- Article type: paper.
- Date text written or revised: 16 Nov 2017
- Number of words: 4548
- Number of figures: 15
- Number of tables: 2

Vision-based systems for structural deformation measurement: case studies

Author 1

- Yan Xu* BEng, MEng
- Research Student, Vibration Engineering Section, College of Engineering, Mathematics and Physical Sciences, University of Exeter, UK, EX4 4QF
- ORCID: 0000-0001-5886-9938

Author 2

- James Mark William Brownjohn BSc, PhD, DEng, CEng, FIMechE, FIStructE
- Professor of Structural Dynamics, Vibration Engineering Section, College of Engineering, Mathematics and Physical Sciences, University of Exeter, UK, EX4 4QF
- ORCID: 0000-0003-4946-5901

Full contact details of corresponding author (*)

Vibration Engineering Section, College of Engineering, Mathematics and Physical Sciences, University of Exeter, UK, EX4 4QF

+44(0)1392726421

yx298@exeter.ac.uk

ABSTRACT

Vision-based systems offer a promising way for displacement measurement and receive increased attention in civil structural monitoring. However, the working performance of vision-based systems, especially the measurement accuracy and the robustness to different field conditions is not fully understood. This study reports three cases studies of vision-based monitoring tests including one in a laboratory, one on a short-span bridge and one on a long-span bridge. The tracking accuracy is quantified in laboratory conditions in the range of 0.02 pixel to 0.20 pixel depending on the target patterns as well as the tracking method selected. The measurement performance under several field challenges are investigated including long-range measurement (e.g. camera-to-target distance at 710 m), low-contrast target patterns, changes of target patterns and changes in lighting conditions. Three representative tracking methods for the video processing, i.e. correlation-based template matching, Lucas Kanade (LK) optical flow estimation and scale-invariant feature transform (SIFT) were used for analysis, indicating their advantages and shortcomings for field measurement. One of the main observations in field application is that changes in lighting conditions might cause some low-frequency measurement error that could be misunderstood without the prior knowledge about structural loading conditions.

Keywords

Bridges; Field testing & monitoring; Noise.

1 INTRODUCTION

Bridges provide vital links in transportation networks and must be managed in a manner that minimises risk to public safety and disruption to service. However, the current inspection process mainly relies on visual check that can be subjective and prone to error. Therefore, reliable approaches for bridge condition assessment are required to assist the decision-making and make the best use of limited maintenance budget.

Bridge deformation is a significant metric for bridge condition assessment. For example, measurement of deformation during controlled vehicle load testing helps to estimate bridge load carrying capacity (Wang et al., 2011) (BBC News, 2015). Serviceability is reflected through deformation during normal operation, since extreme values and ranges indicate problems that may limit operational use.

1.1 Review of vision-based approaches

Limitations of more traditional displacement sensing technologies (e.g. LVDT, GPS, double-integration from acceleration etc.) that necessarily require physical access to a structure have driven research in non-contact optical sensing. Vision-based monitoring methods have promising features e.g. simple instrumentation and installation, operation remote from the structure and capacity for multi-point measurement using a single (camera) sensor. Existing studies have indicated the potential of vision-based methods for structural condition assessment, in particular for system identification (Caetano et al., 2007; Oh et al., 2015; Yoon et al., 2016), finite element model calibration (Feng and Feng, 2015), damage detection (Cha et al., 2017) and contribution to bridge WIM system with camera assistance for traffic monitoring (Ojio et al., 2016).

Target tracking is one critical component in a vision-based system, directly influencing the measurement accuracy. Template matching (Brownjohn et al., 2017; Chang and Ji, 2007; Ehrhart and Lienhart, 2015a; Fukuda et al., 2013; Guo and Zhu, 2016; Macdonald et al., 1997; Stephen et al., 1993) and optical flow estimation (Caetano et al., 2011; Cha et al., 2017; Chen, Wu, et al., 2015; Chen et al., 2017; Chen, Wadhwa, et al., 2015; Diamond et al., 2017; Ehrhart and Lienhart, 2015b; Ji and Chang, 2008a; Khaloo and Lattanzi, 2017; Yang et al., 2017; Yoon et al., 2016) are established methods widely used for bridge deformation measurement whereas feature point matching is a relatively new and promising tracking method that is theoretically scale-invariant and rotation-invariant (Ehrhart and Lienhart, 2015a) and has been validated in several short-range measurement tests (Ehrhart and Lienhart, 2015a, 2015b, Khuc and Catbas, 2017a, 2017b). There are also some other methods through tracking the special shapes of target patterns based on edge detection or image thresholding algorithms (Ji and Chang, 2008b; Lee et al., 2006; Ribeiro et al., 2014; Wahbeh et al., 2003) e.g. line-like cables, circular-shaped dots and chessboard, etc. These methods have limitations for application due to the requirement about pattern shapes.

To find the most appropriate tracking method for structural monitoring, it is necessary to evaluate their measurement accuracy and the robustness to different field conditions. Busca et al. (Busca et al., 2014) evaluated three techniques (template matching, edge detection and digital image correlation) on a steel truss railway bridge, concluding that the three techniques provide similar tracking performance while tracking accuracy is slightly poorer for natural targets. Ehrhart and Lienhart (Ehrhart and Lienhart, 2015a) evaluated the performance of three techniques (optical flow, template matching and feature point matching) by tracking structural features of a footbridge and reported that feature point matching is robust to the changes of background condition (i.e. snowfall) whereas drift over time was observed in the measurement by two other methods.

These two existing studies provide some information about the influence of pattern features and pattern changes on measurement results. However, more studies are necessary to evaluate the field performance of vision-based systems since several critical field challenges are not considered yet e.g. robustness to lighting changes and viability for long-range monitoring.

1.2 Purpose of this study

The purpose of this study is to investigate the effectiveness of vision-based systems for displacement measurement in different environmental conditions through three case studies. One laboratory test and two field tests were performed indicating several influential factors on measurement performance, i.e. estimation error in projection transformation, camera-to-target distance, the distinctiveness of target patterns, changes of target patterns and changes in lighting conditions. Three representative tracking methods were considered for video processing, demonstrating their advantages and shortcomings to deal with the observed influential factors. Other error sources in field tests like camera shake, atmospheric refraction and temperature variations were not apparent in test observations and thus are not discussed in this study.

This paper is the first study to apply the feature point matching method in the long-range monitoring test and the measurement was evaluated through comparison with the processing results using classical tracking methods. The field measurement by vision-based systems might carry some low-frequency error due to camera shake or lighting changes. A method to distinguish the main source of low frequency error is proposed in this study.

To that end, section 2 introduces the methodologies for vision-based displacement measurement, in particular, three representative tracking methods. In section 3, a laboratory uniaxial oscillation test used to validate the video processing methods and to study the tracking accuracy in laboratory conditions is described. Section 4 and 5 report the field tests for mid-span deformation measurement in a short-span railway bridge and a long-span suspension bridge, respectively. The performance of vision-based systems with challenging field conditions e.g. low-contrast patterns, changes of target patterns and changes in lighting conditions are investigated, indicating the advantages and shortcomings of the three tracking methods. Finally the main findings from the three tests are summarised.

2 VIDEO PROCESSING METHODOLOGIES

In vision-based systems, the hardware comprises data acquisition devices, portable computer with video processing software and accessories like tripod and artificial targets (optional). Video processing package is the key part which could fit into a three-component framework in Figure 1, i.e. camera calibration, target tracking and displacement calculation.

Camera calibration is aimed at determining the projection transformation from the structural system to the image plane. The projection transformation used in this study is the planar homography that is calibrated based on a few planar point correspondences and is capable to reconstruct the two-dimensional planar displacement.

Target tracking is critical in the video processing package to locate the target regions in the image plane through tracking methods. The tracking methods considered in this study include template matching and optical flow estimation that are established and classic, as well as feature point matching that receives increased attention in structural monitoring.

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 projection transformation (output of camera calibration).

A software package for the post-processing analysis of video files to measure the structural displacement is developed by the authors using C++ language, partly referring to OpenCV library. This custom-developed software is capable to measure the structural displacement of multiple targets simultaneously within the field of view and offers three options of tracking methods (shown in Figure 1) to adapt to different test conditions.

This section provides a description of the methodologies of three target tracking methods used in the custom-developed video-processing software i.e. correlation-based template matching, Lucas Kanade (LK) optical flow estimation and scale-invariant feature transform (SIFT) method (Lowe, 2004).

2.1 Correlation-based template matching

Template matching is a classic technique for target tracking by searching in a new frame for an area most closely resembling a predefined template, following the procedures demonstrated in Figure 2. A target region is selected as the template that is a subset image in the reference frame. A matching criterion is defined to evaluate the similarity degree between the template and the new frame and the criterion used is zero-mean normalised cross correlation coefficient (ZNCC). The target location in the new frame corresponds to the peak location in the similarity matrix that has resolution at pixel level. Subpixel interpolation schemes (Feng et al., 2015) are required to refine the tracking results to sub-pixel level and the interpolation method used in this study is zero-padding in frequency domain using the matrix multiplication form of discrete Fourier transform (Guizar-Sicairos et al., 2008).

