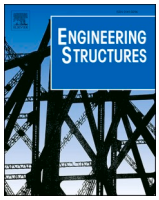
Engineering Structures 319 (2024) 118809

Contents lists available at ScienceDirect

Engineering Structures

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



Smart structural health monitoring using computer vision and edge computing

Zhen Peng [a], Jun Li [a,*], Hong Hao [a,b], Yue Zhong [a]

[a] Centre for Infrastructural Monitoring and Protection, School of Civil and Mechanical Engineering, Curtin University, WA 6102, Australia
[b] Earthquake Engineering Research and Test Center, Guangzhou University, Guangzhou, China

ARTICLE INFO

Keywords:
Edge computing
Computer vision
Displacement measurement
Vibration
Structural health monitoring
Cloud data storage

ABSTRACT

Structural health monitoring (SHM) provides real-time data on the condition and performance of infrastructure, enabling timely and cost-effective maintenance interventions, and hence enhanced safety and extended service life. The computer vision-based non-contact sensor has emerged as a promising alternative to conventional contact-type sensors for structural displacement measurement and SHM. Many of the currently reported vision-based structural displacement measurement systems typically temporarily set up a video camera from a distance to the structure. The collected images or videos are usually stored locally and post-processed offline to obtain structural displacement responses, which is cumbersome and limited to short-term SHM applications. The recent development of technologies empowered by the Internet of Things (IoT) and edge computing has enabled real-time video processing and analysis at the source, minimizing latency, reducing bandwidth requirements, and enabling prompt decision-making, thereby enhancing efficiency and responsiveness compared to traditional offline video recording and processing systems. In this paper, an edge computing vision-based displacement measurement system (EdgeCVDMS) is developed. Video recording, processing, and displacement response identification are entirely performed on an edge device integrated with the vision-based displacement tracking algorithm, thereby greatly reducing the amount of data transmitted to the cloud server. The feasibility and applicability of the developed sensing system are experimentally validated on a laboratory-scaled transmission tower structure. The proposed EdgeCVDMS is cost-effective, easily deployable, and of great potential to be applied for the condition assessment of a larger population of aging civil infrastructure.

1. Introduction

Structural health monitoring (SHM) systems are extensively utilized to offer real-time condition screening of civil engineering structures during both the construction and normal service stages [1–3]. The displacement response of civil engineering structures subjected to operational loads is an important metric for structural condition and performance assessment, as it is directly related to the structure's stiffness [4,5]. For high-rise building structures, the maximum horizontal displacement amplitude and the dynamic displacement responses can be adopted to estimate structural lateral stiffness [6]. Regarding bridge structures, the vertical displacement responses subjected to moving loads contain valuable information about the structural stiffness and load-carrying capacity [7,8]. It is reported that bridge transverse displacements play a crucial role in capturing critical changes in the serviceability of timber railroad bridges [4]. According to a survey of North American railroad bridge structural engineers, bridge deflection under live loading is a critical performance-related parameter used to determine the allocation of the limited maintenance budget to the bridges most in need [9].

Conventional contact or noncontact displacement sensors, like the linear variable differential transducer (LVDT) and laser displacement sensor, face several challenges in the field of SHM implementations [10]. Firstly, they require a stationary reference point, which can be difficult to find in the field [11,12]. Secondly, the distance between the structure and the sensor may be far beyond the measurement range of traditional displacement sensors. Consequently, a temporary scaffold need be set up underneath the structure to provide a fixed reference point for making relative displacement measurements [13]. Owning to innovations in computer vision and the rapid advances in image sensor resolution and efficiency, computer vision-based noncontact sensing has emerged as a promising alternative to conventional contact sensors for SHM [14–17].

* Corresponding author.
E-mail address: junli@curtin.edu.au (J. Li).

https://doi.org/10.1016/j.engstruct.2024.118809
Received 3 December 2023; Received in revised form 27 May 2024; Accepted 13 August 2024
Available online 21 August 2024

Vision systems capture vast amounts of video footage and utilize advanced video processing techniques to identify both surface-level cracks and structural displacement responses [18]. Computer vision-based SHM has significant advantages, including its affordability, ease of setup and operation, and its ability to extract displacement from interested points of the structure by setting up the camera at distances ranging from dozens to hundreds of meters [14].

Typically, the collected images or videos are stored locally for further in-depth analysis. This becomes infeasible and impracticable when applied to the long-term monitoring of large-scale structures. In particular, it can be cumbersome to copy the video and requires specialized knowledge/skills to post-process the video data to obtain bridge displacement responses in practical SHM applications. This issue limits the application of computer vision-based SHM methods to a wider range of infrastructure, especially those located in remote areas. Furthermore, a substantial memory is required to store the video files generated from continuous recording. The Internet of Things (IoT) technique can be a common solution that transmits sensor measurements to a cloud server for centralized processing [19]. Cloud architecture provides access to computation, storage, and even connectivity with ease of access. In SHM applications, structural dynamic response time series such as strain [20] and acceleration [21], as well as environmental variables like temperature, wind speed, and wind direction [22], can be effectively transmitted to the cloud server via wireless networks. However, the wireless network bandwidth may not support continuous capturing and transferring of high-resolution video streams to the cloud server.

To tackle the issue of transferring large volumes of video data over a wireless network, a more viable approach is to perform video and signal processing at the location where the data is generated, i.e., at the edge [23,24]. This entails transmitting only the identified displacement responses and high-level structural condition-related statistical features to the remote host or cloud server. Edge computing refers to the decentralized processing of data near the source, reducing latency and bandwidth usage by bringing computational power closer to the devices generating the data. The integration of edge computing principles into computer vision techniques has the potential to considerably diminish bandwidth requirements and data processing latency, leading to a more practical and cost-effective SHM solution [25]. Edge computing architectures have been accompanied by compact, low-cost embedded devices like Raspberry Pi and NVIDIA Jetson, enabling a lightweight way of data collection, signal processing, and decision-making. Raspberry Pi is a low-cost, single-board computer designed for general-purpose computing, including web browsing, programming, and media playback [26]. It has a relatively low processing power compared to Jetson, which makes it suitable for basic computing tasks. In contrast, Jetson is a family of embedded computing devices specifically designed for artificial intelligence (AI) and machine learning (ML) applications. It is equipped with a high-performance GPU and specialized hardware for accelerating AI and ML workloads, making it suitable for computationally intensive tasks such as object detection/tracking, image classification, and natural language processing. Therefore, the NVIDIA Jetson is more suitable for the in-situ deployment of computer vision-based structural displacement identification and SHM.

In this paper, a computer vision-based real-time structural displacement measurement system using computer vision techniques underlined by edge computing and IoT techniques is proposed. The video recording, computer vision-based structural displacement measurement as well as the signal processing are performed on the developed edge computing vision-based displacement measurement system (EdgeCVDMS). The objective is to maximize the potential of computer vision-based displacement identification techniques and extend the applicability to the SHM of a wider range of structures on a long-term basis by leveraging the advantage of a compact, cost-effective yet computationally powerful edge device. Compared with existing offline computer vision-based structural displacement measurement techniques, the main contributions of this work are summarized as follows:

1. Integration of edge computing and computer vision-based structural displacement identification algorithm: This paper presents a comprehensive framework that leverages an edge computing device to enable real-time processing and analysis of video data for displacement measurement. By moving computational tasks closer to the data source, the system achieves low-latency, real-time displacement measurement while reducing the bandwidth requirements and enabling rapid decision-making.
2. The framework optimizes the use of computational resources at the edge device: several representative computer vision-based displacement tracking algorithms are deployed on the edge device. Suggestions on the appropriate displacement tracking algorithm selection is made based on the accuracy and efficiency of the candidate algorithms.
3. The developed EdgeCVDMS incorporates Amazon Web Services (AWS) for data management and visualization. Structural displacement responses and edge device operating status information are sent to AWS in real-time. End users can remotely access these data and schedule a field structural inspection when an abnormal event is detected.
4. Experimental evaluation and validation: comprehensive experimental validations are conducted to validate the accuracy, efficiency, and continuous operating performance of the developed sensing system in a laboratory environment.

