Chapter 3
A Wirelessly Time-Synchronized Vision-Based System Network for Non-Contact Structural Health Monitoring

Miaomin Wang, Ki Young Koo, Fuyou Xu, and James Brownjohn

Abstract Structural health monitoring (SHM) projects often require numerous sensors across entire structures, leading to significant costs and complex installations. Vision-based systems offer a promising alternative by capturing data remotely across a wide field of view. However, these systems are limited by measurement inaccuracies over large areas and in restricted zones due to a shallow depth of field, which complicates their application in tasks requiring high-accuracy displacement measurements of large-scale structures and input-output measurements. This study proposes the use of multiple time-synchronized vision systems to address these challenges. We introduce the Flexible Vision Network (FVN), a wireless network of vision-based systems where each node's clock is synchronized with satellite atomic clocks via GPS modules. This synchronization allows for flexible deployment without limitations on spatial arrangement or node quantity. Each node in the FVN utilizes a template matching algorithm for structural displacement measurements. The FVN's effectiveness was demonstrated through 3D displacement measurement and modal testing on operational bridges.

Keywords Structural health monitoring · Computer vision · Wireless sensor · Time synchronization · Modal testing

Introduction

Vision-based systems provide a contactless alternative to traditional sensors, enabling the collection of input load and structural response data for structural health monitoring (SHM) [1]. They have therefore gained popularity in the SHM field, especially for measuring structural displacements in bridges, buildings, and dams. This is achieved through digital image processing techniques such as template matching [2–4], feature tracking [5], optical flow [6], and motion magnification [7–9].

Another SHM application involves detecting vehicles and pedestrians on bridges to estimate loads. For instance, Pan et al. [10] developed a vision-based framework to detect and track vehicles on the Jiangyin Bridge in China. However, accurately estimating load requires more than just position data. This has led to methods that combine vision-based detection with weigh-in-motion (WIM) systems and tire-road contact models [11, 12].

Single-camera systems face a trade-off between field of view (FOV) and measurement accuracy, which limits their effectiveness in large, multi-task scenarios. Multi-camera systems help address this issue but require time synchronization and reliable data transmission. Basic synchronization methods involve post-processing with audio cues [13] or overlapping FOVs [14], while more precise techniques use wireless triggers like radio frequency or stroboscopic lights. Unfortunately, temperature-sensitive crystal oscillators in cameras can introduce timing uncertainties [15]. Wired camera systems using GPIO and Ethernet cables allow synchronized data transmission but are limited by the processing capacity of the master computer and cable length [16].

Wireless networks like the SmartVision system [17] eliminate cable restrictions by using radio waves for synchronization. However, scaling these networks can be challenging due to bandwidth constraints and spatial limitations on node placement.

Miaomin Wang · Ki Young Koo · James Brownjohn
Faculty of Environment, Science and EconomyUniversity of Exeter, UK
e-mail: miaominwang@hotmail.com; k.y.koo@exeter.ac.uk; j.brownjohn@exeter.ac.uk

Fuyou Xu
School of Infrastructure EngineeringDalian University of Technology, China
e-mail: fuyouxu@hotmail.com

Janko Slavic et al. (eds.), *Computer Vision & Laser Vibrometry, Vol. 6*, Conference Proceedings of the Society for Experimental Mechanics Series, https://doi.org/10.13052/97887-438-0151-1_3

Synchronizing with other sensor types (e.g., accelerometers) is also difficult because of differences in time protocols across systems.

This paper introduces the Flexible Vision Network (FVN), a wireless vision sensor network that uses GNSS-based time synchronization, eliminating master node dependency. It offers flexibility in spatial deployment and supports multi-task SHM applications.

Flexible Vision Network (FVN)

Hardware of an FVN node

In FVN, all nodes share the same basic hardware. As shown in Figure 1, each FVN node comprises an image capture device, an image processing unit, a GNSS module, and other peripheral components. These nodes are designed to be both portable and economical.

