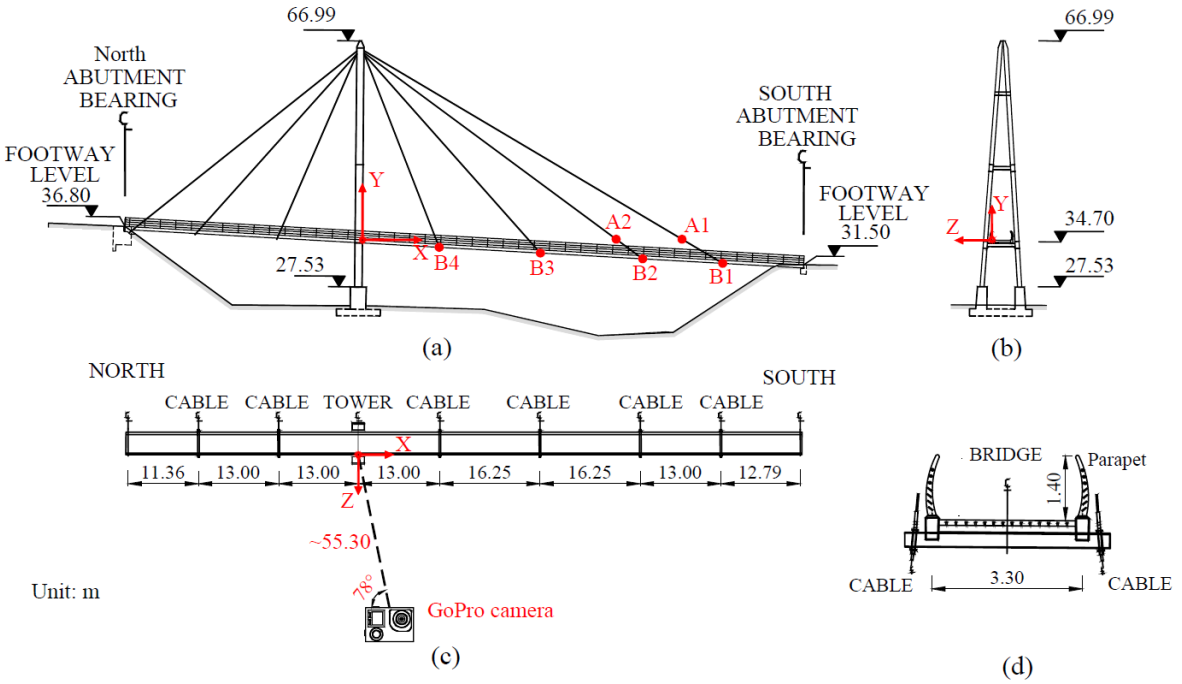


Figure 10 Baker Bridge information and the sensor locations: (a) the west elevation of the bridge with the marked locations of six wireless accelerometer sensors APDM Opal™ (B1~4, A1~2); (b) the tower elevation of the bridge; (c) the deck plan of the bridge with the marked location of the GoPro camera in the west side; and (d) a typical deck section of the bridge.

In a previous ambient modal test [66], four modal frequencies below 2.5 Hz were observed in the vertical direction, i.e. 0.94 Hz, 1.62 Hz, 2.0 Hz and 2.24 Hz. Thus, the bridge has noticeable vibration response due to pedestrian traffic.

4.2 Description of a monitoring test

A four-hour monitoring test was performed from 13:37 to 17:42 on the afternoon of the Exeter Chiefs vs. Newcastle Falcons fixture on 25th Feb 2017, with a crowd of 10,469 spectators [67], many of them arriving by train (Digby and Sowton Station) and crossing the bridge to reach the stadium. Match kick off was at 15:00 with two halves of 40 minutes each plus stoppages and a 10 minute interval. A GoPro Hero 4 Black camera, Figure 10 (c), was mounted on the top of a tripod at the central reservation of the A379 carriageway below and approx. 55.30 m southwest of the bridge tower. The west side of bridge was included in the field of view with one sample frame indicated in Figure 11(a). The nominal frame rate was set as 30 Hz while the actual one was 29.97 Hz. Narrow field of view setting was selected with the corresponding focal length equivalent to 30-34 mm. The image dimensions was 1920 pixels × 1080 pixels. The weather was overcast with little illumination change during recording.

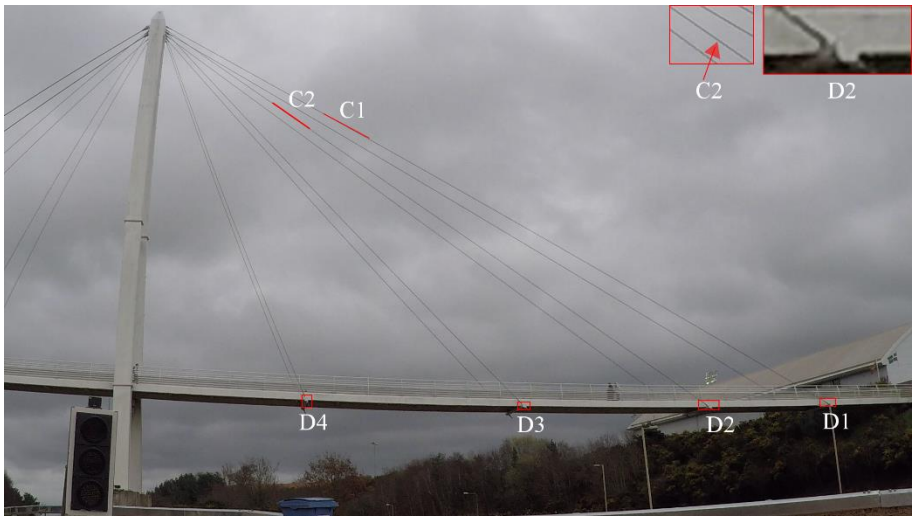
The video processing procedures consist of three main steps similar to those in the laboratory validation test. During the camera calibration, camera intrinsic parameters were determined ahead of the test. A sample corrected frame after removing the influence of lens distortion is shown in Figure 11(b). Camera extrinsic matrix was determined on site based on several pairs of 2D-to-3D point correspondences. The

structural coordinate system was specified with the origin at the deck height of the tower section, the Y axis along the vertical direction and the Z axis along the transverse direction (See Figure 10(a) and (b)). The control points (CP) are marked in Figure 11(b) with known structural coordinates from the as-built drawings provided by the Devon County Council: CP1-4 along the edge of the bridge tower and CP5-11 near the outer-section of the crossbeams to which the cables are secured.

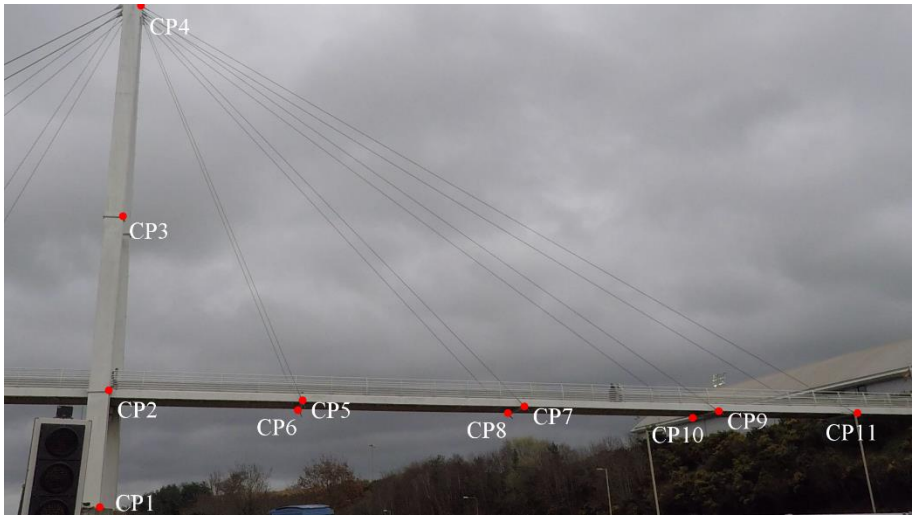
For target tracking, four targets (D1-4) along the deck longitude direction and two targets (C1-2) at the cable edges were chosen for tracking, all at the southwest side of the bridge in Figure 11(a). The pixel dimensions of the selected targets in video frames are indicated in

Table 2. Due to the limited availability of stable features in the bridge deck, the height of these deck targets is approximately 20 pixels, smaller than the suggested value (40 pixels) in the previous study [24].

