



Long-term displacement measurement of full-scale bridges using camera ego-motion compensation



Junhwa Lee^a, Kyoung-Chan Lee^b, Seunghoo Jeong^a, Young-Joo Lee^a, Sung-Han Sim^{c,*}

^a School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, Republic of Korea

^b High-Speed Railroad Systems Research Center, Korea Railroad Research Institute, Uiwang, Republic of Korea

^c School of Civil, Architectural Engineering and Landscape Architecture, Sungkyunkwan University, Suwon, Republic of Korea

ARTICLE INFO

Article history:

Received 23 July 2019

Received in revised form 6 January 2020

Accepted 13 January 2020

Keywords:

Computer vision

Long-term displacement

Motion-independent measurement

Camera ego-motion compensation

ABSTRACT

Civil infrastructures experience long-term deflection as a result of persistent and transitory loadings, including self-weight, pre-stress, traffic, temperature variation, and stress redistributions due to damage. Thus, long-term displacement is an important safety indicator that can be widely employed in structural health monitoring. However, its measurement is challenging because of the errors induced by the ego-motion of sensors. Researchers have attempted to compensate for the motion-induced error in computer vision-based displacement measurement methods for full-scale civil structures by using fixed objects in the background. Because remotely located objects cannot fully provide six degrees-of-freedom (6-DOF) camera motions, further developments are necessary for complete error compensation. This paper proposes a long-term displacement measurement strategy that uses computer vision-based ego-motion compensation. The proposed system consists of main and sub-cameras attached to each other. While the main camera employs the conventional computer vision-based method for displacement measurement, the sub-camera measures the ego-motions of the dual-camera system from which the motion-induced errors are estimated and compensated for. The proposed long-term displacement measurement algorithm was numerically validated, and the sub-camera was found to provide a noticeable error compensation. A laboratory-scale test showed that the motion-induced error was reduced from 44.1 mm to 1.1 mm. A field application conducted upon a newly constructed railway bridge provided continuous long-term displacement measurement data, which were consistent with LiDAR-based displacement measurements and numerical predictions.

© 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Civil engineering structures exhibit long-term deflection throughout their lifetimes. Initiated in the construction phase, the creep and shrinkage in the building materials develop over time due to chemical processes combined with persistent and transitory loadings arising from self-weight, pre-stress, traffic, and temperature fluctuations. In addition, undesirable influences such as natural disasters, collisions, and irregular stress distribution can cause damage that weakens the stiffness of the structure and thus produces deflections. Thus, the inevitable long-term deflection reflects the structural health that can be used for general maintenance purposes [1–6] and early indicators of structural failures [7,8].

* Corresponding author.

E-mail address: ssim@skku.edu (S.-H. Sim).

Most existing approaches for measuring the displacement of bridge structures have a limited applicability to long-term measurement. Contact-type displacement transducers such as the linear variable differential transformer (LVDT) and draw-wire displacement sensor require a supporting structure [9,10], which is often difficult to install and may deform during the measurement. Non-contact sensors such as a global navigation satellite system (GNSS) [11] and a total station [12] have an insufficient measurement resolution for small- to medium-sized bridges. A laser Doppler vibrometer (LDV) [13] can measure the displacement by analyzing the phase shift; however, the discontinuity of the phase at multiples of 2π causes a drift error [14], which is a critical issue for long-term monitoring. Indirect methods estimate the structural displacement from other physical quantities such as the acceleration [15–18], strain [19,20], or a combination of multi-metric data [21–24]. However, such indirect approaches were developed for short-term dynamic displacement sensing, and their validity for long-term measurement has not yet been discussed. Thus, displacement measurement methods in the literature mainly focus on short-term measurement of up to several hours; long-term measurement that can last for several months has not been fully explored in terms of sensor installation, sensing accuracy, and drift error.

Several studies on long-term displacement measurement have been reported; the majority simply employed conventional sensors for long periods of time. Cao et al. [25] measured the deflections of two pre-stressed concrete beams under sustained loads for 470 days by using dial gauge-type displacement transducers to compare the calculated and measured long-term deflections. According to Burdet [26], hydrostatic displacement meters were fabricated for long-term monitoring and implemented on Paudèze and Lutrive viaducts over 35 years. Beltempo et al. [27] reported on a total station that was permanently installed at a stationary location and employed in the long-term deflection monitoring of the Colle Isarco viaduct over a 30 year period. Lee et al. [28] developed a LiDAR-based long-term measurement method that can sparsely record the displacement of a bridge without permanent installation, and this was validated through application to a full-scale railway bridge for about 5 months and a comparison with a finite element model. These long-term displacement measurement approaches show promising practical applications for full-scale civil structures; however, handling of the sensor movement and rigorous validation (e.g., cross-validation with another sensor) have received limited consideration.

Recent advances in computer vision technology have allowed cameras to be employed as displacement sensors [29]. Under the common assumption that the camera is fixed at a stationary point, most computer vision-based methods compute the structural displacement from a video of structural motions [30–36]. This assumption prevents the long-term application of computer vision-based methods, because the complete fixation of a camera for a long period of time is impractical. Researchers have attempted to compensate for camera motion-induced errors by using fixed objects in the background, such as buildings and mountains [37–39]. However, the background rarely remains at the exact same position for a long period of time, which limits the applicability of existing methods to full-scale structures. In addition, remotely installed cameras cannot fully compensate for the error induced by subtle out-of-plane camera motions [40]. Thus, existing computer vision-based displacement measurement methods for full-scale civil structures are vulnerable to the camera's ego-motions, which are difficult to compensate for with a single camera.

This paper proposes a computer vision-based strategy for long-term bridge displacement measurement, which uses a dual-camera system for ego-motion compensation. The dual-camera system consists of main and sub-cameras rigidly attached to each other. The main camera measures the structural displacement via a conventional method [34], which can be disturbed by camera motions. Meanwhile, the sub-camera measures the camera motion in six degrees of freedom (6-DOF), which is used to estimate motion-induced errors. The estimated errors are then subtracted from the displacement measured by the main camera to yield the true displacement. The rest of this paper describes the proposed long-term displacement measurement algorithm in detail. A numerical simulation as well as laboratory and field experiments were performed to validate the applicability of the proposed system to bridge structures.

2. Limitations of conventional computer vision-based displacement measurement

Conventional computer vision-based methods for structural displacement measurement comprise a target, camera, image processing, and coordinate transform algorithm [34]. A target is attached to a point on the structural component for which the displacement needs to be measured. The target is generally designed to have noticeable shapes, called features, for robust tracking of the target motions. A camera is remotely installed to record images of the target movement. The positions are extracted in the form of image coordinates using an image processing technique such as feature detection, digital image correlation, and optical flow. The structural displacement represented in the image space is transformed to the physical space via affine transform [30], calibration-based extrinsic parameter extraction [31], or homography transform [34]. Computer vision-based methods are widely adopted for monitoring the displacements of civil engineering structures because of their advantages of convenient installation, low system cost, and adequate measurement resolution.

When a camera is installed for the long-term monitoring of a full-scale structure, unexpected camera motions are unavoidable. Even if a camera is firmly fixed at a stationary point, its self-weight and temperature fluctuations induce an inevitable and gradual movement of the entire system. In addition, the camera is prone to permanent relocation during long-term operations, often caused by impacts such as earthquakes and unexpected collisions. Thus, the complete fixation of the camera is difficult for long-term monitoring applications.