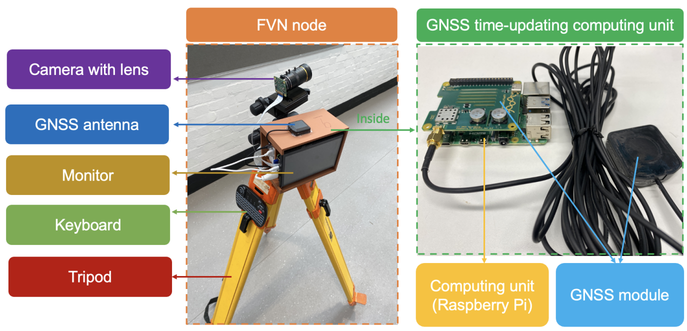


Fig. 1 Hardware of a FVN node.

(1) Image Capture: The system uses the Raspberry Pi High Quality Camera, which provides a high resolution of 4056 $\times$ 3040 pixels at an affordable cost. It includes a variable focal length manual-zoom lens (8–50 mm), allowing customization for different applications. For capturing high-speed motion, the system uses the Raspberry Pi Global Shutter Camera, which avoids the motion blurs of rolling shutter cameras by capturing the entire frame at once.
(2) Processing Unit: The Raspberry Pi 4B is the main computing component, chosen for its balance of processing power and cost-efficiency. This model has a Broadcom BCM2711 Quad-core Cortex-A72 64-bit SoC and 8GB of RAM, which is sufficient for real-time image processing tasks.
(3) GNSS Module: For precise timekeeping, the setup includes a Uputronics Raspberry Pi GPS/RTC Expansion Board with a GNSS antenna. The board connects to the Raspberry Pi through UART and pulse-per-second (PPS) connections.
(4) Display and Power Supply: A compact 7-inch monitor is used for display. A 50,000 mAh power bank supplies power to the Raspberry Pi, GNSS module, and monitor. All components, except the camera, are enclosed in a specially printed case.
(5) Control Interface: Each node is operated using a wireless keyboard and mouse.

Internal clock updating of FVN nodes

Each Raspberry Pi independently aligns its internal clock with accurate atomic clocks in satellites. This setup enables all nodes to synchronize using a common external time reference, ensuring inter-node synchronization.

(1) A GNSS antenna, placed with a clear view of the sky, receives signals from multiple satellites. These signals include time data encoded in National Marine Electronics Association (NMEA) sentences and pulse-per-second (PPS) signals.
(2) These GNSS signals are then sent to the GNSS board via a cable, where the NMEA sentences and PPS signals are extracted.
(3) The NMEA sentences are transmitted to the Raspberry Pi through the UART pin, while the PPS signals are delivered via the PPS pin.
(4) Within the Raspberry Pi, a Network Time Protocol (NTP) daemon processes the NMEA sentences and PPS signals to obtain time information.

(5) The NTP daemon compares the time data with the Raspberry Pi's local clock and, if necessary, adjusts the local clock to match the GNSS time.

Through this configuration, each Raspberry Pi in the FVN nodes synchronizes with the satellite clocks, allowing the nodes to remain synchronized with each other. A consistent process is followed for image capture and timestamping: when a camera image arrives at the Raspberry Pi, it is timestamped and encoded into a video. This setup ensures that each node records videos with precise timestamps aligned to the satellite clocks. During data extraction from these videos, each piece of extracted data is time-stamped accurately. The data are then resampled to achieve alignment.

Image processing of FVN nodes

FVN nodes are primarily designed to measure structural displacement. The camera needs to be placed nearly perpendicular to the target on the structure. A scaling factor (SF) is applied to convert pixel displacement into actual physical displacement. This factor is calculated by dividing a known physical length by the pixel length in the image of the target.

To obtain structural displacement from the video, a zero-mean normalized cross-correlation coefficient (ZNCC) algorithm is used. The authors' previous paper describes this algorithm's workflow in detail [2].

Verification On Time Synchronisation Accuracy

In this test, two FVN nodes were used to measure the time difference between displacement signals from nodes with different setups. This was done to evaluate their synchronization performance. Two types of cameras, rolling shutter and global shutter, were used, with specifications listed in Table 1. The test setup is shown in Figure 2, and the node configurations are in Table 2.

Table 1 The specifications of the two cameras used in this study.

Camera	**Sensor**	**Shutter type**	**Used resolution**	**Pixel size**	**Price (as of 2024)**
Raspberry High Quality Camera	Sony IMX477	Rolling shutter	1280*720	1.55 μm	£50
Raspberry Global Shutter Camera	Sony IMX296	Global shutter	1280*720	3.45 μm	£50

Table 2 FVN node configurations in this validation test.

