Multiple targets displacement response identification in bridges via an enhanced monocular vision-based method

Weizhu Zhu [a,b], Zimeng Shen [c,d], Tengjiao Jiang [e,*], Tong Wu [f], Zhixiang Zhou [a,b]

[a] College of Civil and Transportation Engineering, Shenzhen University, Shenzhen 518060, China
[b] National Key Laboratory of Green and Long-life Road Engineering in Extreme Environment, Shenzhen 518060, China
[c] School of Civil Engineering, Chongqing Jiaotong University, Chongqing 400074, China
[d] State Key Laboratory of Mountain Bridge and Tunnel Engineering, Chongqing 400074, China
[e] Department of Structural Engineering, Norwegian University of Science and Technology, Rich. Birkelands vei 1A, Trondheim 7491, Norway
[f] School of Architecture and Civil Engineering, Xihua University, Chengdu 610039, China

ARTICLE INFO

ABSTRACT

Routine displacement measurement of bridge structures is essential for accurate evaluation of structural safety. Vision-based displacement measurement methods, as a representative of non-contact techniques, offer distinct advantages over traditional contact-based approaches However, existing vision-based methods face challenges such as limited target coverage, high time costs, and complex operational requirements. To address these issues, this study introduces a simplified multi-target displacement response identification (MTDRI) method. The MTDRI integrates a subpixel displacement extraction technique with a novel monocular vision-based displacement measurement method. It requires only a single setup and three external parameters of the camera, *i.e.,* the nearest perpendicular distance (L_H), pitch (θ), and yaw (β) angle, to directly and automatically calculate the physical displacement of all targets, enhancing its applicability to field bridges. Static crossover validation tests and three sets of dynamic tests were conducted to evaluate the performance of the proposed method. The results show that the maximum error in the static displacements was 2.55 %, while the dynamic displacement responses achieved an RMSE of 0.43 mm and an NRSEM of 2.02 %. These findings confirm that the MTDRI outperforms traditional vision-based measurement methods and can accurately identify multi-target displacement responses of bridge structures.

1. Introduction

Civil engineering structures experience long-term deflection throughout their lifespans due to external loads [1–3]. In particular, prolonged exposure to high temperatures, humidity, heavy loads, and high-speed traffic can accelerate damage to long-lived bridge structures [4]. Consequently, bridge structural health monitoring (BSHM) for in-service bridges is critical to ensure their safe maintenance. Vertical displacement measurements provide valuable insights into both the static and dynamic behaviors of bridge structures [5–7]. These measurements serve as key indicators of structural response, enabling BSHM to assess bridge structural safety and detect potential risks under various external loads [8]. Thus, routine displacement measurement during structural inspections and condition assessments is crucial to identify deteriorating bridges and prevent catastrophic failures.

Traditional BSHM systems typically adopt contact sensors, such as accelerometers, linear variable differential transformers (LVDTs), and Global Positioning System (GPS), etc., to identify bridge responses. Accelerometers capture only the dynamic displacement, struggling with the pseudo-static component crucial for overall displacement [9]. LVDTs measure relative displacement but require physical access to the structure and a fixed reference point [10,11]. For urban bridges, especially those with vehicle traffic or navigation demands, installing dedicated support facilities for contact sensors presents significant challenges. Furthermore, GPS is primarily used for large-scale bridge structures to provide absolute displacement data by receiving signals from satellites. However, its lower accuracy and sampling rate its effectiveness for precise measurement [12,13]. In contrast, non-contact measurement techniques have gained increasing attention for identifying structural displacement responses [14–16]. These strategies acquire data without direct physical interaction with the structure, offering significant

---

Nomenclature

APVE	absolute peak value error
BSHM	bridge structural health monitoring
FOV	field of view
f	the focus lengths of the optical lens
GB-radar	ground-based radar interferometer
GPS	the global positioning system
LAC	a laser-and -camera
LDV	Laser Doppler Vibrometer
L_H	the horizontal distance between the structures and the camera
L_H^i	the distance between optical center and the target
LVDTs	linear variable differential transformers
IMU	Inertial Measurement Unit
MTDRI	multi-target displacement identification
NRMSE	the normalized root mean squared error
O	optical center
PBOF	phase-based optical flow
P_s^i	the targets in structural plane before deformation
P_s^t	the targets in structural plane after deformation
$P_s^i P_s^t$	the physical displacement of bridge structures
P_v^i	the targets in image plane before deformation
P_v^t	the targets in image plane before deformation
RMSE	the root mean square error
SFs	scale factors
S_c^x	x-the physical pixel size of the image sensor
S_c^y	y-the physical pixel size of the image sensor
x_v^c	x-image center coordination
y_v^c	y-image center coordination
x_v^i	x-image coordination of the targets before deformation
y_v^i	y-image coordination of the targets before deformation
x_v^t	x-image coordination of the targets after deformation
y_v^t	y-image coordination of the targets after deformation
UAVs	unmanned aerial vehicles
α	angle $\angle P_s^i O P_s^t$
β	the yaw angle
γ	angle $\angle P_c^o O P_v^1$
θ	the pitch angle
μ	angle $\angle O P_o^i P_o^t$
η	the chief ray angle
τ	angle $\angle O P_v^i P_v^c$
ψ	angle $\angle P_s^i P_s^t O$
φ	angle $\angle P_v^c P_v^1 O$
ϕ	angle $\angle P_v^i O P_v^1$
ω	angle $\angle O P_s^i P_s^t$

advantages in BSHM, such as enhanced spatial resolution and avoidance of mass-loading effects [17,18]. However, despite advancements in non-contact sensors, challenges remain. For instance, techniques like laser scanners can be time-consuming [19], ground-based radar (GB-radar) interferometer is costly [20], and Laser Doppler Vibrometer (LDV) requires complex operation [21], all of which limit their practical application in the field.

As an alternative non-contact measurement method, vision-based displacement measurement strategies offer distinct advantages over the methods above, making them a key area of intensive research for BSHM applications.

1.1. State-of-the-art (SOTA) review

1.1.1. Vison-based measurement strategies

Vision-based measurement strategies have gained widespread used for displacement measurement, driven by advancements in computer vision technologies [22–24]. Ge et al. [25] developed a vision-based displacement influence line (DIL) measurement system that integrates a Go-Pro camera and a consumer-grade camera to capture vehicle positioning and mid-span displacement, respectively. The proposed approach was validated through laboratory testing, demonstrating that the maximum standard deviation of the measured DILs was less than 0.08 mm. Han et al. [26] accurately identified the mid-span displacement response of a bridge by using unmanned aerial vehicles (UAVs), with a laser spot used as a reference to eliminate UVA motion. Field tests indicated that the normalized root mean square error (NRMSE) of the proposed method was 1.038 %. Xu et al. [27] integrated a Siamese network-based tracker and correlation-based template matching to enhance the precision of vision-based mid-span displacement measurements. A long-span road bridge test demonstrated that the proposed method has the capability of handling challenging scenarios, such as partial occlusion and illumination changes. Jeon et al. [28] applied a convolutional neural network (CNN) and modified Kanade-Lucas-Tomasi (KLT) algorithm to address the loss of feature points in traditional approaches, improving the robustness of cable displacement measurement. Lu et al. [29] integrated the YOLO v5 and digital image processing to mitigate the effects of ambient light changes, with experimental results showing maximum and minimum relative error of 0.4122 % and 0.0024 %, respectively. However, these strategies primarily focus on single-point displacement measurement, the limited targets result in incomplete response data, making it challenging to completely characterize structural response under loads. Therefore, developing techniques that enable direct measurement of structural displacement with high spatial resolution, *i.e.,* multi-target displacement measurement.

1.1.2. Multi-target displacement measurement strategies

Multi-target displacement response identification is essential for accurately assessing the health of bridge structures [30]. Typically, multi-camera approaches are utilized in field applications. Han et al. [31] developed a laser-and-camera (LAC) system for multi-target displacement measurement, utilizing multiple portable cameras to track laser spots projected from laser emitters. Validation tests demonstrated that the mean absolute percentage errors (MAPEs) were all under 5 %. Hu et al. [32] implemented a camera network consisting of several double-head camera stations to measure the multi-target displacement of bridges. An in-service bridge test showed that all correlation coefficients between the points in the main span exceeded 0.9. To reduce the number of cameras required and decrease system costs, multi-setup approaches have been explored. Wang et al. [33] proposed a vision-based multi-setup strategy that employs three time-synchronized cameras to capture multi-target displacement data and torsional modes of real-life bridges. Similarly, Tian et al. [34] proposed a mobile impact testing method, dividing the structure into multiple sub-regions by moving a single camera. Displacements were calculated using the Gradient-Based Hough Transform (GHT), identifying the displacement responses of six targets on a simply supported beam. However, multi-camera approaches require multiple cameras and a fixed reference point, while multi-setup approaches rely on manual calibration of each setup. As a result, both strategies involve high costs, complexity, and time consumption, posing challenges for real-time and automated measurement of bridge structures.
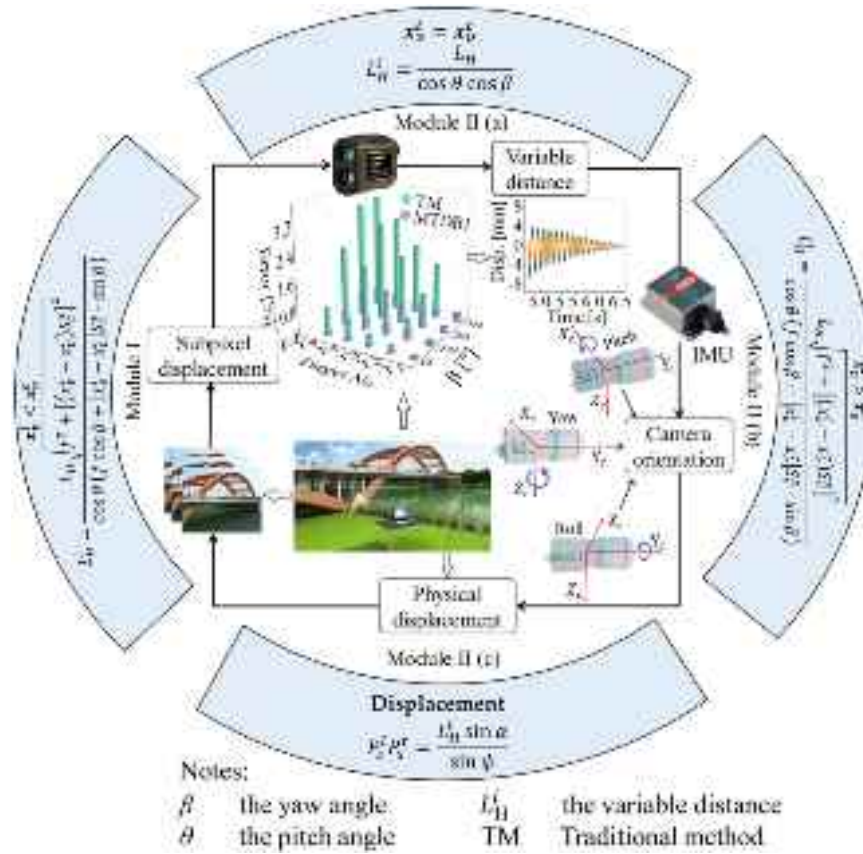
Fig. 1. Framework of multi-target displacement response identification (MTDRI).