Camera motions can induce considerable false measurements, particularly when computer vision is utilized for monitoring the displacement of full-scale civil structures. Fig. 1 shows a camera system in which the target is fixed while the camera

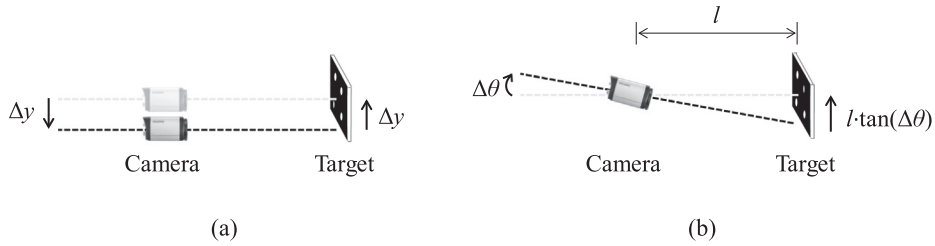


Fig. 1. Examples of (a) translational and (b) rotational camera motions, and the resulting false displacements.

is displaced in the translational and rotational directions by $-\Delta y$ and $\Delta\theta$, respectively. The resulting false displacements along the vertical direction are Δy and $l \cdot \tan(\Delta\theta)$ for each camera motion. The false displacement from the translational motion becomes significant when the true target displacement is comparable to Δy . A considerable false displacement can be induced by the rotational motion if the camera is installed at a distance from the target (see Fig. 1(b)). In practice, the minute movement of the civil engineering structures and remotely installed cameras for full-scale applications mean that camera motions can produce significant false displacements that should be properly considered for accurate measurements. For example, 10 mm of false displacement can be induced by a $\Delta\theta$ of approximately 1 mrad when the camera is positioned 10 m away from the target, which can be larger than the actual displacement of small- to medium-sized bridges.

The false displacement induced by 6-DOF camera motions is difficult to remove through conventional methods, which use a stationary background to compensate for the error [37,38] or estimate the camera pose [39]. Kim et al. [37] and Chen et al. [38] utilized the background motion to compensate for the motion-induced error. As an example, Fig. 2 shows a measurement case where the camera is translated and rotated by $-\Delta y$ and $\Delta\theta$, respectively, which induces the false displacement of $\Delta y + l_1 \cdot \tan(\Delta\theta)$. Here, the background appears to shift by $\Delta y + l_2 \cdot \tan(\Delta\theta)$ when the camera moves, which is subtracted from the original measurement. However, the false displacement cannot be fully removed by subtracting the background motion unless l_1 is close to l_2 . Alternatively, Yoon et al. [39] estimated 6-DOF camera motions using the stationary background to remove the false displacement. The precise measurement of the 6-DOF camera motions was challenging because a remotely installed camera becomes an affine camera [40], for which subtle out-of-plane motions are neglected. In Fig. 2, if the camera motion is estimated by the background motion of $\Delta y + l_2 \cdot \tan(\Delta\theta)$ with a large l_2 and small $\Delta\theta$, it will be estimated as a pure translation of $-\Delta y - l_2 \cdot \tan(\Delta\theta)$ because the out-of-plane rotation $\Delta\theta$ is neglected. Conventional methods can compensate for false displacement in the limited situation that $l_2 - l_1$ is small compared to l_1 .

The conventional computer vision-based displacement measurement methods face issues when being applied to the long-term measurement of full-scale civil structures. Because the error induced by the camera motion is difficult to remove, a better approach for compensating for the false displacement is necessary for the long-term operation of computer vision-based displacement measurement methods.

3. Long-term displacement measurement with the dual-camera system

3.1. Coordinate systems of the dual-camera configuration

The hardware configuration of the dual-camera system consists of two targets and two cameras that are tightly attached to each other, as shown in Fig. 3. When the main target is deployed at a measurement point on a structure, the main camera records the movements of the target. The recorded video is generally used to calculate the structural displacement; however, the measured displacement can include an error from the camera motions. The sub-camera measures its 6-DOF motion rel-

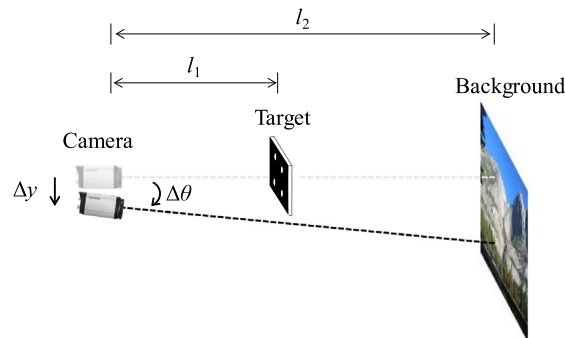


Fig. 2. Common situation where a false displacement is compensated for by the background.

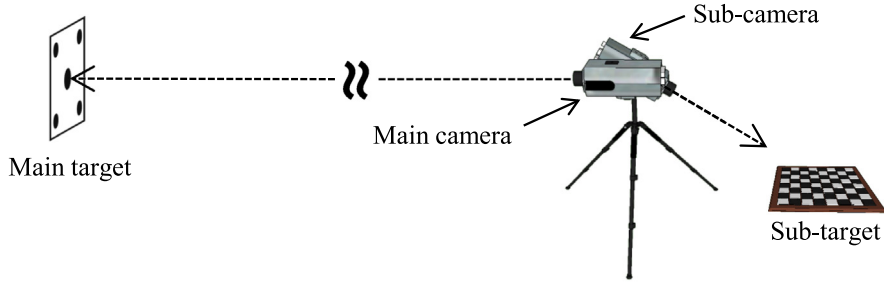


Fig. 3. Hardware configuration of the dual-camera system.

ative to the sub-target, which is placed at a stationary reference point (e.g., the bridge pier). Because the two cameras are fixed to each other, the false displacement can be compensated for by using the identified motions of the sub-camera.

The dual-camera system has three different types of coordinate systems: world, camera, and image, as shown in Fig. 4. The world coordinate system is introduced to express the position of either the main or sub-target. Once an arbitrary point is selected on the target as the origin of the world coordinate system, two basis vectors are defined as two orthogonal unit vectors within the target plane, which can be specified by user-defined patterns on the target. The cross-product of the two basis vectors yields the third basis vector, which is used to complete the three-dimensional world coordinate system. The world coordinate system remains at the initial position even when the target plane moves. The camera coordinate system is provided by the camera employed. The optical axis of the camera is assigned as the z -axis, and unit vectors orthogonal to the z -axis are defined as the x - and y -axes. The camera coordinate system varies as the camera moves in space, in contrast to the world coordinate system. The image coordinate system is two-dimensional and is used for captured images. The image coordinates are in units of pixels, with an origin at the upper left of the image. The three coordinate systems can fully represent the physical and image spaces associated with the camera.

The dual-camera system comprises six coordinate systems in total because the main and sub-cameras each have their own world, image, and camera coordinates, as shown in Fig. 5. Hereafter, the prefixes W , C , and I represent the world, camera, and image coordinate systems, respectively, while the suffixes M and S denote the main and sub-camera systems, respectively. WM and WS are the world coordinate systems defined by the initial positions of the main and sub-targets, respectively. $CM(j)$ and $CS(j)$ are the camera coordinate systems defined by the positions of the main and sub-cameras in the j -th frame of the recorded images, respectively. $IM(j)$ and $IS(j)$ are the image coordinate systems of the main and sub-cameras in the j -th frame, respectively. CM , CS , IM , and IS are time-varying coordinate systems due to the camera motions, while WM and WS are time-invariant.

