



Variational mode decomposition-based multirate data-fusion framework for estimating structural dynamic displacement by integrating vision- and acceleration-based measurements

Zhenfen Jin ^a, Guyuan Chen ^a, Yanbo Niu ^a, Congguang Zhang ^a, Xiaowu Zhang ^a,
Jiangpeng Shu ^{a,b,*}

^a College of Civil Engineering and Architecture, Zhejiang University, Hangzhou 310058, PR China

^b Innovation Center of Yangtze River Delta, Zhejiang University, PR China



ARTICLE INFO

Keywords:

Computer vision
Data fusion
Dynamic displacement measurement
Signal processing
Structural health monitoring
Variational mode decomposition (VMD)

ABSTRACT

Dynamic displacement is a crucial parameter in structural health monitoring (SHM) for assessing the safety, dependability, and suitability of structures under various types of excitations. Computer vision-based methods for dynamic displacement estimation have attracted much interest owing to their cost-effectiveness and convenience. However, these methods are limited by their low sampling rates and high data sensitivity. To compensate for these limitations, methods for combining data obtained from other sensors have been proposed. In this study, an experimental data-fusion framework for displacement estimation based on variational mode decomposition (VMD) was developed to leverage the advantages of vision- and acceleration-based measurements. The measurements were decomposed into ensembles of modes and recomposed to reconstruct the displacement with a higher accuracy and over a wider frequency range. An optimal mode recombination method was proposed to achieve optimal mode combinations. Furthermore, this study introduced an improved vision-based displacement measurement method and a VMD-based indirect acceleration measurement method. The proposed framework was validated through four-story RC structure tests, which demonstrated that the method could enhance the accuracy of displacement estimation and extend the feasible frequency range compared with single-source displacement measurements. The method provides a promising solution for more effective health monitoring of modern structures subjected to a wide variety of dynamic loads.

1. Introduction

The safety of concrete structure during their service life is significantly affected by external loads such as seismic motion, wind loads, and human-induced excitations. The vibrations induced by these loads may lead to severe structural damage and eventually complete failure. Monitoring the dynamic response of structures under external loading is critical for structural health monitoring (SHM) [1,53]. Among the dynamic responses, the time history of the dynamic displacement is vital because it contains direct information for assessing the safety, serviceability, and integrity of structures. Various methods have been proposed to measure the

* Corresponding author.

E-mail address: jpesan@zju.edu.cn (J. Shu).

structural dynamic displacement with the aim of monitoring the health of civil infrastructure and avoiding catastrophic failures.

Conventional approaches for measuring the dynamic displacement of structures mostly rely on sensors such as linear variable differential transformers (LVDTs), fiber optic sensors, accelerometers, strain gauges, piezoelectric sensors, and ultrasonic wave sensors [2–5]. These methods are contact sensor-based and require the installation of monitoring devices in the structures. They take advantage of the high resolution and high sampling rate, providing a relatively reliable accuracy and abundant high-frequency dynamic information for displacement measurements. However, there are several limitations. First, they do not satisfy the demand for a full-field response analysis. Second, sensor installation requires a significant amount of time, cost, and labor resources, not to mention its unsuitability for large-scale civil infrastructure [6–9,54,55]. Hence, many next-generation approaches, particularly noncontact-based approaches, have been developed for measuring the dynamic displacement of civil infrastructure.

Vision-based displacement measurement methods, which are contact-free, inexpensive, and easy to operate, have received widespread attention. These methods have seen considerable improvements owing to developments in computer vision technology. Various methodologies, such as the frame difference, feature point matching, particle image velocimetry (PIV), digital image correlation (DIC), and optical flow, have been considered to measure the structural dynamic displacement [10–23]. The frame difference method computes the displacement by calculating the difference between frames and then performing a subtraction operation. This method can provide a comparatively higher accuracy and computation speed; however, it is sensitive to environmental noise and can only obtain 1D displacement measurements [24]. In the feature point matching method, feature points are considered virtual markers that can be used for nontarget measurements [11]. Although this method takes advantage of the computational efficiency, some challenges include noise interference and uncertainty of the feature point extraction algorithms. Particle image velocimetry (PIV) is widely used to measure 2D flow structures by dividing images into a series of subregions and then measuring the optical flow of each subregion. However, the current PIV algorithm cannot be adapted to map the pattern of model-building motions [25]. The DIC-based method measures the structural displacement by finding the regions with the best correlation in a series of images. The main problem with this method is that the full-field displacement response is acquired by tracking optical textures created by speckle paint; therefore, the camera must be placed near the region of interest for a relatively accurate result [26]. The optical flow method computes the displacement by measuring the optical flow (velocity or displacement) in an image sequence. Current optical flow algorithms, such as the Horn–Schunck (HS) approach [27] and Lucas–Kanade (LK) approach [28], are efficient and can identify complex motion patterns in images. However, these algorithms are associated with the following problem: because the displacement is computed by motion tracking, the errors accumulated between video frames can cause a low-frequency drift [15,29].

Most vision-based dynamic displacement methods have two main challenges: 1) Commercial digital cameras typically have a limited frame rate, which prohibits the collection of high-frequency dynamic motion data, leading to a loss of high-frequency information [30]. 2) To obtain displacement data on a real-world scale, the 2D displacement in pixel units must be converted to meter units; if the structural plane is not parallel to the image plane, scale factors or camera calibration are required [31]. Hence, researchers have started exploring methods to combine data obtained from other approaches with vision-based measurements to compensate for their drawbacks.

Compared with direct displacement measurements, an indirect method based on the integration of acceleration data is easier to implement because it does not require a fixed reference [32]. This method provides more dynamic information than vision-based methods [33,34]. However, a critical challenge must be taken into consideration: the acceleration integration mostly has a low-frequency offset, moving the baseline of displacement away from the correct trajectory. As a result, baseline adjustments are desirable, depending on the two cases below: 1) For cases where the final displacements are close to zero, high-pass filtering techniques are often used to adjust baselines. 2) For cases where non-zero residual displacements are expected (e.g. the cases in which structures have non-linear behaviors), techniques other than high-pass filtering are mandatory. [35,36]. Yet the parameters of high-pass filters need to be manually determined, taking into account the ratio of signal-to-noise in the frequency domain of the Fourier spectrum. Other methods of baseline adjustment (e.g. Using pre-event samples as constraints) should depend on the specific cases, which is not generic and automatic [37].

The strengths and drawbacks of vision- and acceleration-based methods have motivated researchers to develop data-fusion approaches. Various data-fusion methodologies have been proposed, such as multirate Kalman filtering [38–40], complementary filtering [19,21], and mode decomposition [41,52]. Smyth and Wu proposed a multirate Kalman filtering and smoothing technique to accurately estimate the velocity and displacement from noise-contaminated measurements of the displacement and acceleration [38]. Chang et al. proposed an integrated technique that can estimate the structural displacement by utilizing a multirate Kalman filter to combine vision-based and acceleration-based methods; their experimental results revealed that the method outperformed single vision-based approaches [39]. Kim et al. also presented an approach using multirate Kalman filtering to combine the acceleration and intermittent displacement to accurately and automatically estimate the dynamic displacement [40]. Niu et al. proposed an approach based on the Kalman filter, which integrated acceleration and multi-GNSS measurements. The approach was verified by measuring the dynamic displacement of a long-span bridge. However, Kalman filter-based data-fusion methods have limitations in terms of their robustness, accuracy, and sensitivity to the initial states (displacement, velocity, and acceleration) [42]. Hong et al. presented a multirate technique that can combine the measured acceleration and displacement data from GPS-like devices to reconstruct the displacement [32]. Park et al. presented a complementary filter technique to combine vision-based pseudo-static displacement and acceleration-based dynamic displacement. The method was tested both numerically and experimentally, showing a good accuracy [30]. The complementary filter method is insensitive to the initial states; however, it is difficult to define the filter design parameters. Chan et al. proposed a technique that combined GPS measurements and acceleration data using empirical mode decomposition (EMD) and adaptive filtering [41]. Furthermore, based on the EMD, He et al. presented a reconstruction method for the structural dynamic response using remote sensors in the time domain [43]. Such data-fusion methods using mode decomposition are highly efficient in

terms of the computational cost and robustness against measurement noise under certain circumstances. However, the EMD algorithm lacks a mathematical theory, and the decomposition result largely depends on the methods used to find the extreme points, which means that the EMD performance is highly influenced by the different signal envelopes. Therefore, these methods may not be robust to the various structural dynamic responses encountered in practice [44,45].

Considering the strengths and limitations of the current methods, we developed a data-fusion approach to estimate the structural dynamic displacement based on variational mode decomposition (VMD) [46]. The displacement data obtained through vision- and acceleration-based methods were decomposed into a predetermined number of modal responses represented by intrinsic mode functions (IMFs). The vision-based method has a lower sampling rate but a higher measurement accuracy for low-frequency displacements, whereas the acceleration-based method capitalizes on the abundant dynamic components but is susceptible to noise and low-frequency drift. By recomposing the modes of the two measurements, the proposed method can help more effectively reconstruct the displacement in terms of both the measurement accuracy and frequency range.

2. Theory

For illustrating the data fusion methodology and acquisition of vision- and acceleration-based structural displacement, this section provides a theoretical introduction about VMD and traditional frameworks of vision- and acceleration-based displacement measurement.

2.1. Variational mode decomposition

The VMD by Dragomiretskiy and Zosso is a fully intrinsic and adaptive signal processing method widely used in time-frequency analysis, filtering and denoising. The method aims to decompose a signal into different modes, as an alternative to EMD [47]. However, compared with EMD, VMD replaces the most recent definition of IMFs with the corresponding narrowband property instead of modeling the individual modes as signals with explicit IMFs. VMD provides a formula relating the explicit AM-FM parameters with the estimated signal bandwidth, making the algorithm more robust to different signals generated in real scenarios [46].

VMD decomposes a signal into a series of modes u_k with a specific bandwidth in the spectral domain. During the decomposition process, each mode is concentrated around a determined center frequency ω_k . To obtain the sparsity of each mode, the mode bandwidth in the spectral domain is required. The steps for obtaining the mode bandwidth are as follows: 1) For each mode, the Hilbert transform is applied to obtain a unilateral frequency spectrum; 2) The frequency spectrum of the mode is shifted to the baseband by applying exponential tuning to the estimated center frequency; 3) The mode bandwidth is estimated by subjecting the demodulated signal to Gaussian smoothness. The variational problem can then be constructed and expressed as follows:

$$\begin{aligned} \min_{u_k, \omega_k} & \left\{ \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \otimes u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s.t. } & \sum_{k=1}^K u_k = f(t) \end{aligned} \quad (1)$$

where $f(t)$ represents the original signal, t represents the time script, K implicates the number of modes that need to be decomposed, u_k implicates the k -th mode, $\delta(t)$ denotes the Dirac distribution, ω_k denotes the center frequency of each mode, and \otimes represents the convolution operation.

Eq. (1) can be converted into an unconstrained problem by applying a penalty term and the Lagrangian multiplier λ , which can be expressed as follows:

$$L(u_k, \omega_k, \lambda) := \alpha \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) \otimes u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^K u_k(t) \right\rangle \quad (2)$$

Here, α denotes the balancing parameter for the required data fidelity constraint.

Eq. (2) can be solved using the alternating direction method of multipliers (ADMM), which can determine the saddle point of an augmented Lagrangian. Each mode u_k and its center frequency ω_k can then be updated to realize complete the VMD analysis process. Eq. (3) expresses the solution to the optimization problem for u_k :

$$\hat{u}_k^{n+1} = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \quad (3)$$

where n denotes the number of iterations. $\hat{f}(\omega)$, $\hat{u}_i(\omega)$, $\hat{\lambda}(\omega)$, and \hat{u}_k^{n+1} denote the Fourier transforms of $f(t)$, $u_i(t)$, $\lambda(t)$, and u_k^{n+1} , respectively.