For deck targets (D1-4), the structural displacement along the longitude and vertical direction was estimated based on camera calibration and target tracking results. For the cable targets (C1-2), the cable motion estimated in the target tracking step was directly outputted.



(a)



(b)

Figure 11 Sample frames from the video records: (a) one raw frame (including apparent lens distortion) with markers on deck and cable targets for video tracking; and (b) one corrected frame (removing lens distortion influence) with markers on control points (CP1~11) used for the calibration of camera extrinsic matrix.

As well as the vision-based system, six tri-axial wireless accelerometers (APDM Opal^{TM}) were installed on the bridge: four (B1~4) on the deck parapet and two (A1~2) on the cables with locations marked in Figure 10(a). The purpose of the Opal sensors was to corroborate the identification of modal parameters

of bridge deck and cables obtained using the vision-based system. The Opal sensors B1~4 corresponded to the target regions D1~4 in the vision-based system while the sensors A1~2 were collocated to the same cable sequences as the target regions C1~2. The sample rate was set to 128 Hz.

5 MEASUREMENT AND ANALYSIS RESULTS

In this section, the measurement results obtained by the vision-based system are illustrated in time and frequency domains. The time interval for analysis from 16:39 to 17:14 (35 min) thus included periods when large crowds of spectators crossed the bridge on the way home after the match. The measured data from the vision-based system was analysed to investigate the dynamic properties of the bridge including the changing modal frequencies under varying pedestrian loads. Section 5.1 and 5.2 demonstrate the measurement and analysis results of bridge deck displacement and cable vibration, respectively.

5.1 Measurement and analysis of deck displacement

The vertical displacements of the four deck targets along the bridge span are described in this section. The measured data are presented in time domain (section 5.1.1) and frequency domain (section 5.1.2), respectively.

5.1.1 Time history measurement of vertical displacement

Four deck targets (D1~4) were tracked with the time histories of vertical displacement shown in Figure 12 and four extracted frames from the video files in Figure 13. During the recording, the bridge changed from almost empty (Figure 13(a)) to almost full (Figure 13(b)), then reverting to a trickle of pedestrians (Figure 13(d)).

In Figure 12, an obvious downward trend of the bridge deck is observed from 800 s to 1250 s in the measured data at D1 and D2 with the maximum deformation value reaching 72.58 mm and 64.10 mm, respectively. A quick deformation recovery is seen at approximately 1300 s from the measurement at D1~D3 that should correspond to a sudden reduction in bridge loading. The captured frame at 1315 s (Figure 13(c)) shows a clear gap (approx. 16.5 m) between two groups of pedestrians, which accords with observations from the measured data.

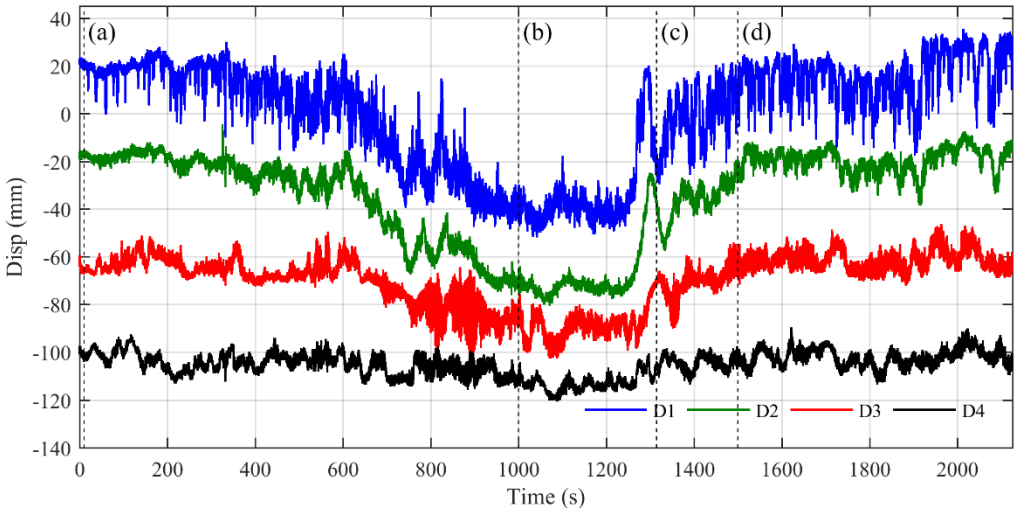


Figure 12 Time histories of vertical displacement at four deck targets D1~D3 by the vision-based system

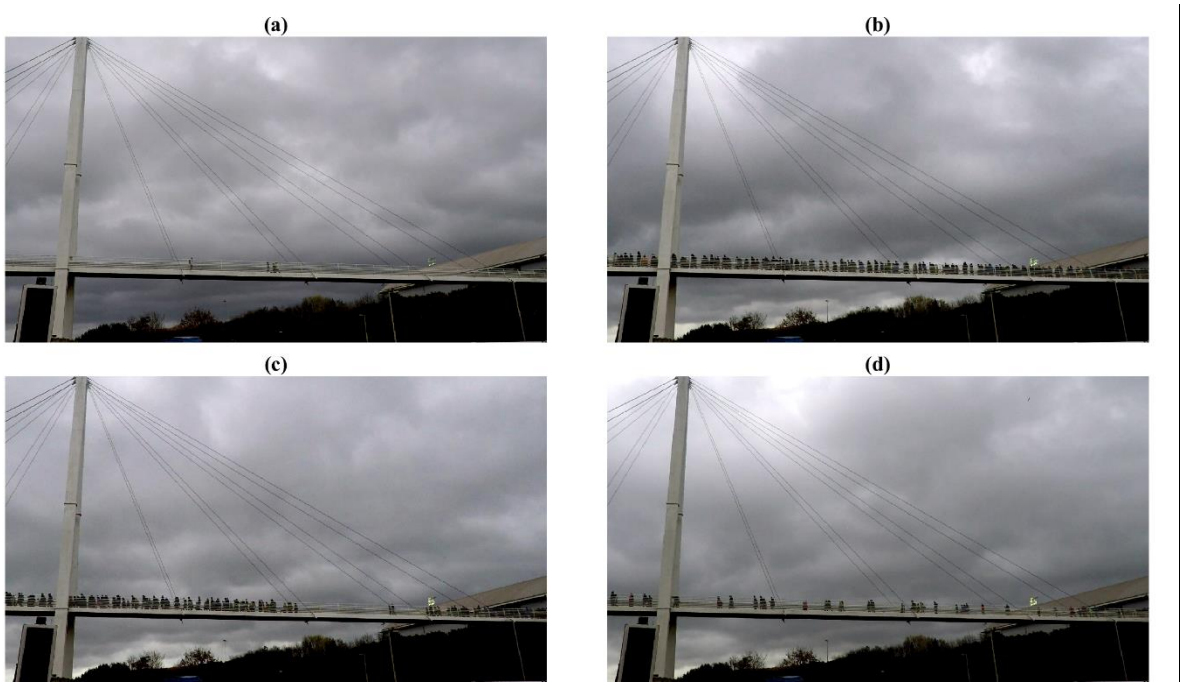


Figure 13 Four captured frames for the recorded video files corresponding to the time step: (a) at 10 s; (b) at 1000 s; (c) at 1315 s and (d) at 1500 s.

Frequent local sharp peaks are observed only in the vertical displacement of D1 in Figure 12. Figure 14(a) zooms into 60 s of signal, and inspection of the video recording shows that the sharp peak of 14.89 mm in the time interval from 34 s to 41 s was induced by the passing of two small groups of pedestrians from the opposite directions shown in Figure 14(b) and (c).