2. Design of EdgeCVDMS for real-time displacement measurement

An EdgeCVDMS for real-time displacement measurement is developed by employing an NVIDIA Jetson Nano edge device, a high-resolution camera with a zoom lens, a 4 G LTE internet dongle, a portable power station, as well as the AWS platform that enables data management and visualization. The general scheme of the developed EdgeCVDMS with hardware and software configuration is shown in Fig. 1. Fig. 1 consists of three phases. In phase 1, the developed edge device will be mounted on a static reference point and targeted to the region of interest (ROI) of the structure for acquiring the video stream. It should be noted that the presented approach is target-free, and uses the natural features as the measurement locations for tracking and matching to obtain the dynamic displacement responses. The selection of the optimal ROI or measurement locations is similar as the optimal sensor placement problem in SHM. Certain criteria, such as obtaining the strongest vibration response intensity, achieving the most independent mode shapes, and minimizing errors in system parameter identification, can be adopted [7]. For example, deflections at the quarter span, mid-span and three-quarter span of bridges or the horizontal displacements of building floors can be selected as the measurement locations. Camera calibration parameters, such as the distance and tilt angle between camera and target point are measured in phase 1. More detailed procedures for camera calibration and scale factor determination are introduced in Section 2.2.1. In phase 2, the computer vision-based displacement tracking algorithm is deployed at the edge device. In particular, the NVIDIA Jetson Nano is primarily responsible for capturing and analyzing the video to identify structural dynamic displacement and transmitting data to AWS for data management and visualization. The main advantage of the developed edge computing-based system over existing computer vision-based displacement measurement studies is that the recorded videos are processed in real-time on the edge device, thus ensuring immediate analysis. Consequently, the videos are not stored on board after processing. The AWS architecture design presented in phase 3 enables the asset owners to remotely access real-time measurements from bridges equipped with the proposed system and receive an alert when an abnormal event

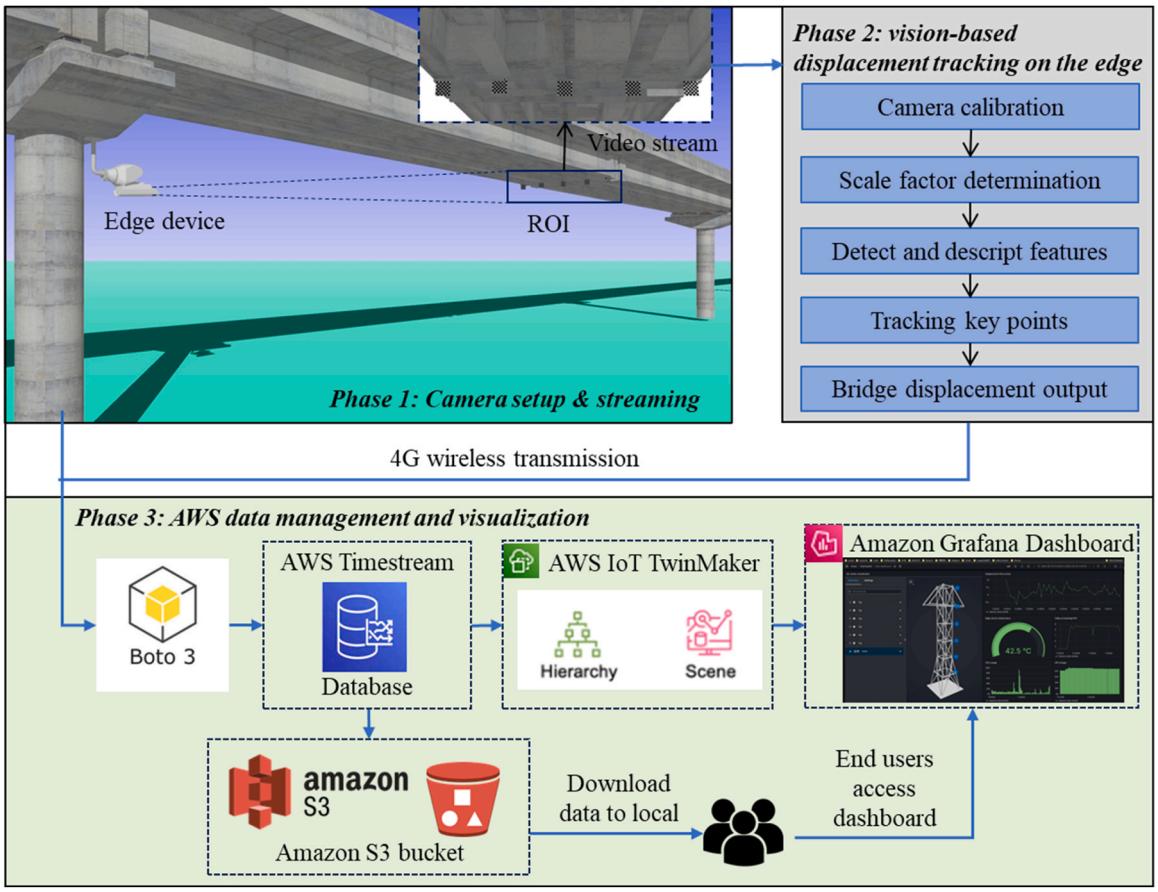


Fig. 1. The high-level workflow of the EdgeCVDMS for SHM.

occurs. Wireless transmission is achieved via 4 G signal provided by a USB internet dongle. The whole system is powered by a portable power station. In field applications, alternative sustainable power sources, such as solar panels and rechargeable batteries, can be adopted. As monitored using the command-line tool JTOP, the average energy consumption of the developed EdgeCVDMS during normal operation is about 5.76 W. The key technical issues in the design of the EdgeCVDMS, including hardware selection, motion tracking algorithm selection, cloud data management, and visualization, are particularly addressed in this section.

2.1. Hardware integration of the EdgeCVDMS

Fig. 2 shows the hardware configuration of the edge device prototype that is primarily responsible for video stream acquisition, video processing, and data transmission to the AWS. Major components include a video camera module, embedded computing boards (i.e., Nvidia Jetson Nano), a Micro SD flash memory card, and a USB 4 G LTE dongle.

The Nvidia Jetson Nano edge device is the most important hardware component of the proposed system, with the role of performing all the computations related to video recording, computer vision-based displacement identification, and signal processing. It is equipped with a quad-core ARM Cortex-A57 CPU running at 1.43 GHz, coupled with a 128-core Nvidia Maxwell GPU. It has 4 GB of LPDDR4 RAM, a microSD card slot for storage, and multiple I/O ports, including USB 3.0, USB 2.0, HDMI, Ethernet, and GPIO pins. The most distinct difference between Nvidia Jetson Nano and other available single-board computers, i.e., Raspberry Pi, is its powerful AI capabilities with its GPU architecture and dedicated hardware accelerators, enabling efficient processing of complex computation workloads directly at the edge.

The ArduCam IMX477 image sensor employed in this study is

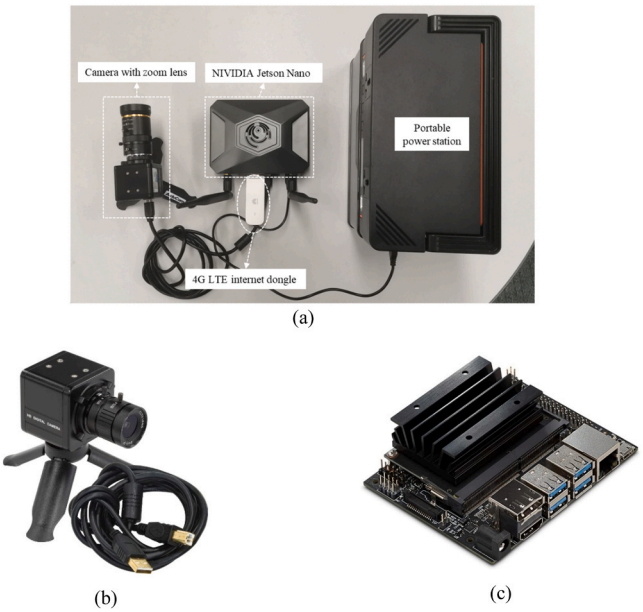


Fig. 2. Overview of the hardware of the EdgeCVDMS: (a) the developed system; (b) the used camera; and (c) the used Nvidia Jetson Nano.

manufactured by Sony company, which has 12.3 Megapixels and supports video streaming at frame rates of 10 fps@ 12MP, 20 fps@ 4 K, 20 fps@ 2 K, and 60 fps@ 1080 P. The pixel size of the ArduCam IMX477 is 1.55 μm × 1.55 μm. The ArduCam IMX477 camera is compatible with Nvidia Jetson Nano and is featured by a high resolution and a high

sensitivity for low-light environments. The camera is mounted in a customized industrial metal enclosure and comes with a 2 ft USB2.0 cable and a mini tripod stand. An Arducam zoom lens with an 8–50 mm adjustable focal length and C-mount is assembled with the ArduCam IMX477 camera for capturing and tracking structure vibration at different distances.