In the frequency domain, the solution to the optimization problem of ω_k can be expressed as follows:

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_k^{n+1}(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k^{n+1}(\omega)|^2 d\omega} \quad (4)$$

As every u_k converges, an ensemble of modes, the sum of which is the original signal, is obtained.

2.2. Vision-based displacement measurement using optical flow

Optical flow refers to the distribution of the apparent velocities of the movement of brightness patterns in an image across the image plane. It describes the motion information of an object [27]. Conventionally, given a frame sequence, the optical flow-based displacement estimation algorithm has the following steps: 1) A series of regions of interest (ROIs) corresponding to different essential image areas are established. 2) Points representing relevant moving objects can be detected using common feature descriptors (e.g., FAST, SURF, SIFT, Harris, and Shi-Tomasi). 3) Optical flow tracking algorithms, such as the Lucas–Kanade (LK) method, can be used to compute the optical flow between consecutive image frames. 4) The 2D displacement can be calculated using the optical flow records of the feature points.

For 2D-to-3D displacement conversion, a pinhole camera model was introduced to estimate the parameters of the camera, including the intrinsic, extrinsic, and distortion information. Conventionally, a pinhole camera model is used to estimate the camera parameters, expressed as follows:

$$p = K[RT]P_w \quad (5)$$

where p denotes the 2D coordinates of a pixel in the image plane, P_w denotes the 3D coordinates of the point in the world coordinate system, and K denotes the intrinsic matrix. The extrinsic matrix related to the camera position in the world coordinates is determined by the rotation matrix R and translation vector T . The camera parameters can be obtained through calibration. Thus, the 2D displacement of the feature points can be converted into 3D displacement of the objects.

The optical-flow-based method has a relatively high accuracy and robustness over other vision-based methods [24]. However, the unfixed optical flow-based method is prone to errors owing to the accumulation of frame distortion or incorrect selection of feature points, and time-consuming manual calibration is an inevitable step.

2.3. Acceleration-based displacement measurement using simple integration

In the time domain, the relationship between the displacement, velocity, and acceleration can be expressed as follows:

$$\dot{x} = \int \ddot{x} dt + c_1 \quad (6)$$

$$x = \int \dot{x} dt + c_1 t + c_2 \quad (7)$$

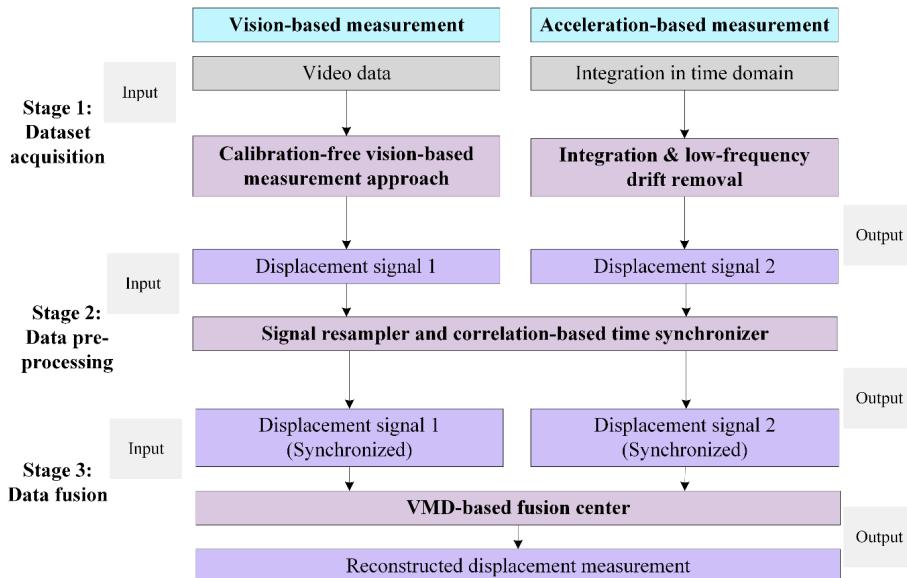


Fig. 1. Overall illustration of the proposed method.

where x , \dot{x} , and \ddot{x} denote the displacement, velocity, and acceleration, respectively. The constraints of the numerical integration are c_1 and c_2 . Given the time history of the structural acceleration, the dynamic displacement can be estimated through integration. Yet in practice, the low-frequency error due to unavoidable computational truncation and rounding errors, as well as measurement noise, inevitably causes distortions in the estimation results. More importantly, for a complicated signal, such as a vibration signal, the drift of the estimated displacement is not linear, is comparatively high, and cannot be solved by either parameter fitting or using a single filter.

3. Methodology

Fig. 1 illustrates the proposed framework, which comprises four sections: (1) A calibration-free vision-based measurement approach to calculate the displacement signal of the structure from video data; (2) A drift-free indirect displacement measurement from the acceleration to calculate another displacement signal, the source of which is the accelerometer records; (3) A data preprocessing procedure including a resampler and time synchronizer; (4) VMD-based data-fusion center that decomposes and recomposes the synchronized signals to reconstruct the displacement measurement results.

Compared with currently available techniques, the proposed methodology aims to obtain more reliable displacement measurements from different sources and present an adaptive and automated data fusion technique. The proposed vision-based measurement approach aims to obtain reliable displacement and apply 2D-3D conversion without calibration, while traditional vision-based approaches need a calibration process and usually error-prone because of image distortion. In the next stage, the drift-free acceleration-based measurement approach is presented as an alternative way to baseline adjustment without high-pass filters. The purpose of these steps is to provide more reliable sources to the data fusion stage. The data preprocessing procedure is to uniformize sampling rates and

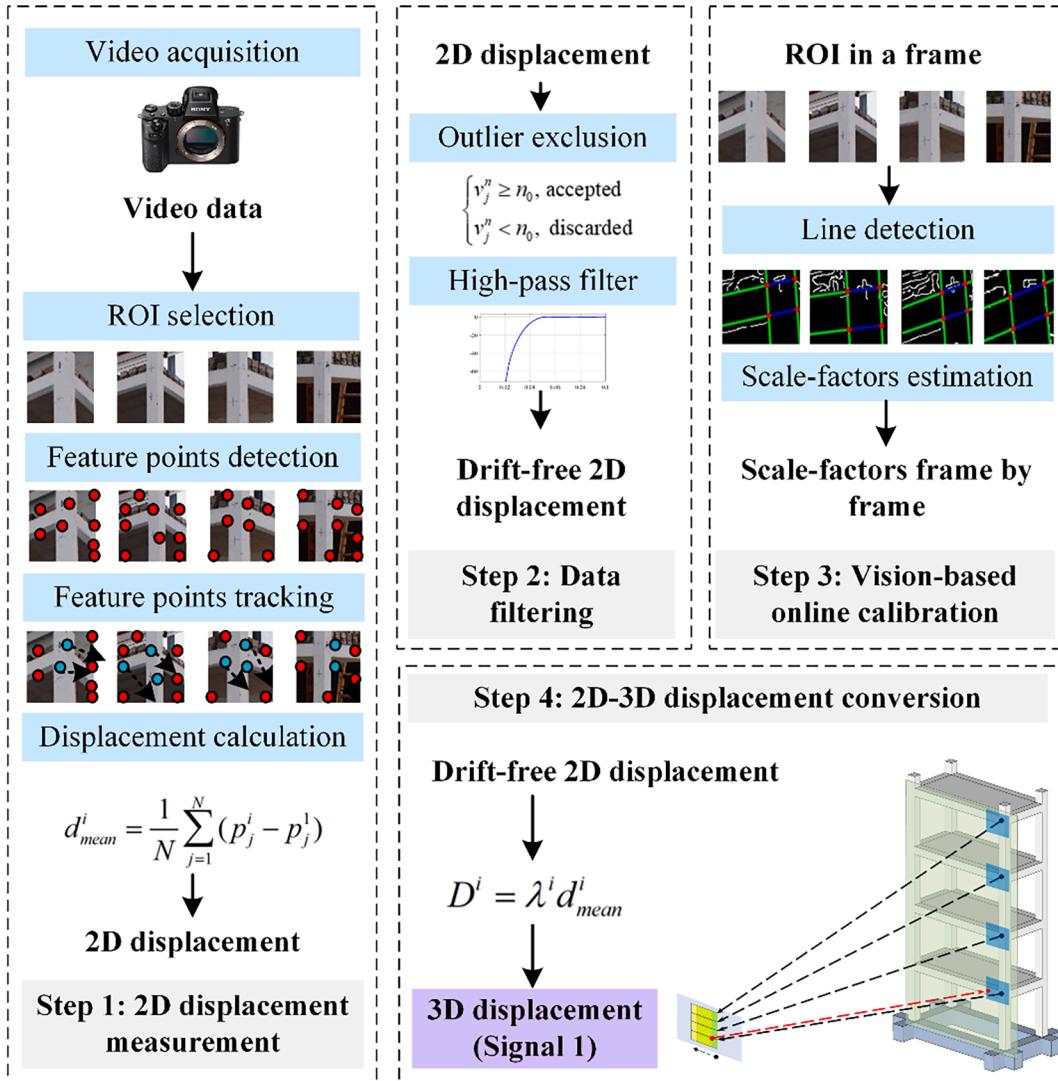


Fig. 2. Framework of the estimation process of improved vision-based dynamic displacements.

remove the offsets of the displacement from the two sources. Finally, the VMD-based data-fusion center aims to present an automated method for integrating measurements to improve measurement accuracy and expand viable frequency range.

3.1. Calibration-free vision-based displacement measurement approach

As shown in Fig. 2, the proposed approach comprises four essential steps: (1) 2D (Two-dimensional) displacement measurement; (2) Data filtering; (3) Vision-based online calibration; (4) 2D-to-3D conversion of the displacement.

First of all, a camera for video acquisition should be positioned where the entire structure can be captured within its frame range. As the video data were acquired, the ROIs were located at the nodes of each floor. In each ROI, feature points p_j ($j = 1, 2, \dots, N$) can be detected using the Shi-Tomasi corner descriptor [48]. Next, the LK algorithm was implemented to track the optical flow of the detected feature points from the ROIs. Subsequently, the position of each feature point was recorded. Considering that the structural displacement is 1D, the 2D displacements can be calculated through a subtraction between x coordinates of the same point in different frames. The 2D displacement of the floor can then be calculated using the mean displacement of the adjacent feature points:

$$d_{mean}^i = \frac{1}{N} \sum_{j=1}^N (p_j^i - p_j^1) \quad (8)$$

where p_j^i denotes the x coordinate of the j -th feature point in the i -th frame ($i = 1, 2, \dots, n$); d_{mean}^i denotes the 2D structural displacement; and N denotes the total number of detected feature points.

To ensure that most of the points in the ROI represent structural displacements, it is important to ensure that they have similar values. A voting-based filter was proposed to discard outliers whose displacements drift away from the average trajectory. In the i -th frame, the standard deviation of the 2D displacement of the point set σ^i can be calculated using Eq. (9):

$$\sigma^i = \sqrt{\frac{1}{N} \sum_{j=1}^N (d_j^i - d_{mean}^i)^2 / N} \quad (9)$$

where d_j^i represents the 2D displacement of the j -th feature point. As expressed in Eq. (10), in the i -th frame, if the difference between d_j^i and d_{mean}^i is lower than σ^i , the vote value v_j^i is added by one; otherwise, it is equal to zero. Finally, by comparing the sum of the votes of each point v_j^n and the given threshold n_0 , if v_j^n is greater than n_0 , the point is considered an accepted point; otherwise, it is considered a discarded point, as expressed in Eq. (11). This method not only filters incorrect feature points with unreliable displacements but also removes the lost tracking feature points from the point set. Survival points were set to calculate a more reliable trajectory of the structural displacement.

$$v_j^i = \begin{cases} v_j^{i-1} + 1, & d_j^i - d_{mean}^i < \sigma^i \\ v_j^{i-1}, & d_{mean}^i - \sigma^i < d_j^i < d_{mean}^i + \sigma^i \\ v_j^{i-1}, & d_j^i > d_{mean}^i + \sigma^i \end{cases} \quad (10)$$

$$\begin{cases} v_j^n \geq n_0, & \text{accepted} \\ v_j^n < n_0, & \text{discarded} \end{cases} \quad (11)$$

Fig. 3 illustrated the result of outlier exclusion using a period of displacement signal from a case study, which was used for experimental verification in section 4. As shown in the figure, the proposed data filter corrects the trajectory of the displacement compared with the unfixed signal.