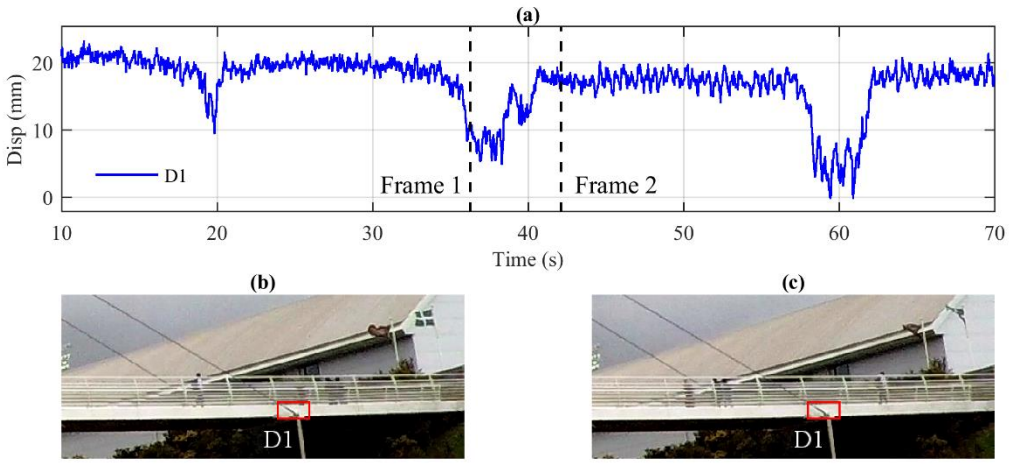


Figure 14 A 60-s time history signal of deck displacement in vertical direction at the target D1 with passing pedestrians: (a) time history of vertical displacement at D1; (b) Frame 1 at 36 s; and (c) Frame 2 at 42 s.

5.1.2 Frequency components of vertical displacement

The power spectral densities (PSD) of vertical displacement and acceleration measurement were estimated using Welch's method with the window length of one minute duration and a 50% overlap. Figure 15 illustrates the estimation results for the signals recorded during three time intervals i.e. ([0, 400] s, [800, 1200] s and [1600, 2000] s). The first time interval ([0, 400] s) was at the end of the match (during stoppage time) and thus few pedestrians crossed the bridge; the second duration ([800, 1200] s) was after the Rugby match and the bridge was almost fully occupied by pedestrians; and the third time interval was after most spectators had left and still a few pedestrians were crossing the bridge. Acceleration data of the deck at B3 was not available due to a faulty battery.

In Figure 15(a), four apparent modal frequencies are identified using peak-picking with the values of 0.92 Hz, 1.61 Hz, 2.00 Hz and 2.23 Hz, which match well with the results from acceleration measurement in (b). The displacement measurement at the deck point D1 contains significant quasi-static response due to the local deformations resulting from passing pedestrians (See Figure 14), preventing identification of the first modal frequency at 0.92 Hz.

In the second time interval shown in (c), only the second modal frequency is clearly indicated, with the value decreased to 1.48 Hz (from 1.61 Hz). The signal power near this frequency value is increased sharply compared with the data in the other two periods. The shift of the second modal frequency is also observed in (d) from acceleration data.

In the third time interval shown in (e), the second mode still contains the highest power with frequency value shifted back to 1.59 Hz. The first and third modal frequencies are identified with the same values (0.92 Hz and 2.00 Hz) as in the first time interval. The observations match well with the analysis results of the acceleration data shown in (f).

Through the analysis, it indicates that:

- The measured data by vision-based system captures the modal frequencies of the bridge deck accurately through the comparison with the acceleration data;
- The second mode of the bridge deck is very sensitive to the occupation status of the bridge and the frequency value reduced from 1.61 Hz to 1.48 Hz with full pedestrian occupancy, corresponding to a reduction of 8%.

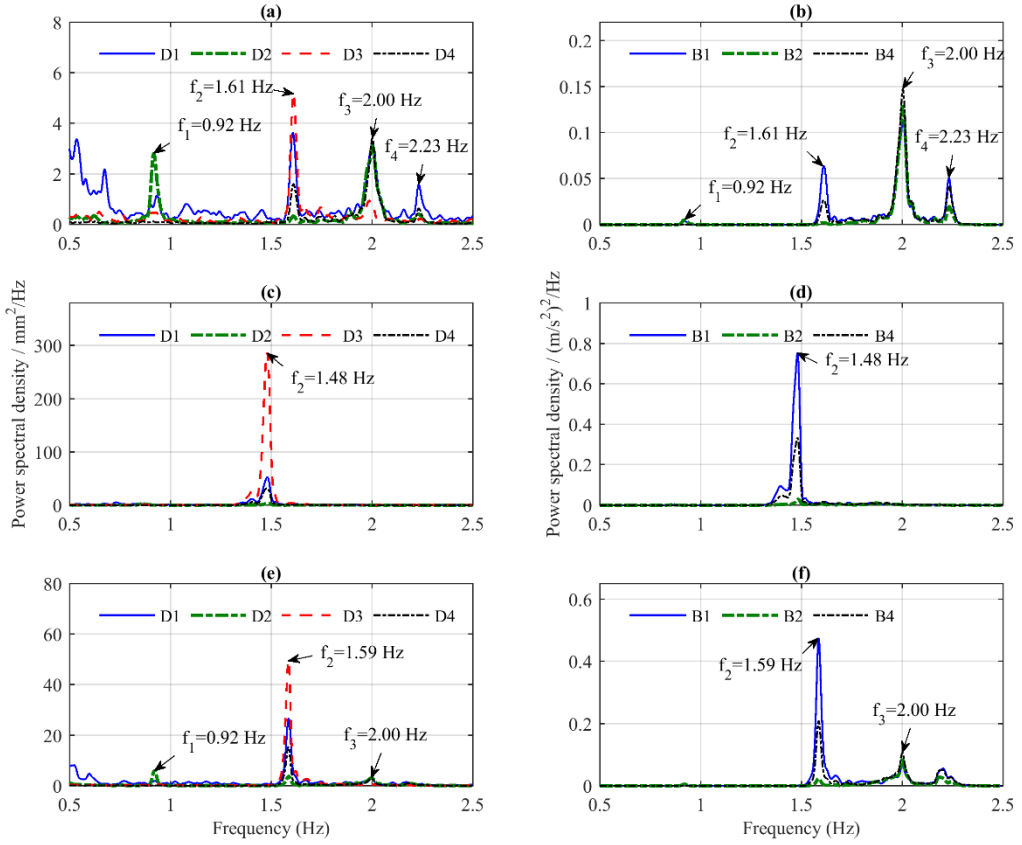


Figure 15 Power spectral densities (PSD) of time history signals with the three rows representing the three time intervals, the left column representing the PSD of displacement measured by vision-based system and the right column representing the PSD of acceleration measured by Opals: (a) PSD of vertical displacement at the time range of [0, 400] s; (b) PSD of vertical acceleration at the time range of time range of [0, 400] s; (c) PSD of vertical displacement at the time range of [800, 1200] s; (d) PSD of vertical acceleration at the time range of [800, 1200] s; (e) PSD of vertical displacement at the time range of [1600, 2000] s; and (c) PSD of vertical acceleration at the time range [1600, 2000] s.

The SSI method [59] was used to identify the modal frequencies and mode shapes from the collected data in the third time interval (i.e. [1600, 2000] s) and the analysis results were compared with those observed from a previous ambient modal test using APDM opal sensors [66]. Figure 16 compares results for the two bending modes in vertical direction:

- The second modal frequency estimated by displacement data is 1.58 Hz, lower than the value (1.62 Hz) reported in [66]. This is due to more frequent crossing pedestrians on the test day.

- The third modal frequency (2.00 Hz) estimated by displacement data matches the value in [66].
- For these two bending modes, the mode shape ordinates (red circular dots) at the points D1~4 predicted by the vision-based measurement match well with the mode shapes (blue curves) previously estimated in [66].

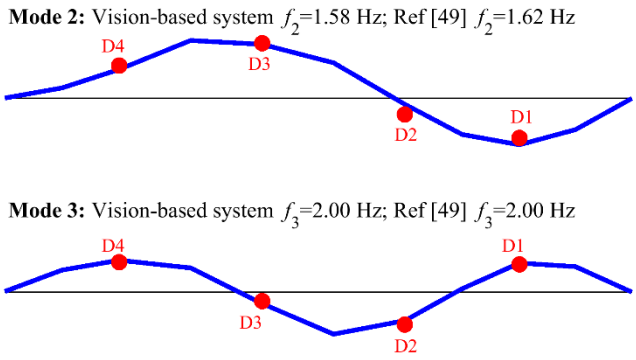


Figure 16 Mode shapes and frequency estimates of the bridge longer span: blue curves represent the mode shapes extracted from the previous ambient modal test [66] corresponding to the longer span closest to the stadium; and red dots represent the mode shapes extracted from displacement data measured by vision-based system.