1.1.3. Displacement transformation methods

Displacement transformation plays a pivotal role in converting pixel displacement into physical displacement of structures [35]. Typically, scale factors (SFs), an intermediate variable, are derived using the pinhole camera model to calculate physical displacement. However, the SFs for different targets within a constant field of view (FOV) vary and are sensitive to camera orientation [36]. Consequently, research into displacement transformation methods for multi-target displacement measurement has intensified. Wu et al. [37] employed a linear fitting method to establish the relationship between the SFs and the distances to the center target. Validation tests showed that the reconstructed displacement achieved a normalized root mean square error (NRMSE) of 0.24 % and an absolute peak value error (APVE) of 3.32 %. Similarly, Tian et al. [38] proposed an off-axis digital image correlation for multi-target displacement measurement, utilizing a spatial straight-line fitting scheme to determine SFs of all targets based on at least three distance sets. Yu et al. [39] simplified Tian's method by deriving the SFs of all targets using a curve fitting method and only a single distance measurement from the optical center to the object. However, these methods primarily rely on curve fitting to derive SFs from limited distance data of equidistance targets, rather than directly converting measurements to physical displacement, thus increasing the risk of errors.

1.2. Research gaps

Advancements in non-contact sensor technologies offer great potential for precise displacement measurement of bridge structures. However, there is still a lack of efficient, low-cost, and agile systems capable of accurately identifying structural displacement responses. Existing vision-based methods primarily focus on single or limited targe displacement measurements, which hinders the complete assessment of structural characteristics during the operation phase due to insufficient response data. Furthermore, previous studies highlight the challenges in automating multi-target displacement with current approaches. These gaps emphasize the urgent need for focused research to bridge the gap between theoretical developments and practice applications, delivering effective and tangible solutions.

1.3. Aim and structure of the study

This study introduces a monocular vision-based multi-target displacement response identification (MTDRI) approach for bridge structures, which is designed to directly calculate the physical displacement of all targets using only three external parameters. Both static and dynamic experiments are conducted for performance validation of the proposed method. The remainder of this study is organized as follows: Section 2 outlines the framework of MTDRI. Section 3 details the methodologies of the subpixel displacement extraction method and multi-target displacement transformation method. The static and dynamic experiments to validate the performance of MTDRI are described in Section 4. Finally, Section 5 concludes the major findings, along with inspiration for future works.

2. Framework of MTDRI

2.1. Conceptual framework

Aiming to resolve the shortcomings of the existing vision-based displacement measurement systems, this study proposes an improved vision-based multi-target dis-placement measurement system for bridge structures. The system comprises two major modules, *i.e.,* module I – subpixel multi-target displacement extraction, and module II –multi-target physical displacement transformation, as shown in Fig. 1.
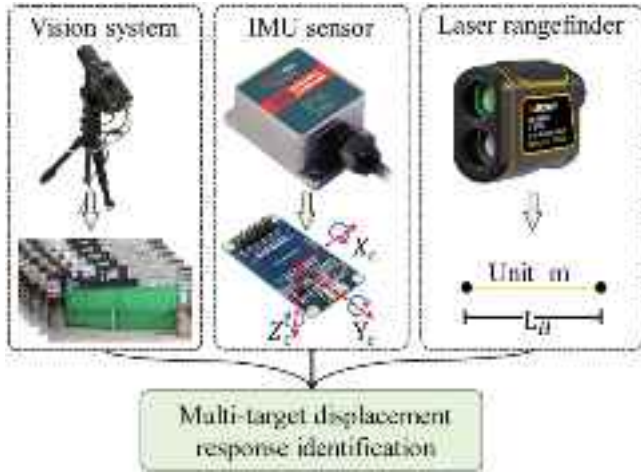
Fig. 2. Components of the proposed system.

2.2. System composition

In this study, the proposed vision-based multi-target displacement measurement system primarily requires only three measurement devices for high-precision measurement of bridge structures, as shown in Fig. 2. First, the image sequences including structural displacement response data are collected utilizing the high-resolution camera. Next, the orientation of the camera, *i.e.,* pitch, and yaw, are significant parameters for displacement calculation, which can be measured via an IMU sensor. Finally, the nearest perpendicular distance (L_H) between the optical center and the object is captured using a low-cost consumer-grade laser rangefinder. Notably, adjustments to the roll angle are generally unnecessary in field applications, as minor changes do not significantly affect the FOV [40], so this study does not address it.

3. Strategies of MTDRI

3.1. Subpixel displacement extraction

The accuracy of pixel displacement for each target on the bridge significantly impacts the measurement efficiency of the vision-based system. The pixel displacement is typically extracted using feature tracking algorithms, which track the changes in target coordinates between consecutive frames. The traditional feature tracking algorithms are restricted to extracting solely integer pixel displacements. However, in field applications, deformations smaller than an integer pixel, which are often barely visible, can occur in bridge structures with high stiffness, leading to poor precision in vision-based measurements. To address the issue, this study employs a robust phase-based optical flow (PBOF) algorithm, which focuses on the phase information of image signals. PBOF is more effective than traditional methods, particularly in capturing structural sub-pixel displacement.

The PBOF method is grounded in the assumption of brightness constancy over time, meaning that the motion information of an edge in an image is embedded in the spatial phase shifts between two consecutive frames. The principle is derived from the Fourier shift theorem and is localized utilizing the complex steerable [41]. Consequently, given an image $I(x,y,t)$, where x and y are the pixel coordinates and t denotes time, the image is convolved with a quadrature pair of steerable B_κ. The convolution output is a complex-valued signal $R(x,y,t)$, which can be expressed as:

$$R(x, \quad y, \quad t) = A_\kappa(x, \quad y, \quad t)e^{j\bullet\varnothing_\kappa(x, \quad y, \quad t)} = I(x, \quad y, \quad t) \otimes B_\kappa \quad (1)$$

where $A_\kappa(x,y,t)$ represents the edge strength or image contrast at pixel (x,y) and time t; $\varnothing_\kappa(x,y,t)$ is the spatial phase at pixel location (x,y) in
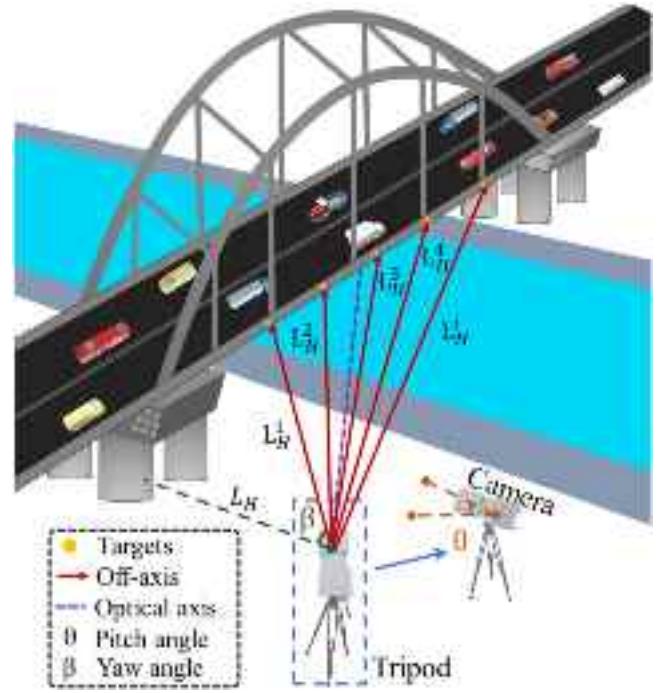
direction κ, at time t.

The spatial phase $\varnothing_\kappa(x,y,t)$ is related to the displacement of targets in the image, which is the core of the motion estimation. The displacement information can be extracted by calculating the phase difference between the reference frame ($t = 0$) and the current frame at time t. The phase difference $\Delta\phi_\kappa(x)$ for a pixel at location x is:

$$\Delta\phi_\kappa(x) = \phi_\kappa(x, \quad t) - \phi_\kappa(x, \quad 0) \quad (2)$$

where $\phi_\kappa(x, 0)$ is the phase of the reference at $t = 0$, and $\phi_\kappa(x,t)$ is the phase at time t.

The phase difference $\Delta\phi_\kappa(x)$ can then be integrated with the gradient of the reference image using an ordinary least squares (OLS) method to estimate the pixel displacement $\delta(t)$:

$$\delta(t) = \left(\frac{\partial\phi_\kappa(x, \quad 0)}{\partial\kappa}\right)^{-1} \bullet \Delta\phi_\kappa \quad (3)$$

3.2. Physical displacement transformation

3.2.1. Distance-solving model

In the vision-based measurement strategy, the distance between the optical center and the target is a vital parameter for calculating the physical displacement of the bridge. In the multi-target measurement, however, the distance varies as the target location changes, as shown in Fig. 3, where $L_H^1, L_H^2, L_H^3, L_H^4$, and L_H^i are unequal. This variation in distance presents a challenge for current vision-based methods to achieve high-precision multi-target displacement measurement. In the existing research, laser rangefinders are commonly utilized to measure these variable distances. This approach divides the structure into multiple subsections based on target regions and measures the distance between the optical center and each target. Thus, it falls under the multi-setup measurement strategy, which increases the complexity of the measurement system. Furthermore, the robustness and accuracy of the laser rangefinder are sensitive to both the illumination environment and the distance to the target. Its efficiency, particularly for long-distance and multi-target measurement, is significantly decreased. Therefore, the laser rangefinder is most effective for accurate short-distance measurements.
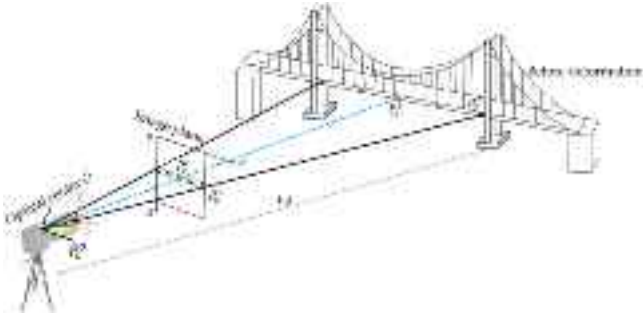

Fig. 3. Illustration of variable distance from the optical center to each target.