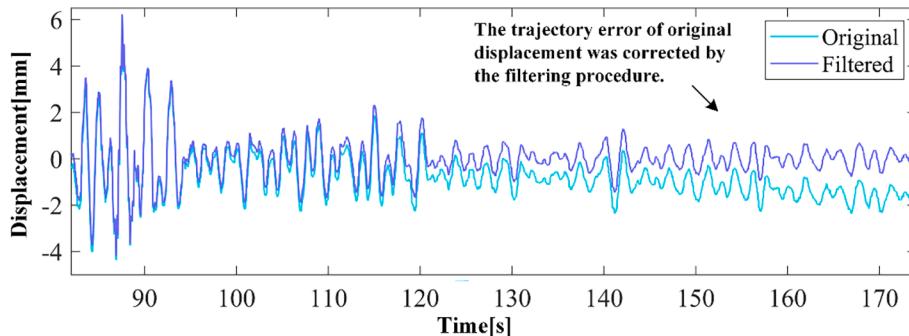


Fig. 3. Demonstration of the data filtering result.

In the next step, a vision-based online calibration method is proposed for subpixel-level scale factor estimation. Considering that the structural displacement is 1D, the displacement in the depth direction can be omitted; thus, the scale factors can be estimated frame-by-frame using 2D geometric features (length, width, diameter, etc.). In this study, the width of the column in each ROI was selected for scale factor estimation. Fig. 4 shows the estimation process of the scale factor in the i -th frame. First, the Canny edge detection algorithm [49] was implemented for edge extraction from the image. The Hough transform was then applied to detect the lines of the column edges, including the two sides of the column, in the image plane. Because of distortion in the image, the two lines representing the edges of the column were not perpendicular to each other. Therefore, the 2D column width was calculated using the average value of the measured lengths of lines l_{1i} and l_{2i} . Because the ROI covers a small area of the image frame, the averaging approach also corrects the image distortion. Given column width L , the i -th frame scale factor λ_i can be calculated using Eq. (12), and the i -th structural displacement in meter unit D^i can be calculated using Eq. (13). The corrected data $[D^1, D^2, \dots, D^n]$ represent the vision-based displacement measurement.

$$\lambda^i = \frac{L}{(l_{1i}^i + l_{2i}^i)/2} \quad (12)$$

$$D^i = \lambda^i d_{mean}^i \quad (13)$$

3.2. Drift-free acceleration-based displacement measurement

Conventionally, the low-frequency drift of a signal is removed using geometric fitting or filters. However, in most cases, acceleration data contain a large amount of noise, which makes the integration process much more complicated than when using geometric parameters, which are difficult to compute, and it is challenging to determine the cut-off frequencies of the filter.

To address these issues, this paper proposes a VMD-based detrending approach for removing low-frequency errors. Considering the inputted signal to be a dynamic vibration response with frequencies higher than 1 Hz, the mode of large low-frequency errors will reach convergence at a particular center frequency under an appropriate bandwidth parameter α . Suppose $u(t)$ is the displacement measured from integrating the acceleration data, and $u_k(t)$ is the k -th individual mode, we have (14) after the signal is decomposed by VMD:

$$u(t) = \sum_{k=1}^K u_k(t) \quad (14)$$

The mode $u_1(t)$ with the lowest center frequency ω_1 is the low-frequency error. By simply removing $u_1(t)$ from $u(t)$, the drift-free displacement measurement can be obtained as follows:

$$u_{drift-free}(t) = u(t) - u_1(t) \quad (15)$$

However, it is important to mention that VMD has an unavoidable boundary effect: in most cases, samples near signal boundaries are distorted after the process, which is the most important limitation of VMD [46]. In order to restrict boundary effect, the displacement obtained from the integration is extended before performing the decomposition. The Arburg self-regression algorithm [50] was implemented to extend the boundary data of the displacement signal. In the next stage, the extended value was used as the input data for the VMD. After the removal of $u_1(t)$, the extended data on the boundaries of $u_{drift-free}(t)$ were deleted, and the initial length of the displacement signal was recovered. Finally, the remaining displacement data represent the drift-free measurement from the acceleration, which is defined as the displacement signal 2. Fig. 5 shows an overall illustration of the proposed method.

Fig. 6 illustrates how the proposed method removes low-frequency drift and reduces VMD boundary effect. The original $u(t)$ was computed from $a(t)$, as shown in Fig. 6(a and b). Next, VMD was implemented to separate the drift content and dynamic displacement, with boundary effects, illustrated in Fig. 6(c – e). In comparison, by implementing VMD to the extended original signal shown in Fig. 6(f), the boundary effect affects the extension part instead of the original signal part. After the extension part was cut, the displacement was recovered and the boundary effect is reduced, which is shown in Fig. 6(j).

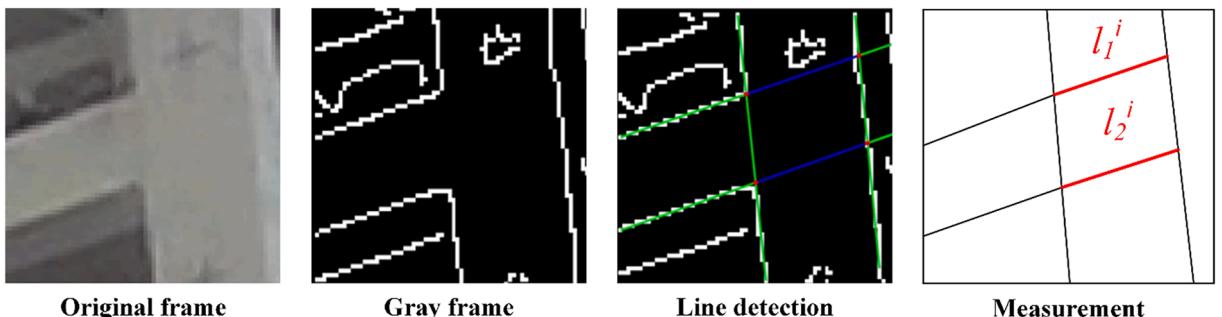


Fig. 4. Illustration of the estimation process of scale factors.

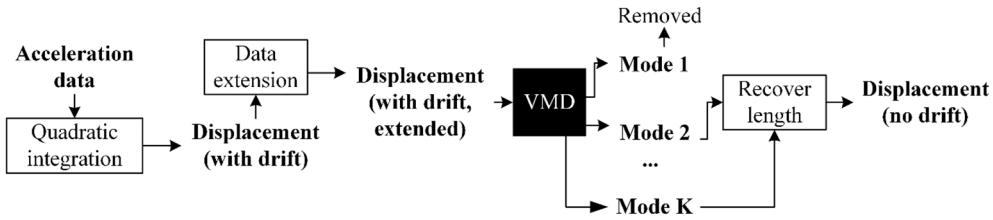


Fig. 5. Proposed acceleration-based measurement approach.

3.3. Preprocessing before data fusion

Before data fusion, it is crucial to preprocess the two signals obtained from different sources because of differences in their sampling rates and offsets. The preprocessing includes two steps: data resampling and synchronization. Assuming the sampling rates of vision- and acceleration-based displacement as p (depends on camera configuration) and q (depends on different sensors determined by the structure under test), since the p is usually different from q , the former should be re-sampled to the same value as q . In this study, the resampling function in MATLAB signal processing toolbox is used for resampling, which performs the following steps: (1) Insert zeros to up-sample the vision-based displacement signal by q . (2) Filter the up-sampled signal by an FIR antialiasing filter. (3) Apply linear interpolation to the signal. (4) Down-sample the filtered signal by p . Fig. 7(a) demonstrates the process of resampling vision-based displacement using a period of the signal, showing that the resampled signal was in agreement with the original one in most parts. Fig. 7(b) shows deviations between resampled and original signals occurred in positions where the signal has rapid changes, yet the deviations were diminutive, indicating that the resampled signal could be considered intact. After resampling, the vision-based displacement is synchronized with the acceleration-based displacement signal via a cross-correlation between the two datasets. The point corresponding to the maximum cross-correlation determines the time lag.

3.4. VMD-based multirate data fusion center

Taking $u_{\text{Low-frequency}}(t)$ as the low-frequency component from the vision-based displacement and $u_{\text{Dynamic}}(t)$ as the high-frequency component from the acceleration-based displacement, the reconstructed displacement $u_{\text{reconstruct}}(t)$ can be integrated using the following two components:

$$u_{\text{Reconstruct}}(t) = u_{\text{Low-frequency}}(t) + u_{\text{Dynamic}}(t) \quad (16)$$

Suppose that the vision-based and acceleration-based displacement data are decomposed into K_1 and K_2 modes by VMD, we have:

$$\begin{aligned} u_{\text{vision}}(t) &= \sum_{i=1}^{K_1} u_{\text{vision}}^i(t) \\ u_{\text{acceleration}}(t) &= \sum_{j=1}^{K_2} u_{\text{acceleration}}^j(t) \end{aligned} \quad (17)$$

$u_{\text{Low-frequency}}(t)$ can be obtained from the low-order modes of the vision-based displacement, and $u_{\text{Dynamic}}(t)$ can be obtained from the high-order modes of the acceleration-based displacement:

$$u_{\text{Reconstruct}}(t) = \sum_{i=1}^{k_1} u_{\text{vision}}^i(t) + \sum_{j=k_2}^{K_2-k_1} u_{\text{acceleration}}^j(t) \quad (18)$$

where k_1 denotes the number of low-order modes of the vision-based displacement, and k_2 denotes the number of high-order modes of acceleration-based displacement. The problem involves determining the appropriate values of K_1 , K_2 , k_1 , and k_2 . In this study, K_1 and K_2 were determined using the converged center frequencies of the individual modes. For an input signal $u(t)$, the center frequency ω_k of a mode u_k is iterated using Eq. (4), and it finally converges until

$$\sum_k \frac{\|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2}{\|\hat{u}_k^n\|_2^2} < \epsilon \quad (19)$$

Here, ϵ denotes the convergence threshold [46]. The values of the center frequencies of each mode correspond to different values of K . Note that if K is too high, a mode in a narrow-band frequency range may be decomposed into two modes having similar center frequencies. This implies that the signal is over-segmented. Subsequently, the best K can be defined from the following steps: 1) Set an initial value of K ; 2) Starting from the first mode, search the center frequencies ω_{k+1} and ω_k of the two adjacent modes; 3) If ω_{k+1} and ω_k satisfy Eq. (20), choose $K - 1$ as the best K ; otherwise, increase K by 1 and proceed to Step 2.