The frequency responses of measured signals in Figure 15 indicate dependency on time, whose study requires time-frequency analysis rather than methods based on the Fourier transform (e.g. Welch's method) that are designed for the analysis of stationary signals. Continuous wavelet transform (CWT) analysis was therefore used to acquire the time-frequency distribution of displacement and acceleration measurement.

Figure 17(a) indicates the CWT results for displacement measurement at the deck target D1 with the frequency range from 1.3 Hz to 1.8 Hz which covers the variations of the second modal frequency of the bridge deck (varying from 1.48 Hz to 1.61 Hz in Figure 15). During the analysis, the two parameters (frequency bandwidth f_b and central frequency f_c) in the complex Morlet wavelet were tuned according to the minimisation of Shannon wavelet entropy, reaching optimal parameters at f_b =4.5 Hz and f_c = 29 Hz. To consider the edge effect, the influenced region was estimated to be 40 s duration according to [61] and the padding scheme of reflecting the signal at two ends was used to mitigate the edge effect. A threshold (e.g. -0.5 in Figure 17(a)) was set for the wavelet transform modulus value during the plotting for a clear visualisation. The instantaneous frequencies were estimated by the modulus maxima at each time step and are shown as the sparse dots in the figure. The results in Figure 17(a) indicate an obvious variation of the second modal frequency during the recording period.

- In the first 500 s, the modal frequency has small deviation with the value over 1.60 Hz.
- During the time interval from 600 s to 1100 s, a sharp decrease of the modal frequency value is observed with the lowest shifting to approx. 1.37 Hz, a reduction of 15%.
- The data after 1500 s reflects a recovery of frequency value to approx. 1.58 Hz.

- These observations match well with the analysis results for acceleration data (B1) shown in Figure 17(b).

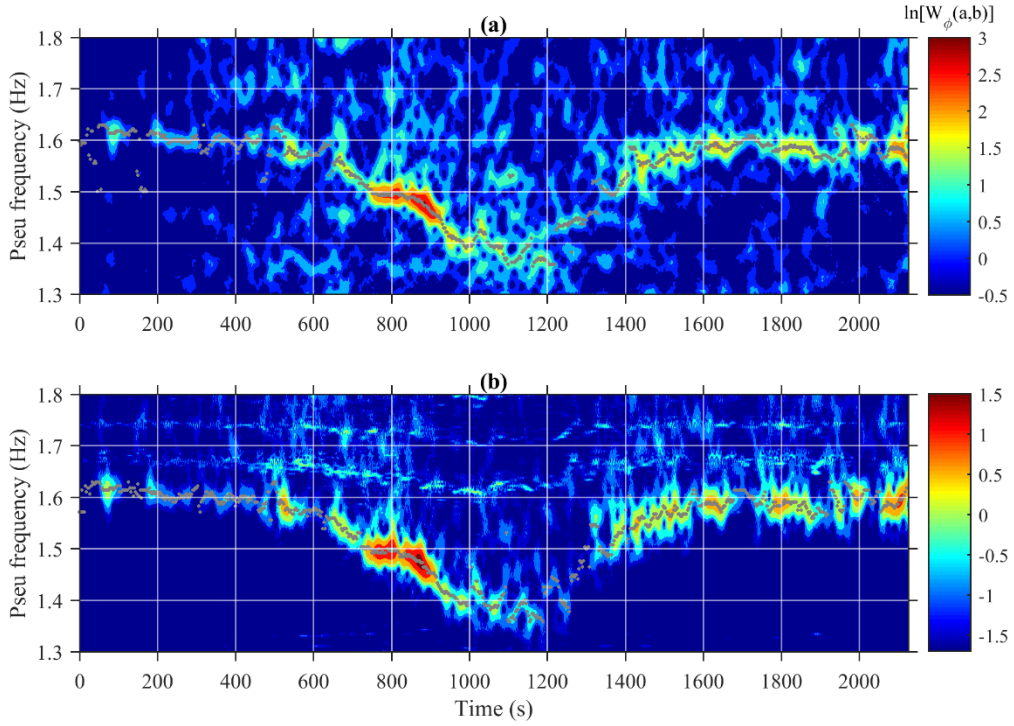


Figure 17 Contour plot of CWT analysis results of displacement measurement (D1) and acceleration measurement (B1): (a) wavelet transform modulus for the measured displacement by vision-based system at the frequency range of [1.3, 1.8] Hz with the estimated instantaneous frequencies marked as sparse dots; and (b) wavelet transform modulus for the measured acceleration at the frequency range of [1.3, 1.8] Hz with the estimated instantaneous frequencies marked as sparse dots.

5.2 Measurement and analysis of cable vibration

This section presents the measurement results of cable vibration using the vision-based system. The measured data were directly used to estimate cable modal frequencies by peak-picking from power spectral densities as well as using SSI. To evaluate the variations of cable modal frequency with changing pedestrian loads, the CWT analysis was performed on the measured data to identify the time-frequency distribution of cable vibration that was compared with the observations from acceleration measurement.

Two cable targets C1 and C2 shown in Figure 11(a) were tracked, with the time histories of cable motion shown in Figure 18(a) and (c). The cable motion here corresponds to the motion of cable projection in the image in pixel units. The power spectral densities of the cable motions during the three time intervals (i.e. ([0, 400] s, [800, 1200] s and [1600, 2000] s) are indicated in Figure 18 (b) and (d) for the cable targets C1 and C2, respectively.

- The modal frequency of the cable C1 is approximately 1.66 Hz for the first and third time intervals and slightly increased for the second time window.

- For the cable C2, the modal frequency could be identified as approximately 2.10 Hz during the first and third time intervals while the analysis result for the data from the second time interval indicates no obvious peak frequency, but rather a frequency range with higher energy near 2.2 Hz.