Fig. 4. Schematic of the variable distance solving method.

This study introduces a novel distance calculation method that incorporates the nearest perpendicular distance (L_H) from the optical flow to the optical center, along with the external parameters (*i.e.,* the pitch and yaw angle) of the camera, to automatically resolve the variable distance and simplify the operation procedures, as shown in Fig. 4.

The P_s^1 and P_s^i are the initial targets on the bridge, respectively, and the corresponding image coordinates are P_v^1 and P_v^i, respectively. The P_v^c denotes the image center coordinate. Additionally, the pith angle of the camera orientation is inevitable in field applications. Therefore, according to the Pythagorean theorem, the variable distance L_H^i can be expressed as:

$$L_H^i = \frac{L_H}{\cos\theta\sin(\phi + \gamma)} \tag{4}$$

where L_H denotes the nearest perpendicular distance from the optical center to the object, typically referring to the distance from the camera to the adjacent abutment. θ denotes the pitch angle of the camera, ϕ and γ represent the $\angle P_v^i O P_v^1$ and $\angle P_v^1 O P_c^o$ in the distance-solving model, respectively.

The value of L_H can be easily measured using a low-cost consumer-grade laser rangefinder, and the camera orientation parameters (θ) can be captured via the IMU sensor attached to the camera. However, the values of ϕ and γ are unknown and challenging to obtain, constricting the applications of this method. To overcome this, in this study, the equivalent substitution method was employed to eliminate the unknown parameters. According to the homothetic theory, the $\angle OP_v^1 P_v^i = \angle OP_o^1 P_c^s$ and $\angle P_v^1 OP_c^o = \angle P_o^1 OP_s^1$, which can be represented by φ and γ, respectively. The $\angle OP_v^1 P_s^s$ is the exterior angle of the $\Delta P_o^1 P_s^1 P_c^s$, so, the γ can be written as:

$$\gamma = \varphi - \beta \tag{5}$$

where β denotes the yaw angle of the camera, which also can be measured via the IMU sensor.

The value of γ is constant and is determined through the initial camera orientation. However, the ϕ is the internal parameter of the camera, which is a core factor for variable distance and varies with the location of the target changes. Therefore, the distance-solving model can be divided into three parts based on the image coordinate of the target. Firstly, if the target region is on the left in the image, i.e., $x_v^i < x_v^c$, as shown in Fig. 5(a), the x-coordinate of the target is less than the x-coordinate of the image center, as shown in Fig. 5(b). The ϕ is the interior angle of the $\Delta OP_v^i P_v^1$, thus, the ϕ can be expressed as:

$$\phi = \pi - \varphi - \tau \tag{6}$$

where τ represents the $\angle OP_v^i P_v^1$ of the $\Delta OP_v^i P_v^1$ and is also an internal parameter of the camera.

Secondly, if the target on the center in the image, as shown in Fig. 5 (c), the x-coordinate of the target coincides with the x-coordinate of the image center, i.e., $x_v^i = x_v^c$, as shown in Fig. 5(d). The $\Delta OP_v^i P_v^1$ is a right-

angled triangle due to OP_v^i is the optical axis, i.e., $\angle OP_v^i P_v^1 = \pi/2$. Thus, the ϕ can be expressed as:

$$\phi = \pi/2 - \varphi \tag{7}$$

Finally, if the target on the right in the image, as shown in Fig. 5(e), the x-coordinate of the target is larger than the x-coordinate of the image center, i.e., $x_v^i > x_v^c$, as shown in Fig. 5(f). The τ is the exterior angle of the $\Delta OP_v^i P_v^1$, thus, the ϕ can be expressed as:

$$\phi = \tau - \varphi \tag{8}$$

Consequently, according to the derivations above, the unknown parameters ϕ and γ are eliminated by the internal (τ) and external (β) parameters, which can be derived. Thus, Eq. (4) can be written as:

$$L_H^i = \begin{cases} \dfrac{L_H}{\cos\theta\sin(\tau + \beta)}, & x_v^i < x_v^c \\[2mm] \dfrac{L_H}{\cos\theta\cos\beta}, & x_v^i = x_v^c \\[2mm] \dfrac{L_H}{\cos\theta\sin(\tau - \beta)}, & x_v^i > x_v^c \end{cases} \tag{9}$$

where x_v^i and x_v^c represent the x-coordinate of the target and image center, respectively.

The $\Delta OP_v^c P_v^i$ is a right-angled triangle due to the optical axis OP_v^c perpendicular to the image plane, so, the $\angle OP_v^c P_v^i = \pi/2$ that is presented by the ρ. According to the sine low, the τ can be expressed as:

$$\tau = \sin^{-1}\frac{OP_v^c \bullet \sin\rho}{OP_v^i} = \sin^{-1}\left(\frac{f}{\sqrt{f^2 + \left[(x_v^i - x_v^c)S_c^x\right]^2 + \left[(y_v^i - y_v^c)S_c^y\right]^2}}\right) \tag{10}$$

where f denotes the focal length; (x_v^i, y_v^i) and (x_v^c, y_v^c) represent the image coordinate of the target and image center; S_c^x and S_c^y are lateral and longitudinal pixel size of the optical sensors, respectively, which can be obtained from the product manual provided by the manufacturer.

The proposed distance-solving model assumes that the targets are aligned along the horizontal center line of the image, meaning their y-coordinate are equal to the image center. This introduces the relationship $y_v^i - y_v^c = 0$. In summary, the variable distance L_H^i can be expressed as:

$$L_H^i = \begin{cases} \dfrac{L_H\sqrt{f^2 + \left[(x_v^i - x_v^c)S_c^x\right]^2}}{\cos\theta(f\cos\beta + |x_v^i - x_v^c|S_c^x \bullet \sin\beta)}, & x_v^i < x_v^c \\[3mm] \dfrac{L_H}{\cos\theta\cos\beta}, & x_v^i = x_v^c \\[3mm] \dfrac{L_H\sqrt{f^2 + \left[(x_v^i - x_v^c)S_c^x\right]^2}}{\cos\theta(f\cos\beta - |x_v^i - x_v^c|S_c^x \bullet \sin\beta)}, & x_v^i > x_v^c \end{cases} \tag{11}$$

3.2.2. Vertical displacement calculation method

BSHM typically focuses on vertical displacement under the external loads, as shown in Fig. 6(a). The simplified calculation model is expressed in Fig. 6(b).
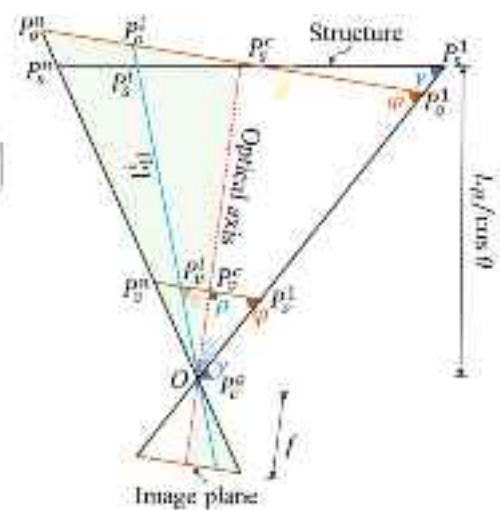
The initial target P_s^i on the structure moves to a new location P_s^t under external loads, causing the corresponding image coordinate to change from P_v^i to P_v^t. According to the cosines law, the vertical displacement $P_s^i P_s^t$ can be presented as:

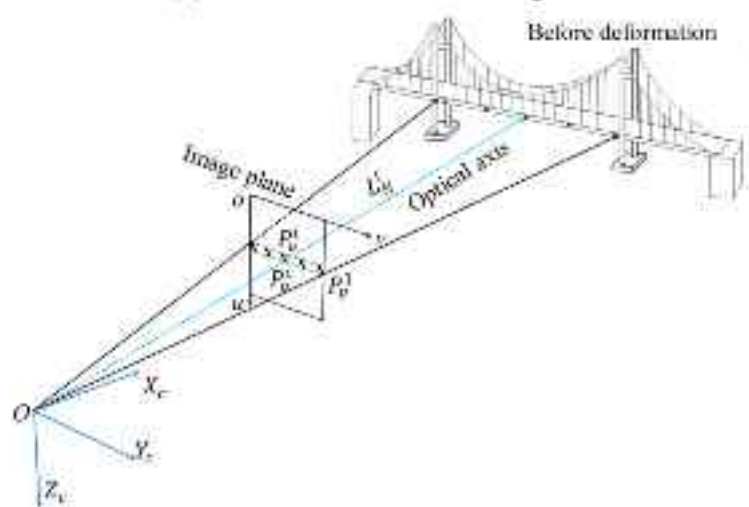$$P_s^i P_s^t = \frac{L_H^i \sin\alpha}{\sin\psi} \tag{12}$$