Case	**Sensor**	**Frame rate**	**Processing mode**
1	Rolling shutter	20 Hz	Real time
2	Rolling shutter	100 Hz	Post processing
3	Global shutter	50 Hz	Post processing

The two nodes measured the dynamic displacement of a shared target on a laboratory cantilever. Although indoors, the nodes received GNSS signals through the lab's partially glass roof. The camera frame rate was set to 20 Hz for real-time processing in Cases 1 and increased to 100 Hz for video recording in Cases 2. In Cases 3, the global shutter camera operated at 50 Hz.

The nodes began measuring at different times, and then the cantilever was manually excited to induce free decay vibration. The time difference between the displacement signals from the two nodes was calculated, and results are shown in Table 3. This test was repeated five times.

Table 3 Time difference between displacement signals from the two nodes in different cases.

Case	**Time difference in various test trials (Unit: ms)**					**Mean**
	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	
1	1.77	2.58	2.27	2.37	2.15	2.23
2	2.26	1.87	2.76	2.67	2.41	2.39
3	0.01	0.12	0.24	0.06	0.10	0.11

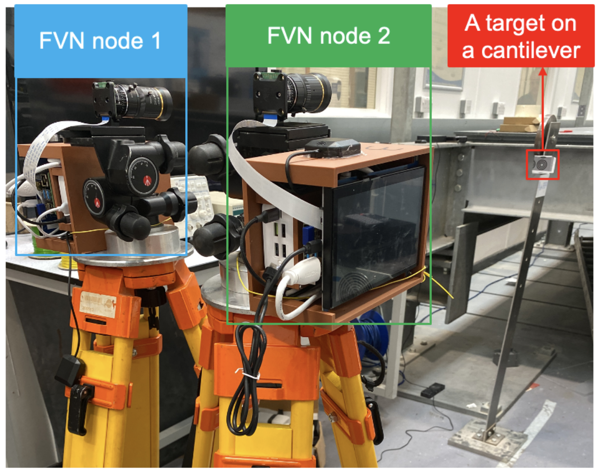


Fig. 2 The setup of the validation test.

Results showed that the camera type significantly affected the time difference between signals output by the two nodes. With the rolling shutter camera, the maximum time difference was 2.76 ms, and the average time differences were over 2 ms. With the global shutter camera, time differences were under 0.2 ms, and the maximum time difference was 0.24 ms.

Applications Of FVN

3D displacement measurement of a cross-sea bridge

Measuring 3D displacement of bridges typically requires a stereo vision system. However, as the measurement distance increases, accuracy in depth decreases. For long-span bridges, like measuring at midspan, a large baseline between the two cameras is necessary to maintain accuracy in depth. This setup poses significant challenges for camera communication with each other and with the computing unit.

Using FVN nodes, this task can be simplified. In this approach, two FVN nodes are used: one node measures horizontal and vertical displacements at midspan, and the other measures the longitudinal displacement of the bridge girder.

This method was applied to measure 3D displacement at the midspan of the Xinghaiwan Bridge, a cross-sea bridge in Dalian, China. It is a double-layer steel truss suspension bridge, with a main span of 460 m and side spans of 180 m, as shown in Figure 16a. The bridge tower has a height of 114.31 m.

During the measurement, the lower crossbeam of one bridge tower was used as a stationary area. Two FVN nodes were placed on this crossbeam, as shown in Figures 3(a) and 3(b). One node measured horizontal and vertical displacements, while the other measured longitudinal displacement at the end of the bridge girder. The longitudinal displacement along the girder was assumed to be consistent, as the rigid body movement was much larger than any extension in the longitudinal direction.

FVN Node 1, equipped with a 300 mm lens, was positioned 227 m from the midspan, with its optical axis aligned parallel to the longitudinal direction. The scaling factor was determined to be 3.718 mm/pixel. Since it was not feasible to install artificial markers under the main girder, the white drainage pipe at the midspan was selected as the target, as shown in Figure 3(b).

FVN Node 2, configured with a 120 mm lens, had its optical axis perpendicular to the bridge's longitudinal direction. It measured the longitudinal displacement of the main girder using bolts near the bridge girder's support as the target. This node was located 6.414 m from the target, with a calculated scaling factor of 0.263 mm/pixel.