Template matching has been applied in structural monitoring since the earliest work on the Humber and Severn Bridges in 1990s (Macdonald et al., 1997; Stephen et al., 1993), and the recent applications include displacement monitoring tests in a railway bridge (Feng et al., 2015), a long-span bridge (Ye et al., 2013) and a high-rise building (Liao et al., 2010).

2.2 Lucas Kanade optical flow estimation

Lucas Kanade optical flow estimation detects the motions in an image from the brightness pattern shift (Beauchemin and Barron, 1995). The calculation process imposes one temporal constraint on image properties and one spatial constraint on motion consistency, i.e. that the pixel intensities of an object do not change between consecutive frames and that neighbouring pixels have similar motion. The image motions are derived using the following equation,

$$\begin{bmatrix} dx \\ dy \end{bmatrix} = \begin{bmatrix} \sum_i I_{xi}^2 & \sum_i I_{xi} I_{yi} \\ \sum_i I_{xi} I_{yi} & \sum_i I_{yi}^2 \end{bmatrix}^{-1} \begin{bmatrix} -\sum_i I_{xi} I_{ti} \\ -\sum_i I_{yi} I_{ti} \end{bmatrix} \tag{1}$$

where dx and dy denote the motions in image plane; I_x, I_y and I_t represent the spatial and temporal gradients of image intensities; and i denotes the *i*th pixel location in a 3×3 patch around a feature point (x, y).

The procedures are demonstrated in Figure 3. The most prominent corners are detected by Shi-Tomasi corner detector (Jianbo Shi and Tomasi, 1994) in the reference frame with their locations updated in the new frame by Eq.(1). Backward estimation is then performed from the new frame to the reference frame in order to refine the point correspondences based on the error between the initially detected points and the backward-estimated points. Point correspondences are further refined based on geometric alignment using the least median of squares method (Massart et al., 1986).

The LK optical flow estimation was applied in a laboratory test of a multi-storey metal tower (Yoon et al., 2016) for system identification, for field application of a footbridge deformation monitoring (Ehrhart and Lienhart, 2015b) and for bridge stay-cable vibration measurement (Caetano et al., 2011; Ji and Chang, 2008a).

2.3 Scale-invariant feature transform (SIFT)

Feature point matching is an efficient tracking technique through matching feature points in consecutive images based on their local appearance. Several robust feature extractors and descriptors are reported in literature (Alahi et al., 2012; Bay et al., 2008; Calonder et al., 2010; Lowe, 2004; Rublee et al., 2011) and the one used in this study is the SIFT method.

The procedures are demonstrated in Figure 4. Keypoints are extracted from the local extremes in the Difference of Gaussian images (differences of Gaussian filtered images with varied blur level) and described by vectors using the gradient magnitudes and orientations of neighbouring pixels. Keypoints between two images are matched by identifying their nearest neighbours evaluated using the Euclidean distances between keypoint descriptor vectors. The outliers in keypoint correspondences are removed based on geometric alignment using the least median of squares method (Massart et al., 1986).

The SIFT method has been validated in the deformation measurement test of a railway bridge (Khuc and Catbas, 2017a). Other feature point matching methods are also validated only in short-range monitoring tests (Ehrhart and Lienhart, 2015a, 2015b; Khuc and Catbas, 2017b).

3 VALIDATION TEST IN CONTROLLED ENVIRONMENT

This section investigates the performance of three tracking methods (i.e. correlation-based template matching, LK optical flow estimation and SIFT method) in controlled environmental conditions through a laboratory uniaxial oscillation test.

3.1 Test configuration

An APS 400 Electrodynamic shaker was set vertically on solid ground shown in Figure 5(a) and driven by an input of chirp signal with frequency range 0.5 Hz to 2 Hz. The test was run twice, with no artificial target in Run 1 and with the chessboard pattern attached to shaker mass (shown in Figure 5(b)) in Run 2 in order to increase the feature salience.

One GoPro Hero4 camera was fixed on the ground looking at the oscillating shaker mass; sample frames in two runs are given in (c) and (d). The frame rate was set as nominally 60 Hz with narrow field-of-view option. A Balluff micropulse transducer was attached to the shaker mass shown in (b) to provide an accurate reference displacement measurement sampled at 256 Hz.

The recorded video files were post-processed using the custom-developed video processing software to derive the displacement data of shaker mass. Camera calibration was performed according to the dimensions of shaker mass indicated in Figure 5(b) to estimate the planar homography; the target locations in the image were estimated by the three tracking methods, respectively; and the horizontal and vertical displacement was acquired through transforming the image locations to the structural locations via the planar homography.

3.2 Measurement results

The measurement results using the GoPro camera and Balluff transducer in two runs are illustrated in Figure 6, indicating that,

- Vision-based system based on a consumer-grade camera is qualified to measure the vertical oscillation of shaker mass by tracking either natural features or artificial target patterns with the cross-correlation coefficient (compared with the Balluff measurement) reaching over 99%.
- Oscillation occurred only in the vertical direction and any displacement measurement in the horizontal direction corresponds to measurement error. In Run 1, correlation-based template

matching provides the best measurement performance while the SIFT measurement includes the highest noise. Measurement error in Run 2 is much smaller than in Run 1 after increasing the target salience.

- During the oscillation period in Run 2, the horizontal measurement indicates a chirp signal with a similar shape as the vertical measurement but with much smaller amplitude. It might be caused by the error in projection transformation, making the dominant vibration leak to minor motion direction.

To investigate the tracking accuracy of the three methods, the data collected during the stationary period were taken into account with the estimation results summarised in Table 1. This indicates that in laboratory conditions, correlation-based template matching is the most accurate method while the SIFT achieves the poorest accuracy with the highest sensitivity to the distinctiveness of target patterns.

Laboratory evaluation avoids numerous difficulties that can diminish performance in field conditions due to environmental influences such as camera-to-target distance, unstable target patterns and lighting conditions, etc. In section 4 and 5, two field tests on bridge deformation measurement are reported to illustrate the real-world working performance of vision-based systems.

4 DEFORMATION MEASUREMENT TEST ON A SHORT-SPAN BRIDGE

The vision-based system applied in the two field tests is the Imetrum Dynamic Monitoring Station (DMS) originating from research at the University of Bristol and commercialised via the university spin-out company Imetrum formed in 2003. The system includes one or more GigE high performance cameras for data acquisition and the software 'Video Gauge' for the real-time video processing.

The target tracking algorithms used in the software are proprietary extensions of correlation-based template matching techniques which enable better than 1/100 pixel resolution at sample rates beyond 100 Hz in field applications. The system used by University of Exeter has been trialled in several one-day field campaigns on a number of bridges in the UK, indicating comparable or even better measurement accuracy compared to an LVDT for short-range measurement (Hester et al., 2017) and the GPS for long-range measurement (Xu et al., 2017).

In this study, the Imetrum system was used in field tests mainly for two functions, 1) as a data acquisition device to record the video files of the bridges that would be analysed using the custom-developed video processing software, and 2) to provide (using the Imetrum proprietary video processing software) the reference data of bridge deformation for the evaluation of the measurement results provided by the custom-developed video processing software.

This section reports a case study using the vision-based system for the deformation measurement of a short-span railway bridge. Section 4.1 and 4.2 below describe the test setup and the results obtained, respectively.

4.1 Test configuration

The Mineral Line Bridge is a steel girder bridge with 14 m span, carrying the West Somerset Railway near Watchet, UK. Figure 7(a) indicates the test setup with an Imetrum camera mounted at the top of a tripod approximately 12.5 m away from the mid-span of the bridge. One sample frame captured by the Imetrum camera is illustrated in Figure 7(b) with the target region for measurement marked by a rectangular box.

Since the fundamental frequency of the bridge was estimated to be above 10 Hz, a frame rate of 30 Hz was set for the Imetrum system. The camera calibration was performed using the dimensions of the bridge girder and the artificial target at mid-span. The displacement data along the vertical and longitudinal directions was measured directly by the Imetrum system and also extracted from the custom-developed post-processing software using the three tracking methods.

4.2 Measurement results

Time history data of displacement measurement are illustrated in Figure 8, indicating that,

- The three tracking methods all capture the vertical deformation of the bridge induced by the passing of a steam train with one steam locomotive and seven passenger carriages with the maximum deformation at 6.4 mm.
- A similar deformation pattern during the train passing occurs in the longitudinal measurement with much smaller amplitudes (less than 0.56 mm) apart from the measurement by SIFT method due to

the high noise. It might be caused by the error of projection transformation, making the vertical deflection data leaked to the horizontal direction.

- An apparent low-frequency motion trend in the horizontal direction is observed after 30 s for all the four measurement that should be an error since the bridge was empty with no heavy loading. During this period, the four methods provide different amplitude values and the SIFT measurement is the most noisy one but with the smallest amplitude. The LK optical flow estimation method failed to measure in some frames possibly due to the large brightness changes. In terms of the image motion, the maximum drift in the image horizontal direction reaches approx. 0.3 pixel, larger than the estimated tracking accuracy in laboratory conditions. It indicates that the tracking accuracy becomes poorer in field conditions for any of the three tracking methods.