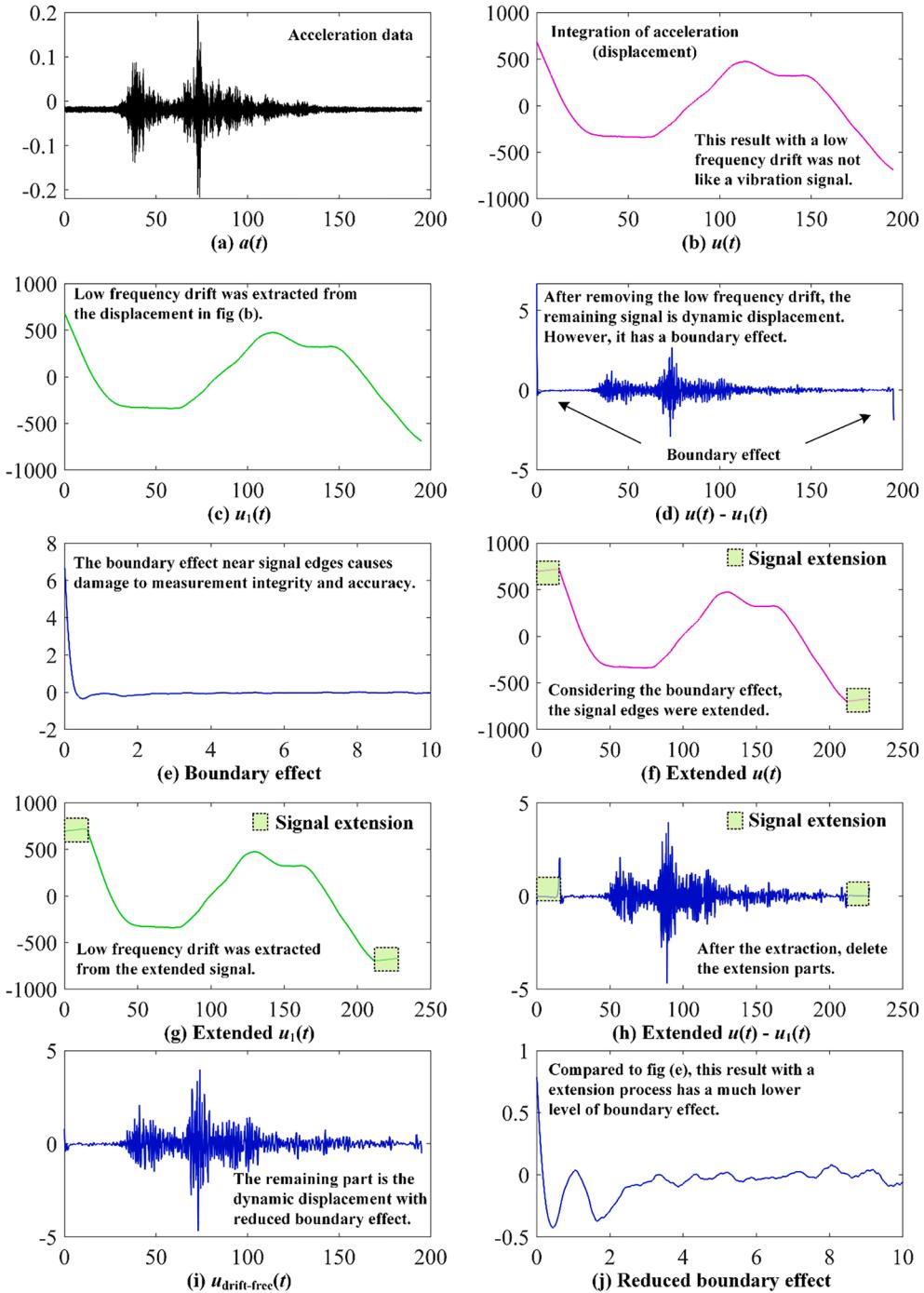


Fig. 6. Illustration of obtaining drift-free acceleration-based displacement.

$$\frac{\omega_{k+1} - \omega_k}{\beta k} < \varepsilon_1 \quad (20)$$

where ε_1 denotes the tolerance of overbinning criterion, β denotes the penalty coefficient of the mode order.

The selection of k_1 and k_2 can be considered a problem when selecting the cutoff frequency of a complementary filter. A complementary filter is a combination of high- and low-pass filters. In this case, the extraction of low- or high-frequency components from a signal can be approximated as a low-pass or high-pass filtering process. The cutoff frequency of the filter is determined using the method proposed by Park et al. [30]. This method determines the cutoff frequency based on the power spectral density of the noise.

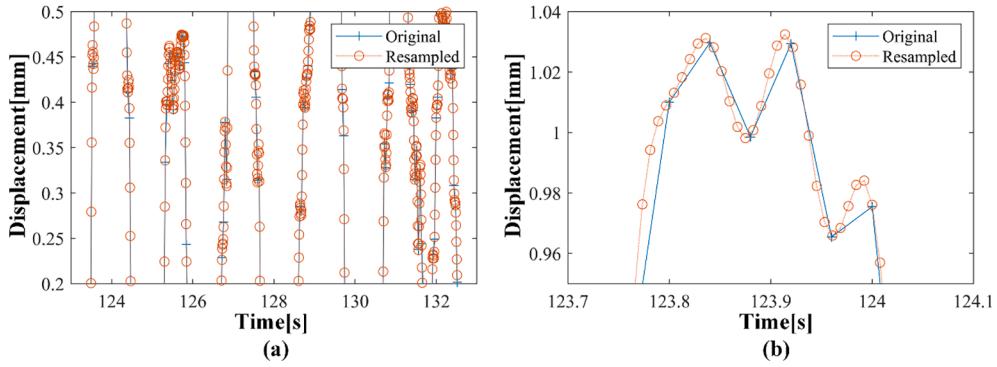


Fig. 7. (a) Illustration of data resampling; (b) Deviations between the original and resampled signal.

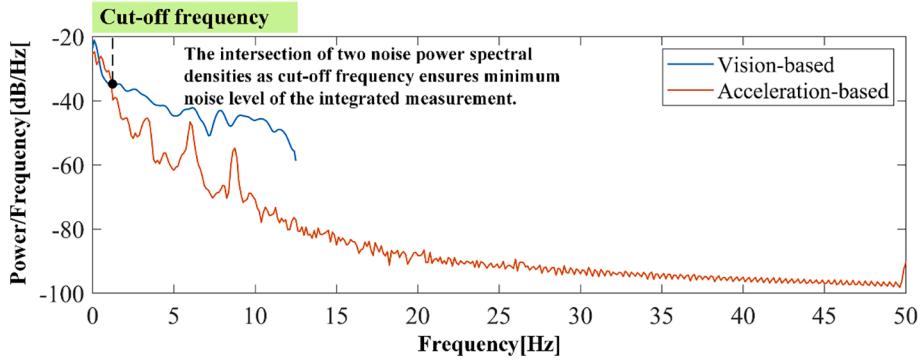


Fig. 8. Noise power spectral densities of vision-based and acceleration-based measurements.

Fig. 8 illustrates the noise power spectral density of the vision- and acceleration-based displacements from the case study. The optimal cutoff frequency f_c is selected at the location where the minimum overall noise power spectral density occurs at the intersection between the two noise power spectral densities:

$$f_c = \frac{1}{2\pi} \sqrt{\frac{N_{acceleration}}{N_{vision}}} \quad (21)$$

where N_{vision} and $N_{acceleration}$ denote the noise densities of the vision- and acceleration-based displacements, respectively. After mode decomposition, the cutoff frequency f_c corresponds to a particular mode order D , the center frequency ω_D of which is close to f_c . Since f_c often lies in low frequency area, and f_c and ω_D are very close, the evaluation criterion is represented by the root of the difference between them:

$$\sqrt{|\omega_D - f_c|} < \varepsilon_2 \quad (22)$$

where ε_2 denotes the threshold. Subsequently, k_1 and k_2 can be determined as follows:

$$\begin{cases} k_1 = D \\ k_2 = D + 1 \end{cases} \quad (23)$$

As the two displacements were measured from the same structure under the same excitation, they should be similar in the frequency domain; therefore, each mode frequency after their decomposition should also be relatively similar. As a result, the selection of D from vision-based or acceleration-based displacement modes should produce a similar result. Another essential parameter of the VMD implementation is the balancing parameter of the data fidelity constraint α . The parameter α has a significant impact on the bandwidth of each mode. As α increases, the bandwidth of each mode tends to decrease. According to experience, the value of α works well among [1000, 4000] [51]. In this study, α was set to 2500.

4. Experimental verification

4.1. Experimental setup

A dynamic experiment was conducted to evaluate the performance of the proposed data-fusion displacement estimation strategy. A four-story reinforced concrete (RC) structure was erected on a 25-ton shaking table measuring $4.0\text{ m} \times 4.0\text{ m}$ at Tongji University. Fig. 9 shows the actual image and schematic of the setup [56]. The shaking table was used to perform a dynamic experiment with unidirectional earthquake inputs from three sources. Under these excitations, the RC frame vibrated one-dimensionally in a plane parallel to the floor. Table 1 presents the excitation source information.

The experiment video was captured by a SONY HDR-PJ200 camera (resolution: 1920×1080 , sampling rate: 25 Hz). Acceleration data (sampling rate: 256 Hz) was measured using accelerometers. Also, wire-type displacement sensors (sampling rate: 128 Hz) were attached on structure joints to measure the ground-truth of dynamic displacement, for reference. The positions of these sensors are illustrated in Fig. 9(b), where D1 to D4 represent the positions of wire-type displacement sensors, and A0 to A4 represent the accelerometers. In this study, the structural displacement of Floor 1 in Case 1 was used for experimental verification. As the experiment started, the RC structure vibrated and all sensors began to record.

To evaluate the performance of the proposed methodology, two single-source displacement measurement approaches, namely purely vision-based (sampling rate: 25 Hz) and acceleration-based (sampling rate: 256 Hz), were introduced for comparison. Structural displacement measured by wire-type displacement sensors were used as reference for evaluating the accuracies of different methods. For the purely vision-based method, we used the initial LK optical flow-tracking algorithm provided in the MATLAB CV-Toolbox, and for the purely acceleration-based method, we used the integration-detrending strategy described in Section 3.2, with no data fusion. Three methods were implemented independently to measure the displacement of ROI I in Case 1. The validation details are illustrated in Table 2.

4.2. Experimental results

Fig. 10(a) shows the initial frame of the video data, containing an overall view of the RC frame. The frame shows that the column closer to our camera viewpoint is not straight because of radial image distortion. The distortion should be fixed to ensure an accurate pixel-to-meter conversion. Based on the positions of the nodes to be observed, four regions of interest (ROIs) were selected for the vision-based measurements. The lengths and widths of the ROIs were set to approximately 80 and 60 pixels, respectively. Both feature point detection and scale factor estimation were implemented within the range of ROIs.

Fig. 10(b) shows the positions of the ROIs and the detected feature points. As the excitation was activated, the extracted feature points vibrated, and their displacements were recorded. Table 3 presents the number of detected feature points, residual feature points after data filtering, and the corresponding ROI information. Based on the results, the ROIs representing higher floors were more affected by image distortion and incorrect point selection, resulting in more discarded feature points and lower point survival rates.