The experiment took place on the morning. FVN Node 1 started at 09:29:36, while FVN Node 2 started at 09:17:33. Both systems were time-synchronized via their GNSS antennas. Data from the overlapping 100-min period were analyzed, with displacement signals shown in Figure 4. The test results confirmed the feasibility of using two FVN nodes for 3D displacement measurement on long-span bridges.

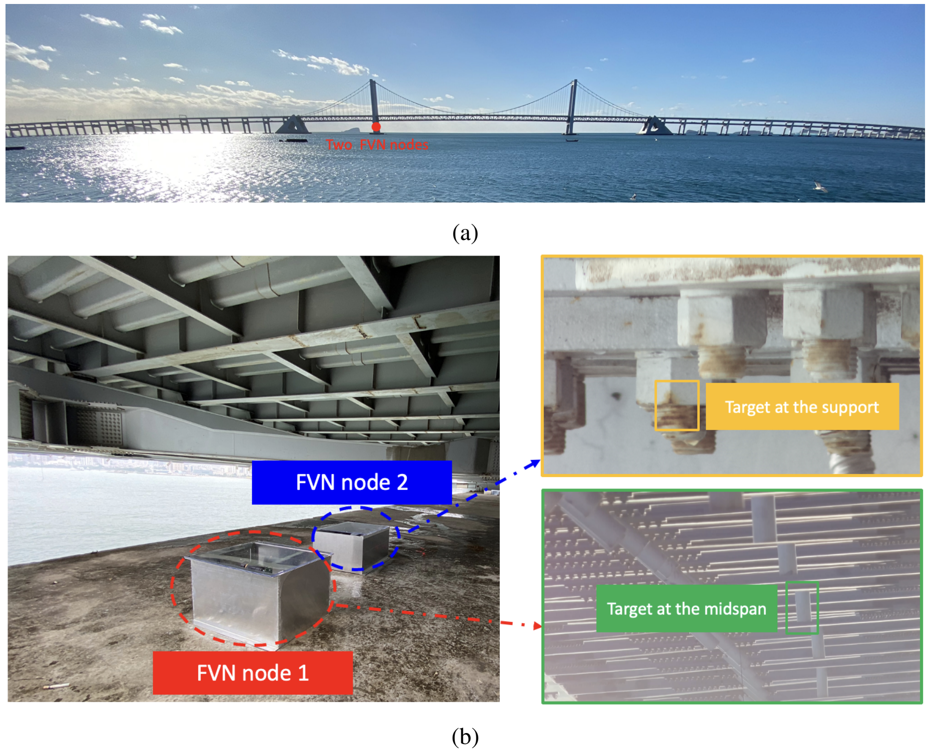


Fig. 3 The setup of the Xinghaiwan Bridge measurement.