The notation $\mathbf{B}_{\Omega(j) \rightarrow \Psi(k)}$ is introduced to express coordinate transforms from $\Omega(j)$ to $\Psi(k)$, which are two coordinate systems selected from WM , WS , CM , CS , IM , and IS . For example, $\mathbf{B}_{CS(1) \rightarrow IM(2)}$ is the coordinate transform from $CS(1)$ to $IM(2)$. To

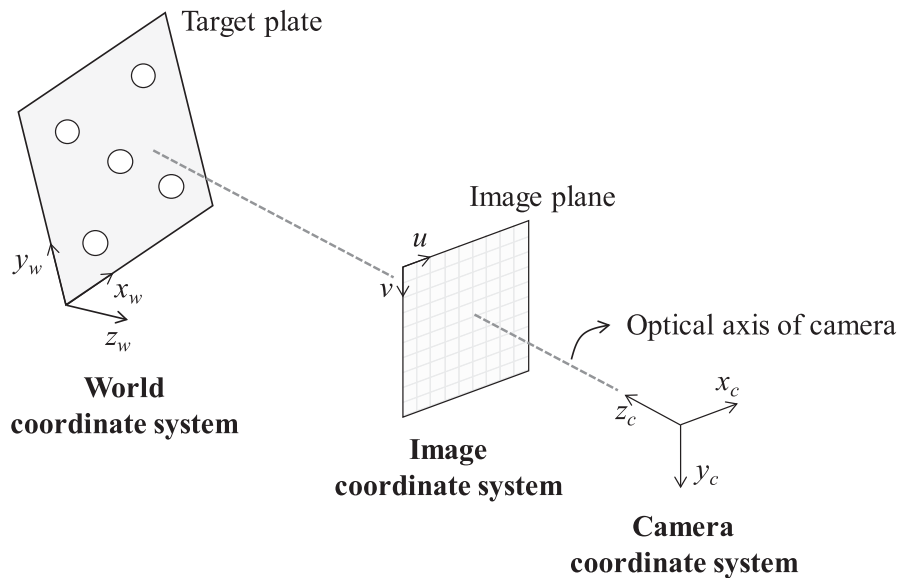


Fig. 4. Coordinate systems of the dual-camera system.

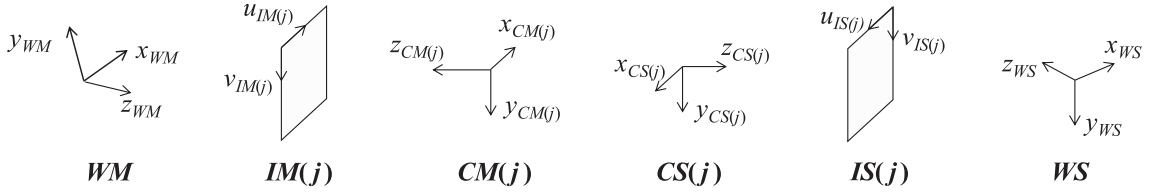


Fig. 5. Coordinate systems of the dual-camera system.

include translation, rotation, shearing, scaling, and perspective transforms in $\mathbf{B}_{\Omega(j) \rightarrow \Psi(k)}$, a homogeneous coordinate representation is adopted, in which an additional 1 is attached to the end of the coordinate vector. Hence, the coordinate transform $\mathbf{B}_{\Omega(j) \rightarrow \Psi(k)}$ is either a 4×4 , 3×4 , or 3×3 matrix depending on the source and destination coordinates to be transformed. The coordinate systems and transforms defined hereafter are used to explain the proposed displacement calculation algorithm given in Section 3.2.

3.2. Displacement calculation algorithm

3.2.1. Formulation for the true displacement

The proposed displacement calculation algorithm determines the true structural displacement by compensating for the false displacement caused by camera motions. Consider a main target with N_M feature points, and let $\{x_i(k) \ y_i(k) \ z_i(k) \ 1\}^T_{WM}$ be the position of the i -th feature point at the k -th frame in WM. The image coordinate of the feature point at the k -th frame in $IM(j)$, which is denoted by $\{u_i(k) \ v_i(k) \ 1\}^T_{IM(j)}$, can be expressed by the coordinate transform of $\{x_i(k) \ y_i(k) \ z_i(k) \ 1\}^T_{WM}$ as follows:

$$\begin{Bmatrix} u_i(k) \\ v_i(k) \\ 1 \end{Bmatrix}_{IM(j)} = \mathbf{B}_{WM \rightarrow IM(j)} \begin{Bmatrix} x_i(k) \\ y_i(k) \\ 0 \\ 1 \end{Bmatrix}_{WM} \quad (i = 1, 2, \dots, N_M) \quad (1)$$

where $\mathbf{B}_{WM \rightarrow IM(j)}$ is a projective transform from WM to $IM(j)$. Note that $z_i(k) = 0$ because the origin of WM is on the main target, which is assumed to undergo in-plane motions only. Even though the different frame indices k and j appear in $\{u_i(k) \ v_i(k) \ 1\}^T_{IM(j)}$, the image coordinate can only be obtained when $k = j$, which is the actual measurement of the main camera. On the other hand, image coordinates with different k and j can be used to represent the false displacement induced by camera motions. The false displacement is viewed as the initial coordinates of a feature point in the j -th image coordinate system $IM(j)$, which reflects the camera motions. Thus, the false displacement $\{u_i(1) \ v_i(1) \ 1\}^T_{IM(j)}$ can be formulated using Eq. (1) with $k = 1$ as:

$$\begin{Bmatrix} u_i(1) \\ v_i(1) \\ 1 \end{Bmatrix}_{IM(j)} = \mathbf{B}_{WM \rightarrow IM(j)} \begin{Bmatrix} x_i(1) \\ y_i(1) \\ 0 \\ 1 \end{Bmatrix}_{WM} \quad (2)$$

The true displacement can be obtained by compensating for the false displacement in Eq. (2) with the current measurement at the j -th frame from the main camera (i.e., $\{u_i(j) \ v_i(j) \ 1\}^T_{IM(j)}$). Subtracting Eq. (2) from Eq. (1) with $k = j$ leads to

$$\begin{Bmatrix} u_i(j) - u_i(1) \\ v_i(j) - v_i(1) \\ 0 \end{Bmatrix}_{IM(j)} = \mathbf{B}_{WM \rightarrow IM(j)} \begin{Bmatrix} x_i(j) - x_i(1) \\ y_i(j) - y_i(1) \\ 0 \\ 0 \end{Bmatrix}_{WM} \quad (3)$$

where $\{u_i(j) - u_i(1) \ v_i(j) - v_i(1) \ 0\}^T_{IM(j)}$ and $\{x_i(j) - x_i(1) \ y_i(j) - y_i(1) \ 0 \ 0\}^T_{WM}$ are the true displacements expressed in $IM(j)$ and WM, respectively. Note that the desired displacement is the change in position relative to the initial frame. Eq. (3) can be simplified by considering the first two rows only as:

$$\begin{Bmatrix} u_i(j) - u_i(1) \\ v_i(j) - v_i(1) \end{Bmatrix}_{IM(j)} = \tilde{\mathbf{B}}_{WM \rightarrow IM(j)} \begin{Bmatrix} x_i(j) - x_i(1) \\ y_i(j) - y_i(1) \end{Bmatrix}_{WM} \quad (4)$$

where $\tilde{\mathbf{B}}_{WM \rightarrow IM(j)}$ is the 2×2 submatrix of $\mathbf{B}_{WM \rightarrow IM(j)}$ obtained by deleting the third and fourth columns and the third row. Knowing that $\{u_i(j) \ v_i(j)\}^T_{IM(j)}$ can be directly obtained from the main camera, $\tilde{\mathbf{B}}_{WM \rightarrow IM(j)}$ and $\{u_i(1) \ v_i(1)\}^T_{IM(j)}$ need to be deter-

mined to calculate the true displacement expressed in WM (i.e., $\{x_i(j) - x_i(1) \ y_i(j) - y_i(1)\}^T_{WM}$). Detailed formulations for determining $\tilde{\mathbf{B}}_{WM \rightarrow IM(j)}$ and $\{u_i(1) \ v_i(1)\}^T_{IM(j)}$ are provided in Sections 3.2.2 and 3.2.3, respectively.