According to the authors' test experiences and the literature, the low-frequency error in vision-based measurement could be caused by either the camera motion (Zhao et al., 2017) or changes in lighting conditions (Brownjohn et al., 2017). The error induced by camera motion should be consistent at those pixels corresponding to the stationary objects in the field of view. The camera motion is believed not to be the main error source in this case because Figure 8 is for one among several targets tracked in the bridge girder (in Figure 9(a)), and these show image motions inconsistent in both the amplitude and direction as shown in Figure 9(b) and (c).

To quantify the influence of lighting changes, mean pixel intensity at the initially selected target region (T0) were calculated as shown in Figure 10(a). This indicates a growth of averaged brightness from 28.5 s to 37 s followed by a gentle decrease, which has a trend similar to the measurement error in the longitudinal direction in Figure 8(a). The initial frame and the frame at the time step 37 s are shown in Figure 10(b) and (c) for visual comparison of lighting changes. In the second frame, the bolts within the rectangular region (T0) are more distinctive against the background due to the improved lighting. Therefore, it is believed that the low-frequency error in the longitudinal direction is caused by the lighting changes.

None of the three tracking methods is robust to large lighting changes in field tests. The SIFT method experiences the least influence in measurement while the LK optical flow might fail to identify the features under apparent lighting changes. The low-frequency error due to lighting changes might mislead the users about the bridge loading condition if no prior knowledge is available.

A sharp brightness increase is observed at time step 18.57 s in Figure 10(a) corresponding to an outlier in the measurement by correlation-based template matching and the Imetrum system in Figure 8. The cause of these anomalies is changes in target pattern due to a bird flying in front of the target, as shown in Figure 11. This indicates that the LK optical flow estimation and SIFT method are not sensitive to small pattern changes whereas the correlation-based template matching method might fail to track.

The demonstration of vision-based monitoring in a short-span bridge indicates that the vision-based system is capable to measure the bridge deflection under traffic loads using any of the three tracking methods, although the measurement might become unstable when suffering from apparent changes in lighting conditions or target patterns.

In the next section, a similar test was performed on a long-span bridge to investigate the viability of vision-based system for the long-range monitoring.

5 DEFORMATION MEASUREMENT TEST ON A LONG-SPAN BRIDGE

This section describes a case study of using a vision-based system measuring deformation of the Humber Bridge, UK, a suspension bridge with (at the time of writing) the world's eighth longest span. Section 5.1 and 5.2 describe the test setup and the results obtained, respectively. The Imetrum system was used in field for video acquisition and the video files were post-processed by the custom-developed software to evaluate the performance of three tracking methods for the long-range measurement.

5.1 Test configuration

The Humber Bridge, with main span of 1410 m links the towns of Hessle and Barton across the Humber estuary. A single day of field testing using the Imetrum system on 22nd July 2015 was used to measure the displacement at mid-span of the bridge which has been reported in (Brownjohn et al., 2017; Xu et al., 2017). The camera, equipped with a lens of 300 mm focal length, was located at the base of the north tower shown in Figure 12(a), 710 m from the mid-span of the bridge. An artificial target with the pattern of concentric rings in Figure 12(b) was attached to the parapet at the mid-span. One sample frame by vision-based system is shown in Figure 12(c) with the annotations at the two target regions, T1 covering the artificial target region and T2 under the deck soffit with low-contrast patterns.

The previous modal test (Brownjohn et al., 2010) indicates 14 modes in the vertical direction with the modal frequencies lower than 1 Hz and the fundamental frequency was 0.117 Hz. Since the frequency components of interest were lower than 1 Hz, the frame rate of the Imetrum system was set as 10 Hz. The camera calibration was performed according to the square dimension (1 m) of the artificial target. The displacement data along the vertical and transverse directions was measured directly by the Imetrum system and also extracted from the custom-developed post-processing software using the three tracking methods.

5.2 Measurement results

An 80-second signal of vertical displacement at the target region T1 recorded at approx. 19.22 PM (BST) is illustrated in Figure 13. The maximum displacement reaches 160.5 mm at approximately 31 s by the Imetrum measurement while some data points using the three post-processing methods are missing for about 0.9 s when the displacement values reach their maxima. Two frames during this period are shown in Figure 14 indicating a big change in target pattern. Due to the low sun elevation in the west, the target panel on the east side was initially partially in the shadow of the bridge railing, shown in (a). When one tall vehicle passed the mid-span of the bridge between the sun and the target, sunlight was completely blocked, making the whole target pattern visible in the image. This indicates that the three tracking methods are all not robust to large changes in target patterns. Imetrum system, with its proprietary algorithms, is more robust in this case.

The target region T2 is located at the deck soffit that is less salient with smaller spatial changes in target patterns compared with the target region T1. In Figure 15, the maximum displacement measured at T2 is 128.2 mm and 130.7 mm by the correlation-based template matching and SIFT method, respectively while the LK optical flow method failed to track several frames including the period reaching the maximum displacement. This indicates that the LK optical flow method has higher requirements on salience and stability of target patterns. The displacement measured at T2 is smaller than that measured at T1 because 1) the two targets would experience different motion as the bridge rotates about its longitudinal axis due to eccentric traffic loading; and 2) the bridge axis directions projected in the image plane and the projection transformation were determined according to the artificial target panel and thus were not perfectly aligned with those for the feature target T2 under the deck soffit.

6 CONCLUSIONS

This paper investigates the performance of vision-based systems for displacement measurement in laboratory and field tests. Three representative tracking methods (i.e. correlation-based template matching, LK optical flow estimation and SIFT method) were used for the post processing of test records with their performance compared with a commercial vision-based system, Imetrum DMS.

In laboratory conditions, the tracking accuracies for the two methods, correlation-based template matching and LK optical flow estimation are close varied from 0.02 pixel to 0.10 pixel depending on target patterns while the accuracy of SIFT method is poorer in the range between 0.03 pixel and 0.20 pixel.

The working performance of three tracking methods in field tests are summarised in Table 2.

- All the three tracking methods are effective for either short range or long range measurement (e.g. camera-to-target distance at 710 m) with the displacement varying from several millimetres to ten centimetres. However, the tracking accuracy becomes poorer than that achieved in laboratory conditions.
- The salience of target patterns has a direct influence on the measurement accuracy and high-contrast patterns are preferred for tracking. LK optical flow estimation has the highest requirement about the distinctiveness and stability of target patterns and might fail to track when the other two methods work fine.
- Changes to target patterns due to object obstruction or daytime shadows might lead to missing data. Correlation-based template matching is the method most sensitive to the changes of target patterns while the other two tracking methods are also influenced when facing large changes on target patterns.
- Changes of lighting conditions might cause some low-frequency measurement error using any of the three tracking methods, which could be misunderstood without the prior knowledge of structural

loading. SIFT method is influenced by the lighting changes but provides smaller measurement error compared with the other two methods.

It is indicated that apart from constraining the test conditions, e.g. testing in overcast weather or selecting the sheltered target patterns, how to deal with varying lighting conditions in the field is still an open question for vision-based measurement.

Another important observation is that although the two-dimensional displacement measurement is provided by the vision-based system, the measurement along the minor deformed direction might not be reliable. This is because the error in projection transformation might lead to the leakage of dominant deformation to the minor deformed direction. Thus special attention should be given to interpret the measured displacement along the minor deformed direction.

AKNOWLEDGEMENT

We would like to thank the West Somerset Railway and the Humber Bridge Board for permission to use their bridges and for assistance they provided. Also thanks to Farhad Huseynov and Karen Faulkner for organising the field test in the Mineral Line Bridge. Finally the authors would like to thank the two anonymous reviewers for their constructive comments.

REFENRENCES

Alahi, A., Ortiz, R. and Vandergheynst, P. (2012), "Freak: Fast retina keypoint", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 510–517.

Bay, H., Ess, A., Tuytelaars, T. and Van Gool, L. (2008), "Speeded-up robust features (SURF)", *Computer Vision and Image Understanding*, Vol. 110 No. 3, pp. 346–359.

BBC News. (2015), "Lorry tests as monitors installed in Forth Road Bridge first", *BBC News Web Site,* available at: http://www.bbc.co.uk/news/uk-scotland-edinburgh-east-fife-35151289 (accessed 21 December 2015).

Beauchemin, S.S. and Barron, J.L. (1995), "The computation of optical flow", *ACM Computing Surveys*, Vol. 27 No. 3, pp. 433–466.

Brownjohn, J.M.W., Magalhaes, F., Caetano, E. and Cunha, A. (2010), "Ambient vibration re-testing and operational modal analysis of the Humber Bridge", *Engineering Structures*, Vol. 32 No. 8, pp. 2003–2018.