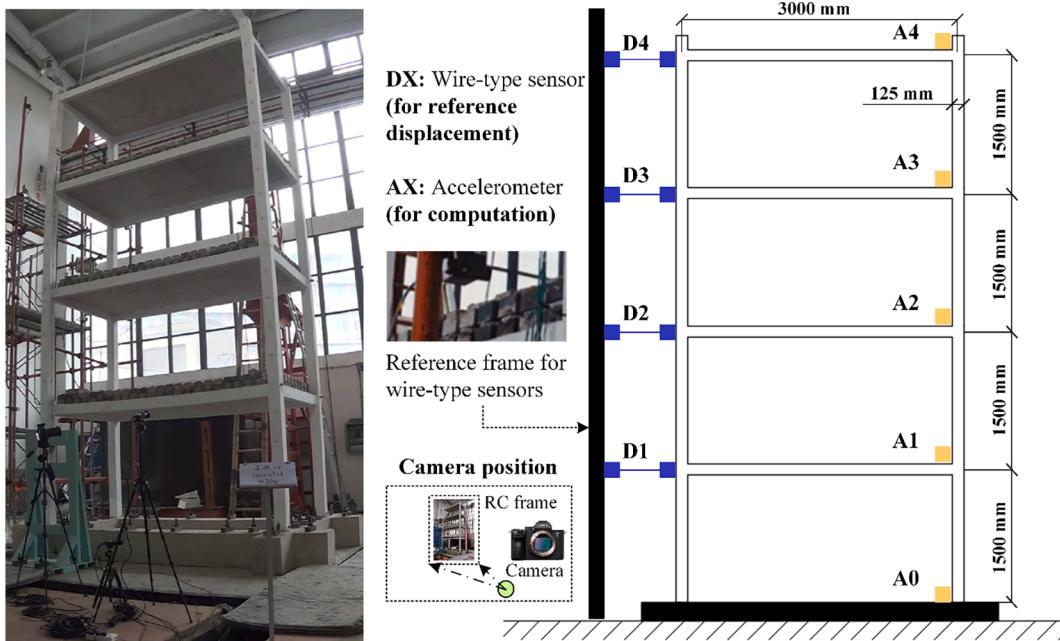


Fig. 9. (a) Panoramic schematic of the RC structure; (b) Design diagram of the structure with essential geometric information and sensor positions.

Table 1

Information about excitation sources employed in the experiment.

Case	Excitation source
1	North-south (NS) component of the 1940 Imperial Valley earthquake measured at El Centro station
2	NS component of the 1994 Northridge earthquake measured at the Sylmar station
3	Simulated Shanghai wave provided in a seismic design code for Shanghai

Table 2

Illustration of measurement methods for validation.

Measurement method	Source	Description	Sampling rate
Vision-based	Video data		25 Hz
Acceleration-based	Accelerometer		256 Hz
Proposed data fusion-based	Video + Acceleration		256 Hz
Vibration-based	Wire-type sensor	As $u_{\text{reference}}$	128 Hz

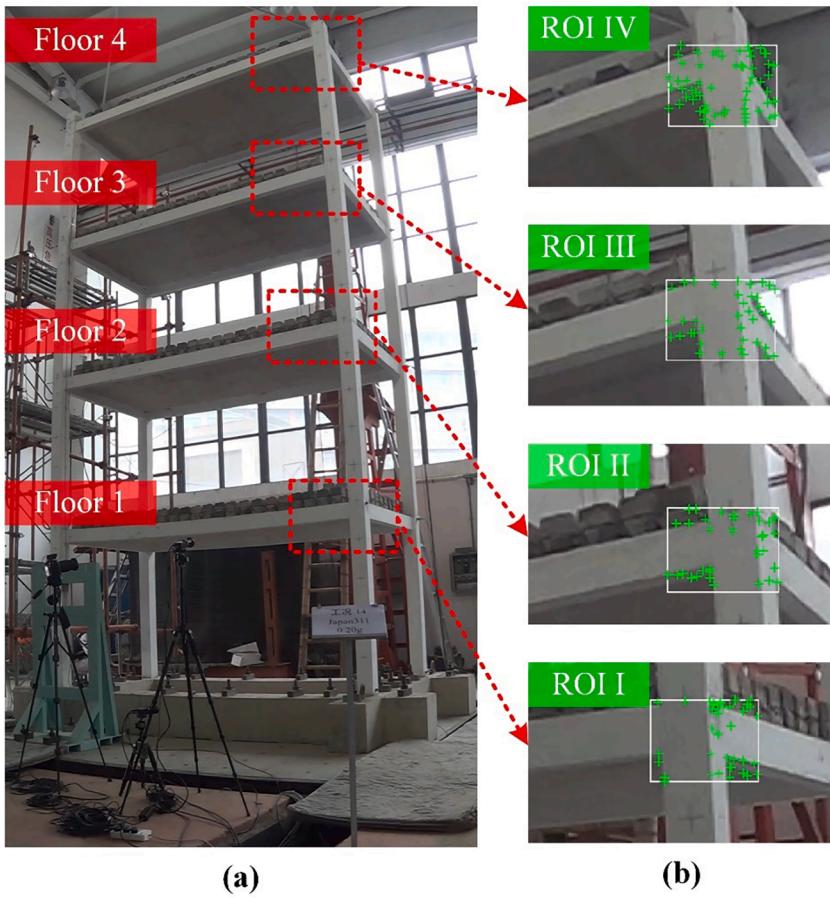
**Fig. 10.** Extracted feature points from each floor and each case of the RC frame setup.

Fig. 11 shows the time histories of the scale factors calculated for ROI I. The mean values of the scale factors from ROIs I to IV were 0.2838, 0.2716, 0.2528, and 0.2348 mm/pixel, respectively.

As the scale factors were calculated, the vision-based displacement $u_{\text{vision}}(t)$ was obtained. Next, the acceleration-based displacement, $u_{\text{acceleration}}(t)$, was calculated using Eq. (15). Before data fusion, the vision-based displacement was resampled from 25 Hz to 256 Hz and synchronized with acceleration-based displacement. In the data-fusion stage, by setting the penalty coefficient β to 2 and threshold ϵ_1, ϵ_2 to 0.5, the parameters K_1 and K_2 were both set to 5 using Eq. (20). Based on the noise density of the two signals, the cutoff frequency f_c was set to 1.82 Hz using Eq. (21). The recomposition parameter D was set to 2 based on Eqs. (22) and (23). Finally, the reconstructed displacement was calculated using Eq. (18).

Table 3

Information of feature points and scale factors for obtaining displacement dataset 1.

Floor	1	2	3	4
ROI size (pixels)	80×60	80×60	80×60	80×60
Detected feature points (number)	36	38	38	60
Residual feature points (number)	34	35	34	50
Point survival rate (%)	94.4	92.1	89.5	83.3
Average value of scale factors (mm/pixel)	0.2838	0.2716	0.2528	0.2348
Standard deviation of scale factors	0.0046	0.0014	0.0019	0.0017

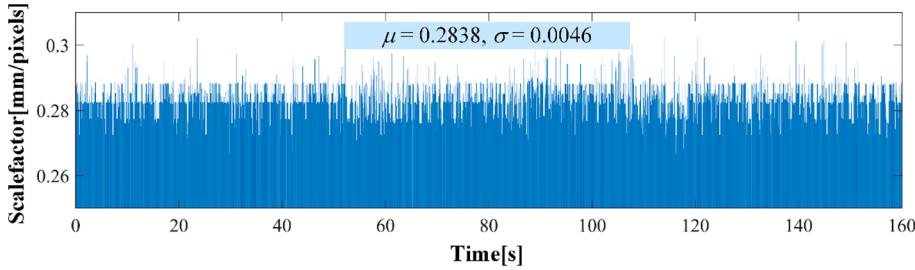
**Fig. 11.** Estimated scale factors for ROI I.

Fig. 12(a) shows the measurement results. The trajectory of the measurement made using our proposed method is in general agreement with the reference. **Fig. 12(b)** shows the details of a short period in the displacement history. As shown in the figure, the measurement made by the purely vision-based method deviates from the ground-truth trajectory because of its low-frequency error. **Fig. 12(c)** show the measurement errors of the three methods, reflecting their absolute accuracies. The acceleration-based method has the highest maximum error due to its error accumulation from the high noise level in the integration process.

The measurement errors were analyzed using four metrics: maximum error, average error (μ), root-mean-square error (RMSE), and $\pm 2 \sigma$, where σ denotes the standard deviation of the error distribution. The maximum error reflected the absolute accuracy of the measurements. μ reflects the drift of the measurement away from the ground truth. The RMSE was calculated using Eq. (24), where u_{measured} denotes the displacement measured by either the vision-based, acceleration-based, or proposed data-fusion-based method, $u_{\text{reference}}$ denotes the vibration-based measurement, and n denotes the number of sample points. This metric reflects the deviations in the measurement and the corresponding ground truth. $\pm 2 \sigma$ corresponds to a 95 % confidence level of the measurement error. The method with the lowest error range had a lower 2σ . It must be mentioned that systematical errors at a small level brought by sensor computational truncation, sample rounding during measurement procedures (especially in the reference measurement obtained by the wire-type sensor) unavoidably affects the evaluation of precision, leading the precisions of values to ten-thousandths of a millimetre. As a result, the precisions of errors were rounded to hundredths of a millimetre, followed by a range of ± 0.01 mm.

$$RMSE = \sqrt{\sum_i (u_{\text{measured}} - u_{\text{reference}})^2 / n} \quad (24)$$

Table 4 shows the calculated metrics. Compared with conventional vision-based and acceleration-based methods, the proposed data-fusion method produced the lowest maximum error of 1.92 ± 0.01 mm and an average error of -0.03 ± 0.01 mm. The vision-based method produced the lowest $\pm 2 \sigma$; however, its μ was relatively higher due to the low-frequency drift. In contrast, the acceleration-based method produced the widest measurement error range due to the high noise level in the low-frequency domain. This also affected the error range of the proposed method because the combined displacement had a high-frequency component from the acceleration-based measurement. **Fig. 13** shows the measurement error distributions in the ROI I for the vision, acceleration, and proposed methods.

The RMSE values calculated using Eq. (24) for the vision-based, acceleration-based, and proposed methods were 0.37 ± 0.01 , 0.60 ± 0.01 , and 0.30 ± 0.01 mm, respectively. This result suggests that the proposed data-fusion method can reduce noise more effectively.

Fig. 14 shows a comparison of the measurements in the frequency domain. The purely vision-based method has a restricted frequency range because of its inherently low sampling rate, in which higher frequencies cannot be detected correctly. In contrast, the proposed data-fusion method covers frequency components ranging from the DC to the Nyquist frequency of the acceleration measurement (128 Hz); thus, the modal frequencies from the first order to the fourth order could be correctly detected. As the high-frequency component of the vision-based measurement was filtered, the noise level reduced significantly. This result demonstrates the anti-noise capability of the proposed method. **Fig. 15** illustrates the comparisons in measurements in each mode. The acceleration-based and proposed data-fusion-based methods had obtained great results in detecting all modal frequencies, yet the vision-based method failed to detect the fourth mode, indicating its low performance for detecting high modal frequencies.

The accuracy for frequency spectra of these measurement methods was measured by the error of identified modal frequency from

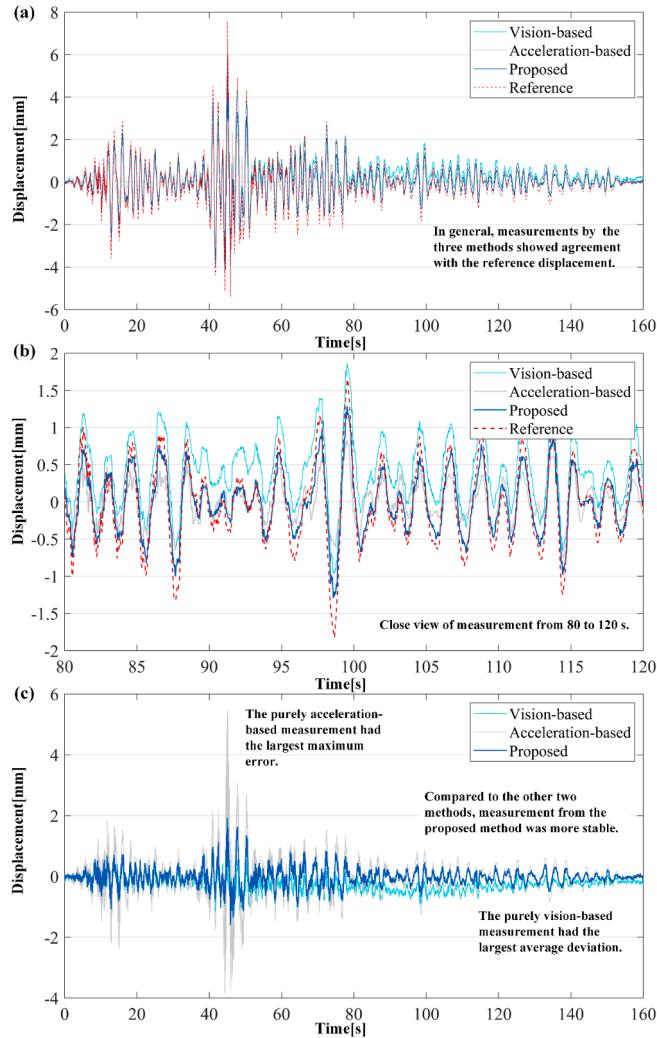


Fig. 12. Comparison of the displacements measured by three different methods in the experiment.