The component retail prices of the developed EdgeCVDMS are listed in Table 1. The overall hardware cost is about AU$700. The annual ongoing cost of AWS service and 4 G data is approximately AU$400 in the current market.

2.2. Computer vision-based structural displacement measurement algorithm

The video stream of the structure's ROI is captured by an associated camera module and is then processed by the computer vision-based displacement identification algorithm deployed on the edge device to obtain the displacement measurement results. The displacement identification algorithm is composed of four main steps [27]:

Step 1: Source video pre-processing and scale factor determination.

Step 2: Features detection and description in the ROI of the first frame. In the first frame, an ROI of the target structure with distinct corners and textures that stand out from the surrounding background is selected.

Step 3: Features matching and tracking in all subsequent frames. The feature points extracted from the subsequent frames are matched with the feature points in the first frame based on the similarity of the descriptors.

Step 4: Displacement conversion and post-processing. The displacement identification results are converted from subpixel values to physical displacements using the scale factor determined in Step 1. Interpolation is used to replace the very few '*NaN*' values for a better resolution.

The video processing is performed on a Linux system using the Python language and the OpenCV library.

2.2.1. Camera calibration and scale factor determination

Before motion tracking, it is essential to perform camera calibration to remove the radial distortion (especially in the case of a wide-angle lens in use) and determine the pixel scale factor SF. In this work, the camera calibration method proposed by Zhang [28] is adopted to obtain the camera lens distortion matrix γ. This can be conducted in the laboratory by placing a photo of a chessboard at several different angles to the camera's optical axis.

Given the fact that the measurement target on the civil structure is difficult to access in the field, the following two practical pixel scale factor *SF* determination methods are adopted:

(1) When the distance D and tilt angle θ between camera and target point are known, the pixel scale factor can be approximated by:

$$SF_1 = \frac{D}{f\cos^2\theta} d_{pixel} \tag{1}$$

where f and d_{pixel} denote the camera focal length and per pixel length (e.g., in μm/pixel), respectively. The pixel size of ArduCam IMX477 adopted in this study is 1.55 μm× 1.55 μm. In the situation of camera perpendicular to the target point, $\cos^2\theta = 1$ can be eliminated from Eq. (1). In field test, the distance D and tilt angle θ can be measured from laser rangefinder. This method is suitable for the camera equipped with a fixed focal length.

(2) When the vibration of the target on the structure is dominated by in-plane motion and there is pattern on structural surface with known dimension, a homography transformation can be adopted to establish a mapping between the image plane and the target plane. The homography transformation matrix H can be estimated with a minimum of 4 pairs of points on both planes. Then, the pixel scale factor can be approximated by:

$$SF_2 = \frac{d_{known}}{I_{known}} \tag{2}$$

where d_{known} and I_{known} represent the actual physical length of the pattern on structural surface and the corresponding pixel length at the image plane after homography transformation. This method is applicable for camera equipped with adjustable zoom lens [29].

2.2.2. Feature points detection and tracking

For in-situ civil structures under operational conditions, the target's movement remains relatively small in comparison to the extensive field of view (FOV). As a result, ensuring the precision of the computer vision-based displacement identification algorithm is crucial. Table 2 lists some of the representative feature point detection algorithms included in Python OpenCV 4.6.0. In Table 2, Scale-Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF), Accelerated-Kaze (AKAZE), and Binary Robust Invariant Scalable Keypoints (BRISK) algorithms can output sub-pixel level feature points detection accuracy. The Features from Accelerated Segment Test (FAST), Oriented FAST and Rotated BRIEF (ORB), Harris, and Shi-Tomas algorithms output pixel-level feature points or corners detection accuracy. In this study, the first four sub-pixel level algorithms are implemented to evaluate the accuracy and efficiency. For consistent and fair comparison, the Brute-Force (BF) matcher is employed to match the feature points detected from the first and subsequent video frames. It is noted that the accuracy of these pixel-level algorithms can be further refined by calculating the centroids of a bunch of pixels around the detected feature point. Some feature point detection algorithms, including SIFT, AKAZE and BRISK are not GPU-supported in OpenCV 4.6.0 and are not evaluated in this study. In the first frame, the detected feature points are sorted in the descending order according to the corresponding response value, and the highest quality feature point is selected to match with the feature points detected from subsequent frames.

2.2.3. displacement conversion

The displacement identification results are converted from subpixel image coordinates to physical displacements using the scale factor determined in Section 2.2.1. Interpolation is used to replace the very few NaN values for better resolution. The exception '*NaN*' occurs when the computer vision algorithm cannot match the feature points in a specific

Table 1
Price breakdown of the developed EdgeCVDMS in 2023.

Item	One-off expenses			Ongoing expenses	
	Nvidia Jetson Nano	Video camera module	4 G transmission module	Estimated AWS service	Data plan
Cost (AU $)	354	312	30	200 /year	200 /year

Table 2
Summary of feature point detection algorithms included in python OpenCV 4.6.0.

Feature point detection algorithms	CPU-support	GPU-support	Accuracy
SIFT	yes	no	Sub-pixel
SURF	yes	yes	Sub-pixel
AKAZE	yes	no	Sub-pixel
BRISK	yes	no	Sub-pixel
FAST	yes	yes	pixel
ORB	yes	yes	pixel
Harris	yes	yes	pixel
Shi-Tomas	yes	yes	Pixel

frame with the highest-quality feature point detected from the first frame. The root mean square error (RMSE) and normalized root mean square error (NRMSE) defined in Eqs. (3–4) are employed to quantify the accuracy of the computer vision-based displacement identification method [30]. RMSE and NRMSE are calculated as follows,

$$RMSE = \sqrt{\frac{\sum_{1}^{N} (d_{vision} - d_{laser})^2}{N}} \tag{3}$$

$$NRMSE = \frac{RMSE}{d_{laser}^{\max} - d_{laser}^{\min}} \tag{4}$$

RMSE is a quantitative measure of the average magnitude of error, with the same unit of the displacement measurement. The NRMSE is normalized by the difference between maximum and minimum displacement value.

2.3. AWS for data management and visualization

A large portion of infrastructures, i.e., bridges are located in remote area, which require wireless signal transmission to connect the data to the internet for remote access. AWS for edge computing provides scalable cloud resources, cost-effectiveness and global reach, complementing edge devices and enabling efficient data processing and analysis at the edge. It offers robust security features and seamless integration between edge devices and the cloud, enhancing performance and enabling sophisticated edge applications. Based on the above-mentioned advantage, AWS is employed in this study for data management and visualization. As indicated in Fig. 1, 4 G wireless communication is adopted to transmit the data from the edge device to the cloud server. There are other well-established wireless communication technologies, such as Wi-Fi, Bluetooth, ZigBee, and LoRa. However, considering the transmission range and distribution density of the base station, 4 G is the optimal solution among them.

In AWS, the Timestream service is utilized to create and configure a cloud database. The displacement identification results, along with the edge device's CPU usage, GPU usage, device temperature, and video processing speed information, are transmitted to the Timestream database via the Boto 3 software development kit. A 3D scene of the monitored structure is established in AWS IoT TwinMaker, where the Lambda function reads values from the Timestream database to the AWS IoT TwinMaker workspace. For data visualization, an Amazon Managed Grafana dashboard is created to provide real-time visualization of the 3D scene and data. The end user can remotely access the dashboard using the link provided in the page at the Amazon Managed Grafana workspace URL. Additionally, the end user can access and download data to a CSV file for further in-depth analysis via an Amazon S3 bucket. The web-based Amazon Managed Grafana dashboard is presented in Fig. 3, where the left side displays the 3D model of the transmission tower, and the blue icons along each floor represent the displacement output points. Thresholds, such as the maximum lateral displacement threshold, can be defined and attributed to each icon. When the measured displacement exceeds the pre-defined threshold, the icon color changes and sends a warning alarm to the end user. Apart from the displacement time series, the dashboard also visualizes device temperature, video processing FPS, CPU, and GPU usage information to monitor the work status of the edge device.

It is important to note that the AWS data management and visualization architecture detailed in Phase 3 of Fig. 1 is highly scalable. This scalability is achieved by deploying the developed EdgeCVDMS to a population of infrastructure, and the data sourced from different structures can be grouped into the AWS Timestream database for storage and visualization.

Fig. 3. Dashboard created to visualize displacement and edge device status data.