Brownjohn, J.M.W., Xu, Y. and Hester, D. (2017), "Vision-Based Bridge Deformation Monitoring", *Frontiers in Built Environment*, Vol. 3, pp. 1–16.

Busca, G., Cigada, A., Mazzoleni, P. and Zappa, E. (2014), "Vibration Monitoring of Multiple Bridge Points by Means of a Unique Vision-Based Measuring System", *Experimental Mechanics*, Vol. 54, pp. 255–271.

Caetano, E., Silva, S. and Bateira, J. (2007), "Application of a vision system to the monitoring of cable structures", *Seventh International Symposium on Cable Dynamics*, pp. 225–236.

Caetano, E., Silva, S. and Bateira, J. (2011), "A vision system for vibration monitoring of civil engineering structures", *Experimental Techniques*, Vol. 35 No. 4, pp. 74–82.

Calonder, M., Lepetit, V., Strecha, C. and Fua, P. (2010), "Brief: Binary robust independent elementary features", *European Conference on Computer Vision*, Crete, Greece, pp. 778–792.

Cha, Y.-J., Chen, J.G. and Büyüköztürk, O. (2017), "Output-only computer vision based damage detection using phase-based optical flow and unscented Kalman filters", *Engineering Structures*, Vol. 132, pp. 300–313.

Chang, C.C. and Ji, Y.F. (2007), "Flexible Videogrammetric Technique for Three-Dimensional Structural Vibration Measurement", *Journal of Engineering Mechanics*, Vol. 133 No. 6, pp. 656–664.

Chen, C.-C., Wu, W.-H., Tseng, H.-Z., Chen, C.-H. and Lai, G. (2015), "Application of digital photogrammetry techniques in identifying the mode shape ratios of stay cables with multiple camcorders", *Measurement*, Vol. 75, pp. 134–146.

Chen, J.G., Davis, A., Wadhwa, N., Durand, F., Freeman, W.T. and Buyukozturk, O. (2017), "Video Camera-Based Vibration Measurement for Civil Infrastructure Applications", *Journal of Infrastructure Systems*, Vol. 23 No. 3, p. 11.

Chen, J.G., Wadhwa, N., Cha, Y.-J., Durand, F., Freeman, W.T. and Buyukozturk, O. (2015), "Modal identification of simple structures with high-speed video using motion magnification", *Journal of Sound and Vibration*, Vol. 345, pp. 58–71.

Diamond, D.H., Heyns, P.S. and Oberholster, A.J. (2017), "Accuracy evaluation of sub-pixel structural vibration measurements through optical flow analysis of a video sequence", *Measurement*, Vol. 95, pp. 166–172.

Ehrhart, M. and Lienhart, W. (2015a), "Monitoring of Civil Engineering Structures using a State-of-the-art Image Assisted Total Station", *Journal of Applied Geodesy*, Vol. 9 No. 3, pp. 174–182.

Ehrhart, M. and Lienhart, W. (2015b), "Development and evaluation of a long range image-based monitoring system for civil engineering structures", in Shull, P.J. (Ed.), *Proc. SPIE Structural Health Monitoring and Inspection of Advanced Materials, Aerospace, and Civil Infrastructure*, Vol. 9437, San Diego, California, United States, p. 94370K.

Feng, D., Feng, M., Ozer, E. and Fukuda, Y. (2015), "A Vision-Based Sensor for Noncontact Structural Displacement Measurement", *Sensors*, Vol. 15 No. 7, pp. 16557–16575.

Feng, D. and Feng, M.Q. (2015), "Model Updating of Railway Bridge Using In Situ Dynamic Displacement Measurement under Trainloads", *Journal of Bridge Engineering*, Vol. 20 No. 12, pp. 1–12.

Fukuda, Y., Feng, M.Q., Narita, Y., Kaneko, S. and Tanaka, T. (2013), "Vision-Based Displacement Sensor for Monitoring Dynamic Response Using Robust Object Search Algorithm", *IEEE Sensors Journal*, Vol. 13 No. 12, pp. 4725–4732.

Guizar-Sicairos, M., Thurman, S.T. and Fienup, J.R. (2008), "Efficient subpixel image registration algorithms", *Optics Letters*, Vol. 33 No. 2, pp. 156–158.

Guo, J. and Zhu, C. (2016), "Dynamic displacement measurement of large-scale structures based on the Lucas–Kanade template tracking algorithm", *Mechanical Systems and Signal Processing*, Vol. 66–67, pp. 425–436.

Hester, D., Brownjohn, J., Bocian, M. and Xu, Y. (2017), "Low cost bridge load test: calculating bridge displacement from acceleration for load assessment calculations", *Eng. Struct.*, Vol. 143, pp. 358–374.

Ji, Y.F. and Chang, C.C. (2008a), "Nontarget Image-Based Technique for Small Cable Vibration Measurement", *Journal of Bridge Engineering*, Vol. 13 No. 1, pp. 34–42.

Ji, Y.F. and Chang, C.C. (2008b), "Nontarget Stereo Vision Technique for Spatiotemporal Response

Measurement of Line-Like Structures”, *Journal of Engineering Mechanics*, Vol. 134 No. 6, pp. 466–474.

Jianbo Shi and Tomasi. (1994), “Good features to track”, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94*, IEEE Comput. Soc. Press, pp. 593–600.

Khaloo, A. and Lattanzi, D. (2017), “Pixel-wise structural motion tracking from rectified repurposed videos”, *Structural Control and Health Monitoring*, Vol. 24 No. 11, p. e2009.

Khuc, T. and Catbas, F.N. (2017a), “Computer vision-based displacement and vibration monitoring without using physical target on structures”, *Structure and Infrastructure Engineering*, Vol. 13 No. 4, pp. 505–516.

Khuc, T. and Catbas, F.N. (2017b), “Completely contactless structural health monitoring of real-life structures using cameras and computer vision”, *Structural Control and Health Monitoring*, Vol. 24 No. 1, p. e1852.

Lee, J.J., Cho, S., Shinozuka, M., Yun, C., Lee, C.-G. and Lee, W.-T. (2006), “Evaluation of Bridge Load Carrying Capacity Based on Dynamic Displacement Measurement Using Real-time Image Processing Techniques”, *Steel Structures*, Vol. 6 No. 5, pp. 377–385.

Liao, W.Y., Chen, W.H., Ni, Y.Q. and Xia, Y. (2010), “Development of a vision-based real-time displacement measurement system for Guangzhou New TV Tower”, in Casciati, F. and Giordano M. (Eds.), *Proceedings of the 5th European Workshop on Structural Health Monitoring*, Sorrento, Naples, Italy, pp. 450–455.

Lowe, D.G. (2004), “Distinctive image features from scale-invariant keypoints”, *International Journal of Computer Vision*, Springer, Vol. 60 No. 2, pp. 91–110.

Macdonald, J.H.G., Dagless, E.L., Thomas, B.T. and Taylor, C.A. (1997), “Dynamic measurements of the Second Severn Crossing”, *Proceedings of the Institution of Civil Engineers - Transport*, Vol. 123 No. 4, pp. 241–248.

Massart, D.L., Kaufman, L., Rousseeuw, P.J. and Leroy, A. (1986), “Least median of squares: a robust method for outlier and model error detection in regression and calibration”, *Analytica Chimica Acta*, Vol. 187 No. C, pp. 171–179.

Oh, B.K., Hwang, J.W., Kim, Y., Cho, T. and Park, H.S. (2015), “Vision-based system identification technique for building structures using a motion capture system", *Journal of Sound and Vibration*, Elsevier, Vol. 356, pp. 72–85.

Ojio, T., Carey, C.H., OBrien, E.J., Doherty, C. and Taylor, S.E. (2016), "Contactless Bridge Weigh-in-Motion", *Journal of Bridge Engineering*, American Society of Civil Engineers, Vol. 21 No. 7, p. 4016032.

Ribeiro, D., Calçada, R., Ferreira, J. and Martins, T. (2014), "Non-contact measurement of the dynamic displacement of railway bridges using an advanced video-based system", *Engineering Structures*, Vol. 75, pp. 164–180.

Rublee, E., Rabaud, V., Konolige, K. and Bradski, G. (2011), "ORB: An efficient alternative to SIFT or SURF", *Proceedings of the IEEE International Conference on Computer Vision*, IEEE, pp. 2564–2571.

Stephen, G.A., Brownjohn, J.M.W. and Taylor, C.A. (1993), "Measurements of static and dynamic displacement from visual monitoring of the Humber Bridge", *Engineering Structures*, Elsevier, Vol. 15 No. 3, pp. 197–208.

Wahbeh, A.M., Caffrey, J.P. and Masri, S.F. (2003), "A vision-based approach for the direct measurement of displacements in vibrating systems", *Smart Materials and Structures*, Vol. 12 No. 5, pp. 785–794.

Wang, N., O'Malley, C., Ellingwood, B.R. and Zureick, A.-H. (2011), "Bridge Rating Using System Reliability Assessment. I: Assessment and Verification by Load Testing", *Journal of Bridge Engineering*, Vol. 16 No. 6, pp. 854–862.