Table 4

Results of the accuracy of vision-based, acceleration-based, and proposed data-fusion methods in time domain.

Method	Maximum error (1 ± 0.01 mm)	μ (1 ± 0.01 mm)	RMSE (1 ± 0.01 mm)	$\pm 2\sigma$ (1 ± 0.01 mm)
Vision-based	2.46	0.28	0.37	0.49
Acceleration-based	5.45	-0.03	0.60	1.20
Proposed data-fusion	1.92	-0.03	0.30	0.59

the measurements:

$$\text{Accuracy} = \left(1 - \frac{|f_{\text{measured}} - f_{\text{reference}}|}{f_{\text{reference}}} \right) \times 100\% \quad (25)$$

where f_{measured} refers to modal frequency detected by the purely vision-based, acceleration-based or proposed data-fusion method, $f_{\text{reference}}$ denotes modal frequency detected by reference displacement from wire-type sensors. The accuracy for frequency spectra of these methods is shown in Table 5. Among these methods, the purely acceleration-based method attained the best accuracy for frequency spectra, as its accuracy for the first, second and fourth modal frequency reached 100 %. The proposed data-fusion method also detected all four modal frequencies, with the accuracy of 100 %, 100 %, 99.41 % and 99.79 % from the first to fourth mode. The purely vision-based method was less accurate, with the accuracy of 99.57 %, 99.71 %, 99.85 % from the first to third mode, and failed to detect the fourth modal frequency.

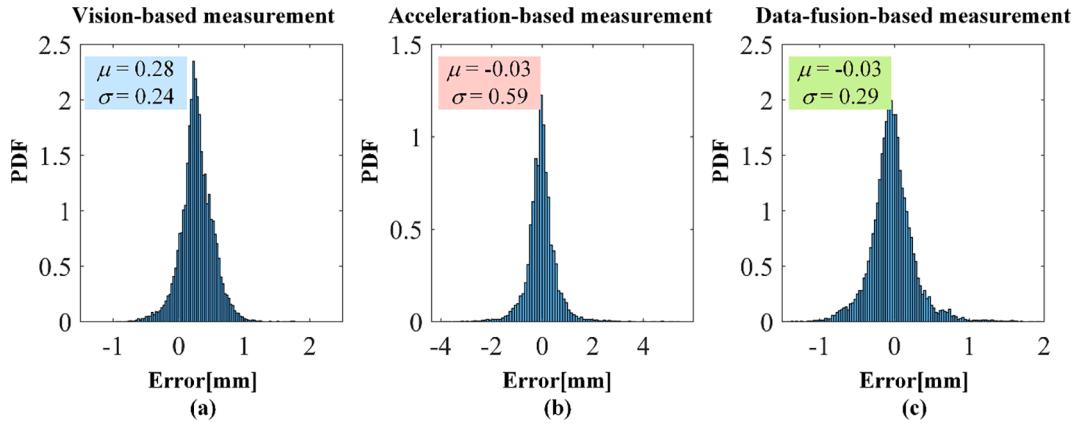


Fig. 13. Measurement error distributions of (a) Purely vision-based method, (b) Acceleration-based method, and (c) Proposed data-fusion method.

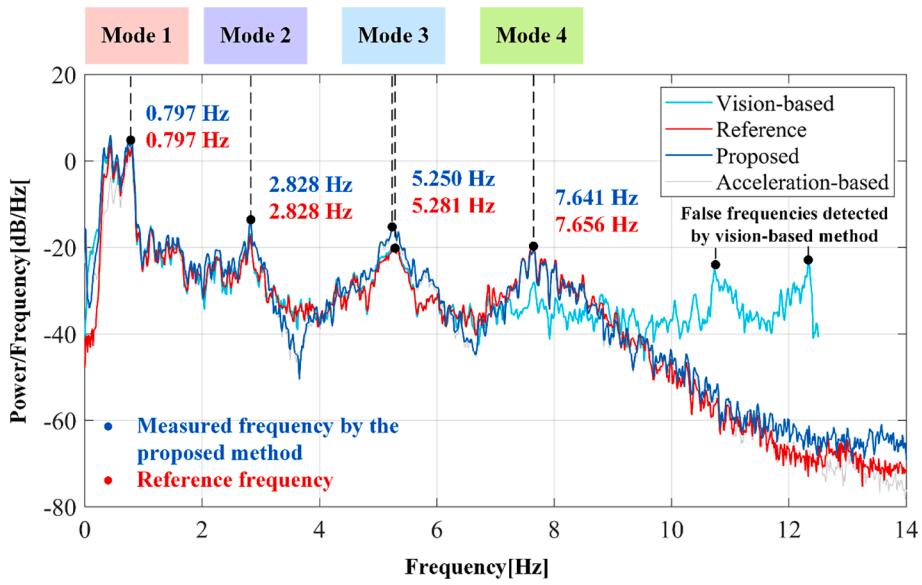


Fig. 14. Power spectral density of the measurement results.

5. Discussion

5.1. Optimal K for mode decomposition

For this study, it is important to discuss how the parameter K affect the accuracy of displacement measurement. The determination of K is based on overbinning evaluation. Overbinning means that the important parts of input signal are shared by two or more modes [31]. Table 6 shows the center frequencies of each mode decomposed from the vision- and acceleration-based measurements, corresponding to different K . Take signal 1 as an example, when $K = 6$, the center frequencies of the 5th and 6th mode are very close, which means the signal is over segmented. Therefore, the optimal K for both signals is 5.

Note that the measurement should be similar to the true displacement in frequency domain, the optimal K should correspond to the decomposition result similar to true displacement. From Table 6 and the power spectral density shown in Fig. 14, we can see when $K = 5$, each modal frequency is close to the ground-truth. The result is the same as determining K by overbinning evaluation. This indicates that the proposed method is able to get a series of modes representing real structural frequencies, and also proves that parameters of overbinning evaluation (e.g., ϵ_1 , β) proposed in section 4.4 can be told in advance, based on prior knowledge of the excitation or the structure form. However, the specific method to do so is beyond this manuscript. Fig. 16 shows the result of mode decomposition of vision-based measurement. As the optimal K is determined, frequency information is more reliable since low- and high-order modes are clearly divided, which ensures a good result for mode recomposition.

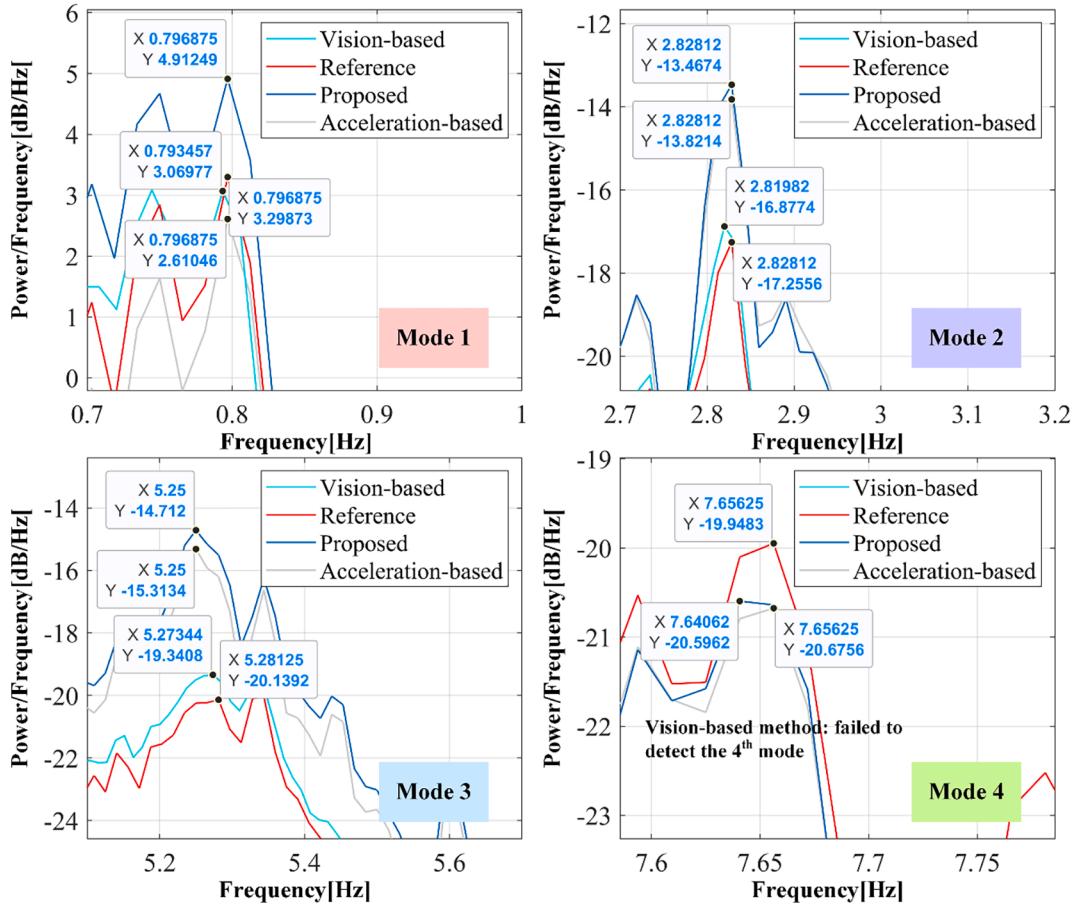


Fig. 15. Comparisons of methods to measuring different modes.

Table 5

Results of the accuracy of vision-based, acceleration-based, and proposed data-fusion methods for frequency spectra.