where α and ψ presents the $\angle P_s^i OP_s^t$ and $\angle P_s^i P_s^t O$, respectively, which are
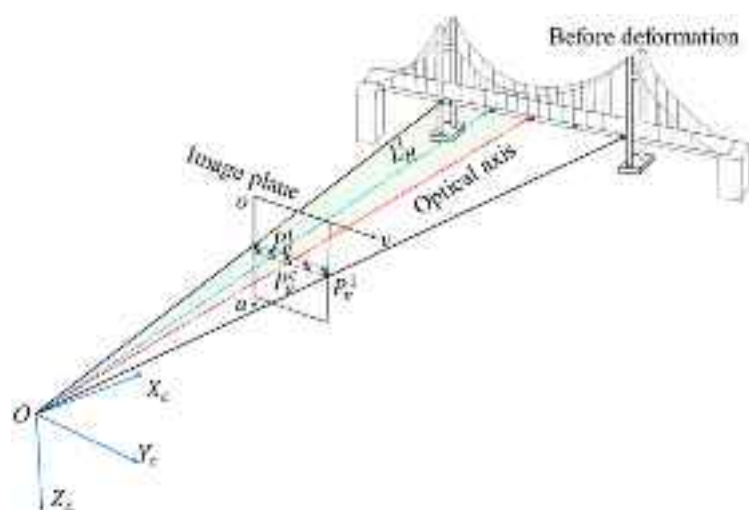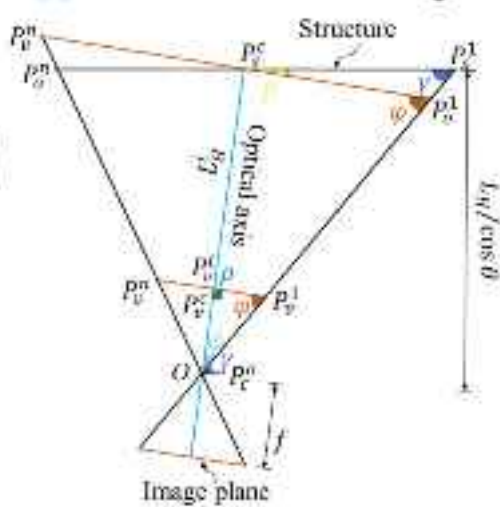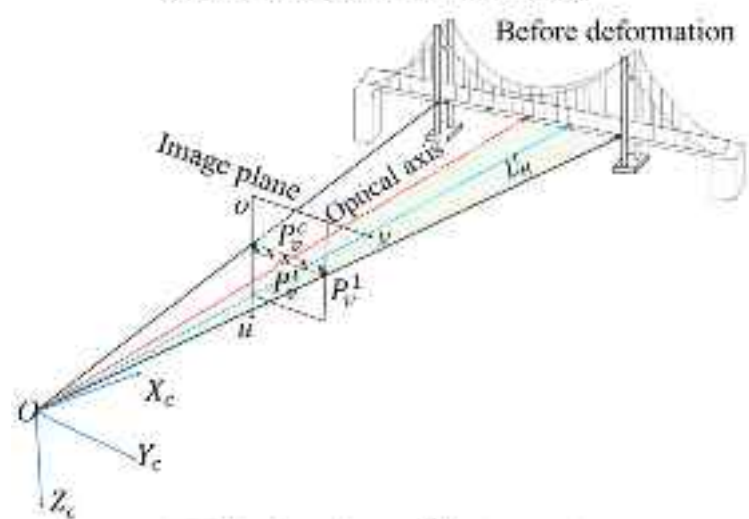
(a) the location of the targets

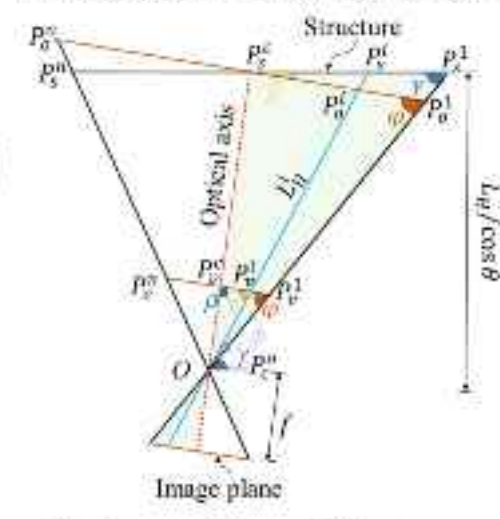(b) the coordinate of the targets

(c) the location of the targets

(d) the coordinate of the targets

(e) the location of the targets

(f) the coordinate of the targets

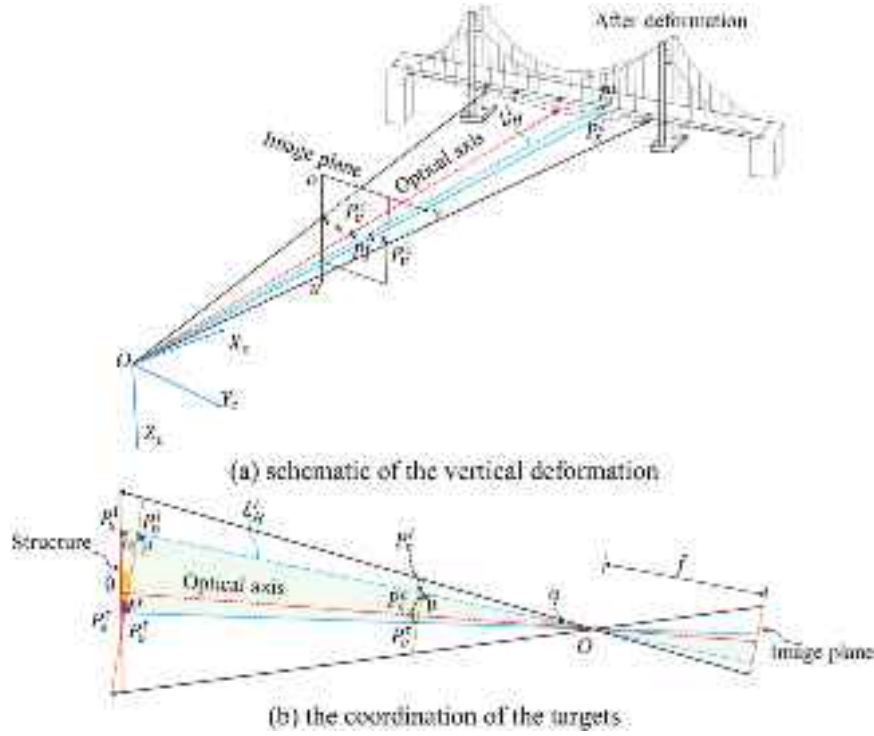**Fig. 5.** Schematic of distance-solving model.

Fig. 6. Illustration of the vertical displacement solving model.

also unknown parameters.

The α, ψ, and ω are the interior angles of the $\Delta OP_s^i P_s^t$, so, the ψ can be expressed as:

$$\psi = \pi - \omega - \alpha \tag{13}$$

According to the homothetic theory, $\angle P_v^t P_v^i O = \angle P_o^t P_o^i O$, represented by μ. Additionally, the $\angle P_o^t P_o^i O$ is the exterior angle of the $\Delta P_s^i P_s^m P_o^i$. Therefore, the ω can be written as:

$$\omega = \mu - \theta \tag{14}$$

Integrating the aforementioned distance-solving model, the vertical displacement $P_s^i P_s^t$ can be represented as:

$$P_s^i P_s^t = \begin{cases} \dfrac{L_H \sin\alpha \sqrt{f^2 + \left[(x_v^i - x_c^c)S_c^x\right]^2}}{(f\cos\beta + |x_v^i - x_v^c|S_c^x \bullet \sin\beta)\cos\theta\sin(\alpha + \mu - \theta)}, & x_v^i < x_v^c \\[3ex] \dfrac{L_H \sin\alpha}{\cos\theta\cos\beta\sin(\alpha + \mu - \theta)}, & x_v^i = x_v^c \\[3ex] \dfrac{L_H \sin\alpha \sqrt{f^2 + \left[(x_v^i - x_v^c)S_c^x\right]^2}}{(f\cos\beta - |x_v^i - x_v^c|S_c^x \bullet \sin\beta)\cos\theta\sin(\alpha + \mu - \theta)}, & x_v^i > x_v^c \end{cases} \tag{15}$$

The α and μ are the interior angles of the $\Delta OP_v^i P_v^t$, according to the cosine law, the α and μ can be expressed as:

$$\alpha = \cos^{-1}\left(\frac{OP_v^{i\,2} + OP_v^{t\,2} - P_v^i P_v^{t\,2}}{2OP_v^i \bullet OP_v^t}\right)$$

$$\mu = \cos^{-1}\left(\frac{P_v^i P_v^{t\,2} + OP_v^{i\,2} - OP_v^{t\,2}}{2P_v^i P_v^t \bullet OP_v^i}\right) \tag{16}$$

The OP_v^c is the optical axis of the camera, perpendicular to the image plane. Therefore, the $\Delta OP_v^i P_v^c$ and $\Delta OP_v^t P_v^c$ are right-angled triangles, and

$\angle OP_v^c P_v^i = \angle OP_v^c P_v^t = \pi/2$. Consequently, the $P_v^i P_v^t$, OP_v^i and OP_v^t can be expressed as:

$$P_v^i P_v^t = \sqrt{\left[(x_v^i - x_v^t)S_c^x\right]^2 + \left[(y_v^i - y_v^t)S_c^y\right]^2}$$

$$OP_v^i = \sqrt{OP_v^{c2} + P_v^c P_v^{i\,2}} = \sqrt{\left[(x_v^i - x_v^c)S_c^x\right]^2 + \left[(y_v^i - y_v^c)S_c^y\right]^2 + f^2}$$

$$OP_v^t = \sqrt{OP_v^{c2} + P_v^c P_v^{t\,2}} = \sqrt{\left[(x_v^t - x_v^c)S_c^x\right]^2 + \left[(y_v^t - y_v^c)S_c^y\right]^2 + f^2} \tag{17}$$

where (x_v^t, y_v^t) represents the image coordinate of the target after deformation.

According to the assumption of the distance-solving model and the target merely produces vertical displacement, so, Eq. (17) can be simplified as:

$$P_v^i P_v^t = |y_v^i - y_v^t|S_c^x$$

$$OP_v^i = \sqrt{\left[(x_v^i - x_v^c)S_c^x\right]^2 + f^2}$$

$$OP_v^t = \sqrt{\left[(x_v^t - x_v^c)S_c^x\right]^2 + \left[(y_v^t - y_v^c)S_c^y\right]^2 + f^2} \tag{18}$$

Thus, by incorporating the Eqs. (16) and (18), the physical displacement $P_s^i P_s^t$ can be derived from the Eq. (15). This demonstrates that the proposed MTDRI only requires three external parameters, *i.e.*, L_H, θ, and β, enhancing the practicability of vision-based measurement systems for field applications.

3.3. Algorithm – pseudo code version

Once the videos of structural motion and external parameters of the camera are capture by the measurement system, the displacement outcomes are computed using Algorithm 1.

Algorithm 1. Vision-based multi-target displacement response identification of bridges

---

Input: A video data $\mathbb{X}$, the nearest perpendicular distance L_H, the pitch θ and yaw β angles, the focal length f, the lateral S_c^x and longitudinal S_c^y pixel size of optical sensors.

Output: The physical displacement of all targets $\hat{\mathbb{y}} = \left[[\hat{\mathcal{Y}}_1], \cdots, [\hat{\mathcal{Y}}_i] \right]$.