3.2.2. Formulation of the frame-invariant transform matrix $\tilde{\mathbf{B}}_{WM \rightarrow IM(j)}$ with the main camera

The main camera can be assumed to be an affine camera; this generally holds when the image coordinates with and without the affine camera assumption are effectively identical [40]. Let x be the image coordinate and x_{affine} be the image coordinate with the affine camera assumption. When the difference between x and x_{affine} is less than the minimum achievable pixel resolution (e.g., 1 pixel without a subpixel algorithm), an affine camera can be assumed. The difference is formulated based on the camera geometry [40] as:

$$\|x - x_{affine}\|_2 = \left\| \frac{\Delta z}{d_0} (x - x_0) \right\|_2 \quad (5)$$

where $\|\cdot\|_2$ denotes the two-norm, Δz is the depth motion (i.e., translational motion parallel to the optical axis), d_0 is the initial distance between the target and focal point, and x_0 is the camera's principal point which is generally near the center of the image. $\|x - x_{affine}\|_2$ can be less than 1 pixel when both $\Delta z/d_0$ and $x - x_0$ are small. Hence, for the affine camera assumption to hold, the depth motion Δz should be small compared to the initial target distance d_0 , and the image coordinate of the feature should be near the center of the image. Knowing that the displacement of full-scale bridges along the longitudinal direction is significantly less than their span length, it is reasonable to assume that $\Delta z/d_0$ is nearly zero [35]. The main camera has a negligible $\|x - x_{affine}\|_2$ and thus can be considered as an affine camera.

$\tilde{\mathbf{B}}_{WM \rightarrow IM(j)}$ can be obtained with the main camera under the affine camera assumption. If the main camera is affine, the third row of the transform matrix $\mathbf{B}_{WM \rightarrow IM(j)}$ in Eq. (3) is constrained to $[0 \ 0 \ 0 \ 1]$. Furthermore, if a small camera rotation is assumed, the first three columns pertaining to rotation [40] are independent of the image frame. Thus, $\mathbf{B}_{WM \rightarrow IM(j)}$ can be expressed as

$$\mathbf{B}_{WM \rightarrow IM(j)} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14}(j) \\ b_{21} & b_{22} & b_{23} & b_{24}(j) \\ 0 & 0 & 0 & 1 \end{bmatrix}_{WM \rightarrow IM(j)}, \quad (6)$$

where b_{qr} is an entry in the first and second rows of $\mathbf{B}_{WM \rightarrow IM(j)}$ ($q = 1, 2; r = 1, \dots, 4$). Because the third column in $\mathbf{B}_{WM \rightarrow IM(j)}$ has no contribution due to the 0 in $\{x_i(k) \ y_i(k) \ 0 \ 1\}^T_{WM}$, Eq. (1) can be rewritten in the simplified form as:

$$\begin{Bmatrix} u_i(k) \\ v_i(k) \\ 1 \end{Bmatrix}_{IM(j)} = \begin{bmatrix} b_{11} & b_{12} & b_{14}(j) \\ b_{21} & b_{22} & b_{24}(j) \\ 0 & 0 & 1 \end{bmatrix}_{WM \rightarrow IM(j)} \begin{Bmatrix} x_i(k) \\ y_i(k) \\ 1 \end{Bmatrix}_{WM} \quad (7)$$

The positions of all feature points in the main target can be used to arrange Eq. (7) as

$$\begin{Bmatrix} b_{11} \\ b_{12} \\ b_{14}(j) \\ b_{21} \\ b_{22} \\ b_{24}(j) \end{Bmatrix} = \begin{bmatrix} x_1(k) & y_1(k) & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1(k) & y_1(k) & 1 \\ & \vdots & & \vdots & & \\ x_{N_M}(k) & y_{N_M}(k) & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_{N_M}(k) & y_{N_M}(k) & 1 \end{bmatrix}_{WM}^+ \begin{Bmatrix} u_1(k) \\ v_1(k) \\ \vdots \\ u_{N_M}(k) \\ v_{N_M}(k) \end{Bmatrix}_{IM(j)} \quad (8)$$

where $+$ denotes the Penrose–Moore pseudo inverse [41]. Because the initial positions of the features in WM are prior information, the pseudo inversed matrix in Eq. (8) can be built with the known initial coordinates $\{x_i(1) \ y_i(1) \ 1\}^T_{WM}$. In addition, $\{u_i(1) \ v_i(1) \ 1\}^T_{IM(1)}$ can be obtained as a measurement from the main camera in the initial frame, as discussed in Section 3.2.1. The known coordinates of the feature points at the initial frame (i.e., $\{x_i(1) \ y_i(1) \ 1\}^T_{WM}$ and $\{u_i(1) \ v_i(1) \ 1\}^T_{IM(1)}$) can be used to determine the transform matrix in Eq. (7) with $j = 1$, which also includes $\tilde{\mathbf{B}}_{WM \rightarrow IM(j)}$ in Eq. (4). The main target needs to have at least three features to solve for the six unknown parameters (i.e., $b_{11}, b_{12}, b_{21}, b_{22}, b_{14}(1)$, and $b_{24}(1)$) in the transform matrix.

3.2.3. Formulation of the camera motion-induced error with the sub-camera

The sub-camera measures its 6-DOF camera motions with respect to the initial position, expressed as $\mathbf{B}_{CS(1) \rightarrow CS(j)}$, by using the planar target-based calibration method proposed by Zhang [42]. Consider that the sub-camera is placed close to the fixed sub-target containing N_S feature points. In contrast to the main camera, which is assumed to be an affine camera, the sub-camera conforms to the perspective camera model, enabling sensitive measurement of the 6-DOF motion. The position of the i -th feature point in $IS(j)$ (i.e., $\{\bar{u}_i \ \bar{v}_i \ 1\}^T_{IS(j)}$) and in WS (i.e., $\{\bar{x}_i \ \bar{y}_i \ 0 \ 1\}^T_{WS}$) can be expressed with the notation defined in Section 3.1 as follows:

$$\begin{Bmatrix} \bar{u}_i \\ \bar{v}_i \\ \bar{w}_i \end{Bmatrix}_{IS(j)} = \mathbf{B}_{CS(j) \rightarrow IS(j)} \mathbf{B}_{WS \rightarrow CS(j)} \begin{Bmatrix} \bar{x}_i \\ \bar{y}_i \\ 0 \\ 1 \end{Bmatrix}_{WS} \quad (i = 1, 2, \dots, N_S) \quad (9)$$

where $\mathbf{B}_{CS(j) \rightarrow IS(j)}$ is a projective transform from $CS(j)$ to $IS(j)$ and $\mathbf{B}_{WS \rightarrow CS(j)}$ is a similarity transform from WS to $CS(j)$. $\mathbf{B}_{CS(j) \rightarrow IS(j)}$ consists of intrinsic camera parameters as:

$$\mathbf{B}_{CS(j) \rightarrow IS(j)} = \begin{bmatrix} \alpha & \gamma & \mu & 0 \\ 0 & \beta & \nu & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}_{CS(j) \rightarrow IS(j)}. \quad (10)$$

$\mathbf{B}_{WS \rightarrow CS(j)}$ has the extrinsic camera parameters of the rotation matrix $\mathbf{R}_{3 \times 3}(j)$ and translation vector $\mathbf{T}_{3 \times 1}(j)$ as:

$$\mathbf{B}_{WS \rightarrow CS(j)} = \begin{bmatrix} \mathbf{R}_{3 \times 3}(j) & \mathbf{T}_{3 \times 1}(j) \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}_{WS \rightarrow CS(j)}. \quad (11)$$

Eqs. (10) and (11) can be computed from the image of the sub-target via the calibration method. Once $\mathbf{B}_{WS \rightarrow CS(j)}$ is obtained for every frame, the 6-DOF sub-camera motion, with respect to its initial position, is computed as follows:

$$\mathbf{B}_{CS(1) \rightarrow CS(j)} = \mathbf{B}_{WS \rightarrow CS(j)} \mathbf{B}_{WS \rightarrow CS(1)}^{-1} \quad (12)$$

Thus, the 6-DOF motion of the sub-camera with respect to initial position, expressed as $\mathbf{B}_{CS(1) \rightarrow CS(j)}$, can be determined in every measurement frame.