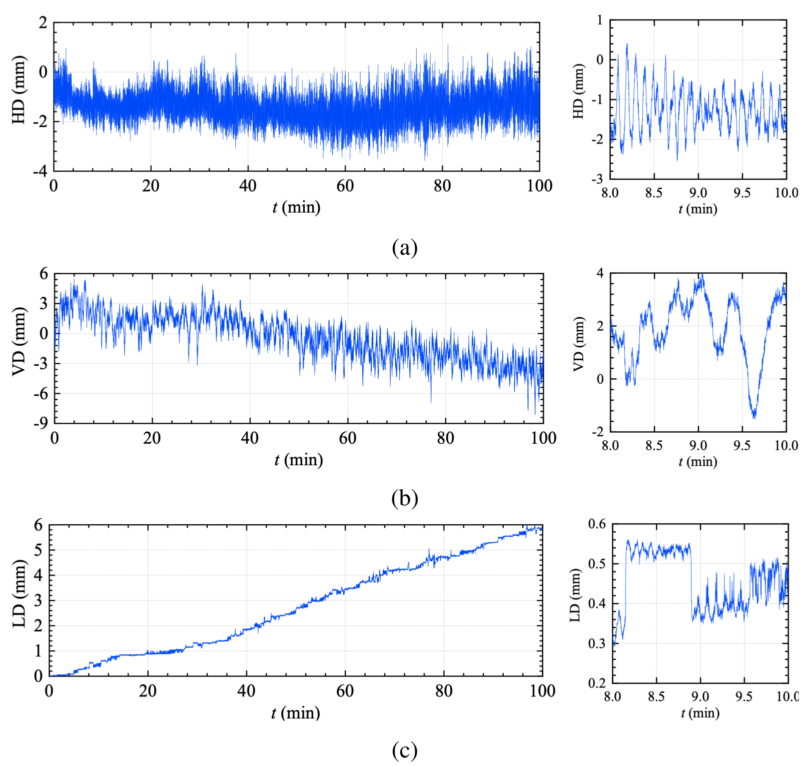
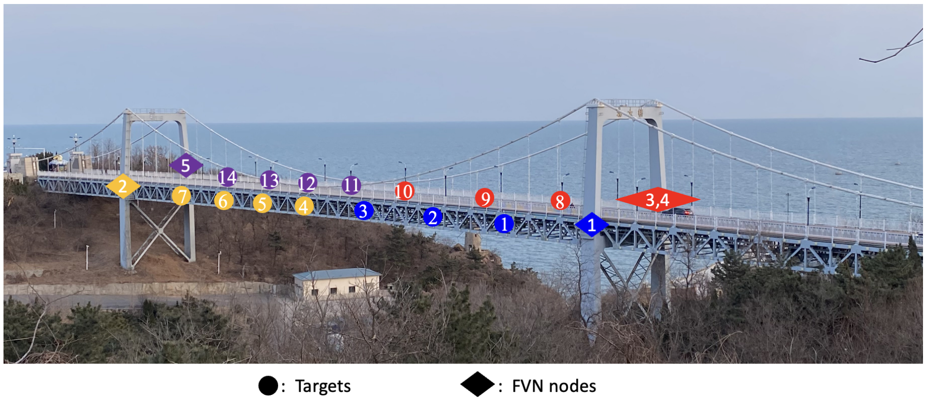


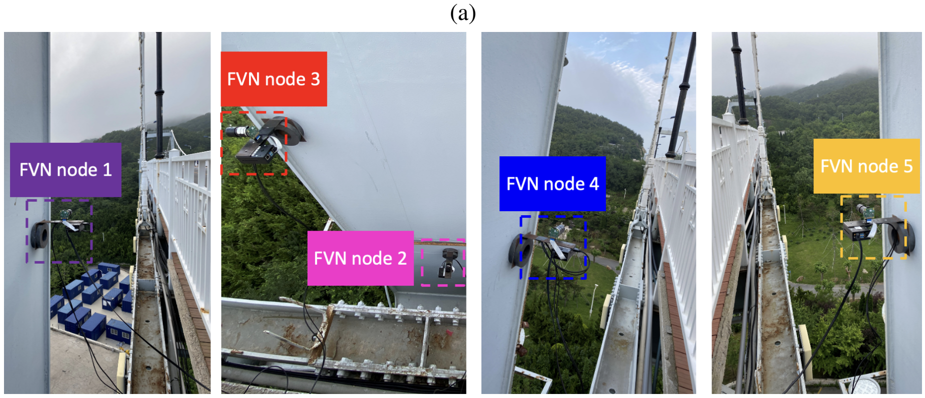
Fig. 4 Measurement result of the Xinghaiwan Bridge: (a) horizontal displacement (HD), (b) vertical displacement (VD), and (c) longitudinal displacement (LD).