1    $t_j[[\ ]]$←Obtain the pixel displacement of all targets by Eq. (3)
2    $t_j\left[t_1[x_v^i, y_v^i], \cdots, t_j[x_v^i, y_v^i] \right]$←Define the targets coordinates of first frame as the initial coordinates
3    $t_c[x_v^c, y_v^c]$←Obtain the center coordinate of frames
4   **for** $t_j[x_v^t, y_v^t] \in t_j[[\ ]]$ **do**
5   **for** $y_v^t \in t_j[\ ]$ **do**
6   Obtain intermediate variable OP_v^i, OP_v^t, and $P_v^i P_v^t$ by Eq. (17)
7   Calculate core parameters α, μ, and τ by Eq. (15) and Eq. (9), respectively
8   **if** $x_v^i < x_v^c$ **then**
9   Capture variable distance from the optical center to the targets L_H^i by Eq. (10)
10   Calculate displacement of targets $P_s^i P_s^t$ by Eq. (14)
11   **return** $\hat{\mathbb{y}} \leftarrow \left[\hat{\mathcal{Y}}_i \right]_{n \times 1}$
12   **elif** $x_v^i = x_v^c$ **then**
13   Capture variable distance from the optical center to the targets L_H^i by Eq. (10)
14   Calculate displacement of targets $P_s^i P_s^t$ by Eq. (14)
15   **return** $\hat{\mathbb{y}} \leftarrow \left[\hat{\mathcal{Y}}_i \right]_{n \times 1}$
16   **else**
17   Capture variable distance from the optical center to the targets L_H^i by Eq. (10)
18   Calculate displacement of targets $P_s^i P_s^t$ by Eq. (14)
19   **return** $\hat{\mathbb{y}} \leftarrow \left[\hat{\mathcal{Y}}_i \right]_{n \times 1}$
20   **end if**
21   **end for**
22 **end for**

---

4. Experimental validation

The robustness and efficiency of the proposed method were validated through static and dynamic tests conducted at the National Key Laboratory of Green and Long-life Road Engineering in Extreme Environment, Guangdong, China. The displacement measurements obtained by the proposed method were compared with the traditional method from literature [35] and the actual displacement from the laser rangefinder.

4.1. Static tests

4.1.1. Experiment setup

The uniformly distributed anchor holes in the reaction wall were employed to simulate the vertical displacement of multiple targets in the structure under external loads, as shown in Fig. 7. Five anchor holes located along the same row were selected to represent measurement points on the structure. Assuming a downward movement of the structure, the column corresponding to each anchor hole could be regarded as the vertical displacement resulting from the applied load. The size and interval were 200 mm and 283 mm, respectively. Consequently, the reference values for the actual displacement could be determined by counting the number of downward shifts of the anchor holes.

A FUJIFILM GFX100S camera with a resolution of 11648 × 8736 pixels and a sensor size of 43.8 × 32.9 mm, paired with a GF100–200 mm F5.6 R LM OIS WR lens, was employed as the vision sensor. The nearest perpendicular distance (L_H) from the optical center to the surface of the reaction wall was measured through a laser rangefinder (SNDWAY® SW-600A [42]), which has a measurement
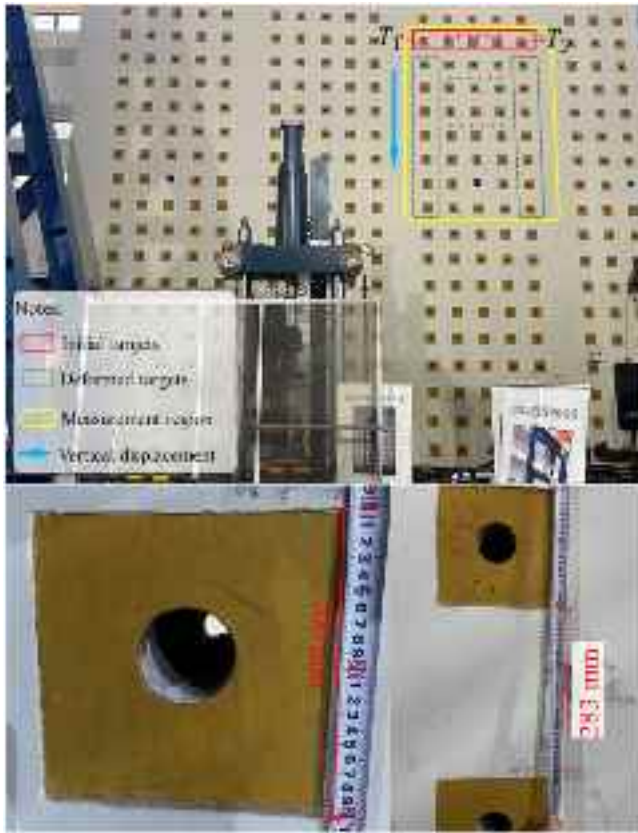
Fig. 7. Illustration of the anchor holes in the reaction wall.

accuracy of ±1 m. Additionally, a high-precision IMU (RION® DDM350B [43]) was utilized to capture the camera orientation data (*i.e.,* pitch and yaw angles). The yaw angle accuracy of this sensor is 0.8°, while the static and dynamic accuracies of the pitch angle are 0.2° and 0.5°, respectively. The camera orientation was controlled via a pan-tilt (ZIFON® YT1000).

The robustness of the proposed method to the camera orientation was verified by conducting the cross-over validation tests. The tests varied the pitch from 0° to 20° and the yaw angle from 0° to 30°, with step sizes of 10°. It should be noted that the maximum pitch and yaw angles were limited to 20° and 30°, respectively, due to laboratory space constraints. Additionally, when adjusting the yaw angle of the camera, the
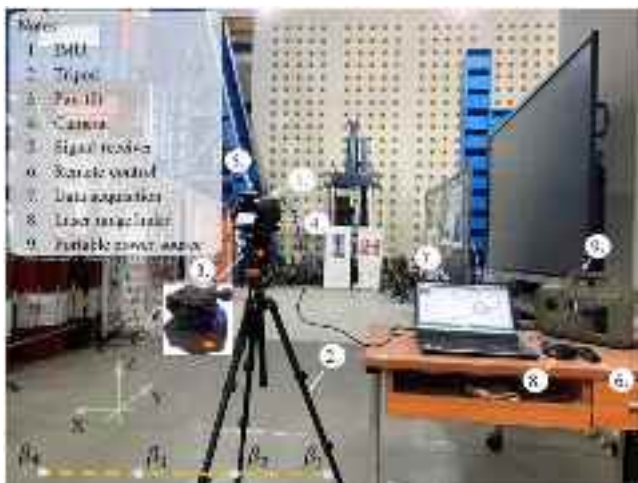

Fig. 8. Overall setup of the vision-based system.

camera location along the x-axis was also adjusted to maintain a consistent FOV and L_H. The overall device setup of the measurement system is exhibited in Fig. 8.

4.1.2. Experiment results

According to the traditional displacement method introduced in [35], the scale factor of the T_1 target was employed as the benchmark of the fixed scale factor at each test. The Percentage Error was utilized as the performance evaluation criteria to qualify the accuracy of the proposed multi-target displacement calculation method, which is expressed as follows:

$$PE = \frac{\left| Y_{ij}^m - Y_{ij}^n \right|}{Y_{ij}^n} \times 100\% \tag{19}$$

where Y_{ij}^m and Y_{ij}^n denote the vertical displacement calculated by the vision-based method and the reference, respectively. The i and j present the i-*th* pitch angle and j-*th* yaw angle, respectively.

The errors of two vision-based displacement measurement methods compared to the reference are shown in Fig. 9. The traditional method exhibited its best performance at C_{11}, with errors below 0.813 %. However, as the target coordinates changed, the error for each target increased, primarily due to the perspective effect, which caused the SFs of each target to vary. Additionally, the performances of traditional method worsened as the pitch and yaw angles increased, amplifying the perspective distortion. The maximum error for the traditional method reached 5.386 %, while the corresponding maximum error for the proposed method was 2.552 %. This demonstrated that the proposed multi-target displacement method outperformed the traditional method and effectively rectified the displacement errors induced by the camera orientation.

4.2. Dynamic tests

4.2.1. Scale of the bridge prototype

To validate the accuracy and robustness of the proposed MTDRI in identifying dynamic displacement responses of structures under external loads, a simply supported beam model with a span of 5600 mm was fabricated in the laboratory, as shown in Fig. 10(a). The bridge was constructed using three Q235 rectangular steel tubes, as illustrated in Fig. 10(b). Each steel tube has a length of 6000 mm, a width of 100 mm, a height of 50 mm, and a wall thickness of 5 mm, as shown in Fig. 11.