The 6-DOF sub-camera motion $\mathbf{B}_{CS(1) \rightarrow CS(j)}$ is employed to formulate the false displacement. First, Eq. (1) is expanded with $\mathbf{B}_{CS(1) \rightarrow CS(j)}$ as follows:

$$\begin{Bmatrix} u_i(k) \\ v_i(k) \\ 1 \end{Bmatrix}_{IM(j)} = \mathbf{B}_{CS(j) \rightarrow IM(j)} \mathbf{B}_{CS(1) \rightarrow CS(j)} \mathbf{B}_{WM \rightarrow CS(1)} \begin{Bmatrix} x_i(k) \\ y_i(k) \\ 0 \\ 1 \end{Bmatrix}_{WM} = \mathbf{B}_{CS(j) \rightarrow IM(j)} \mathbf{B}_{CS(1) \rightarrow CS(j)} \begin{Bmatrix} x_i(k) \\ y_i(k) \\ z_i(k) \\ 1 \end{Bmatrix}_{CS(1)}. \quad (13)$$

Note that the z-coordinate is generally nonzero in $CS(1)$. $\mathbf{B}_{CS(1) \rightarrow CS(j)}$ is a rigid transformation that can be expressed as the rotation $\bar{\mathbf{R}}_{3 \times 3}(j)$ and the translation $\bar{\mathbf{T}}_{3 \times 1}(j)$. Thus, Eq. (13) can be further expanded as follows:

$$\begin{Bmatrix} u_i(k) \\ v_i(k) \\ 1 \end{Bmatrix}_{IM(j)} = \begin{bmatrix} \bar{b}_{11} & \bar{b}_{12} & \bar{b}_{13} & \bar{b}_{14} \\ \bar{b}_{21} & \bar{b}_{22} & \bar{b}_{23} & \bar{b}_{24} \\ 0 & 0 & 0 & 1 \end{bmatrix}_{CS(j) \rightarrow IM(j)} \begin{bmatrix} \bar{\mathbf{R}}_{3 \times 3}(j) & \bar{\mathbf{T}}_{3 \times 1}(j) \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}_{CS(1) \rightarrow CS(j)} \begin{Bmatrix} x_i(k) \\ y_i(k) \\ z_i(k) \\ 1 \end{Bmatrix}_{CS(1)}. \quad (14)$$

As discussed for Eq. (2), the false displacement at the j -th frame is the initial position of the main target represented in $IM(j)$, which is Eq. (14) with $k = 1$:

$$\begin{Bmatrix} u_i(1) \\ v_i(1) \\ 1 \end{Bmatrix}_{IM(j)} = \begin{bmatrix} \bar{b}_{11} & \bar{b}_{12} & \bar{b}_{13} & \bar{b}_{14} \\ \bar{b}_{21} & \bar{b}_{22} & \bar{b}_{23} & \bar{b}_{24} \\ 0 & 0 & 0 & 1 \end{bmatrix}_{CS(j) \rightarrow IM(j)} \begin{bmatrix} \mathbf{R}_{3 \times 3}(j) & \mathbf{T}_{3 \times 1}(j) \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}_{CS(1) \rightarrow CS(j)} \begin{Bmatrix} x_i(1) \\ y_i(1) \\ z_i(1) \\ 1 \end{Bmatrix}_{CS(1)}. \quad (15)$$

The unknown constants $\bar{b}_{11}, \bar{b}_{12}, \bar{b}_{13}, \bar{b}_{14}, \bar{b}_{21}, \bar{b}_{22}, \bar{b}_{23}, \bar{b}_{24}$, and $\{x_i(1) \ y_i(1) \ z_i(1)\}_{CS(1)}^T$ need to be determined to calculate the false displacement at the j -th frame $\{u_i(1) \ v_i(1) \ 1\}_{IM(j)}^T$.

The dual-camera system requires another initial calibration process to find the unknown constants in Eq. (15) before the actual measurement. The calibration process involves artificially moving the cameras h times assuming that both targets do not move from their original positions. Thus, $\{x_i(k) \ y_i(k) \ z_i(k) \ 1\}_{CS(j)}^T$ is the same as $\{x_i(1) \ y_i(1) \ z_i(1)\}_{CS(j)}^T$ during the calibration process. The false displacement with respect to the initial frame can be obtained as the difference between the positions $\{u_i(1) \ v_i(1) \ 1\}_{IM(j)}^T$ and $\{u_i(1) \ v_i(1) \ 1\}_{IM(1)}^T$:

$$\begin{aligned} & \begin{Bmatrix} u_i(1) \\ v_i(1) \end{Bmatrix}_{IM(j)} - \begin{Bmatrix} u_i(1) \\ v_i(1) \end{Bmatrix}_{IM(1)} \\ &= \begin{bmatrix} \bar{b}_{11} & \bar{b}_{12} & \bar{b}_{13} & \bar{b}_{14} \\ \bar{b}_{21} & \bar{b}_{22} & \bar{b}_{23} & \bar{b}_{24} \end{bmatrix}_{CS(j) \rightarrow IM(j)} \left(\begin{bmatrix} \bar{\mathbf{R}}_{3 \times 3}(j) & \bar{\mathbf{T}}_{3 \times 1}(j) \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}_{CS(1) \rightarrow CS(j)} - \begin{bmatrix} \bar{\mathbf{R}}_{3 \times 3}(1) & \bar{\mathbf{T}}_{3 \times 1}(1) \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}_{CS(1) \rightarrow CS(1)} \right) \begin{Bmatrix} x_i(1) \\ y_i(1) \\ z_i(1) \\ 1 \end{Bmatrix}_{CS(1)}. \quad (16) \end{aligned}$$

Because the transform from $CS(1)$ to $CS(j)$ exhibits an identity matrix (i.e., $\bar{\mathbf{R}}_{3 \times 3}(1)$ is an identity matrix and $\bar{\mathbf{T}}_{3 \times 1}(1)$ is a zero vector), Eq. (16) can be simplified to

$$\begin{Bmatrix} u_i(1) \\ v_i(1) \end{Bmatrix}_{IM(j)} - \begin{Bmatrix} u_i(1) \\ v_i(1) \end{Bmatrix}_{IM(1)} = \begin{bmatrix} \bar{b}_{11} & \bar{b}_{12} & \bar{b}_{13} \\ \bar{b}_{21} & \bar{b}_{22} & \bar{b}_{23} \end{bmatrix}_{CS(j) \rightarrow IM(j)} \left[\bar{\mathbf{R}}_{3 \times 3}(j) - \mathbf{I}_{3 \times 3} \quad \bar{\mathbf{T}}_{3 \times 1}(j) \right] \begin{Bmatrix} x_i(1) \\ y_i(1) \\ z_i(1) \\ 1 \end{Bmatrix}_{CS(1)}. \quad (17)$$

The nine unknown parameters $\bar{b}_{11}, \bar{b}_{12}, \bar{b}_{13}, \bar{b}_{21}, \bar{b}_{22}, \bar{b}_{23}$, and $\{x_i(1) y_i(1) z_i(1)\}_{CS(1)}^T$ are determined by solving the minimization problem in Eq. (18) as:

$$\underset{\substack{\bar{b}_{11}, \bar{b}_{12}, \bar{b}_{13} \\ \bar{b}_{21}, \bar{b}_{22}, \bar{b}_{23} \\ \{x_i(1) y_i(1) z_i(1)\}_{CS(1)}^T}}{\operatorname{argmin}} \sum_{j=1, \dots, h} \left\| \begin{Bmatrix} u_i(1) \\ v_i(1) \end{Bmatrix}_{IM(j)} - \begin{Bmatrix} u_i(1) \\ v_i(1) \end{Bmatrix}_{IM(1)} - \begin{bmatrix} \bar{b}_{11} & \bar{b}_{12} & \bar{b}_{13} \\ \bar{b}_{21} & \bar{b}_{22} & \bar{b}_{23} \end{bmatrix}_{CS(j) \rightarrow IM(j)} \left[\bar{\mathbf{R}}_{3 \times 3}(j) - \mathbf{I}_{3 \times 3} \quad \bar{\mathbf{T}}_{3 \times 1}(j) \right] \begin{Bmatrix} x_i(1) \\ y_i(1) \\ z_i(1) \\ 1 \end{Bmatrix}_{CS(1)} \right\|_2. \quad (18)$$