3. Experiment validation of the EdgeCVDMS

3.1. Experiment setup

In this section, a laboratory experiment is conducted to evaluate the accuracy and computation efficiency of the feature point detection algorithms listed in Table 2. The overall experiment setup is presented in Fig. 4, where a six-storey scaled transmission tower model is fixed on a shake table. The tower's overall height is 3.6 m with an even storey height of 0.6 m. Two Keyence IL-600 laser displacement sensors (#1∼#2) are placed on the left side to measure the first and second floor horizontal displacement responses of the transmission tower as the ground truth. The resolution of the Keyence IL-600 laser displacement sensor is provided as 50 μm in the product specifications. Additionally, six accelerometers are attached to each floor of the structure. The sampling rate of the laser displacement sensors and accelerometers is set as 200 Hz.

The video camera of the developed EdgeCVDMS is fixed on a tripod, approximately perpendicular to the structure at a distance of about 3.5 m. The camera's FOV is adjusted to cover the entire transmission tower. The region between the second and third floors of the structure, where the displacement ground truth is measured from the laser displacement sensor, is selected as the ROI. To excite the structure, ground motion with peak ground accelerations of 0.1 g and 0.2 g is respectively generated by the shake table. Videos with a resolution of 1920 × 1080 are captured and processed by the motion tracking algorithm deployed on the edge device. Since the dimension of the transmission tower is available, the pixel scale factor can be conveniently obtained from Eq. (2). The pixel scale factor of this camera setup is 1.9710 mm/pixel.

3.2. Video processing efficiency evaluation

To evaluate the computation efficiency, the video is cropped into five square ROIs with side lengths of 200, 300, 400, 500, and 600 pixels, respectively. The average video processing FPS of different motion tracking algorithms is presented in Fig. 5. It is found that SURF-CPU is the most efficient with respect to other video processing algorithms when the ROI size is 200 × 200. The efficiency of the SURF algorithm implemented on the GPU (SURF-GPU) of Nvidia Jetson Nano is slightly

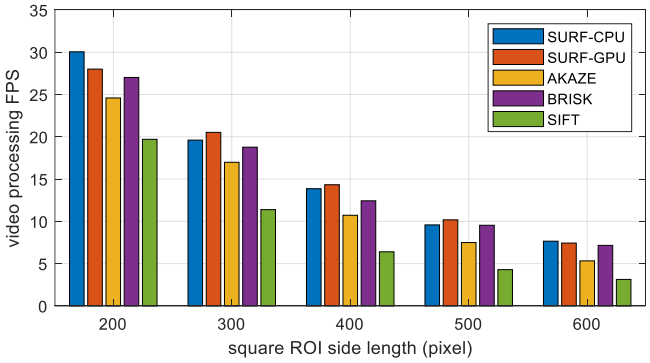


Fig. 5. Comparison of the video processing FPS of different tracking algorithms.

superior to the SURF-CPU and other three candidate algorithms when the ROI size is larger than 200 × 200. It is interesting to find out that the video processing speed of SURF-GPU does not substantially outperform SURF-CPU. One reason is that the system takes time to download the detected keypoint and motion tracking results from GPU to CPU during each iteration. Overall, it can be concluded that the video processing efficiency of the SURF algorithm substantially outperforms the other three candidate algorithms. It can process the video stream to obtain structural displacement responses with a rate higher than 25 Hz when the ROI size is 200 × 200 pixels.

3.3. Displacement measurement accuracy evaluation

The RMSE and correlation coefficient between the displacement measured from the laser displacement sensor #2 (ground truth) and different motion tracking algorithms are presented in Fig. 6 to quantify the accuracy. Overall, using SIFT shows the most accurate results among the other candidate algorithms. The displacement identification results of SURF-CPU and SURF-GPU are exactly the same. The performances of SIFT, SURF, and AKAZE algorithms are comparable and substantially better than the BRISK algorithm. The comparison results suggest that SURF-CPU and SURF-GPU are suitable to be deployed on the edge device with consideration of accuracy and efficiency. The video processing speed is above 25 FPS when the ROI is set to 200 × 200, which means

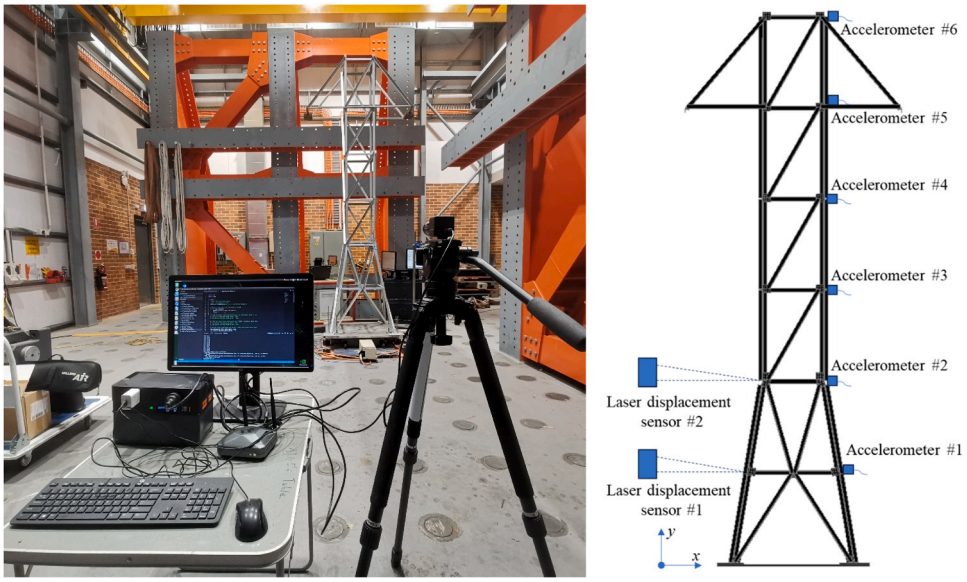


Fig. 4. (a) overall view of experiment setup; (b) laser displacement sensors and accelerometers placement.

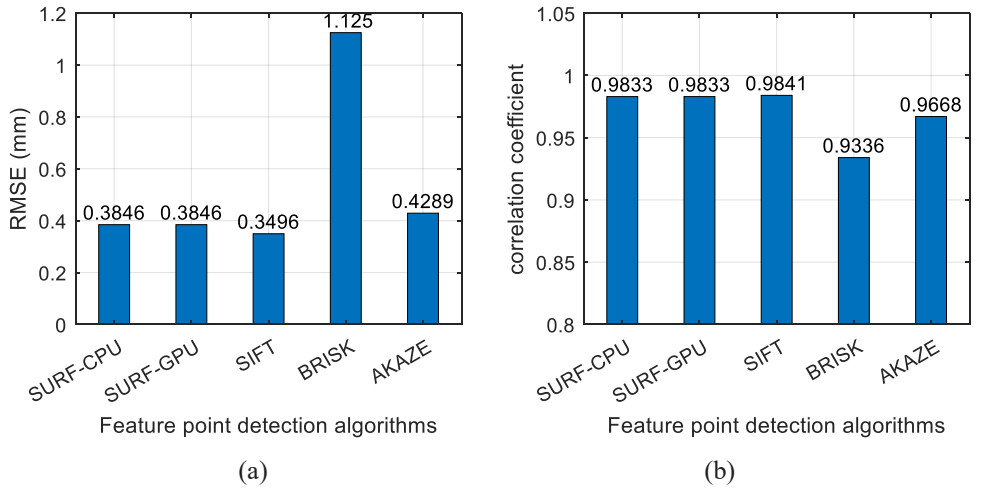


Fig. 6. Comparison of (a) the RMSE and (b) the correlation coefficient of different tracking algorithms.

that the structural static displacement and dynamic displacement information within 12.5 Hz can be captured by the edge device.

The displacement time history measured from the laser displacement sensor, and the SURF-CPU method is presented in Fig. 7. Since both systems work independently, a time-shift exists. In this study, cross-correlation analysis is used to determine and synchronize the time-shift between both signals. The results show a small variation between computer vision-based method and laser displacement sensor measurement. In particular, the correlation coefficient and RMSE between both time history is 0.9833 and 0.3846 mm, respectively; and the NRMSE is about 3.493 %. Thus, it can be concluded that the SURF-based displacement identification algorithm can be employed for accurately identifying the structural displacement. As mentioned in Section 3.1, the pixel scale factor of this camera setup is 1.9710 mm/pixel, which means that the RMSE of SURF algorithm employed in this study is within 0.2 (0.38426/1.9710 ≈0.1950) pixel. The accuracy of the developed EdgeCVDMS can be further improved by adjusting the zoom lens to limit the FOV from the whole structure to a smaller region, which will be further demonstrated in Section 3.4.