4.2.2. Experiment setup

For the indoor experiment on a simply supported bridge scale model, the Canon EOS 5DS R camera (resolution: 1096 ×1080 pixels, sensor size: 36 mm×24 mm, sampling rate: 25 fps) was selected as the image data acquisition device, based on the dimensions of the model and monitoring accuracy requirements. To ensure that the FOV covered the entire bridge model, the camera was equipped with an EF 24–70 mm f/ 2.8 L II USM zoom lens and positioned 5.2 m away. Additionally, LVDTs (DH Model# 5G103) were placed along the longitudinal axis of the model's bottom surface at sections L/5, 2 L/5, 3 L/5, and 4 L/5 to measure the displacement responses of the bridge under load. These four measurement targets were labeled T1, T2, T3, and T4 from left to right. The LVDTs, with a sensitivity of 0.38 mV/mm and a range of 0–50 mm, provided reference data for visual displacement measurements. Data from the contact sensors were collected using an 8-channel dynamic signal acquisition system (DH 5902 N) at a frequency of 100 Hz. To prevent vibration errors from manual shutter control, the camera was remotely triggered using a YouPro® YP-870 Wireless Rx remote control. Furthermore, a high-precision IMU (RION® DDM350B) was used to capture the initial camera orientation data for structural displacement response identification. Prior to the experiment, the camera was calibrated using Zhang's method to correct lens distortion. The
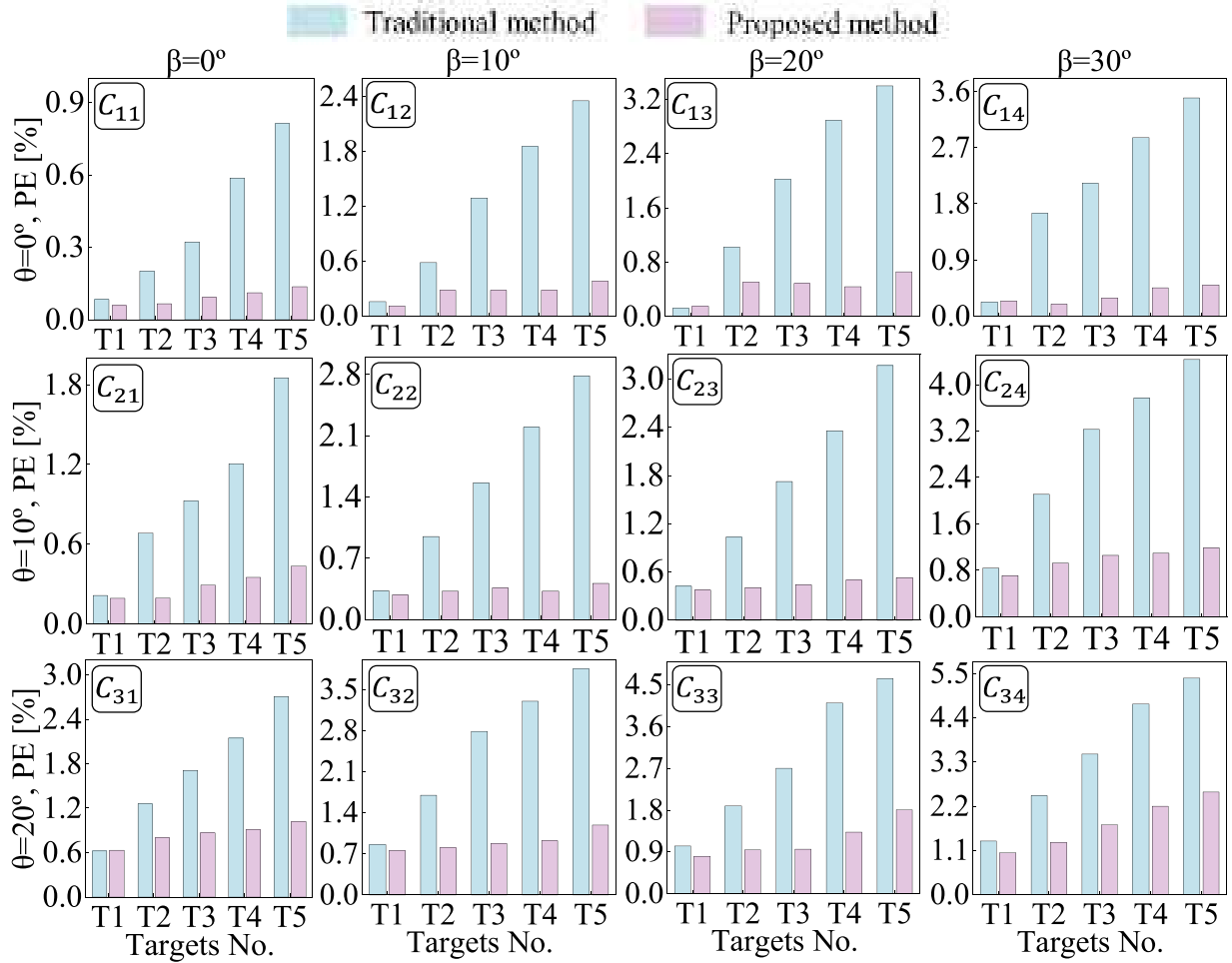
Fig. 9. Errors of two vision-based dipslacement measurement methods.

experimental setup is shown in Fig. 12.

4.2.3. Experiment design

To validate the effectiveness of MTDRI, the yaw angle (β) of the camera was divided into three groups, with a 10° interval between each angle, as shown in Fig. 13(a). The chief ray angle (η), defined as the angle between the measurement target and the optical axis, is related to τ in the distance-solving model and is expressed by $\eta = \pi/2 - \tau$. The pitch angle (θ) was fixed at approximately 5° to maintain consistent FOV

due to laboratory space constraints, as shown in Fig. 13(b). The specific angles are listed in Table 1. During the experiment, structural responses were induced by manually applying a load at the center of the model beam using a rubber hammer. It is important to note that η varies with the location of the measurement target and has a significant impact on measurement precision. The maximum chief ray angle (η_{max}) in this study was calculated as 35°, using the optical sensor size (d) and lens focal length (f) with the formula $\eta_{max} = \tan^{-1}(d/2f)$.

4.2.4. Experiment results

Three displacement response identification methods were employed

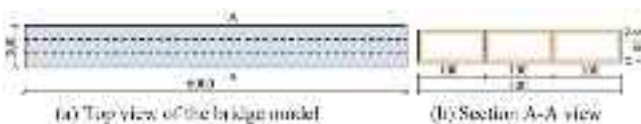**Fig. 10.** Schematic representation of a simply supported beam model bridge.

Fig. 11. Configurations of the experimental model (Unit: mm).

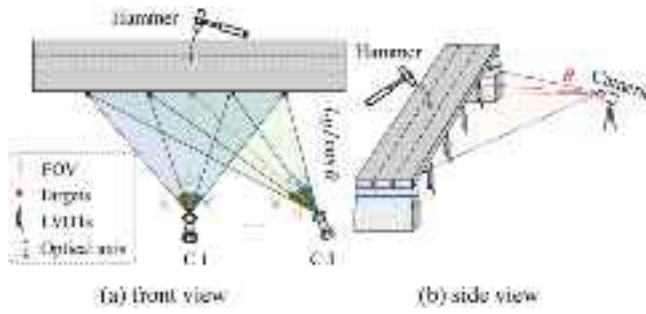**Fig. 12.** Overall experimental design and layout.

Fig. 13. Schematic representation of the experimental design.

Table 1
Design of Test Conditions.

No.	Load type	Camera orientation ()	
		Pitch angle (θ)	Yaw angle (β)
C1		5.02	0
C2	Hammer	5.06	−15.27
C3		5.06	−30.31

Notes: *i*) $\theta > 0$ indicates the camera is tilted upwards, and vice versa, while $\theta < 0$ indicates it is tilted downwards;
ii) $\beta > 0$ presents a clockwise rotation of the camera, while $\beta < 0$ indicates a counterclockwise rotation.

during the tests: *i*) The outcomes captured using contact sensors (LVDTs) served as reference values; *ii*) The results obtained through the traditional method (*i.e.,* the fixed scale factor method) were utilized as comparison values; and *iii*) The displacement extracted by the proposed method were considered as measured values.

The displacement response identification results from case C1 were selected for comparative analysis due to the space constraints, as shown in Fig. 14. The analysis revealed that the displacements of all four targets determined by the MTDRI closely matched the results obtained from the

LVDTs, demonstrating the effectiveness of the proposed method. In contrast, the traditional method showed good agreement with the LVDTs only for targets T2 and T3, as depicted in Fig. 14(b) and (c). However, as the distance of targets T1 and T4 from the midspan increased, the agreement between the traditional method and LVDTs significantly deteriorated, as shown in Fig. 14(a) and d).

Linear regression analysis method was employed to reveal the relationships among the three types of displacements for all four targets in case C1, as depicted in Fig. 15. The results demonstrated stronger linear correlations for MTDRI compared to the traditional method. However, the traditional method exhibited better linear correlation for targets T2 and T3 than for T1 and T4, highlighting the superior accuracy of the proposed method in identifying multi-target displacement response.

4.2.5. Errors analysis

To further quantify the accuracy of MTDRI and assess the impact of camera orientation on monitoring precision, a comparative analysis was conducted on displacement data from all measurement targets under different working conditions using both the traditional method and MTDRI. Root Mean Square Error (RMSE) and Normalized Root Mean Square Error (NRMSE) were employed as evaluation metrics, calculated using the following formulas:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(disp_i - disp_i'\right)^2}$$

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(disp_i - disp_i'\right)^2}}{disp_{i,\max}' - disp_{i,\min}'} \times 100\% \tag{20}$$

where n is the number of measurement data; $disp_i$ and $disp_i'$ represent the *i*-th displacement values obtained by vision sensors and the LVDT, respectively; $disp_{i,\max}'$ and $disp_{i,\min}'$ denote the maximum and minimum values of $disp_i'$.

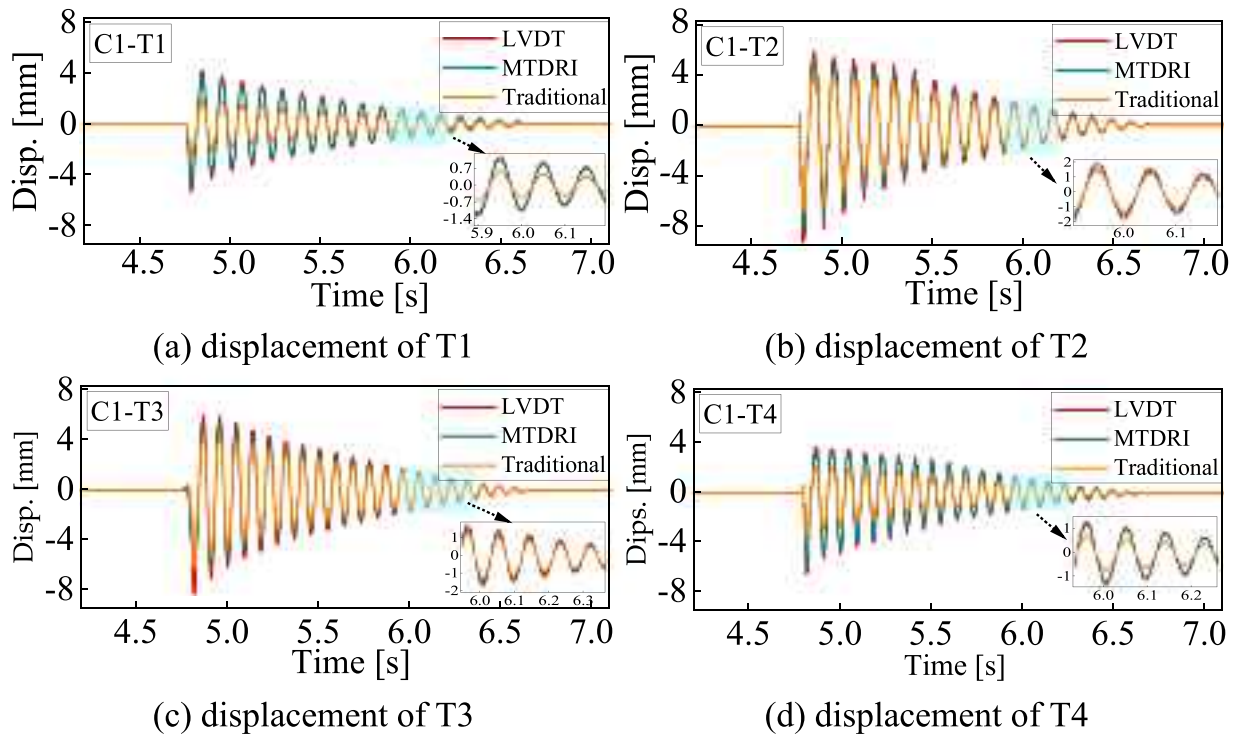The RMSEs and NRMSEs of the displacement measurements under

(a) displacement of T1

(b) displacement of T2

(c) displacement of T3

(d) displacement of T4

Fig. 14. Comparison of the displacements in scenario C1 (θ=5.02, β=0).