3D modal test of a suspension bridge

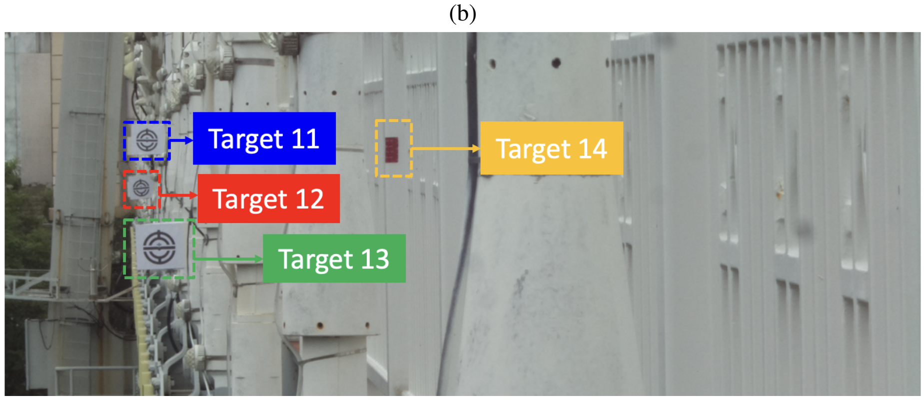
A 3D modal test was conducted on the Beida Bridge, a vehicle suspension bridge in Dalian, China, using five FVN nodes. Measurement points were set on both sides of the main girder at seven 1/8-span sections, as shown in Figure 5(a). Taking advantage of the portability and easy installation of FVN nodes, these five nodes were placed on the bridge tower columns. Each camera was aligned to points of the same color, as illustrated in Figure 5(b). Artificial targets were used in this test, as shown in Figure 5(c). Wireless time synchronization was achieved through the GNSS modules in each system.



(a)



(b)



(c)

Fig. 5 The setup of the modal testing of the Beida Bridge (a) devices layout, (b) FVN nodes, and (c) targets.

A heavy truck crossed the bridge four times, exciting various vibration modes. The truck drove over one of the two lanes, inducing both vertical bending and torsional modes. Using the SSI-COV algorithm, the first five vertical bending modes and the first three torsional modes were identified, with frequencies and mode shapes displayed in Figure 6.

Another modal test on the Beida Bridge was conducted in 2017, providing mode frequencies and shapes for comparison. Table 4 compares the previous test results with the frequencies identified in this study, where “S”, “A”, and “V” denote

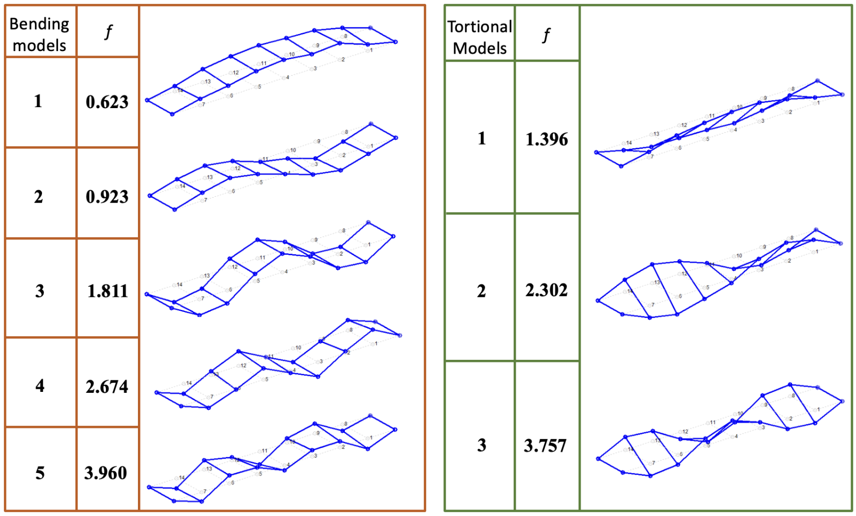


Fig. 6 Identified models of Beida Bridge.

Table 4 Identified mode frequencies based on test data at different time (Unit: Hz).

Models	**Test in 2017**	**This study**	**Difference**
SV1	0.63	0.62	1.59%
AV1	0.88	0.92	4.54%
SV2	–	1.81	–
AV2	–	2.67	–
SV3	–	3.96	–
ST1	–	1.40	–
AT1		2.30	–
ST2	–	3.76	–

symmetric, antisymmetric, and vertical bending of the girder, respectively. The mode frequencies from this test are close to the previous results. These findings demonstrate the effectiveness of FVN in 3D modal analysis of long-span bridges.

Conclusion

This study introduces a wireless vision sensor network, the Flexible Vision Network (FVN), for structural health monitoring (SHM). The key findings are:

(1) When two FVN nodes, each using a rolling shutter camera (Raspberry Pi High Quality camera), measured a common target, the time difference between the displacement signals was less than 2.76 ms. With global shutter cameras (Raspberry Pi Global Shutter cameras), the difference stayed below 0.24 ms. These results indicate significant time-stamping uncertainty in rolling shutter cameras.