Method	Mode 1 (Hz)	Mode 2 (Hz)	Mode 3 (Hz)	Mode 4 (Hz)
Vision-based	0.7935	2.8198	5.2734	Failed
Accuracy	99.57 %	99.71 %	99.85 %	0 %
Acceleration-based	0.7969	2.8281	5.2500	7.6563
Accuracy	100 %	100 %	99.41 %	100 %
Proposed	0.7969	2.8281	5.2500	7.6406
Accuracy	100 %	100 %	99.41 %	99.79 %
Vibration-based (For reference)	0.7969	2.8281	5.2813	7.6563
	-	-	-	-

5.2. Analysis on automation of parameters

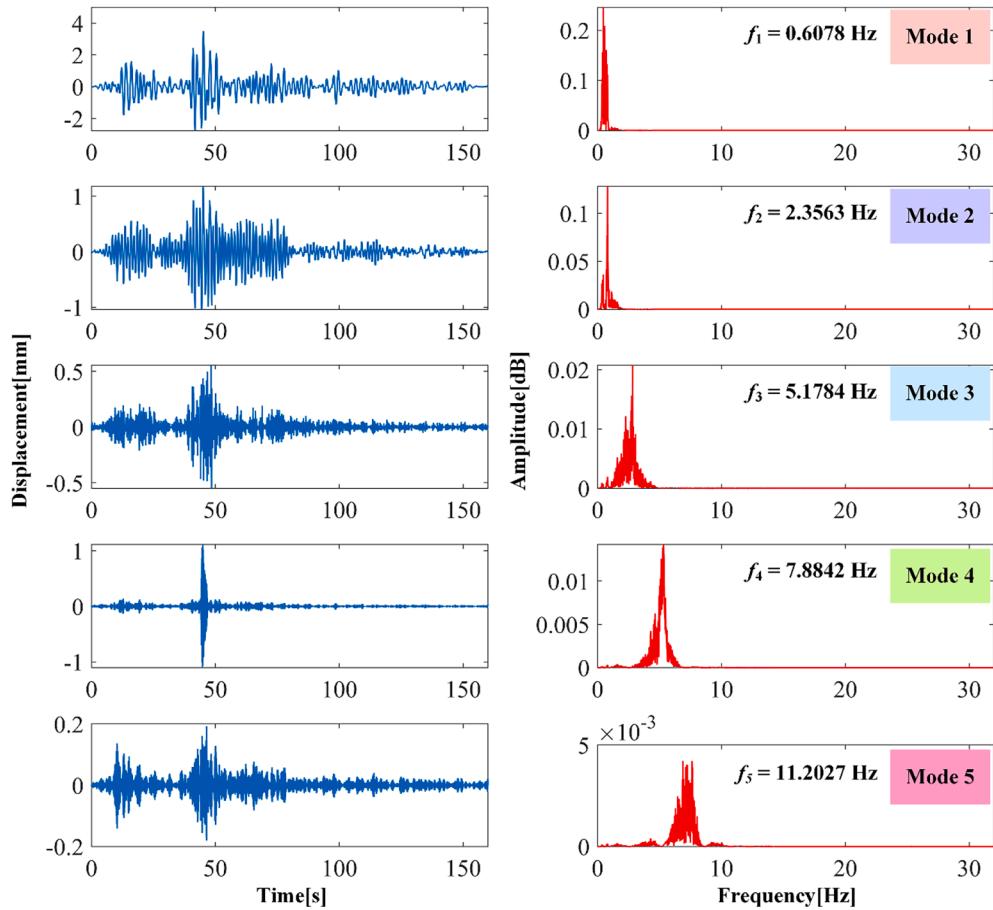
Table 7 provides the level of automation of parameters in the proposed framework. Most parameters of VMD (e.g., α , ε) have reference values. In the proposed method, ε_1 is usually set from 0.1 to 0.5. The smaller threshold leads to smaller tolerance of overbinning. The penalty coefficient β of mode orders should be adjusted according to the excitation or the structure form. In the mode recomposition stage, the close threshold ε_2 is usually set from 0.1 to 1, since the interaction of noise power spectral densities often lies in low-frequency area.

6. Conclusions

We developed a VMD-based multirate data-fusion framework to estimate the dynamic displacement of concrete structures. A scale model experiment was conducted on a four-story RC structure, in which the structural displacements measured using vision-based, acceleration-based, and proposed data-fusion methods were compared. The following conclusions can be drawn from the results.

Table 6Center frequencies of modes corresponding to different K .

Number of modes	Center frequencies of each mode (Hz)							
Vision-based displacement (Signal 1)								
	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
$K_1 = 2$	0.6208	5.4251	–	–	–	–	–	–
$K_1 = 3$	0.6178	4.8300	10.5615	–	–	–	–	–
$K_1 = 4$	0.6080	2.3947	5.3107	10.7822	–	–	–	–
$K_1 = 5$	0.6078	2.3563	5.1784	7.8842	11.2027	–	–	–
$K_1 = 6$	0.6078	2.3526	5.1731	7.7157	10.6573	12.2191	–	–
$K_1 = 7$	0.6076	2.3304	5.1589	7.1804	8.6820	10.7735	12.2581	–
$K_1 = 8$	0.5554	0.7739	2.5899	5.1769	7.1656	8.6471	10.7723	12.2593
Acceleration-based displacement (Signal 2)								
	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
$K_2 = 2$	0.0066	5.4655	–	–	–	–	–	–
$K_2 = 3$	0.0066	0.7066	5.4819	–	–	–	–	–
$K_2 = 4$	0.0066	2.7006	5.0101	7.7009	–	–	–	–
$K_2 = 5$	0.0066	0.6829	2.5812	5.2158	7.7687	–	–	–
$K_2 = 6$	0.0066	0.6836	2.5853	5.2108	7.6388	8.4840	–	–
$K_2 = 7$	0.0066	0.4450	0.6902	2.6278	5.2082	7.6469	8.4790	–
$K_2 = 8$	0.0066	0.3456	0.6893	2.0689	2.5839	5.1880	5.4228	7.7729

**Fig. 16.** Mode decomposition result of vision-based displacement in the case study.

The VMD-based data-fusion framework integrates vision- and acceleration-based measurements to combine the advantages of the two approaches. Vision-based measurements provide reliable low-frequency components, while acceleration-based measurements provide dynamic components over a wider range. Data fusion was performed using mode recomposition. A method for determining the

Table 7

Level of automation of parameters: (+) typically can be determined by experience, (~) automated when other required parameters are determined, (-) manual adjustment is mandatory, according to specific signals.

Stage	Parameters	α	τ	ε
VMD	K (~), can be decided by overbinning evaluation.	(+)	(+)	(+)
Overbinning evaluation	B (-)	ε_1 (+)		
Mode recombination	k_1, k_2 (+), decided by noise density power	ε_2 (+)		

recomposition parameters was also established to obtain an optimal combination of modes. Experimental data has proved that the proposed data-fusion framework improves the accuracy of the purely vision- and acceleration-based methods by leveraging their advantages. In addition, the framework extends the frequency range and improves the detection accuracy of high-order frequencies by introducing dynamic components from acceleration-based displacements.

Besides data fusion, this paper proposed an enhanced calibration-free vision-based displacement estimation approach improving the accuracy of the displacement measurement by automatically filtering incorrect feature points. The vision-based scale factor estimation method can not only perform 2D-to-3D displacement conversion but can also reduce the effect of image distortion. Also, it has been proved that displacement can be indirectly measured from the acceleration using a simple VMD-based integration-detrending technique proposed in this paper, with restricted boundary effects.

Nevertheless, the proposed method has the following limitations:

First, the selection process of some of the recombination parameters, such as the penalty coefficient β and threshold ε , is not perfectly automated. Depending on the case, these parameters should be adjusted to specific values based on the excitation and structure forms.

In terms of efficiency, the computation time for the VMD increases rapidly with increasing number of modes K . If K is excessively high, the search for an optimal K will be time-consuming. Furthermore, the determination of parameter β is beyond the present study. It is necessary to develop more effective and self-adaptive algorithms for parameter search in the future.

Besides, the data fusion involves a resampling procedure which unavoidably causes deviations between original signal and resampled signal. Such deviations should be considered and thoroughly checked in different cases, since they can only be measured qualitatively, instead of quantitatively. In some cases where structure has small responses (e.g., at the level of 10 mm, like the case in this paper), the deviations can be considered diminutive; however, in cases where larger responses are expected, such deviations may affect final measurement accuracy since the absolute values of error increase as signal amplitudes increase. The errors of resampling should be evaluated in different cases otherwise the processed signal may lack of reliability in describing structural response.

As an experimental study, whether the proposed technique can be adopted for other types of structures (e.g., long-span bridges) or not is beyond this manuscript. For further exploration, the discrepancy in structural types should be considered to improve the generality of the proposed framework.

CRediT authorship contribution statement

Zhenfen Jin: Writing – review & editing, Validation, Supervision, Investigation, Conceptualization. **Guyuan Chen:** Writing – original draft, Validation, Software, Methodology, Conceptualization. **Yanbo Niu:** Methodology. **Congguang Zhang:** Methodology. **Xiaowu Zhang:** Methodology. **Jiangpeng Shu:** Writing – review & editing, Project administration, Funding acquisition.

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.

Acknowledgments

This research is financially supported by the Key R&D Program of Zhejiang (2023C01161), Science and Technology Research Project of Zhejiang Province Transportation Department (202217), Zhejiang Provincial Natural Science Foundation (LQ22E080004). The authors would also like to thank the organizations of the 3rd International Competition for Structural Health Monitoring (IC-SHM 2022), ANCRiSST, for their support and generosity in providing valuable data.