One limitation of computer vision-based displacement measurement methods compared to the laser displacement sensor is the sampling rate. Due to the limited computation and data read and write capacity of the edge device, the real-time video recording and processing speed is about 30 FPS when the ROI size is 200 × 200 pixels. To demonstrate that 30

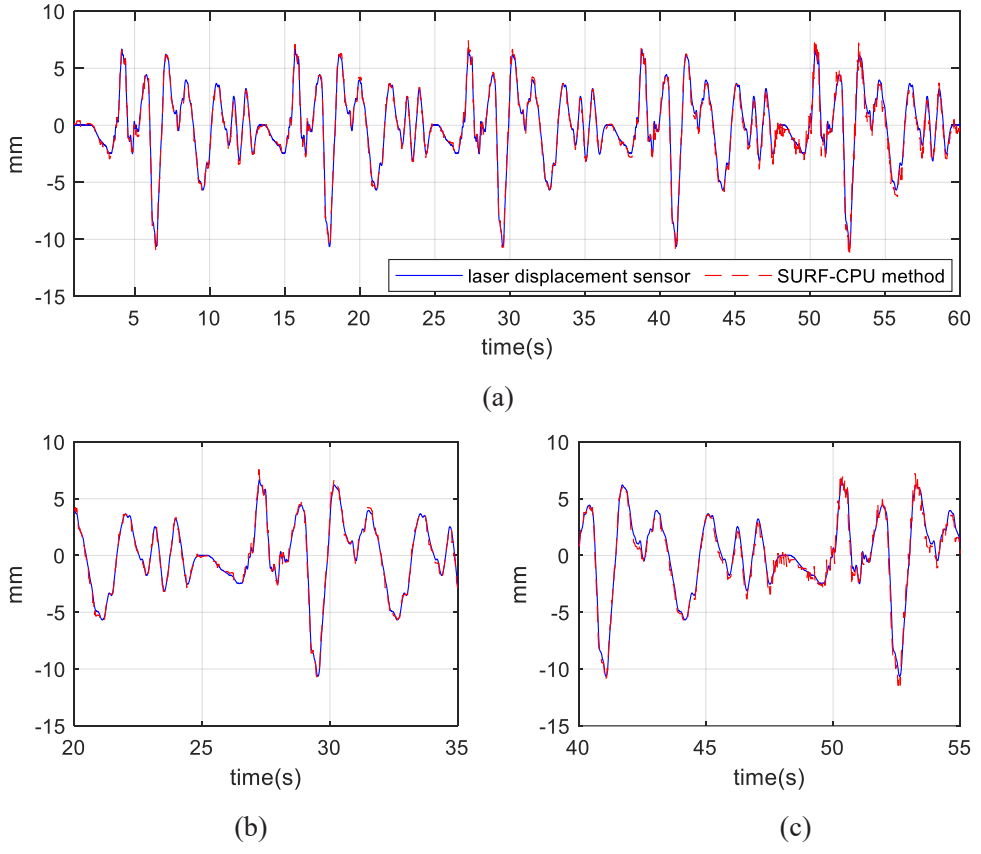


Fig. 7. (a) Displacement time history measured from the laser displacement sensor and the SURF-CPU method and (b-c) zoom-in view.

FPS is sufficient to capture the dynamic displacement of civil structures, Fig. 8 presents the power spectrum of computer vision-based displacement measurement, laser-based displacement measurement, and acceleration response subjected to earthquake excitation with a PGA of 0.2 g. It is observed that the power spectrum of computer vision-based displacement measurement aligns well with the power spectrum of laser-based displacement measurement. It is noted that the energy of displacement response is mainly distributed in the low-frequency range (0–4 Hz). In comparison, the energy of acceleration responses is mainly distributed in the relatively high-frequency range. This is because the transfer function between acceleration and displacement in the frequency domain is $FFT(u,\omega)/FFT(a,\omega) = -1/\omega^2$, where ω represents angular frequency. As a result, the amplitude of the high-frequency displacement component is very small. Considering that the energy of the displacement response mainly distributes in the low-frequency region, it can be concluded that the video processing speed of the developed EdgeCVDMS is able to capture the dynamic displacement of civil structures up to 15 Hz. The Jetson Nano can be replaced by higher-performance edge devices, such as the Jetson AGX Orin, to achieve faster video recording and processing speeds.

It should be noted that the above evaluation is performed in a controlled laboratory condition. The video processing speed of Nvidia Jetson Nano can be affected by the device temperature. Therefore, the practicability of the developed EdgeCVDMS should be further verified in the field on a long-term basis.

3.4. Influential factors analysis

In this section, additional tests are performed to evaluate the effects of lighting condition, shooting angle, and distance between the camera and target on the reliability and accuracy of the developed EdgeCVDMS. In Sections 3.2–3.3, it has been demonstrated that the SURF algorithm is appropriate in terms of accuracy and efficiency. Therefore, the SURF-CPU is employed to further evaluate the reliability of the influential factors. The RMSE and NRMSE are used to evaluate the computer vision-based displacement measurement accuracy. The plan view of the experimental setup is presented in Fig. 9. For camera setup 1, the camera is perpendicular to the target installed on the shake table, with a distance of about 4.2 m. In this test, the shake table is used to generate 20 cycles of sinusoidal excitations with an amplitude of 5 mm and a vibration frequency of 2 Hz. Considering that the lighting condition changes over time during field applications, three different lighting conditions, as shown in Fig. 10, are simulated by adjusting the lights in the laboratory environment.

In practical implementation, it may be difficult to ensure that the camera's optical axis is perpendicular to the target objective. For example, the camera can only be set up on the riverbank for a bridge that crosses the river. Therefore, it is necessary to evaluate the effect of the shooting angle on the displacement measurement accuracy. For camera setups 2 and 3, shooting angles of 30° and 45° are considered, respectively. The excitation in camera setup 2 is the same as camera setup 1.

In camera setup 4, the distance between the camera and target is changed to 15.3 m, which is the maximum allowable distance of the laboratory site condition. In this test, the amplitude of sinusoidal excitation is set to 5 mm, 4 mm, 3 mm, 2 mm, and 1 mm, respectively. The excitation cycles and frequency are the same as camera setups 1 and 2.

3.4.1. Effect of lighting condition on the displacement measurement accuracy

Fig. 11 presents the displacement time histories and the corresponding frequency spectrum measured from the laser displacement sensor, as well as those obtained using the SURF-CPU method, under three different lighting conditions. It can be observed that the displacements measured under different lighting conditions match well with the displacements captured by the laser displacement sensor in time and frequency domains, validating the accuracy of the proposed EdgeCVDMS. The RMSE values of the vibration displacement measurements corresponding to good, moderate, and bad lighting conditions are 0.3894 mm, 0.3782 mm, and 0.4923 mm, respectively. The NRMSE values of the vibration displacement measurements corresponding to good, moderate, and bad lighting conditions are 3.89 %, 3.78 %, and 4.92 %, respectively. The quantitative results indicate that the developed EdgeCVDMS can provide reliable results under different lighting conditions.

In fact, the lighting conditions mainly affect the number of feature points that can be detected. For example, the feature points detected

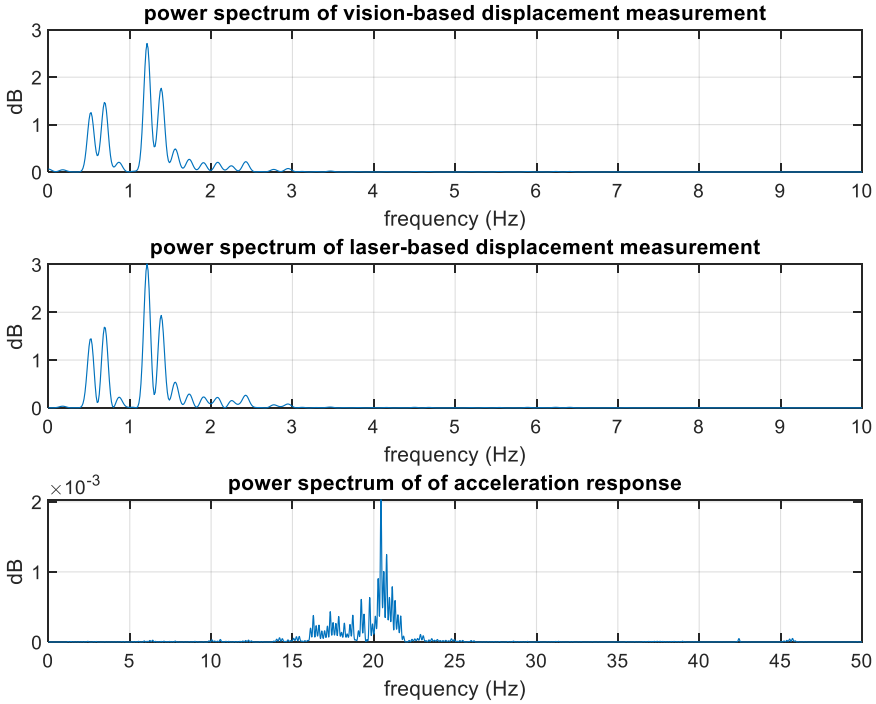


Fig. 8. Power spectrum of computer vision-based displacement measurement, laser-based displacmeent measreuemnt and acceleration response.