Xu, Y., Brownjohn, J.M.W., Hester, D. and Koo, K.Y. (2017), "Long-span bridges: Enhanced data fusion of GPS displacement and deck accelerations", *Engineering Structures*, Vol. 147, pp. 639–651.

Yang, Y., Dorn, C., Mancini, T., Talken, Z., Kenyon, G., Farrar, C. and Mascareñas, D. (2017), "Blind identification of full-field vibration modes from video measurements with phase-based video motion magnification", *Mechanical Systems and Signal Processing*, Elsevier, Vol. 85, pp. 567–590.

Ye, X.W., Ni, Y.Q., Wai, T.T., Wong, K.Y., Zhang, X.M. and Xu, F. (2013), "A vision-based system for dynamic displacement measurement of long-span bridges: algorithm and verification", *Smart Structures and Systems*, Vol. 12 No. 3_4, pp. 363–379.

Yoon, H., Elanwar, H., Choi, H., Golparvar-Fard, M. and Spencer, B.F. (2016), "Target-free approach for vision-based structural system identification using consumer-grade cameras", *Structural Control and Health Monitoring*, Vol. 23 No. 12, pp. 1405–1416.

Zhao, X., Ri, K. and Wang, N. (2017), "Experimental Verification for Cable Force Estimation Using Handheld Shooting of Smartphones", *Journal of Sensors*, Vol. 2017, pp. 1–13.

TABLE CAPTIONS

Table 1 Tracking accuracy at 95% confidence interval for three tracking methods in laboratory conditions.

Tracking accuracy (95% confidence interval)	Correlation-based template matching	LK optical flow estimation	SIFT
Run 1 (pixel)	0.026±0.055	0.029±0.069	0.048±0.149
Run 2 (pixel)	-0.009±0.020	-0.001±0.020	0.011±0.037
Accuracy range (pixel)	[0.01, 0.08]	[0.02, 0.10]	[0.03, 0.20]

Table 2 Working performance of three tracking methods in field tests

Field conditions	Correlation-based template matching	LK optical flow estimation	SIFT
Long-range measurement	✓	✓	✓
Low-contrast patterns	✓	×	✓
Small pattern changes	×	✓	✓
Large pattern changes	×	×	×
Lighting changes	×	×	×

FIGURE CAPTIONS

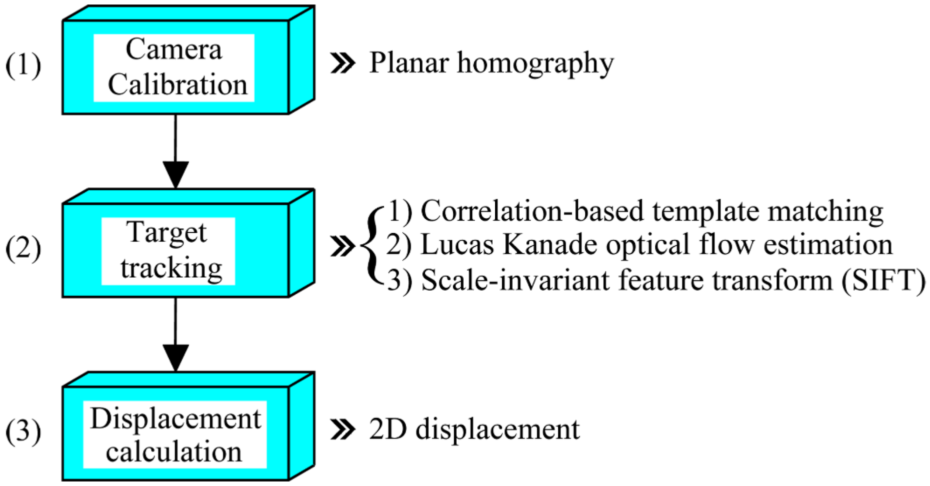

Figure 1 Procedures and methodologies in custom-developed video processing software package for structural displacement measurement.

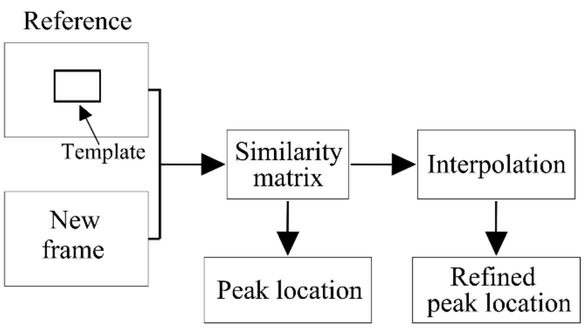

Figure 2 Procedures of one target tracking method: correlation-based template matching.

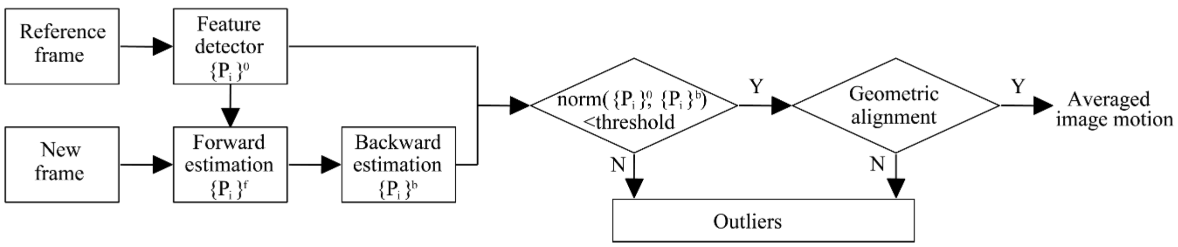

Figure 3 Procedures of one target tracking method: Lucas Kanade (LK) optical flow estimation.

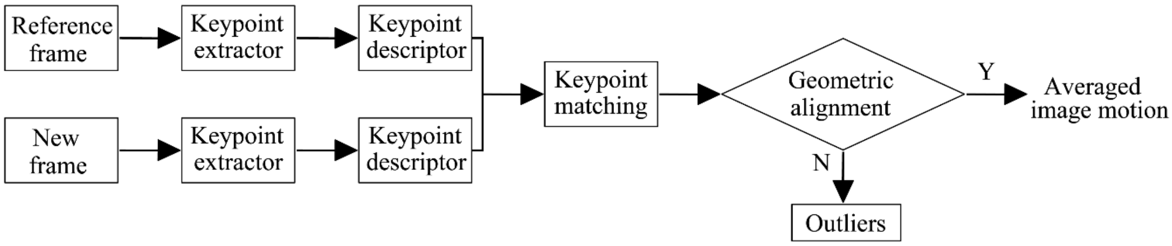

Figure 4 Procedures of one target tracking method: scale-invariant feature transform (SIFT).