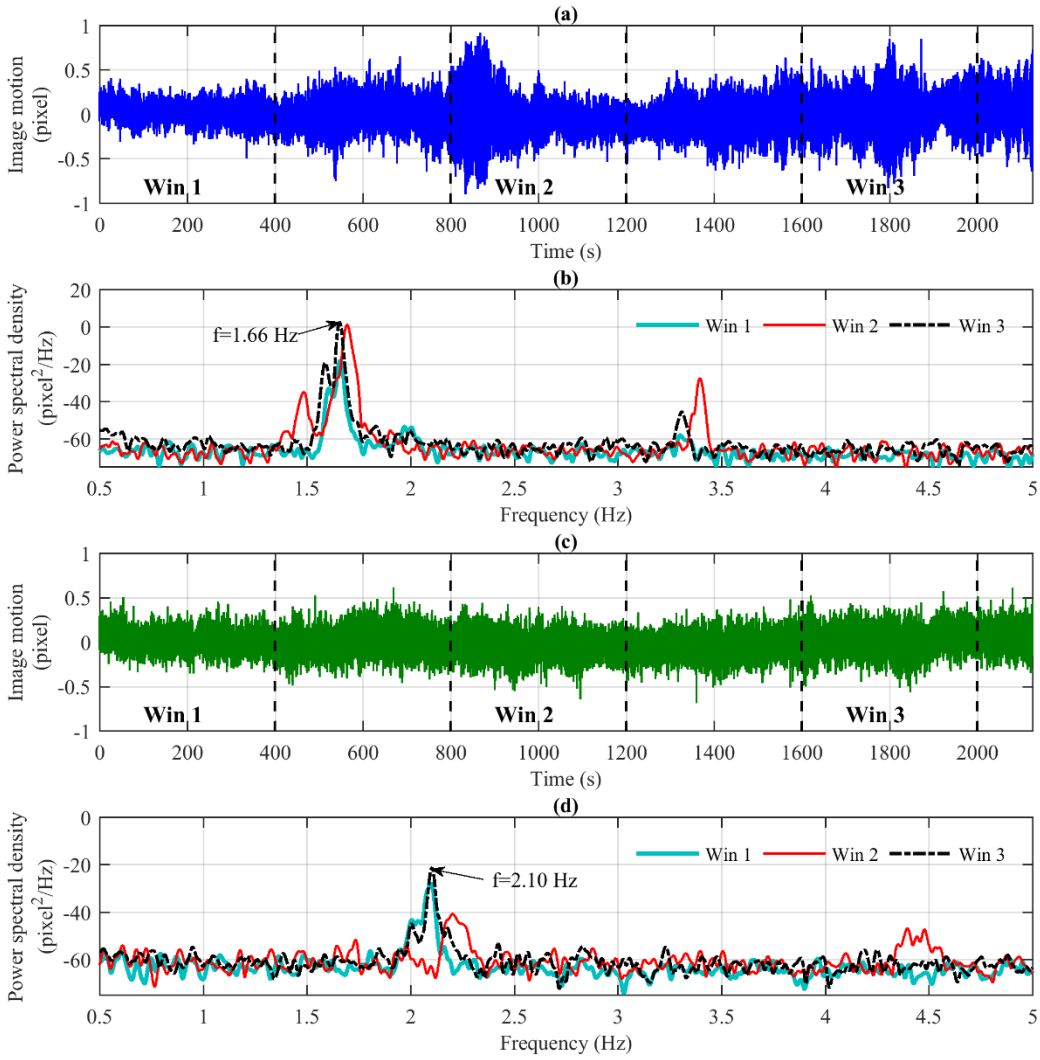


Figure 18 Measurement of cable motions by the vision-based system: (a) image motion at C1; (b) power spectral densities of image motion at C1 in three marked time windows shown in (a) ([0, 400] s, [800, 1200] s and [1600, 2000] s); (c) image motion at C2; and (d) power spectral densities of image motion at C2 in three marked time windows shown in (c) ([0, 400] s, [800, 1200] s and [1600, 2000] s).

The SSI method was used to identify the modal frequencies from the collected displacement and acceleration data during the third time interval (i.e. [1600, 2000] s). Cable target C1 (in Figure 11(a)) and accelerometer A1 (in Figure 10) correspond to the same bridge cable (the longest one, in the southwest side) while the sensor locations were different. C1 was at approximately ¼ span point close

to the bridge tower and A1 was in the lower height close to the cable end where it is attached to the bridge deck. Similarly, C2 and A2 correspond similarly.

The displacement signal at C1 indicates two close modal frequencies at 1.59 Hz and 1.66 Hz while the acceleration signal at A1 captures the first modal frequency at 1.63 Hz as well as the higher modal frequencies at 4.96 Hz, 6.62 Hz and 8.27 Hz. Through comparison, the fundamental frequency of the longest cable (C1) was at approximately 1.66 Hz. The mode at 1.59 Hz identified from the displacement signal might correspond to the second bending mode of the bridge deck. The first modal frequency estimated from the acceleration data is different from the estimated fundamental frequency, which might be due to mixing of frequency responses between the cable fundamental mode and the second bending mode of bridge deck.

The displacement signal at C2 indicates modal frequencies at 2.10 Hz and 4.16 Hz while the acceleration data at A2 capture modal frequencies at 2.12 Hz, 6.30 Hz, 8.36 Hz and 10.48 Hz. Therefore, the fundamental frequency of the second longest cable (C2) was at approximately 2.10 Hz.

The analysis indicates that the vision-based system works better to capture the lower modal frequencies of cables while the accelerometers provide reliable estimations of higher frequency modes.

CWT analysis was performed to acquire the time-frequency distribution of cable vibrations from the vision-based system and accelerometers. The analysis results are shown in Figure 19 and Figure 20 for the cables C1 and C2, respectively. In terms of accelerometer measurement, instead of plotting directly the results near the fundamental frequency of the cable, higher frequency ranges i.e. near the fifth modal frequency for the measurement at A1 and near the third modal frequency for the measurement at A2 are illustrated in Figure 19 (b) and Figure 20 (b), with the corresponding values near the fundamental frequency marked in the right y-axis.

Figure 19 (a) indicates the CWT analysis results for the measurement at the cable C1 by vision-based system. The instantaneous frequencies estimated by the modulus maxima (shown in the figure as sparse dots) initialled at approximately 1.66 Hz, rose to over 1.70 Hz during the time interval from 900 s to 1160 s and then recovered to 1.66 Hz after 1500 s. Since the time interval from 900 s to 1160 s corresponds to the period where the deck points D1 and D2 experienced a large deformation (See Figure 12), the observations indicate that heavy pedestrian loads on the bridge lead to a rise in cable modal frequency by 2.4%, probably by increasing the cable tension. Compared with the analysis result of acceleration measurement shown in Figure 19(b), the time-frequency distribution acquired by vision-based measurement captures the general trend of frequency shift under pedestrian loads over the whole 35 min. However, some details of frequency variation within a short-time range are only identified by acceleration measurement, e.g. a sharp decrease and recovery of cable modal frequency at approx. 1280 s.

As well as the cable C1 modal frequency at approx. 1.66 Hz, another less obvious mode is indicated in Figure 19(a) with frequency value lower than the cable fundamental frequency. This mode is salient in the lighter loading condition e.g. (i) in the time interval from 200 s to 900 s with the modal frequency

decreasing from 1.62 Hz to 1.46 Hz; and (ii) the time interval from 1400 s to 2100 s with the modal frequency increasing from 1.53 Hz to 1.60 Hz. The observed mode shows a similar trend as the variation of the second modal frequency of the bridge deck. Therefore, this mode might be due to forced vibration of the cable by motion of the bridge deck.

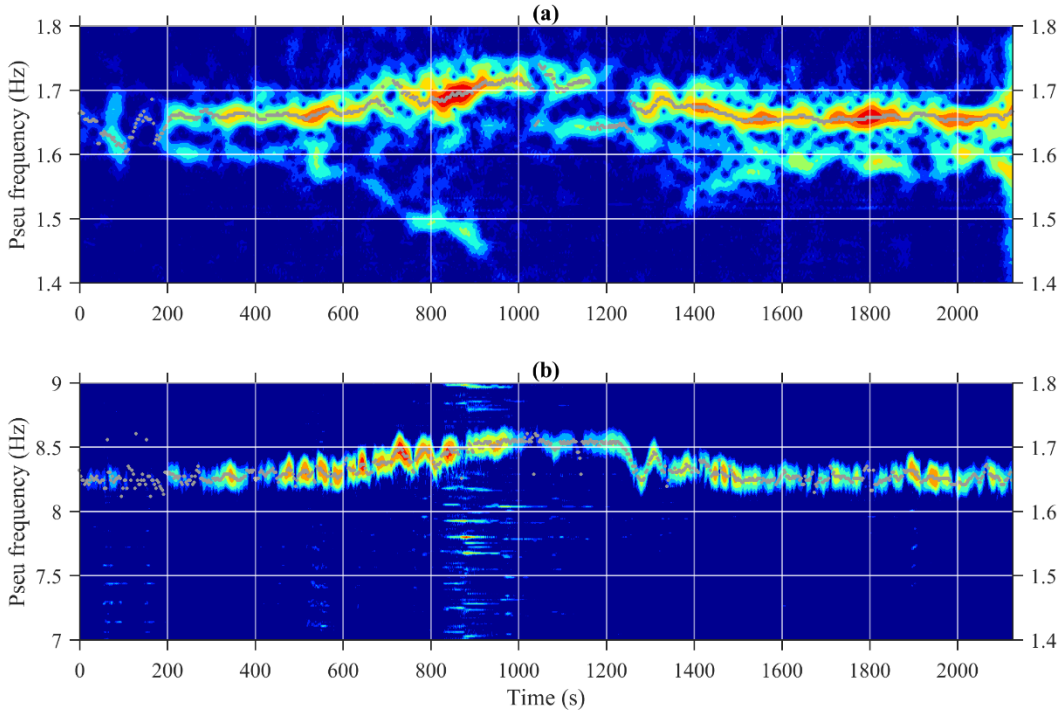


Figure 19 Contour plot of CWT analysis results of cable vibration for the longest cable in the southwest side of the bridge: (a) wavelet transform modulus for the cable motion (C1) measured by vision-based system at the frequency range of [1.4, 1.8] Hz with the estimated instantaneous frequencies marked as sparse dots; and (b) wavelet transform modulus for the cable vibration measured by the accelerometer (A1) at the frequency range of [7, 9] Hz with the estimated instantaneous frequencies marked as sparse dots.