(a) relationships in T1

(b) relationships in T2

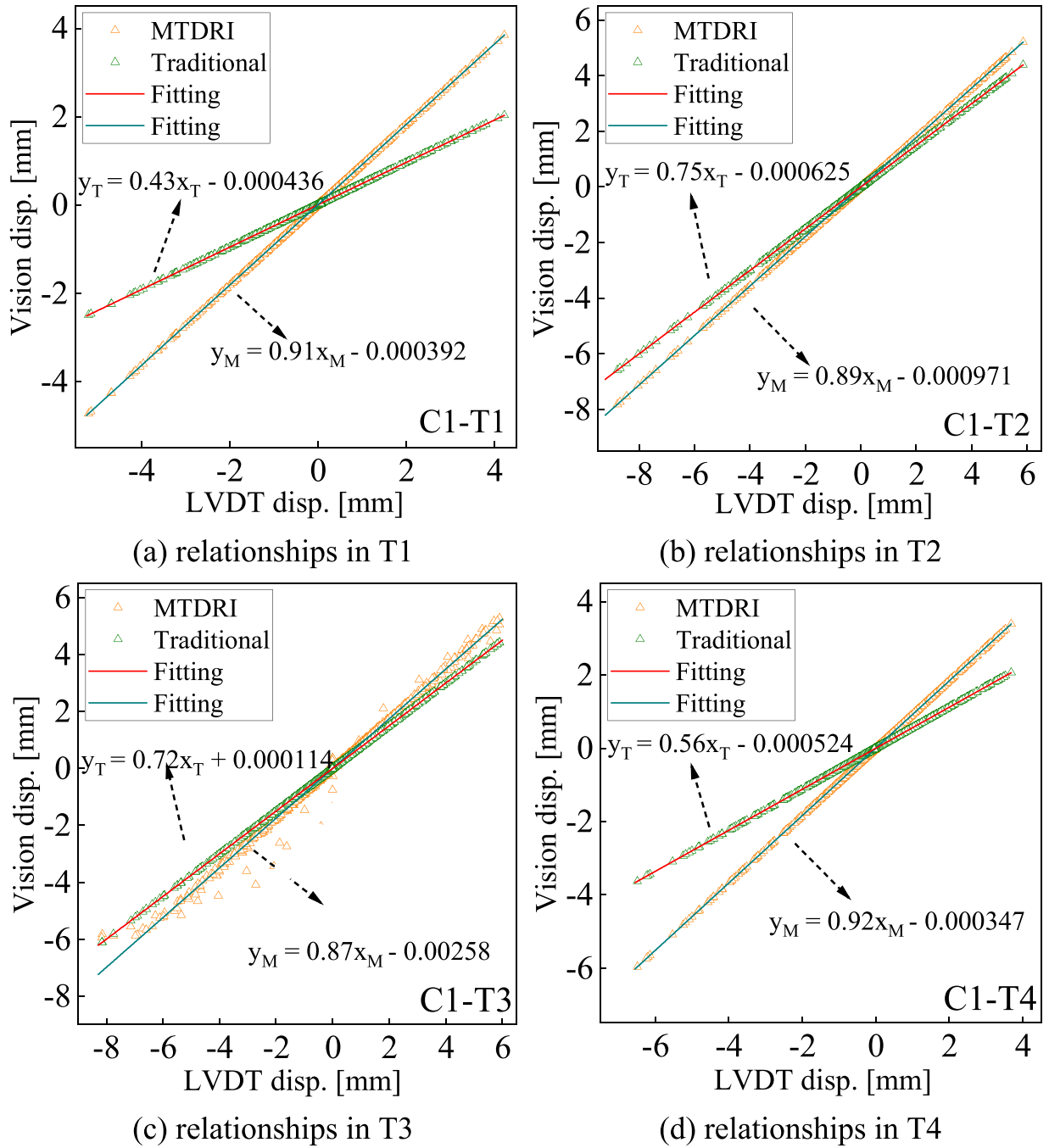(c) relationships in T3

(d) relationships in T4

Fig. 15. Relationships among three types of displacements in C1.

three scenarios are listed in Table 2.

Table 2 presents the displacement errors of the traditional method and MTDRI relative to LVDT across four targets under different conditions. The analysis revealed that, under case C1, the traditional method exhibited smaller displacement errors at T2 and T3 compared to T1 and T4. This discrepancy is primarily attributed to η. Changes in target positions cause variations in the η, with the value of η being larger at T1 and T4 than at T2 and T3. Since the traditional method does not calculate η in real-time, significant errors are observed at T1 and T4, with RMSEs of 0.36 mm and 0.32 mm, and NRMSEs of 3.54 % and 3.39 %, respectively.

Further analysis of conditions C2 and C3 reveals that, under condition C2, the traditional method produces substantially larger errors at

T1, with an RMSE of 0.68 mm and an NRMSE of 4.18 %. This increase is primarily caused by the coupling effects between the β and η, which amplify the displacement error. This coupling effect became more pronounced in condition C3, leading to the RMSE of 1.57 mm and NRMSE of 7.24 % at T1.

In contrast, MTDRI consistently achieves RMSEs and NRMSEs below 0.5 mm and 2.1 %, respectively, across all measurement targets and conditions. These findings demonstrate the ability of MTDRI to effectively correct multi-target displacement measurement errors caused by camera orientation.

Table 2
Comparison of displacement errors.

Case No.	Target No.	η ()	RMSE (mm)		NRMSE (%)	
			Traditional	MTDRI	Traditional	MTDRI
C1	T1	−33.51	0.36	0.06	3.54	0.64
	T2	−12.72	0.28	0.08	1.97	0.55
	T3	12.02	0.25	0.11	1.65	0.72
	T4	32.46	0.32	0.05	3.39	0.58
C2	T1	−31.31	0.68	0.10	4.18	0.63
	T2	−14.73	0.29	0.14	3.28	1.62
	T3	7.55	0.20	0.04	1.62	0.37
	T4	29.74	0.37	0.05	2.73	0.42
C3	T1	−25.75	1.57	0.43	7.24	2.02
	T2	−13.67	0.77	0.19	5.47	1.36
	T3	2.09	0.31	0.10	3.41	1.05
	T4	20.45	0.27	0.03	2.32	0.26

Note: η< 0 signifies the x-coordinate of the target is less than the x-coordinate of the image center, and vice versa.

5. Conclusion

The present work introduced the development of a novel method for accurate multi-target displacement measurement in bridge structures. The MTDRI was proposed to rectify the displacement errors induced by the camera orientation. Comprehensive experiments, including a static and dynamic test, were conducted to investigate its performance. The main contributions of this work are as follow:

- A distance-solving model was proposed, requiring only three external parameters to automatically calculate the distances from the optical center to multiple targets on the structure. This model enhances the practicability of the vision-based multi-target displacement measurement system compared to the traditional methods.
- A multi-target displacement response identification (MTDRI) method was introduced, which considers both external and internal parameters of the camera to rectify the errors caused by camera orientation, thereby improving the accuracy of the multi-target displacement calculations.
- A static cross-over validation test was conducted, with the pitch and yaw angle range 0˚ to 20˚ and 0˚ to 30˚, respectively. The results demonstrated that the MTDRI effectively controlled errors within 2.55 %. Furthermore, the test investigated the impact mechanism of camera orientation on measurement accuracy, revealing that the coupling effect between external and internal parameters significantly degrades displacement calculation precision.
- The efficiency and robustness of the MTDRI were validated through a comprehensive dynamic test. Compared with the traditional method, the RMSE and NRMSE were reduced to below 0.43 mm and 2.02 %, respectively. These results confirm that the proposed MTDRI outperform the traditional approach in multi-target displacement measurement.

The proposed vision-based displacement measurement system is developed to capture high-precision multi-target displacement response data for bridge structures, thereby enhancing the applicability and robustness of vision-based measurement approaches. Despite preliminary results of the proposed MTDRI are promising, further investigations are needed in the following areas: *i*) Addressing camera motion during measurement phase to maintain accuracy; *ii*) Improving the robustness of the feature extraction algorithm under the natural nonhomogeneous background; and *iii*) Ensuring a smooth transition from laboratory experiments to field validation for more comprehensive safety assessments.

CRediT authorship contribution statement

Zhixiang Zhou: Conceptualization, Supervision, Funding acquisition. **Weizhu Zhu:** Writing – original draft, Investigation, Software, Conceptualization, Methodology. **Tengjiao Jiang:** Writing – original draft, Visualization, Supervision. **Zimeng Shen:** Writing – review & editing, Validation, Investigation. **Tong Wu:** Software, 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.