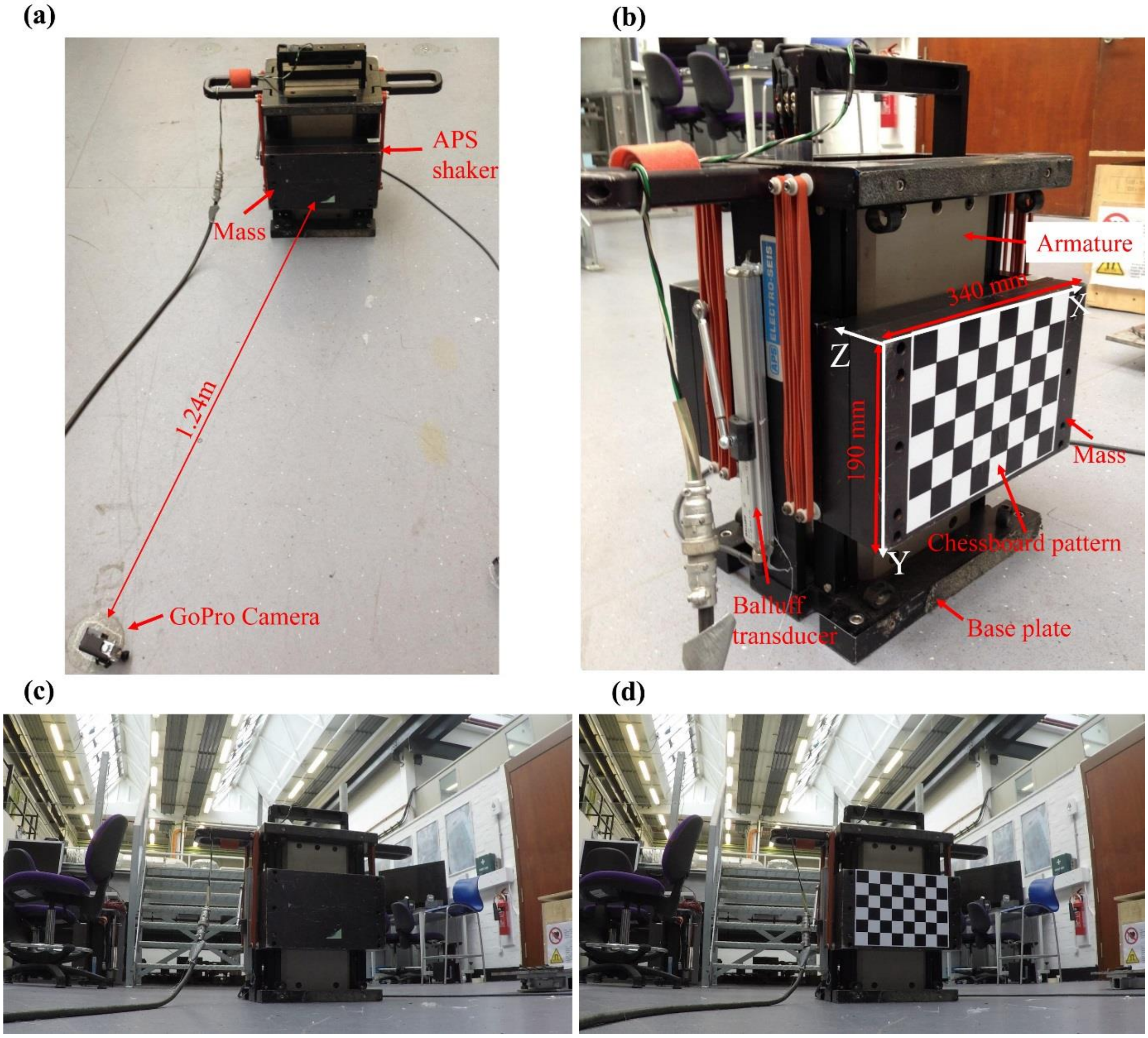

Figure 5 Test configuration of a vision-based system for vertical oscillation measurement of an APS shaker in laboratory: (a) test configuration of shaker and GoPro camera in Run 1; (b) test configuration of shaker in Run 2 with chessboard patterns attached to shaker mass; (c) one sample frame from the GoPro video recorded in Run 1; and (d) one sample frame from the GoPro video recorded in Run 2.

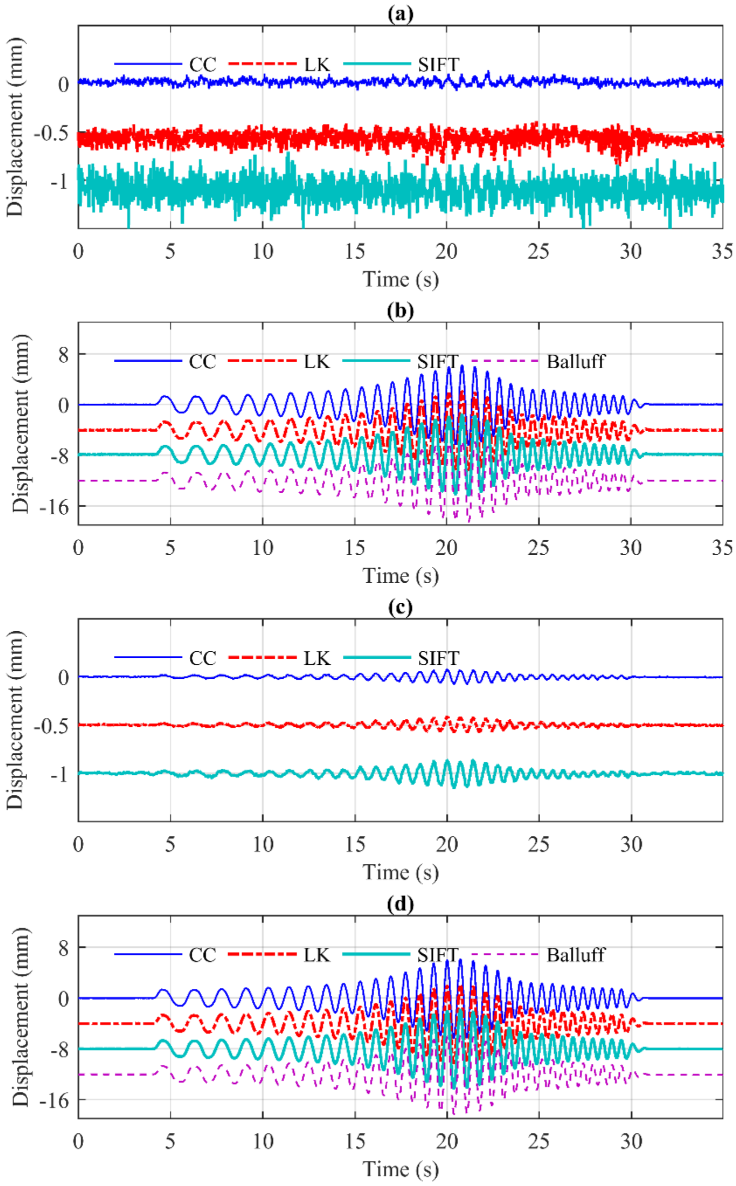

Figure 6 Time histories of displacement measurement of shaker mass acquired by the custom-developed video processing software package: (a) measured displacement in the horizontal direction by vision-based system in Run 1; (b) measured displacement in the vertical direction by vision-based system and Balluff transducer in Run 1; (c) measured displacement in the horizontal direction by vision-based system in Run 2; and (d) measured displacement in the vertical direction by vision-based system and Balluff transducer in Run 2. (Legends CC, LK and SIFT denote the three target tracking methods used for video processing, namely correlation-based template matching, Lucas Kanda optical flow estimation and Scale-invariant feature transform.)

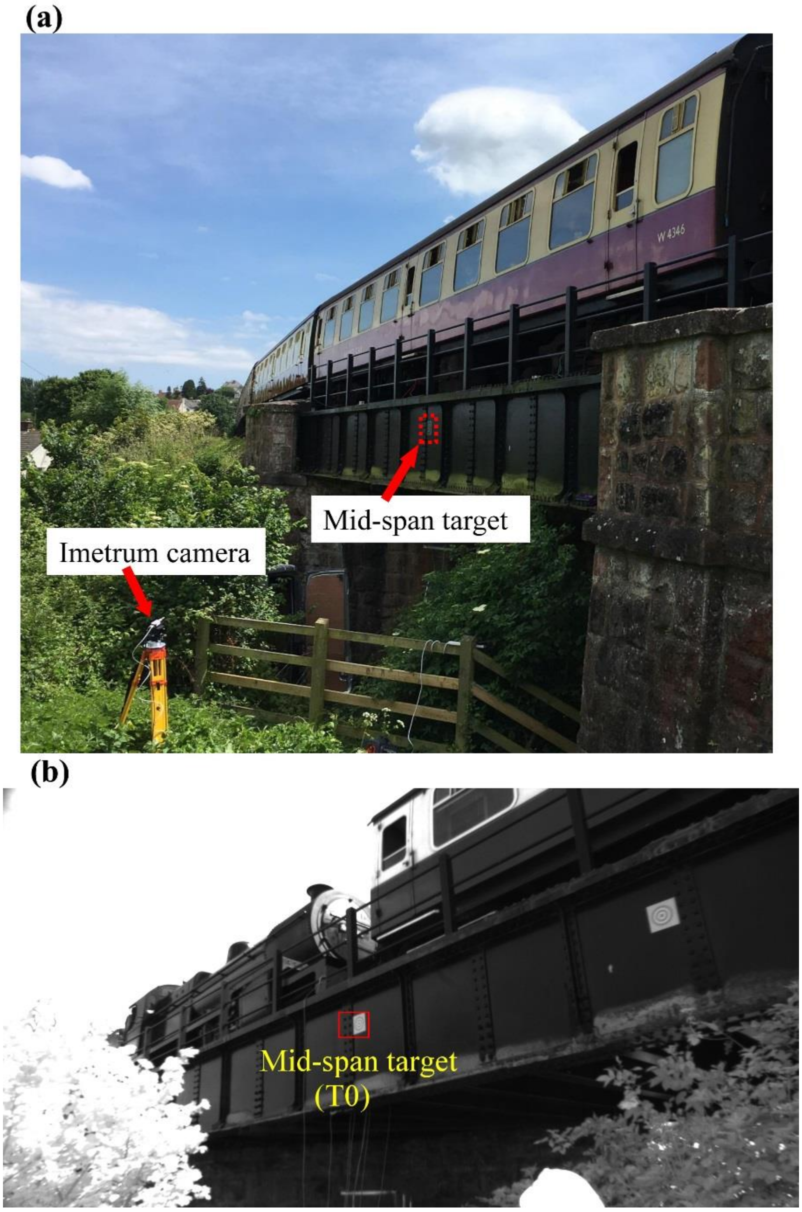

Figure 7 Test configuration of a vision-based system for the mid-span displacement measurement on a railway bridge in Somerset, UK: (a) camera setup near the bridge and the target region at mid-span selected for video tracking; and (b) one sample frame from the recorded video when one steam train passed through the bridge.