The CWT analysis results for the measurement by vision-based system at the cable C2 are indicated in Figure 20 (a) with the estimated instantaneous frequencies shown as sparse dots. The cable modal frequency started off at approx. 2.08 Hz when the bridge was occupied by only a few pedestrians. An obvious rise of modal frequency (exceeding 2.2 Hz) is observed during the time range from 900 s to 1260 s when the bridge was under heavy pedestrian loads. Compared with the quiet period, the maximum shift of cable modal frequency reaches 9.1% with the frequency value reaching 2.27 Hz. In the period after 1500 s, the modal frequency of cable vibration was recovered to approx. 2.1 Hz. These observations match well with those from Figure 20 (b) corresponding to the acceleration measurement (A2) of the same cable. However, the analysis results of acceleration data illustrate better resolution of the variations of cable modal frequency with time, especially during heavy load periods from 900 s to 1260 s.

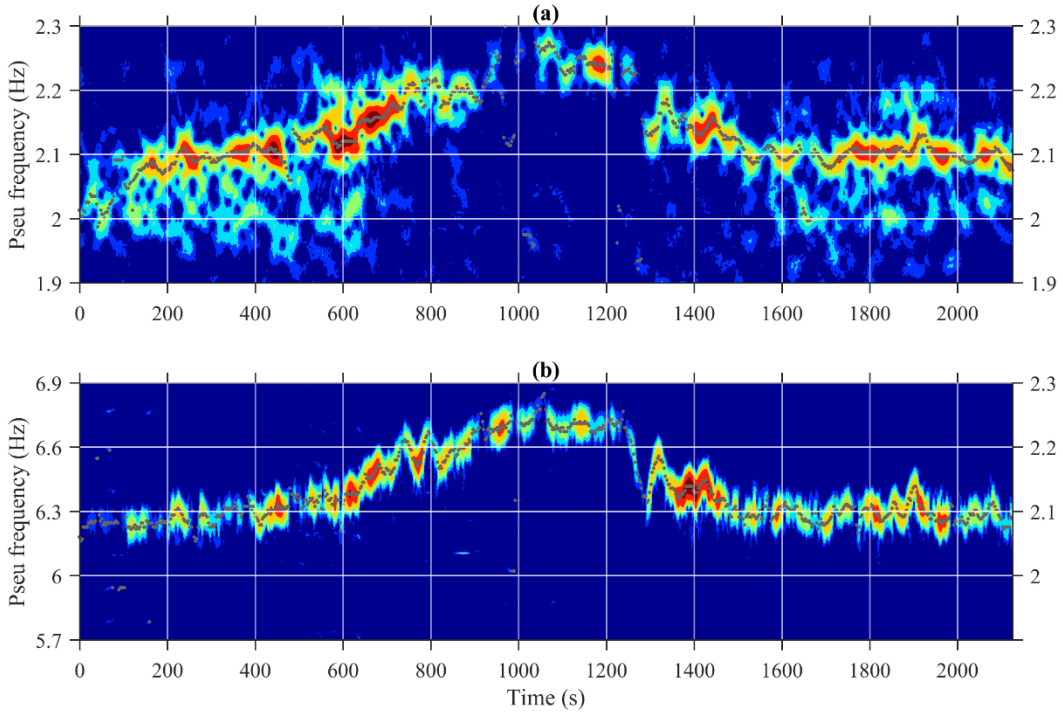


Figure 20 Contour plot of CWT analysis results of cable vibration for the second longest cable in the southwest side of the bridge: (a) wavelet transform modulus for the cable motion (C2) measured by vision-based system at the frequency range of [1.9, 2.3] Hz with the estimated instantaneous frequencies marked as sparse dots; and (b) wavelet transform modulus for the cable vibration measured by the accelerometer (A2) at the frequency range of [7, 9] Hz with the estimated instantaneous frequencies marked as sparse dots.

6 DISCUSSION OF MEASUREMENT ACCURACY OF VISION-BASED SYSTEM

In the field test, the multi-point displacement measurements by the vision-based system were validated to be viable for tracking both deformation induced by passing pedestrians and modal properties of the deck and cables under varying pedestrian loads. Based on this demonstration the procedure would be viable for other (e.g. larger) bridges.

The issue of measurement accuracy of a vision-based system is critical but hard to quantify, especially on site. In the laboratory validation test, the accuracy level was evaluated as measurement of a stationary structure, as well as by comparing measurement using the vision system with accelerometer data, both converted to velocity. However, the measurement accuracy might not be directly comparable to that obtained in other applications such as in the field.

The measurement accuracy of a vision-based system depends on several parameters, e.g. camera-to-target distance [26], estimation of camera intrinsic parameters, dimension information [17] and dispersion of target tracking results in images [68], etc.. Theoretically, displacement measurement using a vision-based system is derived from two parts: i) target tracking results and ii) the transformation metric between the real structure and their projection in image.

- In terms of target tracking, the nominal algorithm resolution can be better than 0.01 pixel with an interpolation scheme while the reported accuracy varies from 0.5 to 0.01 pixel [68]. In this study, the tracking accuracy was quantified to be 0.013 pixel in the laboratory condition while the tracking accuracy in the field test was not evaluated. The ideal image size of the target for correlation-based template matching is suggested to be no less than 40 × 40 pixels [24] to ensure good performance. For field application, determining the camera set-up location should consider the balance between measurement accuracy and the possibility to monitor a large portion of the bridge.
- During the camera calibration in this case, the structural coordinates of control points were derived according to the as-built drawings which might not represent the current condition, e.g. effects of self-weight deflection. Since the calibration process of camera extrinsic matrix is by minimising the total re-projection error between the detected image points and the calculated image projection points based on least-squares optimisation, a better and more stable estimation might be made using more control points.

Other factors on site might influence the measurement accuracy and stability using a vision-based system. For example, Figure 12 shows that the displacement measurement at the deck point D1 did not recover to the initial condition; this might be due to the error caused by camera movement. During the recording, data loss was found due to partial obstruction of the target as well as pattern blur by raindrops. The measurement could also be influenced by atmospheric refraction and turbulence [69].

For a robust sensing system, the measurement accuracy and uncertainty are required for quality assurance and metrological traceability, thus further study is necessary.

7 CONCLUSIONS

A non-contact single-camera vision-based system used for non-contact measurement of bridge displacement provided results comparable to those obtained using an array of wireless accelerometers and offered additional information about quasi-static response to varying pedestrian loads.

In the laboratory validation test, the measurement accuracy of vision-based system was evaluated to be ±0.037 mm under the camera-to-target distance of 5.70 m, but it was not possible to test accuracy directly in the field application, only to compare with another measurement, in this case using the accelerometers.

The multi-point deformation data obtained using the vision system proved to be effective for tracking cable dynamic properties at the same time as bridge deformation, allowing for the effect of varying load on cable tensions to be observed. This provides a powerful diagnostic capability for larger cable-supported structures.

8 ACKNOWLEDGEMENT

We would like to thank Devon County Council for permission to use their bridge and for assistance they provided. Also thanks to Vincent Ao and Jose Capilla for support in the field testing. Finally the authors would like to thank the two anonymous reviewers for their constructive comments.