References

- [1] G.W. Housner, L.A. Bergman, T.K. Caughey, A.G. Chassiakos, R.O. Claus, S.F. Masri, R.E. Skelton, T.T. Soong, B.F. Spencer, J.T.P. Yao, Structural Control: Past, Present, and Future, *J. Eng. Mech.* 123 (1997) 897–971, [https://doi.org/10.1061/\(ASCE\)0733-9399\(1997\)123:9\(897\)](https://doi.org/10.1061/(ASCE)0733-9399(1997)123:9(897)).
- [2] K.S.C. Kuang, S.T. Quek, C.G. Koh, W.J. Cantwell, P.J. Scully, Plastic Optical Fibre Sensors for Structural Health Monitoring: A Review of Recent Progress, *Journal of Sensors*. 2009 (2009) 1–13, <https://doi.org/10.1155/2009/312053>.
- [3] B. LeBlanc, C. Nierzwicki, P. Avitabile, Structural health monitoring of helicopter hard landing using 3D digital image correlation, in: T. Kundu (Ed.), San Diego, California, USA, 2010: p. 76501V. <https://doi.org/10.1117/12.847318>.
- [4] T. Guo, Y.-W. Chen, Field stress/displacement monitoring and fatigue reliability assessment of retrofitted steel bridge details, *Engineering Failure Analysis*. 18 (2011) 354–363, <https://doi.org/10.1016/j.engfailanal.2010.09.014>.
- [5] J. Li, H. Hao, K. Fan, J. Brownjohn, Development and application of a relative displacement sensor for structural health monitoring of composite bridges: A relative displacement sensor for structural health monitoring, *Struct. Control Health Monit.* 22 (2015) 726–742, <https://doi.org/10.1002/stc.1714>.
- [6] Y. Xu, J.M.W. Brownjohn, Review of machine-vision based methodologies for displacement measurement in civil structures, *J Civil Struct Health Monit.* 8 (2018) 91–110, <https://doi.org/10.1007/s13349-017-0261-4>.
- [7] M. Gul, F.N. Catbas, H. Hattori, Image-Based Monitoring of Open Gears of Movable Bridges for Condition Assessment and Maintenance Decision Making, *J. Comput. Civ. Eng.* 29 (2015) 04014034, [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000307](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000307).
- [8] S. SONY, S. Laventure, A. Sadhu, A literature review of next-generation smart sensing technology in structural health monitoring, *Struct Control Health Monit.* 26 (2019) e2321.
- [9] J.-W. Park, J.-J. Lee, H.-J. Jung, H. Myung, Vision-based displacement measurement method for high-rise building structures using partitioning approach, *NDT & E International*. 43 (2010) 642–647, <https://doi.org/10.1016/j.ndteint.2010.06.009>.
- [10] S. Wereley, L. Gui, A correlation-based central difference image correction (CDIC) method and application in a four-roll mill flow PIV measurement, *Exp Fluids*. 34 (2003) 42–51, <https://doi.org/10.1007/s00348-002-0529-1>.
- [11] C.-Z. Dong, F.N. Catbas, A non-target structural displacement measurement method using advanced feature matching strategy, *Advances in Structural Engineering*. 22 (2019) 3461–3472, <https://doi.org/10.1177/1369433219856171>.
- [12] C.E. Willett, M. Gharib, Digital particle image velocimetry, *Experiments in Fluids* (1991).
- [13] S. Acikgoz, M.J. DeJong, K. Soga, Sensing dynamic displacements in masonry rail bridges using 2D digital image correlation, *Struct Control Health Monit.* 25 (2018) e2187.
- [14] P. Ruhnau, T. Kohlberger, C. Schnrr, H. Nobach, Variational optical flow estimation for particle image velocimetry, *Exp Fluids*. 38 (2005) 21–32, <https://doi.org/10.1007/s00348-004-0880-5>.
- [15] J. Shu, C. Zhang, X. Chen, Y. Niu, Model-informed deep learning strategy with vision measurement for damage identification of truss structures, *Mechanical Systems and Signal Processing*. 196 (2023) 110327, <https://doi.org/10.1016/j.ymssp.2023.110327>.
- [16] L. Luo, M.Q. Feng, Z.Y. Wu, Robust vision sensor for multi-point displacement monitoring of bridges in the field, *Engineering Structures*. 163 (2018) 255–266, <https://doi.org/10.1016/j.engstruct.2018.02.014>.
- [17] Y. Xu, J. Brownjohn, D. Kong, A non-contact vision-based system for multipoint displacement monitoring in a cable-stayed footbridge, *Struct Control Health Monit.* 25 (2018) e2155.
- [18] M.A. Kuddus, J. Li, H. Hao, C. Li, K. Bi, Target-free vision-based technique for vibration measurements of structures subjected to out-of-plane movements, *Engineering Structures*. 190 (2019) 210–222, <https://doi.org/10.1016/j.engstruct.2019.04.019>.
- [19] V. Hoskere, J.-W. Park, H. Yoon, B.F. Spencer, Vision-Based Modal Survey of Civil Infrastructure Using Unmanned Aerial Vehicles, *J. Struct. Eng.* 145 (2019) 04019062, [https://doi.org/10.1061/\(ASCE\)ST.1943-541X.0002321](https://doi.org/10.1061/(ASCE)ST.1943-541X.0002321).
- [20] S. Bhowmick, S. Nagarajaiah, Z. Lai, Measurement of full-field displacement time history of a vibrating continuous edge from video, *Mechanical Systems and Signal Processing*. 144 (2020) 106847, <https://doi.org/10.1016/j.ymssp.2020.106847>.
- [21] S. Bhowmick, S. Nagarajaiah, Identification of full-field dynamic modes using continuous displacement response estimated from vibrating edge video, *Journal of Sound and Vibration*. 489 (2020) 115657, <https://doi.org/10.1016/j.jsv.2020.115657>.
- [22] D. Feng, M. Feng, E. Ozer, Y. Fukuda, A Vision-Based Sensor for Noncontact Structural Displacement Measurement, *Sensors*. 15 (2015) 16557–16575, <https://doi.org/10.3390/s150716557>.
- [23] Y. Fukuda, M.Q. Feng, M. Shinozuka, Cost-effective vision-based system for monitoring dynamic response of civil engineering structures, *Struct. Control Health Monit.* 17 (2010) 918–936, <https://doi.org/10.1002/stc.360>.
- [24] Y. Yang, X.B. Yu, Image analyses for video-based remote structure vibration monitoring system, *Front. Struct. Civ. Eng.* 10 (2016) 12–21, <https://doi.org/10.1007/s11709-016-0313-6>.
- [25] C. Igathinathane, L.O. Pordesimo, E.P. Columbus, W.D. Batchelor, S.R. Methuku, Shape identification and particles size distribution from basic shape parameters using ImageJ, *Computers and Electronics in Agriculture*. 63 (2008) 168–182, <https://doi.org/10.1016/j.compag.2008.02.007>.
- [26] S. Bhowmick, S. Nagarajaiah, Spatiotemporal compressive sensing of full-field Lagrangian continuous displacement response from optical flow of edge: Identification of full-field dynamic modes, *Mechanical Systems and Signal Processing*. 164 (2022) 108232, <https://doi.org/10.1016/j.ymssp.2021.108232>.
- [27] B.K.P. Horn, B.G. Schunck, Determining optical flow, *Artificial Intelligence*. 17 (1981) 185–203, [https://doi.org/10.1016/0004-3702\(81\)90024-2](https://doi.org/10.1016/0004-3702(81)90024-2).
- [28] B.D. Lucas, T. Kanade, Iterative image registration technique with an application to stereo vision, in: 1981. 674–679. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-0019647180&partnerID=40&md5=19abd273b2b030069091e2afab0f3ff>.
- [29] J. Zhu, Z. Lu, C. Zhang, A marker-free method for structural dynamic displacement measurement based on optical flow, *Structure and Infrastructure Engineering*. 18 (2022) 84–96, <https://doi.org/10.1080/15732479.2020.1835999>.
- [30] J.-W. Park, D.-S. Moon, H. Yoon, F. Gomez, B.F. Spencer Jr., J.R. Kim, Visual-inertial displacement sensing using data fusion of vision-based displacement with acceleration, *Struct Control Health Monit.* 25 (2018) e2122.
- [31] Y. Shao, L. Li, J. Li, S. An, H. Hao, Computer vision based target-free 3D vibration displacement measurement of structures, *Engineering Structures*. 246 (2021) 113040, <https://doi.org/10.1016/j.engstruct.2021.113040>.
- [32] Y.H. Hong, S.G. Lee, H.S. Lee, Design of the FEM-FIR filter for displacement reconstruction using accelerations and displacements measured at different sampling rates, *Mechanical Systems and Signal Processing*. 38 (2013) 460–481, <https://doi.org/10.1016/j.ymssp.2013.02.007>.
- [33] M. Gindy, H.H. Nassif, J. Velde, Bridge Displacement Estimates from Measured Acceleration Records, *Transportation Research Record*. 2028 (2007) 136–145. <https://doi.org/10.3141/2028-15>.
- [34] J.-W. Park, S.-H. Sim, H.-J. Jung, B. Jr, Development of a Wireless Displacement Measurement System Using Acceleration Responses, *Sensors*. 13 (2013) 8377–8392, <https://doi.org/10.3390/s130708377>.
- [35] D.M. Boore, Effect of Baseline Corrections on Displacements and Response Spectra for Several Recordings of the 1999 Chi-Chi, Taiwan, Earthquake. *Bulletin of the Seismological Society of America*. 91. (2004). 1199–1211. <https://doi.org/10.1785/0120000703>.
- [36] D. Arias-Lara, J. De-la-Colina, Assessment of methodologies to estimate displacements from measured acceleration records, *Measurement*. 114 (2018) 261–273, <https://doi.org/10.1016/j.measurement.2017.09.019>.
- [37] D.M. Boore, Comments on Baseline Correction of Digital Strong-Motion Data: Examples from the 1999 Hector Mine, California, Earthquake. *Bulletin of the Seismological Society of America*. 92. (2002). 1543–1560. <https://doi.org/10.1785/0120000926>.
- [38] A. Smyth, M. Wu, Multi-rate Kalman filtering for the data fusion of displacement and acceleration response measurements in dynamic system monitoring, *Mechanical Systems and Signal Processing*. 21 (2007) 706–723, <https://doi.org/10.1016/j.ymssp.2006.03.005>.
- [39] C.C. Chang, X.H. Xiao, An integrated visual-inertial technique for structural displacement and velocity measurement, *Smart Structures and Systems*. 6 (2010) 1025–1039, <https://doi.org/10.12989/SSS.2010.6.9.1025>.

- [40] J. Kim, K. Kim, H. Sohn, Autonomous dynamic displacement estimation from data fusion of acceleration and intermittent displacement measurements, Mechanical Systems and Signal Processing. 42 (2014) 194–205, <https://doi.org/10.1016/j.ymssp.2013.09.014>.
- [41] W.S. Chan, Y.L. Xu, X.L. Ding, W.J. Dai, An integrated GPS–accelerometer data processing technique for structural deformation monitoring, J Geod. 80 (2006) 705–719, <https://doi.org/10.1007/s00190-006-0092-2>.
- [42] R. Ferrari, F. Pioldi, E. Rizzi, C. Gentile, E.N. Chatzi, E. Serantoni, A. Wieser, Fusion of wireless and non-contact technologies for the dynamic testing of a historic RC bridge, Meas. Sci. Technol. 27 (2016) 124014, <https://doi.org/10.1088/0957-0233/27/12/124014>.
- [43] J. He, X. Guan, Y. Liu, Structural response reconstruction based on empirical mode decomposition in time domain, Mechanical Systems and Signal Processing. 28 (2012) 348–366, <https://doi.org/10.1016/j.ymssp.2011.12.010>.
- [44] Y. Wang, R. Markert, J. Xiang, W. Zheng, Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system, Mechanical Systems and Signal Processing. 60–61 (2015) 243–251, <https://doi.org/10.1016/j.ymssp.2015.02.020>.
- [45] X. Guanlei, W. Xiaotong, X. Xiaogang, Z. Lijia, Improved EMD for the analysis of FM signals, Mechanical Systems and Signal Processing. 33 (2012) 181–196, <https://doi.org/10.1016/j.ymssp.2012.07.003>.
- [46] K. Dragomiretskiy, D. Zosso, Variational Mode Decomposition, IEEE Trans. Signal Process. 62 (2014) 531–544, <https://doi.org/10.1109/TSP.2013.2288675>.
- [47] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.-C. Yen, C.C. Tung, H.H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, Proc. r. Soc. Lond. a. 454 (1998) 903–995, <https://doi.org/10.1098/rspa.1998.0193>.
- [48] D. Ribeiro, R. Calçada, J. Ferreira, T. Martins, Non-contact measurement of the dynamic displacement of railway bridges using an advanced video-based system, Engineering Structures. 75 (2014) 164–180, <https://doi.org/10.1016/j.engstruct.2014.04.051>.
- [49] J. Canny, A Computational Approach to Edge Detection, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-8 (1986) 679–698, <https://doi.org/10.1109/TPAMI.1986.4767851>.
- [50] P.J. Brockwell, R. Dahlhaus, A.A. Trindade, MODIFIED BURG ALGORITHMS FOR MULTIVARIATE SUBSET AUTOREGRESSION, (n.d.).
- [51] X. Jiang, J. Wang, J. Shi, C. Shen, W. Huang, Z. Zhu, A coarse-to-fine decomposing strategy of VMD for extraction of weak repetitive transients in fault diagnosis of rotating machines, Mechanical Systems and Signal Processing. 116 (2019) 668–692, <https://doi.org/10.1016/j.ymssp.2018.07.014>.
- [52] Y Niu, Y Ye, J* Shu, W Zhao, Y Duan, Identifying modal parameters of a multi-span bridge based on high-rate GNSS–RTK measurement using the CEEMD–RDT approach, Journal of Bridge Engineering, ASCE 26 (8) (2021) 04021049.
- [53] J Shu, C Zhang, Y Gao, Y* Niu, A multi-task learning-based automatic blind identification procedure for operational modal analysis, Mechanical Systems and Signal Processing 187 (2023) 109959.
- [54] J* Shu, Z Zhang, I Gonzalez, R Karoumi, The application of a damage detection method using Artificial Neural Network and train-induced vibrations on a simplified railway bridge model, Engineering Structures 52 (2013) 408–421.
- [55] Liu, G, Niu, Y, Zhao, W, Duan, Y, Shu, J*. Data anomaly detection for structural health monitoring using a combination network of GANomaly and CNN, Smart Structures and Systems, 2022, 39(1), 195–206.
- [56] J Wang, J Zhao, Y Liu, J Shan, Vision-based displacement and joint rotation tracking of frame structure using feature mix with single consumer-grade camera, Struct Control Health Monit 28 (12) (2021) e2832, <https://doi.org/10.1002/stc.2832>.