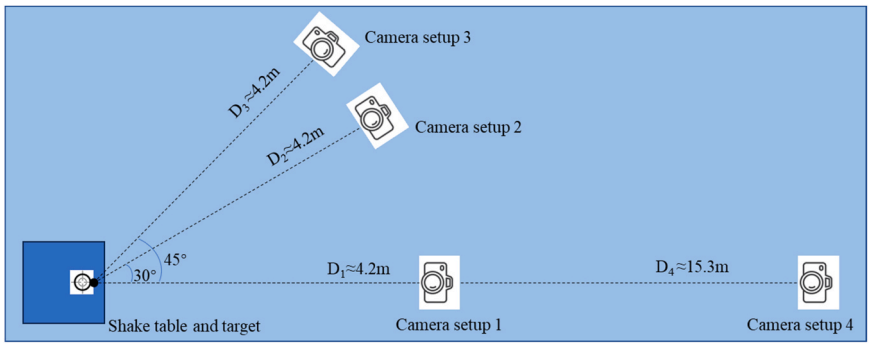


Fig. 9. Plan view of the experiment setup with different angles and distances.

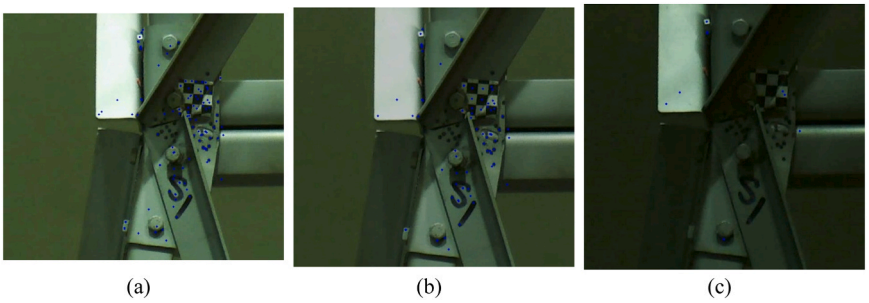


Fig. 10. Pictures of camera FOV under: (a) good; (b) moderate; and (c) bad lighting condition. The numer of feature points are 119, 75 and 19, respectively.

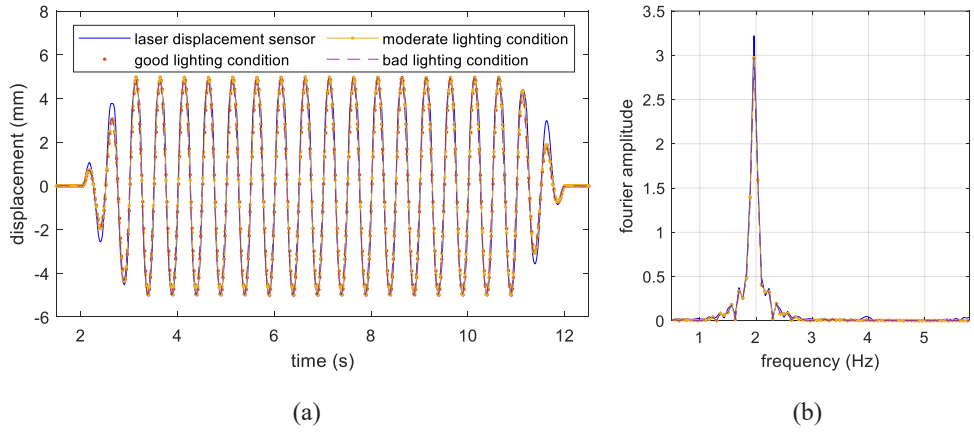


Fig. 11. Displacement (a) time histories and (b) frequency spectrum measured from the laser displacement sensor and the proposed EdgeCVDMS subjected to three different lighting condtion.

from the three frames of Fig. 10 are 119, 75, and 19, respectively. As mentioned in Section 2.2.2, the feature points are sorted in the descending order according to the quality (response value). Only the best quality feature point is selected during the motion tracking process. Typically, the best quality feature point is more robust to the light changes, and consequently, the developed EdgeCVDMS can achieve reliable identification results under different lighting conditions.

3.4.2. Effect of shooting angle on the displacement measurement accuracy

Fig. 12 presents the displacement time histories measured from the laser displacement sensor and the SURF-CPU method corresponding to the three different shooting angles. As shown in Fig. 12, the displacement responses measured from different shooting angles agree well with those measured from laser displacement sensor in time and frequency domains. The RMSE of measurements corresponding to 0°, 30°, and 45° shooting angles are 0.3894 mm, 0.4999 mm, and 0.4250 mm, respectively. The NRMSE of measurements corresponding to 0°, 30° and 45° shooting angles are 4.99 %, 3.78 % and 4.25 %, respectively. The results indicate that the effect of shooting angle on the displacement measurement accuracy is not significant.

3.4.3. Effect of target distance on the displacement measurement accuracy

In the field application, the edge device is installed at a fixed point from a distance to the target on the structural surface. As reported in Ref. [4,29], the displacement responses of bridge structures under operational conditions are at the millimeter level. Therefore, in camera setup 4, the excitation amplitude is respectively set to 5 mm, 4 mm, 3 mm, 2 mm and 1 mm to evaluate the limit of accuracy of the developed EdgeCVDMS. The comparison between the laser displacement sensor and the computer vision-based displacement measurement corresponding to different excitation amplitudes is presented in Fig. 13. It is found that the RMSE corresponding to different excitation amplitude

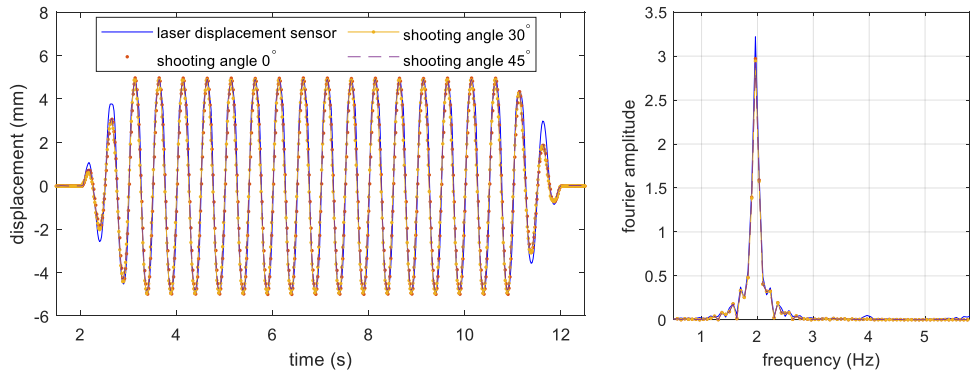


Fig. 12. Displacement (a) time histories and (b) frequency spectrum measured from the laser displacement sensor and the proposed EdgeCVDMS subjected to different shooting angle.