9 REFERENCES

1. Brownjohn JMW, Koo KY, Scullion A, List DI. Operational deformations in long-span bridges. *Structure and Infrastructure Engineering* 2015; **11**(4): 556–574. DOI: 10.1080/15732479.2014.951857.
2. Carden EP, Fanning P. Vibration Based Condition Monitoring: A Review. *Structural Health Monitoring: An International Journal* 2004; **3**(4): 355–377. DOI: 10.1177/1475921704047500.
3. Brownjohn JMW, Moyo P, Omenzetter P, Lu Y. Assessment of highway bridge upgrading by dynamic testing and finite-element model updating. *Journal of Bridge Engineering* 2003; **8**(3): 162–172.
4. Wang N, O'Malley C, Ellingwood BR, Zureick AH. Bridge Rating Using System Reliability Assessment. I: Assessment and Verification by Load Testing. *Journal of Bridge Engineering* 2011; **16**(6): 854–862. DOI: 10.1061/(ASCE)BE.1943-5592.0000172.
5. BBC News. Lorry tests as monitors installed in Forth Road Bridge first. *BBC News Web Site* 2015.
6. Casciati F, Fuggini C. Engineering vibration monitoring by GPS: long duration records. *Earthquake Engineering and Engineering Vibration* 2009; **8**(3): 459–467.
7. Nickitopoulou a, Protopsalti K, Stiros S. Monitoring dynamic and quasi-static deformations of large flexible engineering structures with GPS: Accuracy, limitations and promises. *Engineering Structures* 2006; **28**(10): 1471–1482. DOI: 10.1016/j.engstruct.2006.02.001.
8. Liao WY, Chen WH, Ni YQ, Xia Y. Development of a Vision-based Real-time Displacement Measurement System for Guangzhou New TV Tower. In: Casciati F, editor. *Fifth European Workshop on Structural Health Monitoring 2010*, Itlay: 2010.
9. Feng D, Feng M, Ozer E, Fukuda Y. A Vision-Based Sensor for Noncontact Structural Displacement Measurement. *Sensors* 2015; **15**(7): 16557–16575. DOI: 10.3390/s150716557.
10. Kim SW, Kim NS. Multi-point Displacement Response Measurement of Civil Infrastructures Using Digital Image Processing. *Procedia Engineering* 2011; **14**: 195–203. DOI: 10.1016/j.proeng.2011.07.023.
11. Ribeiro D, Calçada R, Ferreira J, Martins T. Non-contact measurement of the dynamic displacement of railway bridges using an advanced video-based system. *Engineering Structures* 2014; **75**: 164–180. DOI: 10.1016/j.engstruct.2014.04.051.
12. Chang CC, Xiao XH. Three-Dimensional Structural Translation and Rotation Measurement

Using Monocular Videogrammetry. *Journal of Engineering Mechanics* 2010; **136**(7): 840–848.

13. Stephen GA, Brownjohn JMW, Taylor CA. Measurements of static and dynamic displacement from visual monitoring of the Humber Bridge. *Engineering Structures* 1993; **15**(3): 197–208.
14. Macdonald JHG, Dagless E, Thomas B, Taylor C. Dynamic Measurements Of The Second Severn Crossing. *Proc. Inst. Civil Engineers : Transport*, vol. 123, 1997.
15. Wahbeh AM, Caffrey JP, Masri SF. A vision-based approach for the direct measurement of displacements in vibrating systems. *Smart Materials and Structures* 2003; **12**: 785–794. DOI: 10.1088/0964-1726/12/5/016.
16. Ye XW, Ni YQ, Wai TT, Wong KY, Zhang XM, Xu F. A vision-based system for dynamic displacement measurement of long-span bridges: algorithm and verification. *Smart Structures and Systems* 2013; **12**(3_4): 363–379. DOI: 10.12989/sss.2013.12.3_4.363.
17. Martins LL, Rebordão JM, Ribeiro AS. Structural observation of long-span suspension bridges for safety assessment: implementation of an optical displacement measurement system. *Journal of Physics: Conference Series* 2015; **588**: 12004. DOI: 10.1088/1742-6596/588/1/012004.
18. Yoon H, Elanwar H, Choi H, Golparvar-Fard M, Spencer BF. Target-free approach for vision-based structural system identification using consumer-grade cameras. *Structural Control and Health Monitoring* 2016; **23**(12): 1405–1416. DOI: 10.1002/stc.1850.
19. Caetano E, Silva S, Bateira J. Application of a vision system to the monitoring of cable structures. *Seventh International Symposium on Cable Dynamics* 2007: 225–236.
20. Oh BK, Hwang JW, Kim Y, Cho T, Park HS. Vision-based system identification technique for building structures using a motion capture system. *Journal of Sound and Vibration* 2015: 1–14. DOI: 10.1016/j.jsv.2015.07.011.
21. Feng D, Feng MQ. Model Updating of Railway Bridge Using In Situ Dynamic Displacement Measurement under Trainloads. *Journal of Bridge Engineering* 2015; **20**(12): 4015019. DOI: 10.1061/(ASCE)BE.1943-5592.0000765.
22. Cha YJ, Chen JG, Büyüköztürk O. Output-only computer vision based damage detection using phase-based optical flow and unscented Kalman filters. *Engineering Structures* 2017; **132**: 300–313. DOI: 10.1016/j.engstruct.2016.11.038.
23. Ojio T, Carey CH, OBrien EJ, Doherty C, Taylor SE. Contactless Bridge Weigh-in-Motion. *Journal of Bridge Engineering* 2016; **21**(7): 4016032. DOI: 10.1061/(ASCE)BE.1943-5592.0000776.
24. Brownjohn JMW, Xu Y, Hester D. Vision-Based Bridge Deformation Monitoring. *Frontiers in Built Environment* 2017; **3**(April): 1–16. DOI: 10.3389/fbuil.2017.00023.
25. Khuc T, Catbas FN. Computer vision-based displacement and vibration monitoring without using physical target on structures. *Structure and Infrastructure Engineering* 2017; **13**(4): 505–516. DOI: 10.1080/15732479.2016.1164729.
26. Khuc T, Necati Catbas F. Completely contactless structural health monitoring of real-life

structures using cameras and computer vision. *Structural Control and Health Monitoring* 2016.

27. Feng D, Feng MQ. Experimental validation of cost-effective vision-based structural health monitoring. *Mechanical Systems and Signal Processing* 2017; **88**: 199–211. DOI: 10.1016/j.ymssp.2016.11.021.
28. Choi I, Kim J, Kim D. A Target-Less Vision-Based Displacement Sensor Based on Image Convex Hull Optimization for Measuring the Dynamic Response of Building Structures. *Sensors* 2016; **16**(12): 2085. DOI: 10.3390/s16122085.
29. Chen JG, Wadhwa N, Cha YJ, Durand F, Freeman WT, Buyukozturk O. Modal identification of simple structures with high-speed video using motion magnification. *Journal of Sound and Vibration* 2015; **345**: 58–71. DOI: 10.1016/j.jsv.2015.01.024.
30. Yang Y, Dorn C, Mancini T, Talken Z, Kenyon G, Farrar C, *et al.* Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification. *Mechanical Systems and Signal Processing* 2017; **85**: 567–590. DOI: 10.1016/j.ymssp.2016.08.041.
31. Caetano E, Silva S, Bateira J. A vision system for vibration monitoring of civil engineering structures. *Experimental Techniques* 2011; **35**(4): 74–82. DOI: 10.1111/j.1747-1567.2010.00653.x.
32. Kim SW, Jeon BG, Kim NS, Park JC. Vision-based monitoring system for evaluating cable tensile forces on a cable-stayed bridge. *Structural Health Monitoring* 2013; **12**(5–6): 440–456. DOI: 10.1177/1475921713500513.
33. Feng D, Scarangello T, Feng MQ, Ye Q. Cable tension force estimate using novel noncontact vision-based sensor. *Measurement* 2017; **99**: 44–52. DOI: 10.1016/j.measurement.2016.12.020.
34. Ji YF, Chang CC. Nontarget Stereo Vision Technique for Spatiotemporal Response Measurement of Line-Like Structures. *Journal of Engineering Mechanics* 2008; **134**(6): 466–474. DOI: 10.1061/(ASCE)0733-9399(2008)134:6(466).
35. Busca G, Cigada A, Mazzoleni P, Zappa E. Vibration Monitoring of Multiple Bridge Points by Means of a Unique Vision-Based Measuring System. *Experimental Mechanics* 2014; **54**: 255–271. DOI: 10.1007/s11340-013-9784-8.
36. Chen CC, Wu WH, Tseng HZ, Chen CH, Lai G. Application of digital photogrammetry techniques in identifying the mode shape ratios of stay cables with multiple camcorders. *Measurement: Journal of the International Measurement Confederation* 2015; **75**: 134–146. DOI: 10.1016/j.measurement.2015.07.037.
37. Kim SW, Kim NS. Dynamic characteristics of suspension bridge hanger cables using digital image processing. *NDT & E International* 2013; **59**: 25–33. DOI: 10.1016/j.ndteint.2013.05.002.
38. Xu Y, Brownjohn J, Hester D, Koo K. Dynamic displacement measurement of a long span bridge using vision-based system. *8th European Workshop On Structural Health Monitoring*, Bilbao, Spain: 2016.