(2) The use of two FVN nodes enables accurate 3D displacement measurement over long distances, making it practical for monitoring long-span bridges.

(3) It is feasible to use multiple FVN nodes to capture bending and tortional structural models of operational long-span bridges.

Acknowledgments The authors are grateful for the financial support of the Engineering and Physical Sciences Research Council (EPSRC), UK, via the Programme Grant EP/W005816/1.

References

1. M. Wang, W.K. Ao, J. Bownjohn, F.Y. Xu, A novel gradient-based matching via voting technique for vision-based structural displacement measurement, Mech Syst Signal Pr, 171 (2022).
2. M. Wang, K.Y. Koo, C. Liu, F. Xu, Development of a low-cost vision-based real-time displacement system using Raspberry Pi, Eng Struct, 278 (2023).
3. D. Feng, M.Q. Feng, Experimental validation of cost-effective vision-based structural health monitoring, Mech Syst Signal Pr, 88 (2017) 199–211.
4. Y. Xu, J. Zhang, J. Brownjohn, An accurate and distraction-free vision-based structural displacement measurement method integrating Siamese network based tracker and correlation-based template matching, Measurement, 179 (2021) 109506.
5. C. Dong, F.N. Catbas, A non-target structural displacement measurement method using advanced feature matching strategy, Advances in Structural Engineering, 22 (2019) 3461–3472.
6. D. Gorjup, J. Slavič, A. Babnik, M. Boltežar, Still-camera multiview Spectral Optical Flow Imaging for 3D operating-deflection-shape identification, Mech Syst Signal Pr, 152 (2021) 107456.
7. E. Cai, Y. Zhang, X. Lu, P. Li, T. Zhao, G. Lin, W. Guo, A target-free video structural motion estimation method based on multi-path optimization, Mech Syst Signal Pr, 198 (2023) 110452.
8. J.G. Chen, T.M. Adams, H. Sun, E.S. Bell, O. Büyüköztürk, Camera-based vibration measurement of the world war I memorial bridge in Portsmouth, New Hampshire, Journal of Structural Engineering, 144 (2018) 04018207.
9. E. Cai, Y. Zhang, S.T. Quek, Visualizing and quantifying small and nonstationary structural motions in video measurement, Computer-Aided Civil and Infrastructure Engineering, 38 (2023) 135–159.
10. Y. Pan, D. Wang, Y. Dong, B. Peng, A Novel Vision-Based Framework for Identifying Dynamic Vehicle Loads on Long-Span Bridges: A Case Study of Jiangyin Bridge, China, IEEE Transactions on Intelligent Transportation Systems, 23 (2022) 10441–10457.
11. H. Wu, X. Wu, W. He, F. Wang, L. Deng, A computer vision-assisted method for identifying wheel loads of moving vehicles from dynamic bridge responses, Mech Syst Signal Pr, 197 (2023) 110372.
12. X. Kong, J. Zhang, T. Wang, L. Deng, C.S. Cai, Non-contact vehicle weighing method based on tire-road contact model and computer vision techniques, Mech Syst Signal Pr, 174 (2022) 109093.
13. C. Dong, S. Bas, F.N. Catbas, A completely non-contact recognition system for bridge unit influence line using portable cameras and computer vision, Smart Struct. Syst, 24 (2019) 617–630.
14. M. Bolognini, G. Izzo, D. Marchisotti, L. Fagiano, M.P. Limongelli, E. Zappa, Vision-based modal analysis of built environment structures with multiple drones, Automation in Construction, 143 (2022) 104550.
15. J.A. Rice, K. Mechitov, S.-H. Sim, T. Nagayama, S. Jang, R. Kim, B.F. Spencer Jr, G. Agha, Y. Fujino, Flexible smart sensor framework for autonomous structural health monitoring, Smart structures and Systems, 6 (2010) 423–438.
16. M.M. Wang, W.K. Ao, J. Bownjohn, F.Y. Xu, Completely non-contact modal testing of full-scale bridge in challenging conditions using vision sensing systems, Eng Struct, 272 (2022).
17. S.A. V. Shajihan, T. Hoang, K. Mechitov, B.F. Spencer Jr., Wireless SmartVision system for synchronized displacement monitoring of railroad bridges, Comput-Aided Civ Inf, 37 (2022) 1070–1088.