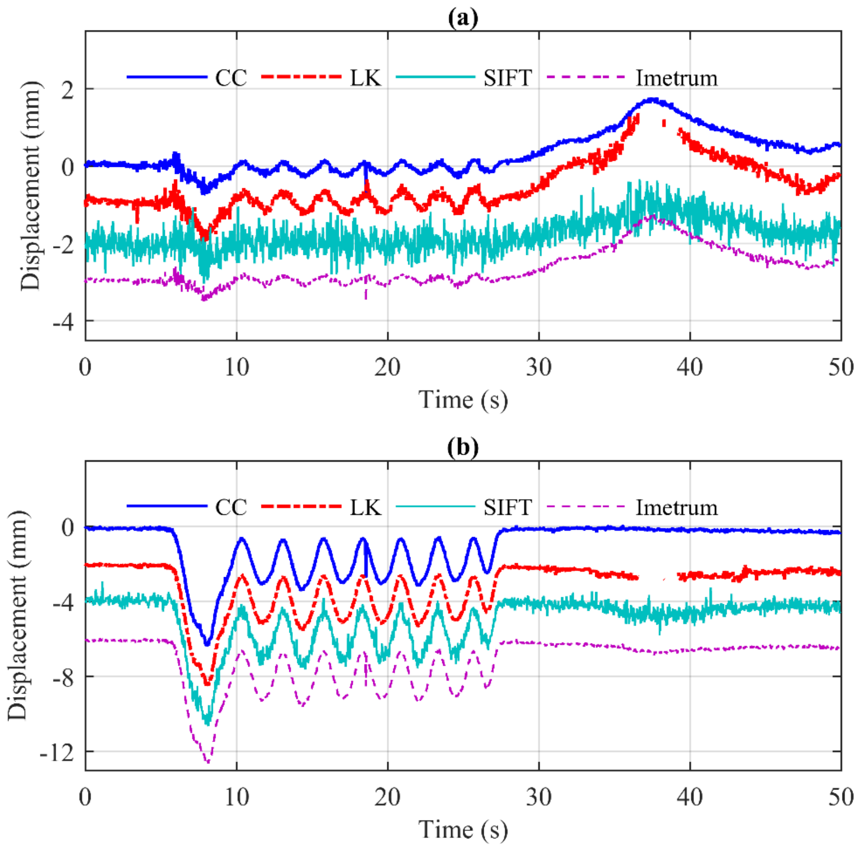

Figure 8 Time histories of displacement measurement acquired by the Imetrum system and by the custom-developed video processing software package using three tracking methods: (a) displacement measurement in the bridge longitudinal direction; and (b) displacement measurement in the vertical direction. (Legends CC, LK and SIFT denote the three target tracking methods used for video processing, namely correlation-based template matching, Lucas Kanda optical flow estimation and Scale-invariant feature transform.)

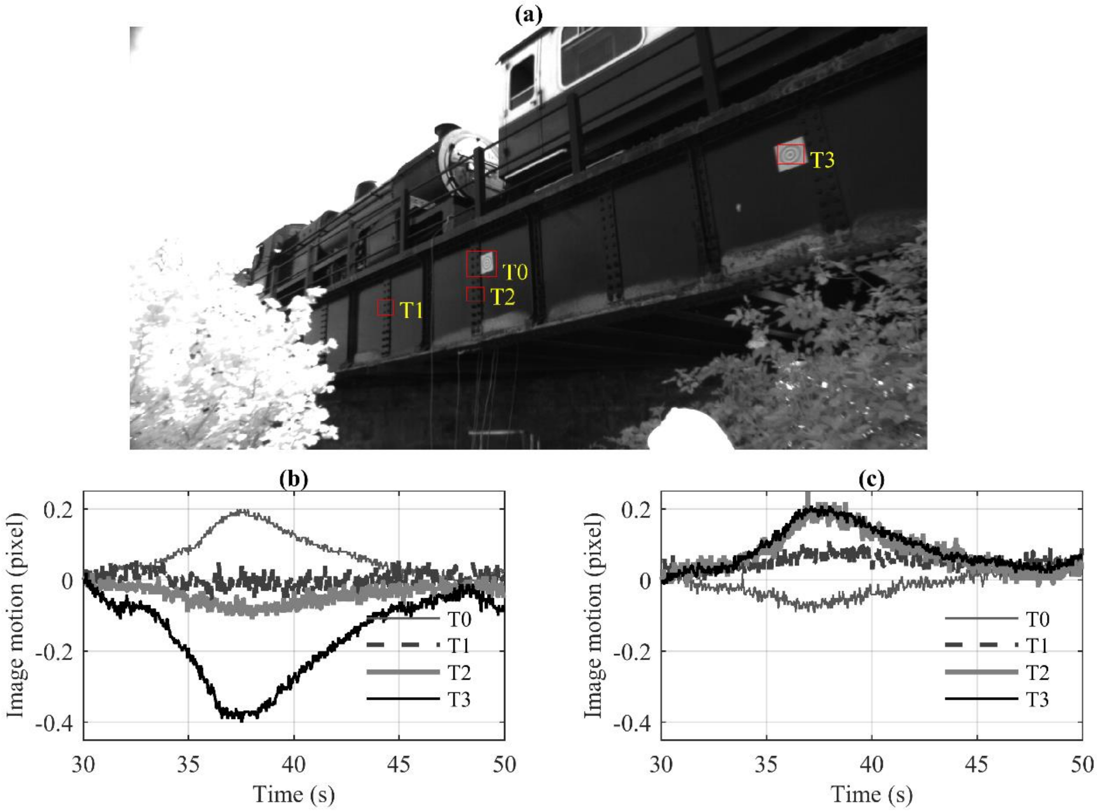

Figure 9 Tracking results of image motions at the selected four target regions (T0~T3) in the bridge girder by the correlation-based template matching: (a) locations of four target regions in the frame selected for tracking; (b) measured image motions at four target regions along the image horizontal direction; and (c) measured image motions at four target regions along the image vertical direction.

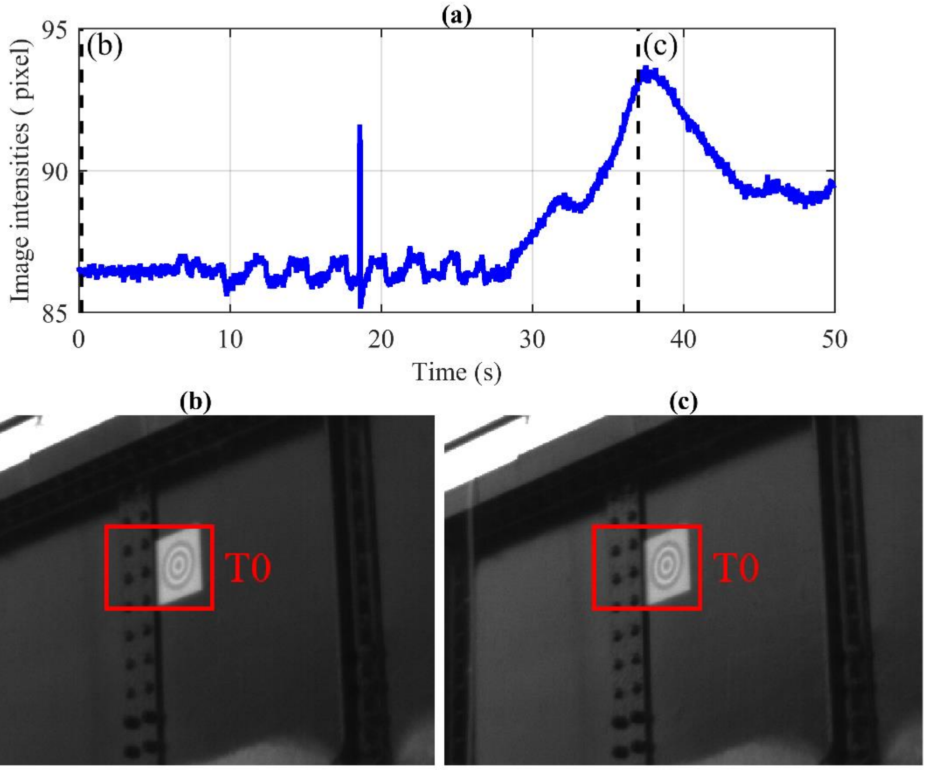

Figure 10 Variations of image brightness at the initially selected target region T0: (a) time history of mean pixel intensity at the target region T0; (b) the truncated initial frame with a rectangular annotation at the target region T0; and (c) the truncated frame at the time step of 37 s with a rectangular annotation at the target region T0.