The nine parameters determined in Eq. (18) can be used to calculate the unknown parameters \bar{b}_{14} and \bar{b}_{24} in Eq. (15) via:

$$\begin{Bmatrix} \bar{b}_{14} \\ \bar{b}_{24} \end{Bmatrix} = \begin{Bmatrix} u_i(1) \\ v_i(1) \end{Bmatrix}_{IM(1)} - \begin{bmatrix} \bar{b}_{11} & \bar{b}_{12} & \bar{b}_{13} \\ \bar{b}_{21} & \bar{b}_{22} & \bar{b}_{23} \end{bmatrix}_{CS(1) \rightarrow IM(1)} \begin{Bmatrix} x_i(1) \\ y_i(1) \\ z_i(1) \end{Bmatrix}_{CS(1)}. \quad (19)$$

Thus, the false displacement in Eq. (15) can be calculated at each frame by using both the camera motions $\bar{\mathbf{R}}_{3 \times 3}(j)$ and $\bar{\mathbf{T}}_{3 \times 1}(j)$ as determined by Eq. (12) and the parameters $\bar{b}_{11}, \bar{b}_{12}, \bar{b}_{13}, \bar{b}_{14}, \bar{b}_{21}, \bar{b}_{22}, \bar{b}_{23}, \bar{b}_{24}$, and $\{x_i(1) y_i(1) z_i(1)\}_{CS(1)}^T$ as determined by Eqs. (18) and (19).

3.2.4. Displacement calculation with error compensation

The true displacement with respect to the initial frame, expressed as $\{x_i(j) - x_i(1) y_i(j) - y_i(1)\}_{WM}^T$, can be obtained in every measurement frame with the dual cameras. If the inverse of $\tilde{\mathbf{B}}_{WM \rightarrow IM(j)}$ is premultiplied with Eq. (4), the true displacement is expressed as

$$\begin{Bmatrix} x_i(j) - x_i(1) \\ y_i(j) - y_i(1) \end{Bmatrix}_{WM} = \tilde{\mathbf{B}}_{WM \rightarrow IM(j)}^{-1} \begin{Bmatrix} u_i(j) - u_i(1) \\ v_i(j) - v_i(1) \end{Bmatrix}_{IM(j)}. \quad (20)$$

Each term on the right-hand side of Eq. (20) can be obtained as discussed in previous sections. The main camera measures $\{u_i(j) v_i(j)\}_{IM(j)}^T$ for each frame, as discussed in Section 3.2.1. $\tilde{\mathbf{B}}_{WM \rightarrow IM(j)}$ can be obtained in the initial measurement, as discussed in Section 3.2.2. The sub-camera helps calculate $\{u_i(1) v_i(1)\}_{IM(j)}^T$ for each frame by using the 6-DOF camera motion, as discussed in Section 3.2.3. By combining information from the main and sub-cameras, the true displacement $\{x_i(j) - x_i(1) y_i(j) - y_i(1)\}_{WM}^T$ can be calculated to offer a camera motion-independent measurement.

The proposed method consists of two stages: calibration and measurement. Table 1 summarizes the parameters that need to be determined in each stage. The tasks in the calibration and measurement stages are as follows.

Calibration stage

1. $\tilde{\mathbf{B}}_{WM \rightarrow IM(j)}$ is computed through the direct linear transform shown in Eq. (8) by using at least three features.
2. The sub-camera is calibrated with the camera calibration method proposed by Zhang [42] to determine the intrinsic camera parameters $\{\alpha, \beta, \gamma, \mu, \nu\}$ in Eq. (10).

Table 1
Summary of the parameters to be determined in the proposed system.

Stage	Measurement provider	Parameter to be determined
Calibration stage	Main camera	$\tilde{\mathbf{B}}_{WM \rightarrow IM(j)}$
	Sub-camera	$\alpha, \beta, \gamma, \mu, \nu$
	Dual cameras	$\bar{b}_{11}, \bar{b}_{12}, \bar{b}_{13}, \bar{b}_{14}, \bar{b}_{21}, \bar{b}_{22}, \bar{b}_{23}, \bar{b}_{24}, \{x_i(1) y_i(1) z_i(1)\}_{CS(1)}^T$
Measurement stage	Main camera	$\{u_i(j) v_i(j)\}_{IM(j)}^T$
	Sub-camera	$\bar{\mathbf{R}}_{3 \times 3}(j), \bar{\mathbf{T}}_{3 \times 1}(j), \{u_i(1) v_i(1)\}_{IM(j)}^T$
	Dual cameras	$\{x_i(j) - x_i(1) y_i(j) - y_i(1)\}_{WM}^T$

3. The dual cameras are calibrated using Eqs. (18) and (19) to determine the unknown parameters $\bar{b}_{11}, \bar{b}_{12}, \bar{b}_{13}, \bar{b}_{14}, \bar{b}_{21}, \bar{b}_{22}, \bar{b}_{23}, \bar{b}_{24}$, and $\{x_i(1) y_i(1) z_i(1)\}^T_{CS(1)}$.

Measurement stage

1. The main camera measures $\{u_i(j) v_i(j)\}^T_{IM(j)}$ for each feature point in the j -th image of the main target.
2. The sub-camera motion at the j -th frame (i.e., $\bar{R}_{3 \times 3}(j)$ and $\bar{T}_{3 \times 1}(j)$ in Eq. (12)) is calculated using the camera calibration method [42] with $\{\alpha, \beta, \gamma, \mu, \nu\}$. Then, $\bar{R}_{3 \times 3}(j)$ and $\bar{T}_{3 \times 1}(j)$ are employed in the false displacement computation in Eq. (15) with $\bar{b}_{11}, \bar{b}_{12}, \bar{b}_{13}, \bar{b}_{14}, \bar{b}_{21}, \bar{b}_{22}, \bar{b}_{23}, \bar{b}_{24}$, and $\{x_i(1) y_i(1) z_i(1)\}^T_{CS(1)}$.
3. The true displacement $\{x_i(j) - x_i(1) y_i(j) - y_i(1)\}^T_{WM}$ in Eq. (20) is calculated by using $\tilde{\mathbf{B}}_{WM \rightarrow IM(j)}$, $\{u_i(1) v_i(1)\}^T_{IM(j)}$, and $\{u_i(j) v_i(j)\}^T_{IM(j)}$.

4. Validation

The proposed long-term displacement measurement method was validated through a numerical simulation, laboratory-scale test, and field test. A numerical simulation was conducted to assess the performance of the dual-camera system in terms of pixel resolution. A dual-camera system was fabricated to meet the requirements and then used in the indoor tests. Artificial camera movements were introduced to generate errors that needed to be compensated for by the proposed algorithm. The proposed dual-camera system was then implemented on a full-scale railway bridge to monitor displacement at the mid-span for 145 days.

4.1. Numerical validation

The dual-camera system was numerically simulated using MATLAB to investigate the performance of the error compensation with regard to pixel resolution. The main and sub-targets were positioned to have a distance of 10 m as shown in Fig. 6. Note that the size of the sub-target needs to be determined based on the distance between the sub-target and the sub-camera to maintain the estimation accuracy of the 6-DOF motion: a larger target is necessary for a longer distance. The main and sub-cameras were assumed to be identical and attached to each other with parallel optical axes in the opposite directions and an 80 mm distance between the origins of CM and CS. The positions of the dual cameras were carefully selected so that the feature points within the main and sub-targets would be properly captured by the cameras. The fields of view of the main and sub-cameras were controlled to be roughly $200 \text{ mm} \times 150 \text{ mm}$ and $150 \text{ mm} \times 100 \text{ mm}$, respectively. The dual cameras were moved with respect to the initial position of the sub-camera CS(1) with various 6-DOF motions as listed in Table 2. Here, the last 100 frames were prepared with uniformly distributed random motions bounded by $\pm 1 \text{ mrad}$ rotations about $x_{CS(1)}, y_{CS(1)}$, and $z_{CS(1)}$ and $\pm 10 \text{ mm}$ translations along $x_{CS(1)}, y_{CS(1)}$, and $z_{CS(1)}$. For each camera motion, image coordinates of the features existing in the main and sub-targets were generated and quantized to simulate four different camera resolutions $1920 \times 1200, 3840 \times 2400, 7680 \times 4800$, and 15360×9600 . Because the original image coordinates can be considered to be an $\infty \times \infty$ resolution, a total of five different cameras were investigated in the simulation.