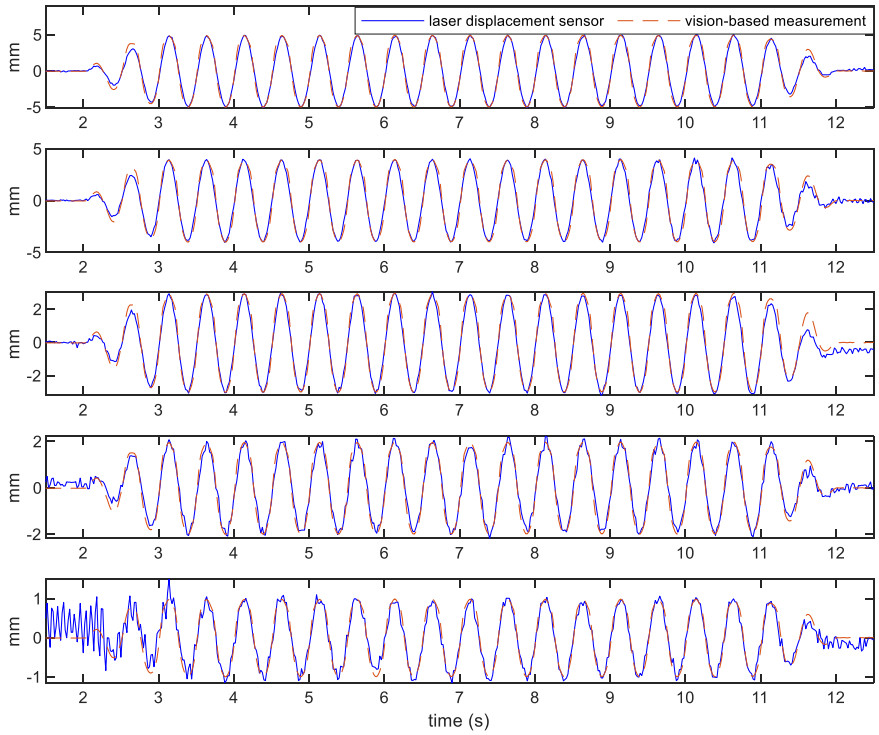


Fig. 13. Displacement time histories measured from the laser displacement sensor and the proposed EdgeCVDMS corresponding to different excitation amplitudes.

levels is between 0.2–0.7 mm. The NRMSE is within 5.6 % when the to-be-measured displacement amplitude is greater than 2 mm.

Table 3 summarizes the error results for displacement measurement corresponding to five different excitation amplitudes using the developed EdgeCVDMS. Overall, a focal length of 50 mm with the present hardware and software configuration allows submillimetre accuracy to be achieved at 15.3 m from the target ROI on the structure. It is noted that the accuracy of computer vision-based displacement measurement is related to the motion tracking algorithm, image sensor performance, zoom lens, and other environmental effects. To further improve the accuracy of the developed EdgeCVDMS, a straightforward way is to change the current 50 mm zoom lens to another lens with a longer focal length. There is a wide range of commercial-grade lenses available with focal length options ranging from several millimetres to several hundred millimetres [31].

Table 3
Summary of the errors for displacement measurement using the developed EdgeCVDMS.

Excitation amplitude	5 mm	4 mm	3 mm	2 mm	1 mm
RMSE (mm)	0.4301	0.5569	0.3735	0.2238	0.6594
NRMSE	4.30 %	6.96 %	6.22 %	5.60 %	32.97 %

Besides the aforementioned influential factors, in field applications, wind-induced camera motions may significantly affect the accuracy of computer vision-based displacement tracking. When the camera is subjected to wind-induced motions, the captured images can be blurry or distorted, making it challenging to track object displacement accurately. The following suggestions are given to alleviate the wind-induced camera motions and displacement measurement errors: (i) using a relatively heavy and solid camera tripod, (ii) using a camera case to cover the video camera, thereby avoiding the effects of wind, and (iii) developing signal processing techniques to remove the displacement identification errors by wind-induced camera motions.

3.5. Continuous operating performance evaluation

In field applications, the edge device will be installed on a fixed point to continuously capture and process video streams to obtain the structural displacement responses. During long periods of continuous operation, prolonged high temperatures can lead to thermal throttling, which may reduce the prescribed video processing frame rate and the working life of the edge device [32]. In this section, the edge device's temperature and average video processing FPS are continuously collected per minute to evaluate the system performance. Fig. 14 and Fig. 15 show the long-term behavior of the edge device in an indoor environment. When the SURF-CPU video processing algorithm is employed, the edge device temperature reaches to 50 °C after 10 min of workload and eventually stabilizes at approximately 52 °C. The video processing speed remains constant at 30 fps almost all the time. When the SURF-GPU video processing algorithm is employed, the edge device temperature reaches to a peak of 53 °C after about 30 min of workload and then stabilizes at about 50 °C. The video processing speed varies between 29–30 fps during the evaluation. During the overall 14 h of operation, the edge device consistently performs without any degradation in performance. The structural displacement corresponding to each video frame has been successfully identified and transmitted to AWS in real-time. The results reveal that the edge device has the potential to be applied for continuous displacement measurement and condition assessment of civil infrastructures.

4. Discussions

The major contribution of this paper is the development of EdgeCVDMS and experimental verifications on the accuracy, efficiency, and stability of this proposed system for structural dynamic displacement measurement and data visualization. There are several issues that should be further investigated before the practical implementations of the proposed EdgeCVDMS.

Firstly, the current version of EdgeCVDMS is applicable to structural single-point or multiple-point displacement measurement within a small area of ROI. For large-scale civil structures, it is desired to simultaneously measure the displacement responses at several different locations. Therefore, a more powerful edge computing sensing system that integrates multiple synchronized cameras can be further developed. For example, higher performance hardware, such as Jetson AGX Orin, that supports six synchronized full HD camera modules along with 65 fps can be adopted.

Secondly, due to the difficulty in setting up the laser displacement sensor in the field test to obtain the displacement ground truth, the performance of the developed EdgeCVDMS is only evaluated in a laboratory-controlled environment. The in-situ applicability should be further validated in future studies.

Thirdly, there are few in-situ deployment technical details, including active cooling strategies, sustainable power supply, sleep, and automatic restart of the EdgeCVDMS in the night and daytime, that should be further developed before field application. In addition, owing to the bandwidth limit, it is challenging and time-consuming to transmit the original videos to the cloud server. A more feasible solution is to continuously store the videos on the edge device. When the storage space is full, the latest video overwrites the oldest one.

Furthermore, this study and most of the existing work focus on the identification of structural dynamic displacement from the computer vision-based method. Given the fact that massive long-term structural displacement measurement is available, how to extract reliable features from the data to support and optimize structural condition assessment and maintenance decision-making should be researched and incorporated into the developed EdgeCVDMS. For example, the displacement measurement can be further used in global level structural system identification [33,34], bridge vibration serviceability evaluation [29], model updating [35] and damage detection [27]. For applications to bridges, the camera should be installed on a fixed reference point, e.g., a bridge abutment or solid ground, to minimize the effect of camera vibrations on displacement measurements.

5. Conclusions

This paper proposes a smart EdgeCVDMS for real-time vibration displacement measurement. The high-definition video camera is integrated with an edge computing device (Nvidia Jetson Nano) to continuously capture, process, and record structural dynamic displacement data. AWS-based data management and visualization are developed to facilitate the scalability of the developed EdgeCVDMS. The key technical issues in the design of the developed EdgeCVDMS, including hardware selection, motion tracking algorithm selection, cloud data storage and visualization, are addressed. A series of experimental studies are conducted to validate the performance and accuracy of the developed EdgeCVDMS. In particular, the accuracy is analysed and compared with results from laser displacement sensors. The video processing efficiency of several candidate computer vision-based displacement tracking algorithms is evaluated. The influential factors, including lighting condition, shooting angle, and target distance on the displacement identification accuracy, are comprehensively analysed. The continuous operating stability and reliability of the developed EdgeCVDMS are evaluated by monitoring the relationship between device temperature and video processing speed. Experimental results demonstrate the promising potential of using the proposed system to perform long-term structural displacement monitoring under operational conditions.

Overall, a focal length of 50 mm with the present hardware and software configuration allows real-time displacement measurement at submillimeter accuracy and 30 FPS when the target ROI on the structure is 15.3 m away from the camera. Compared with existing studies, the significance of this work lies in its capacity to expand the applicability of the computer vision-based displacement measurement technique from short-term and offline modes to long-term and continuous operations.

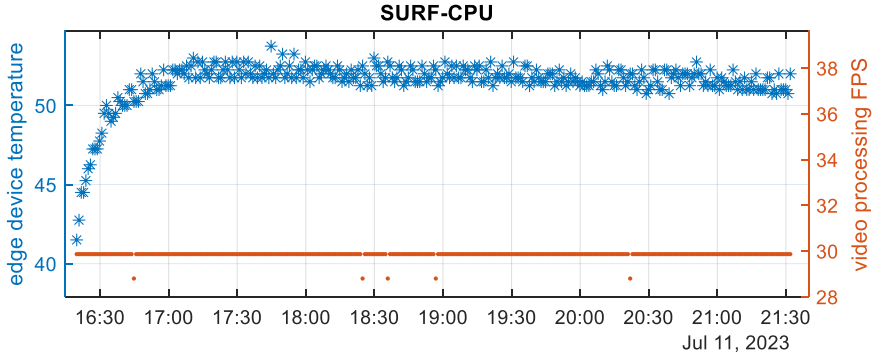


Fig. 14. The variation of edge device temperature and video processing FPS over time when SURF-CPU algorithm is implemented.