39. Lee J hwa, Cho S, Sim S han. Monocular Vision-based Displacement Measurement System Robust to Angle and Distance Using Homography. *6th International Conference on Advances in Experimental Structural Engineering*, University of Illinois, Urbana-Champaign, United States: 2015.
40. Hartley R, Zisserman A. *Multiple view geometry in computer vision*. Cambridge university press; 2003.
41. Park SW, Park HS, Kim JH, Adeli H. 3D displacement measurement model for health monitoring of structures using a motion capture system. *Measurement* 2015; **59**: 352–362. DOI: 10.1016/j.measurement.2014.09.063.
42. Kim SC, Kim HK, Lee CG, Kim SB. A vision system for identifying structural vibration in civil engineering constructions. *2006 SICE-ICASE International Joint Conference*, Bexco, Busan, Korea: 2006. DOI: 10.1109/SICE.2006.315227.
43. Zhang Z. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2000; **22**(11): 1330–1334. DOI: 10.1109/34.888718.
44. Chang CC, Ji YF. Flexible Videogrammetric Technique for Three-Dimensional Structural Vibration Measurement. *Journal of Engineering Mechanics* 2007; **133**(6): 656–664. DOI: 10.1061/(ASCE)0733-9399(2007)133:6(656).
45. Sun D, Roth S, Black MJ. Secrets of optical flow estimation and their principles. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE; 2010. DOI: 10.1109/CVPR.2010.5539939.
46. Tomasi C. *Detection and Tracking of Point Features Technical Report CMU-CS-91-132*. vol. 91. School of Computer Science, Carnegie Mellon Univ. Pittsburgh; 1991. DOI: 10.1016/S0031-3203(03)00234-6.
47. Lee JJ, Shinozuka M. A vision-based system for remote sensing of bridge displacement. *NDT & E International* 2006; **39**(5): 425–431. DOI: 10.1016/j.ndteint.2005.12.003.
48. Lee JJ, Cho S, Shinozuka M. Evaluation of Bridge Load Carrying Capacity Based on Dynamic Displacement Measurement Using Real-time Image Processing Techniques. *Steel Structures* 2006; **6**: 377–385.
49. Guo J, Zhu C. Dynamic displacement measurement of large-scale structures based on the Lucas–Kanade template tracking algorithm. *Mechanical Systems and Signal Processing* 2016; **66**–**67**: 425–436. DOI: 10.1016/j.ymssp.2015.06.004.
50. Ehrhart M, Lienhart W. Monitoring of Civil Engineering Structures using a State-of-the-art Image Assisted Total Station. *Journal of Applied Geodesy* 2015; **9**(3): 174–182. DOI: 10.1515/jag-2015-0005.
51. Szeliski R. *Computer Vision : Algorithms and Applications*. vol. 5. Springer Science & Business Media; 2010. DOI: 10.1007/978-1-84882-935-0.
52. Ehrhart M, Lienhart W. Development and evaluation of a long range image-based monitoring

system for civil engineering structures. In: Shull PJ, editor. *Proc. SPIE*, vol. 9437, 2015. DOI: 10.1117/12.2084221.

53. Pan B, Qian K, Xie H, Asundi A. Two-dimensional digital image correlation for in-plane displacement and strain measurement: a review. *Measurement Science and Technology* 2009; **20**(6): 62001. DOI: 10.1088/0957-0233/20/6/062001.
54. Guizar-Sicairos M, Thurman ST, Fienup JR. Efficient subpixel image registration algorithms. *Optics Letters* 2008; **33**(2): 156. DOI: 10.1364/OL.33.000156.
55. Sobel I. Neighborhood coding of binary images for fast contour following and general binary array processing. *Computer Graphics and Image Processing* 1978; **8**(1): 127–135.
56. Ghosal S, Mehrotra R. Orthogonal moment operators for subpixel edge detection. *Pattern Recognition* 1993; **26**(2): 295–306. DOI: 10.1016/0031-3203(93)90038-X.
57. Ying-Dong Q, Cheng-Song C, San-Ben C, Jin-Quan L. A fast subpixel edge detection method using Sobel-Zernike moments operator. *Image and Vision Computing* 2005; **23**(1): 11–17. DOI: 10.1016/j.imavis.2004.07.003.
58. Welch PD. The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms. *IEEE Transactions on Audio and Electroacoustics* 1967; **15**(2): 70–73. DOI: 10.1109/TAU.1967.1161901.
59. Peeters B, De Roeck G. Reference-based stochastic subspace identification for output-only modal analysis. *Mechanical System & Signal Processing* 1999; **13**(6): 855–878. DOI: 10.1006/mssp.1999.1249.
60. Lin J, Qu L. Feature Extraction Based on Morlet Wavelet and Its Application for Mechanical Fault Diagnosis. *Journal of Sound and Vibration* 2000; **234**(1): 135–148. DOI: 10.1006/jsvi.2000.2864.
61. Yan B, Miyamoto A. A comparative study of modal parameter identification based on wavelet and Hilbert-Huang transforms. *Computer-Aided Civil and Infrastructure Engineering* 2006; **21**(1): 9–23. DOI: 10.1111/j.1467-8667.2005.00413.x.
62. Wang H, Mao JX, Huang JH, Li AQ. Modal Identification of Sutong Cable-Stayed Bridge during Typhoon Haikui Using Wavelet Transform Method. *Journal of Performance of Constructed Facilities* 2016; **30**(5): 1–11. DOI: 10.1061/(ASCE)CF.1943-5509.0000856.
63. Carmona R, Hwang WL, Torresani B. *Practical Time-Frequency Analysis*. vol. 9. Academic Press; 1998. DOI: 10.1016/S1874-608X(98)80033-2.
64. Hester D, Brownjohn J, Bocian M, Xu Y. Low cost bridge load test: Calculating bridge displacement from acceleration for load assessment calculations. *Engineering Structures* 2017; **143**. DOI: 10.1016/j.engstruct.2017.04.021.
65. Stiros SC. Errors in velocities and displacements deduced from accelerographs: An approach based on the theory of error propagation. *Soil Dynamics and Earthquake Engineering* 2008; **28**(5): 415–420. DOI: 10.1016/j.soildyn.2007.07.004.

66. Brownjohn JMW, Bocian M, Hester D, Quattrone A, Hudson W, Moore D, *et al.* Footbridge system identification using wireless inertial measurement units for force and response measurements. *Journal of Sound and Vibration* 2016; **384**: 339–355. DOI: 10.1016/j.jsv.2016.08.008.
67. Exeter Chief News. Chiefs 36 Falcons 14. *Exeter Chiefs Website* 2017.
68. Bing P, Hui-min X, Bo-qin X, Fu-long D. Performance of sub-pixel registration algorithms in digital image correlation. *Measurement Science and Technology* 2006; **17**(6): 1615–1621. DOI: 10.1088/0957-0233/17/6/045.
69. Martins LL, Nunes Vicente Rebordão JM, Silva Ribeiro Á. Thermal Influence on Long-Distance Optical Measurement of Suspension Bridge Displacement. *International Journal of Thermophysics* 2014; **35**: 693–711. DOI: 10.1007/s10765-014-1607-3.

Table 1 Geometric information of control points used to calibrate the projection relation including their coordinates in structural coordinate system and image plane as well the re-projected image coordinates based on the estimated projection relation.

CP No.	**Structural Coordinates (mm)**			**Observed image coordinates (pixel)**		**Re-projected image coordinates (pixel)**		**Re-projection error (pixel)**
	X	Y	Z	U	V	U	V	
1	-2700	-570	0	44	663	45.1	664.9	2.2
2	-2700	0	0	37	466	36.5	469.5	3.5
3	0	0	0	967	471	966.7	460.6	10.4
4	2700	0	0	1859	464	1858.5	469.2	5.3
5	2700	-570	0	1850	655	1850.3	655.7	0.8

Table 2 Image pixel dimensions of the selected targets in video frames

Target no.	**Image size / pixels** (width × height or length**)**
D1	33 ×17
D2	43 × 19
D3	21 × 21
D4	21 × 26
C1	110
C2	96