The resulting vertical displacements without and with the error compensation are shown in Fig. 7(a) and (b)–(f), respectively. Because the main target was fixed, the displacement measured by the main camera without any error compensation was the false displacement due to camera motions, as shown in Fig. 7(a). The error was suppressed using the sub-camera, as

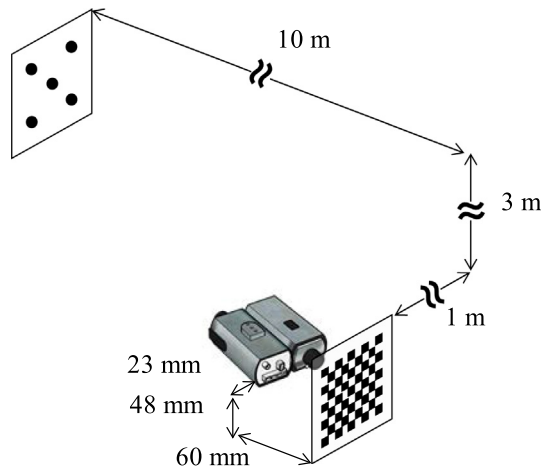


Fig. 6. Simulation setup for cameras and targets.

Table 2
6-DOF camera motions for numerical simulation.

Frame	Camera motions	
	Motion type	Range
1–100	Rotation about $x_{CS(1)}$	[−1 mrad, +1 mrad]
101–200	Rotation about $y_{CS(1)}$	
201–300	Rotation about $z_{CS(1)}$	
301–400	Translation along $x_{CS(1)}$	[−10 mm, +10 mm]
401–500	Translation along $y_{CS(1)}$	
501–600	Translation along $z_{CS(1)}$	
601–700	6-DOF random motions	

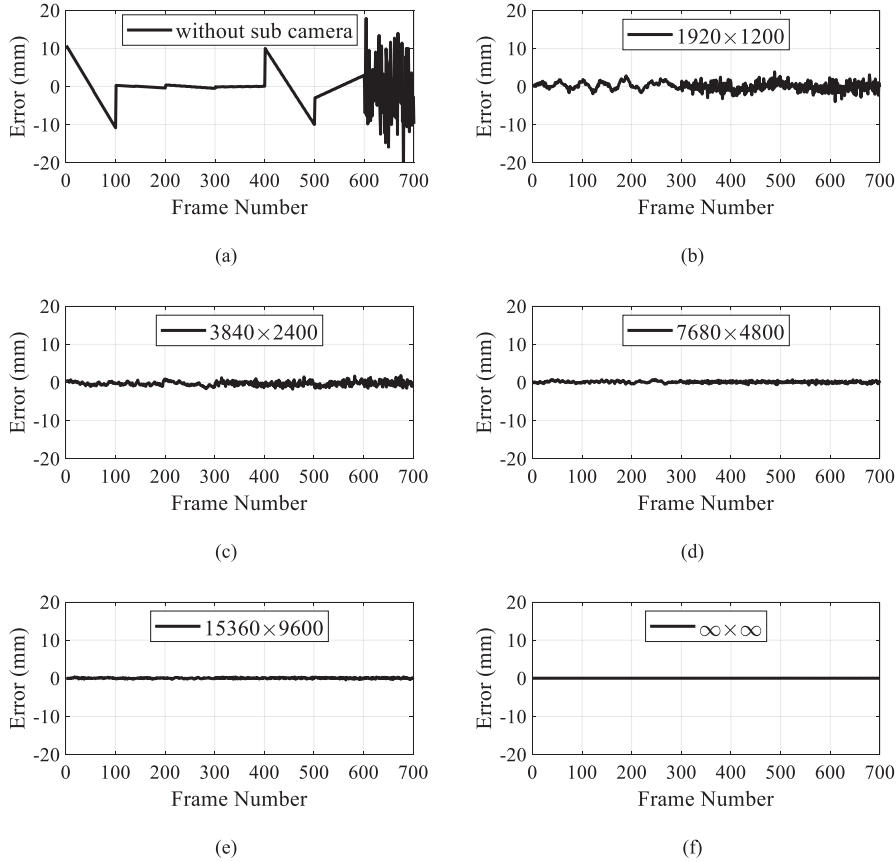


Fig. 7. Effect of the sub-camera and its pixel resolution on the error compensation: (a) no sub-camera, (b) 1920×1200 pixels, (c) 3840×2400 pixels, (d) 7680×4800 pixels, (e) 15360×9600 pixels, and (f) $\infty \times \infty$ pixels.

shown in Fig. 7(b)–(f), which indicates that the performance improved for higher camera resolutions. The error levels can be further lowered by introducing the subpixel algorithm. Thus, the proposed algorithm was demonstrated to be capable of removing camera motion-induced errors.

4.2. Laboratory-scale validation

The proposed method was validated in a laboratory environment with the dual-camera system. The experimental setup and hardware specifications are presented in Fig. 8 and Table 3, respectively. The main target was designed to have five circular points on a black plate that was fixed 11 m away from the camera system. A checkerboard was prepared for the sub-target, which was attached to a stationary wall. Two industrial-grade cameras were rigidly attached to each other by steel plates to construct the dual cameras. The main and sub-cameras were equipped with a telescopic lens and wide-angle lens for resolutions of 0.08 mm/pixel at a 11 m distance and 0.05 mm/pixel at a 50 mm distance, respectively. A pan-tilt motor was installed under the dual cameras (see Fig. 8) and employed for camera calibration to determine the unknown parame-

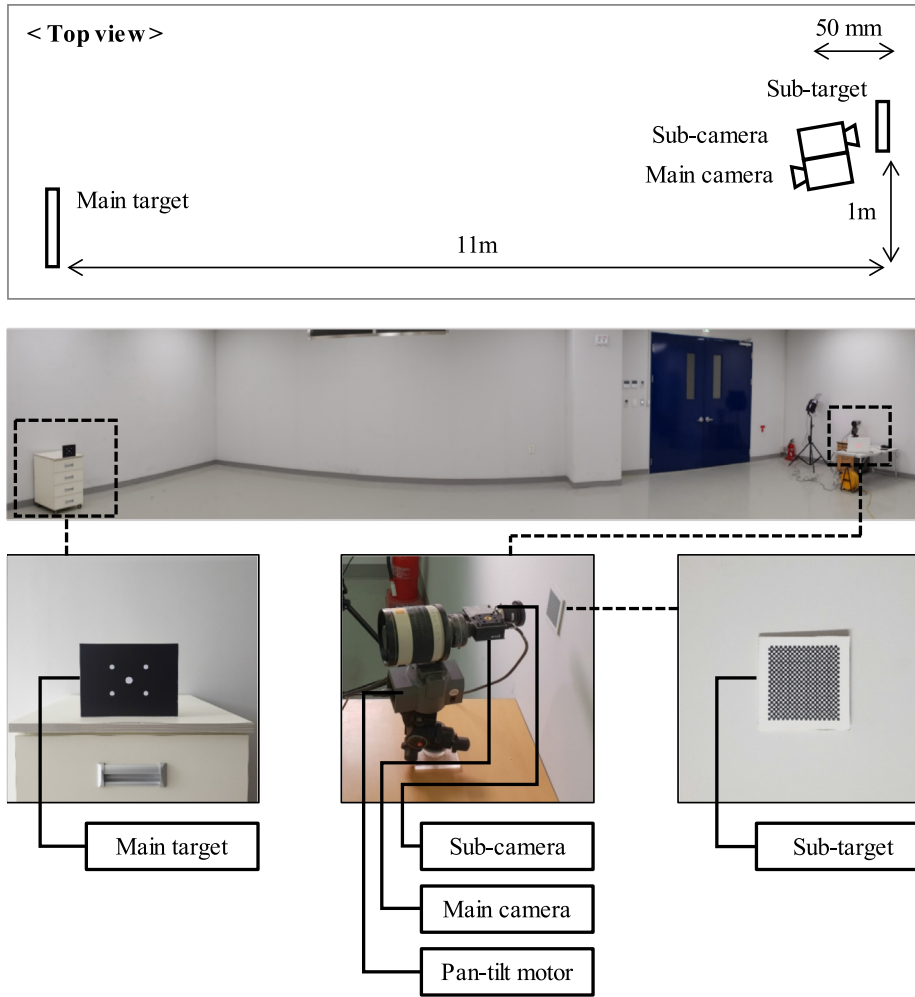


Fig. 8. Laboratory test setup.