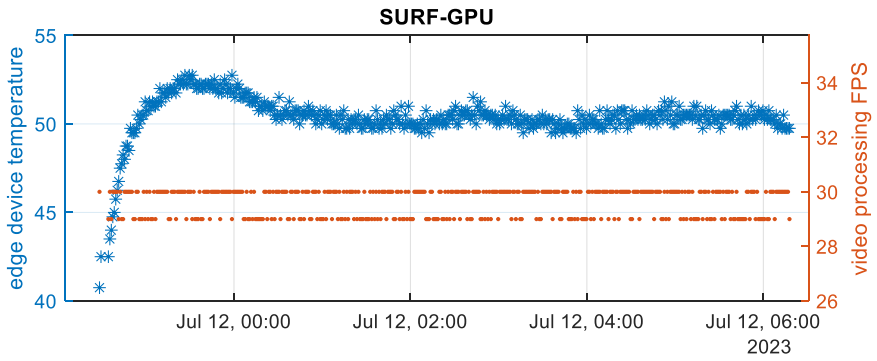


Fig. 15. The variation of edge device temperature and video processing FPS over time when SURF-GPU algorithm is implemented.

The developed EdgeCVDMS is low-cost, easy-deployable, and thus has the potential to be applied in the health monitoring of a population of civil engineering structures.

CRediT authorship contribution statement

Zhen Peng: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Hong Hao:** Writing – review & editing, Supervision, Methodology, Investigation, Conceptualization. **Jun Li:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Yue Zhong:** Validation, Methodology, Investigation.

Declaration of Competing Interest

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

Data availability

Data will be made available on request.

Acknowledgements

The support from Australian Research Council Future Fellowship FT190100801, “Innovative Data Driven Techniques for Structural Condition Monitoring”, is acknowledged.

References

[1] Hao H, Bi K, Chen W, Pham TM, Li J. Towards next generation design of sustainable, durable, multi-hazard resistant, resilient, and smart civil engineering structures. Eng Struct 2023;277:115477.
[2] Peng Z, Li J, Hao H. Data driven structural damage assessment using phase space embedding and Koopman operator under stochastic excitations. Eng Struct 2022;255:113906.
[3] Peng Z, Li J, Hao H, Yang N. Mobile crowdsensing framework for drive-by-based dense spatial-resolution bridge mode shape identification. Eng Struct 2023;292:116515.
[4] Moreu F, Jo H, Li J, Kim RE, Cho S, Kimmle A, et al. Dynamic assessment of timber railroad bridges using displacements. J Bridge Eng 2015;20:04014114.
[5] Ye S, Lai X, Bartoli I, Aktan AE. Technology for condition and performance evaluation of highway bridges. J Civ Struct Health Monit 2020;10:573–94.
[6] Xu YL, Zhang J, Li J, Xia Y. Experimental investigation on statistical moment-based structural damage detection method. Struct Health Monit 2009;8:555–71.
[7] Liu C, Teng J, Peng Z. Optimal sensor placement for bridge damage detection using deflection influence line. Smart Struct Syst, Int J 2020;25:169–81.
[8] Ge L, Koo KY, Wang M, Brownjohn J, Dan D. Bridge damage detection using precise vision-based displacement influence lines and weigh-in-motion devices: experimental validation. Eng Struct 2023;288:116185.
[9] Moreu F, LaFave JM. Current research topics: railroad bridges and structural engineering. Newmark Struct Eng Lab Rep Ser 2012;032.
[10] Xu M, Li J, Wang S, Yang N, Hao H. Damage detection of wind turbine blades by Bayesian multivariate cointegration. Ocean Eng 2022;258:111603.
[11] Peng Z, Yang Z, Tu J. Genetic algorithm based tikhonov regularization method for displacement reconstruction. J Shanghai Jiaotong Univ (Sci) 2019;24:294–8.
[12] Xu M, Au FT, Wang S, Wang Z, Peng Q, Tian H. Dynamic response analysis of a real-world operating offshore wind turbine under earthquake excitations. Ocean Eng 2022;266:112791.
[13] Ko J, Ni YQ. Technology developments in structural health monitoring of large-scale bridges. Eng Struct 2005;27:1715–25.
[14] Tan D, Ding Z, Li J, Hao H. Target-free vision-based approach for vibration measurement and damage identification of truss bridges. Smart Struct Syst 2023;31:421–36.
[15] Shao Y, Li L, Li J, An S, Hao H. Computer vision based target-free 3D vibration displacement measurement of structures. Eng Struct 2021;246:113040.
[16] Wang Y, Hu W, Teng J, Xia Y. Phase-based motion estimation in complex environments using the illumination-invariant log-Gabor filter. Mech Syst Signal Process 2023;186:109847.
[17] Levine NM, Narazaki Y, Spencer Jr BF. Performance-based post-earthquake building evaluations using computer vision-derived damage observations. Adv Struct Eng 2022;25:3425–49.
[18] Xu J, Yuan C, Gu J, Liu J, An J, Kong Q. Innovative synthetic data augmentation for dam crack detection, segmentation, and quantification. Struct Health Monit 2023;22:2402–26.
[19] Peng Z, Li J, Hao H. Development and experimental verification of an IoT sensing system for drive-by bridge health monitoring. Eng Struct 2023;293:116705.
[20] Zhao L, Liu Z, Mbachu J. Development of intelligent prefabs using IoT technology to improve the performance of prefabricated construction projects. Sensors 2019;19:4131.
[21] Scuro C, Lamonaca F, Porzio S, Milani G, Olivito R. Internet of Things (IoT) for masonry structural health monitoring (SHM): overview and examples of innovative systems. Constr Build Mater 2021;290:123092.
[22] Mishra M, Lourenço PB, Ramana GV. Structural health monitoring of civil engineering structures by using the internet of things: a review. J Build Eng 2022;48:103954.
[23] Alsakka F, Assaf S, El-Chami I, Al-Hussein M. Computer vision applications in offsite construction. Autom Constr 2023;154:104980.
[24] V. Shajihan SA, Hoang T, Mechitov K, Spencer BF. Jr, Wireless SmartVision system for synchronized displacement monitoring of railroad bridges. Comput Civ Infrastruct Eng 2022.
[25] Lin Y-C, Hsiao C-Y, Tong J-H, Liao C-P, Song S-T, Tsai H-C, et al. Application of edge computing in structural health monitoring of simply supported PCI girder bridges. Sensors 2022;22:8711.
[26] Meng Q, Zhu S. Developing IoT sensing system for construction-induced vibration monitoring and impact assessment. Sensors 2020;20:6120.
[27] Peng Z, Li J, Hao H. Computer vision-based displacement identification and its application to bridge condition assessment under operational conditions. Smart Constr 2024;1:0003.
[28] Zhang Z. A flexible new technique for camera calibration. IEEE Trans Pattern Anal Mach Intell 2000;22:1330–4.
[29] Dong C-Z, Bas S, Catbas FN. Investigation of vibration serviceability of a footbridge using computer vision-based methods. Eng Struct 2020;224:111224.
[30] Peng Z, Li J, Hao H, Li C. Nonlinear structural damage detection using output-only Volterra series model. Struct Control Health Monit 2021;28:e2802.
[31] Feng D, Feng MQ. Computer vision for SHM of civil infrastructure: from dynamic response measurement to damage detection–A review. Eng Struct 2018;156:105–17.
[32] Benoit-Cattin T, Velasco-Montero D, Fernández-Berni J. Impact of thermal throttling on long-term visual inference in a CPU-based edge device. Electronics 2020;9:2106.
[33] Tan D, Li J, Hao H, Nie Z. Target-free vision-based approach for modal identification of a simply-supported bridge. Eng Struct 2023;279:115586.
[34] Chen Z-W, Ruan X-Z, Liu K-M, Yan W-J, Liu J-T, Ye D-C. Fully automated natural frequency identification based on deep-learning-enhanced computer vision and power spectral density transmissibility. Adv Struct Eng 2022;25:2722–37.
[35] Martini A, Tronci EM, Feng MQ, Leung RY. A computer vision-based method for bridge model updating using displacement influence lines. Eng Struct 2022;259:114129.