References

[1] J. Lee, K.-C. Lee, S. Jeong, Y.J. Lee, S.-H. Sim, Long-term displacement measurement of full-scale bridges using camera ego-motion compensation, Mech. Syst. Signal Process. 140 (2020) 106651, https://doi.org/10.1016/j.ymssp.2020.106651.
[2] S. Ri, J. Ye, N. Toyama, N. Ogura, Drone-based displacement measurement of infrastructures utilizing phase information, Nat. Commun. 15 (2024), https://doi.org/10.1038/s41467-023-44649-2.
[3] S. Ri, Q. Wang, H. Tsuda, H. Shirasaki, K. Kuribayashi, Displacement measurement of concrete bridges by the sampling Moiré method based on phase analysis of repeated pattern, Strain 56 (2020), https://doi.org/10.1111/str.12351.
[4] H. Zhang, J.X. He, X.G. Jiang, Y.F. Gong, T.Y. Hu, T.J. Jiang, J.T. Zhou, Quantitative characterization of surface defects on bridge cable based on improved YOLACT plus, Case Stud. Constr. Mater. 21 (2024), https://doi.org/10.1016/j.cscm.2024.e03953.
[5] K. Kim, C. Jaemook, J. Chung, G. Koo, I.-H. Bae, Structural displacement estimation through multi-rate fusion of accelerometer and RTK-GPS displacement and velocity measurements, Measurement 130 (2018), https://doi.org/10.1016/j.measurement.2018.07.090.
[6] M.J. Jiménez-Martínez, N. Quesada-Olmo, J. Zancajo-Jimeno, T. Mostaza-Pérez, Bridge deformation analysis using time-differenced carrier-phase technique, Remote Sens. 15 (2023), https://doi.org/10.3390/rs15051458.
[7] S. Ri, H. Tsuda, K. Chang, S. Hsu, F. Lo, T. Lee, Dynamic deformation measurement by the sampling Moiré method from video recording and its application to bridge engineering, Exp. Tech. 44 (2020), https://doi.org/10.1007/s40799-019-00358-4.
[8] M.H. Daneshvar, M. Saffarian, H. Jahangir, H. Sarmadi, Damage identification of structural systems by modal strain energy and an optimization-based iterative regularization method, Eng. Comput. 39 (2023) 2067–2087, https://doi.org/10.1007/s00366-021-01567-5.
[9] S.A. V. Shajihan, T. Hoang, K. Mechitov, B.F. Spencer, Wireless SmartVision system for synchronized displacement monitoring of railroad bridges, Comput. Aided Civ. Infrastruct. Eng. 37 (2022) 1070–1088, https://doi.org/10.1111/mice.12846.
[10] C. Dong, O. Celik, N. Catbas, Marker-free monitoring of the grandstand structures and modal identification using computer vision methods, Struct. Health Monit. 18 (2018) 1491–1509, https://doi.org/10.1177/1475921718806895.
[11] S. Ri, M. Fujigaki, Y. Morimoto, Sampling moir, method for accurate small deformation distribution measurement, Exp. Mech. 50 (2009) 501–508, https://doi.org/10.1007/s11340-009-9239-4.
[12] N. Manzini, A. Orcesi, C. Thom, M.A. Brossault, S. Botton, M. Ortiz, J. Dumoulin, Performance analysis of low-cost GNSS stations for structural health monitoring of civil engineering structures, Struct. Infrastruct. Eng. 18 (2022) 595–611, https://doi.org/10.1080/15732479.2020.1849320.
[13] T.H. Yi, H.N. Li, M. Gu, Recent research and applications of GPS-based monitoring technology for high-rise structures, Struct. Control Health Monit. 20 (2013) 649–670, https://doi.org/10.1002/stc.1501.
[14] H. Nam, B. Kim, Y. Kong, Y.N. Miao, G. Park, Enhanced damage detection by dominant wavenumber filtering in steady-state ultrasonic wavefield imaging,

Mech. Syst. Signal Process. 223 (2025), https://doi.org/10.1016/j.ymssp.2024.111814.

[15] S.Y. Liu, L. Yu, W.T. Niu, J.F. Wang, Z. Zhong, J.K. Huang, M.G. Shan, Fast and accurate visual vibration measurement via derivative-enhanced phase-based optical flow, Mech. Syst. Signal Process. 209 (2024), https://doi.org/10.1016/j.ymssp.2023.111089.

[16] P.L. Wang, B. Chen, Z.R. Gao, Y. Fu, S.Z. Chen, X.W. Zhu, High precision full-field vibration measurement by LDV-induced stroboscopic fringe projection, Opt. Lasers Eng. 183 (2024), https://doi.org/10.1016/j.optlaseng.2024.108481.

[17] Y. Yang, C. Dorn, T. Mancini, Z. Talken, S. Nagarajaiah, G. Kenyon, C. Farrar, D. Mascareñas, Blind identification of full-field vibration modes of output-only structures from uniformly-sampled, possibly temporally-aliased (sub-Nyquist), video measurements, J. Sound Vib. 390 (2016), https://doi.org/10.1016/j.jsv.2016.11.034.

[18] T.J. Jiang, A. Ronnquist, Y. Song, G.T. Froseth, P. Navik, A detailed investigation of uplift and damping of a railway catenary span in traffic using a vision-based line-tracking system, J. Sound Vib. 527 (2022), https://doi.org/10.1016/j.jsv.2022.116875.

[19] P. Tysiac, M. Miskiewicz, D. Bruski, Bridge non-destructive measurements using a laser scanning during acceptance testing: case study, Materials 15 (2022), https://doi.org/10.3390/ma15238533.

[20] Y.W. Chen, Q.H. Huang, T.B. Zhang, M. Zhou, L.M. Jiang, Measurement refinements of ground-based radar interferometry in bridge load test monitoring: comprehensive analysis on a multi-span cable-stayed bridge, Remote Sens. 16 (2024), https://doi.org/10.3390/rs16111882.

[21] C.C. de Oliveira, S. Dhakal, R.B. Malla, Deflections and frequencies of long-span steel railroad truss bridges under passenger train excitation using laser Doppler vibrometer, J. Struct. Integr. Maint. 9 (2024), https://doi.org/10.1080/24705314.2024.2411864.

[22] T. Jiang, G. Frøseth, P. Nåvik, A. Rønnquist, Assessment of pantograph-catenary interaction in a railway overlap section via a novel optical-based method, Mech. Mach. Theory 177 (2022) 105045, https://doi.org/10.1016/j.mechmachtheory.2022.105045.

[23] T.J. Jiang, G.T. Froseth, A. Ronnquist, X. Kong, L. Deng, A visual inspection and diagnosis system for bridge rivets based on a convolutional neural network, Comput. Aided Civ. Infrastruct. Eng. (2024), https://doi.org/10.1111/mice.13274.

[24] T.J. Jiang, G. Froseth, S.R. Wang, O.Y.W. Petersen, A. Ronnquist, A 6-DOF camera motion correction method using IMU sensors for photogrammetry and optical measurements, Mech. Syst. Signal Process. 210 (2024), https://doi.org/10.1016/j.ymssp.2024.111148.

[25] L.F. Ge, K.Y. Koo, M.M. Wang, J. Brownjohn, D.H. Dan, Bridge damage detection using precise vision-based displacement influence lines and weigh-in-motion devices: experimental validation, Eng. Struct. 288 (2023), https://doi.org/10.1016/j.engstruct.2023.116185.

[26] Y.T. Han, G. Wu, D.M. Feng, Vision-based displacement measurement using an unmanned aerial vehicle, Struct. Control Health Monit. 29 (2022), https://doi.org/10.1002/stc.3025.

[27] Y. Xu, J. Zhang, J. Brownjohn, An accurate and distraction-free vision-based structural displacement measurement method integrating Siamese network based

[28] G. Jeon, S. Kim, S. Ahn, H.K. Kim, H. Yoon, Vision-based automatic cable displacement measurement using Cable-ROI Net and Uni-KLT, Struct. Control Health Monit. 29 (2022), https://doi.org/10.1002/stc.2977.

[29] B. Lu, B.C. Bai, X.F. Zhao, Vision-based structural displacement measurement under ambient-light changes via deep learning and digital image processing, Measurement 208 (2023), https://doi.org/10.1016/j.measurement.2023.112480.

[30] Y.C. Wang, W.H. Hu, J. Teng, Y. Xia, Full-field displacement measurement of long-span bridges using one camera and robust self-adaptive complex pyramid, Mech. Syst. Signal Process. 215 (2024), https://doi.org/10.1016/j.ymssp.2024.111451.

[31] Y.T. Han, G. Wu, D.M. Feng, Structural modal identification using a portable laser-and-camera measurement system, Measurement 214 (2023), https://doi.org/10.1016/j.measurement.2023.112768.

[32] B. Hu, W. Chan, Y. Zhang, Y. Yihe, Q. Yu, X. Liu, X. Ding, Vision-based multi-point real-time monitoring of dynamic displacement of large-span cable-stayed bridges, Mech. Syst. Signal Process. 2023 (2023) 110790, https://doi.org/10.1016/j.ymssp.2023.110790.

[33] M. Wang, W.K. Ao, J. Bownjohn, F. Xu, Completely non-contact modal testing of full-scale bridge in challenging conditions using vision sensing systems, Eng. Struct. 272 (2022) 114994, https://doi.org/10.1016/j.engstruct.2022.114994.

[34] Y.D. Tian, J. Zhang, S.S. Yu, Vision-based structural scaling factor and flexibility identification through mobile impact testing, Mech. Syst. Signal Process. 122 (2019) 387–402, https://doi.org/10.1016/j.ymssp.2018.12.029.

[35] D. Feng, M. Feng, E. Ozer, Y. Fukuda, A vision-based sensor for noncontact structural displacement measurement, Sensors (2015), https://doi.org/10.3390/s150x0000x.

[36] T.J. Jiang, G.T. Froseth, A. Ronnquist, E. Fagerholt, A robust line-tracking photogrammetry method for uplift measurements of railway catenary systems in noisy backgrounds, Mech. Syst. Signal Process. 144 (2020), https://doi.org/10.1016/j.ymssp.2020.106888.

[37] T. Wu, L. Tang, X.Y. Zhang, Y.J. Liu, X.Y. Li, Z.X. Zhou, An improved structural displacement monitoring approach by acceleration-aided tilt camera measurement, Struct. Control Health Monit. 2023 (2023), https://doi.org/10.1155/2023/6247516.

[38] L. Tian, J.H. Zhao, B. Pan, Z.Y. Wang, Full-field bridge deflection monitoring with off-axis digital image correlation, Sensors 21 (2021), https://doi.org/10.3390/s21155058.

[39] S. Yu, Z. Xu, Z. Su, J. Zhang, Two flexible vision-based methods for remote deflection monitoring of a long-span bridge, Measurement 181 (2021) 109658, https://doi.org/10.1016/j.measurement.2021.109658.

[40] L. Tian, T. Ding, B. Pan, Generalized scale factor calibration method for an off-axis digital image correlation-based video deflectometer, Sensors 22 (2022), https://doi.org/10.3390/s222410010.

[41] S. Collier, T. Dare, Accuracy of phase-based optical flow for vibration extraction, J. Sound Vib. 535 (2022) 117112, https://doi.org/10.1016/j.jsv.2022.117112.

[42] Laserrangefinder, SNDWAY® SW-600A. ⟨https://www.sndway.com/productinfo/1278030.html⟩.

[43] IMU, RION® DDM350B. ⟨https://www.rionsystem.com/list_66/304.html⟩.