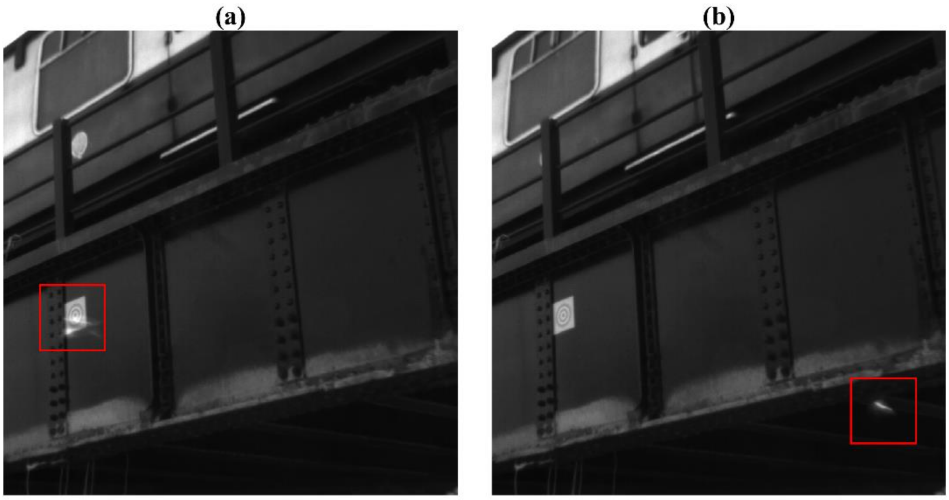

Figure 11 Two consecutive frames from video files at approx. 18.6 s indicating the changes of target pattern due to a flying object (frames truncated and zoomed-in for clarification).

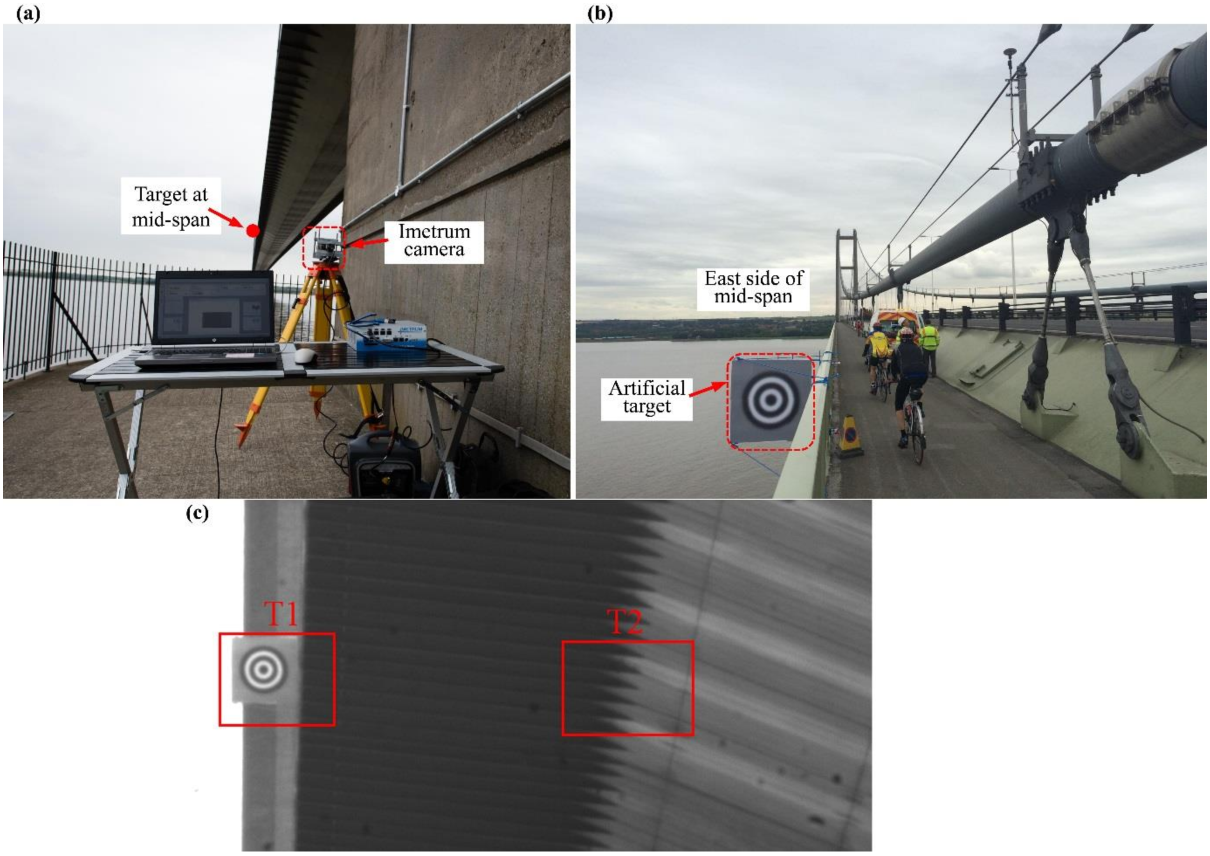

Figure 12 Test configuration of a vision-based system for the mid-span displacement measurement on the Humber Bridge, UK (Xu et al., 2017): (a) camera setup near the bridge tower; (b) an artificial target installed at the east side of mid-span; and (c) one sample frame from the recorded videos with the annotations at two selected target regions for video tracking.

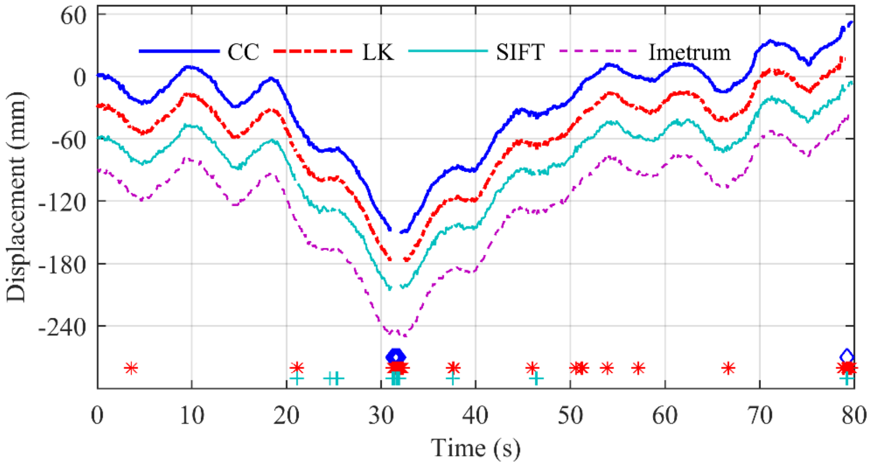

Figure 13 Time histories of displacement measurement in the vertical direction at the target region T1 acquired by Imetrum system and by the custom-developed video processing software package using three tracking methods with the markers indicating the time steps of tracking failure. (Legends have the same meaning as in Figure 8.)

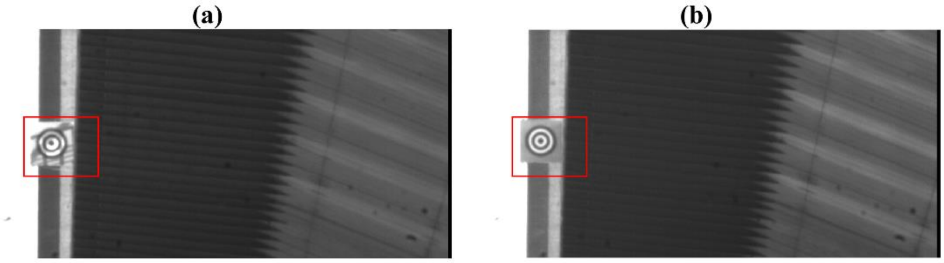

Figure 14 Two frames from the video file at approx. 31.3 s indicating the changes of target pattern due to the passing of one tall vehicle at the mid-span of the bridge that temporally blocked the sunlight, making the whole target pattern visible in the image.

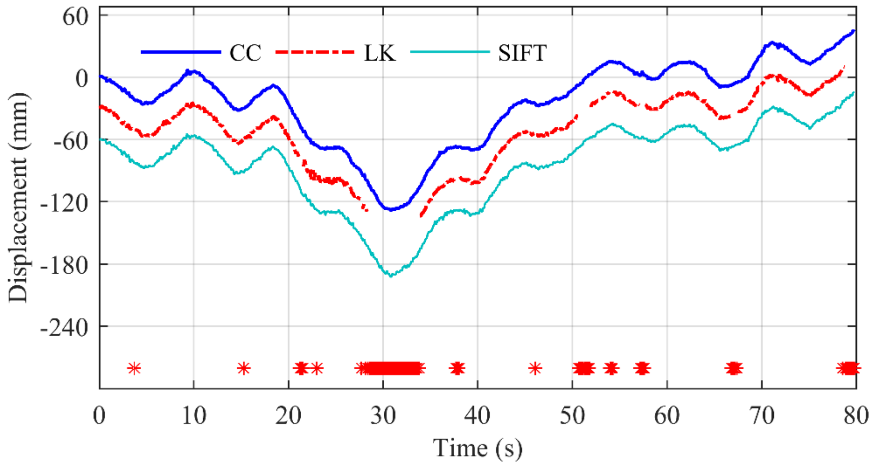

Figure 15 Time histories of displacement measurement in the vertical direction at the target region T2 acquired by the custom-developed video processing software package using three tracking methods with the markers indicating the time steps of tracking failure.