Table 3
Hardware specifications.

Hardware	Specification
Main target	Five white circles on a black background Horizontal/vertical interval: 15 mm/11 mm
Sub-target	32 × 31 checkerboard Interval: 2 mm
Main camera	Model: GS3-U3-23S6C-C (1920 × 1200 pixel resolution, 1/1.2" image sensor) Focal length of the lens: 800 mm 0.08 mm/pixel at 11 m distance
Sub-camera	Model: GS3-U3-23S6M-C (1920 × 1200 pixel resolution, 1/1.2" image sensor) Focal length of the lens: 6 mm 0.05 mm/pixel at 50 mm distance

ters in Eqs. (18) and (19); this allowed for sensitive movement of the camera system. A total of 21 camera motions were generated for the camera calibration. After the calibration, the pan-tilt motor was manipulated to generate gradual three-dimensional rotations and translations of the dual cameras to simulate long-term camera movements.

The dual-camera system was carefully designed to satisfy the requirement of the long-term displacement measurement method described in Section 3. The affine camera assumption was valid for the main camera because $\|x - x_{\text{affine}}\|_2$ was less than 1 pixel, which was calculated based on the working distance (d_0 was 11 m), the depth motion (Δz was less than 5 mm), and the position of the feature point near the center ($x - x_0$ was less than 100 pixels). The sub-camera was a perspective

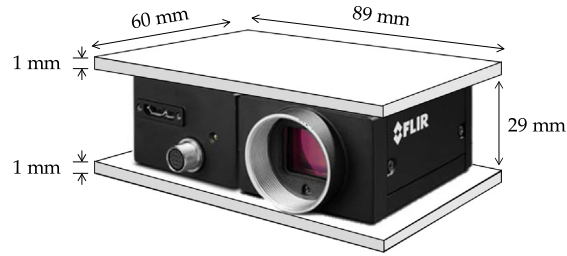


Fig. 9. The steel plates used to constrain the relative positions between the cameras.

camera as the sub-mm of depth motion made visible changes in the captured images. The relative positions between the cameras could be effectively constrained by the steel plates shown in Fig. 9. The steel plates only deform by $12 \mu\text{rad}$ of rotation per 1 N-m of torsional moment and $90 \mu\text{m}$ per 1°C of temperature change; the deformation due to the self-weight of the dual camera system and the thermal expansion under temperature changes are negligible compared to the possible camera ego-motion. This issue will be further discussed in the later section regarding experimental validation. The sub-target, which was made of the foamex polyvinyl chloride, weighed only 11 g and could be tightly attached to the wall using an adhesive (Loctite 401) without causing notable displacement. Thus, we could assume that the error sources except the movements of the whole body of the combined cameras could be handled by appropriately designing the dual-camera system.

The displacement measured by the main camera without the error compensation (denoted as “single camera” hereafter) was compared with that corrected by the dual-camera system. In the calibration stage, 21 sets of images were taken by the main and sub-cameras upon each intentional movement by the pan-tilt motor. Once the unknown parameters had been determined in the calibration process, 6-DOF camera motions were induced using the pan-tilt motor as shown in Fig. 10, which produced a maximum false displacement of 44.1 mm for the main camera as shown in Fig. 10(a). Stepwise camera motions were observed due to the limited performance of the pan-tilt motor. The false displacement was compensated for with the dual-camera system, which resulted in a false displacement of 1.1 mm as shown in Fig. 11(b). Thus, the feasibility of the dual-camera system was demonstrated; it reduced the motion-induced error from 44.1 mm to 1.1 mm.

4.3. Field validation

The dual-camera system was implemented on a full-scale railway bridge to measure its long-term displacement during construction stage. The bridge consists of four pre-stressed concrete girders with identical shapes and dimensions (i.e., $1.8 \text{ m} \times 2.2 \text{ m} \times 40 \text{ m}$), as shown in Fig. 12(a). The hardware listed in Table 3 was employed in the field test; the main target was enlarged twofold because the distance between the target and camera was doubled. The dual-camera body was constrained by steel plates, as shown in Fig. 9, to tightly attach the cameras. The main target was attached to the mid-span of the bridge, where the long-term displacement was desired. The sub-target, made of foamex polyvinyl chloride, was attached to the pier by the adhesive Loctite 401, which assured a permanent bonding during the experiment. The pan-tilt motor shown in Fig. 12(b) was used to adjust the camera view for the initial calibration described in Section 3.2.4. Once the unknown parameters were determined in the calibration, the measurement stage began. Images were taken every minute to obtain the long-term displacement. The displacement measurement commenced on 16 December 2017 and lasted for 145 days.

Two reference displacements were considered for this full-scale experiment: a finite element model and a LiDAR-based long-term displacement measurement method [28]. The finite element model of the test bed bridge was built with MIDAS Civil, which is a commercial software application for finite element analysis [43]. The model considered time-dependent concrete properties including creep and shrinkage under steam curing with 70% relative humidity, as defined in KCI 2012

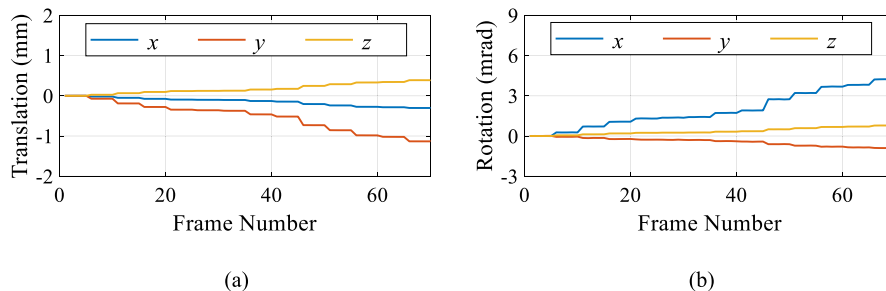


Fig. 10. 6-DOF camera motions generated by the pan-tilt motor: (a) translation and (b) rotation.

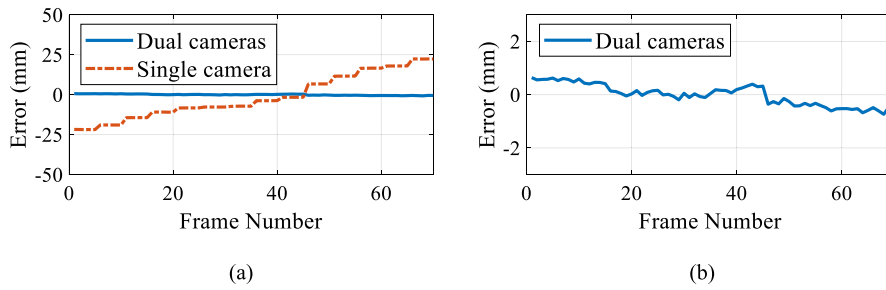


Fig. 11. Results of laboratory tests: (a) comparison between the dual-camera and single-camera methods; (b) magnification of the dual-camera method.

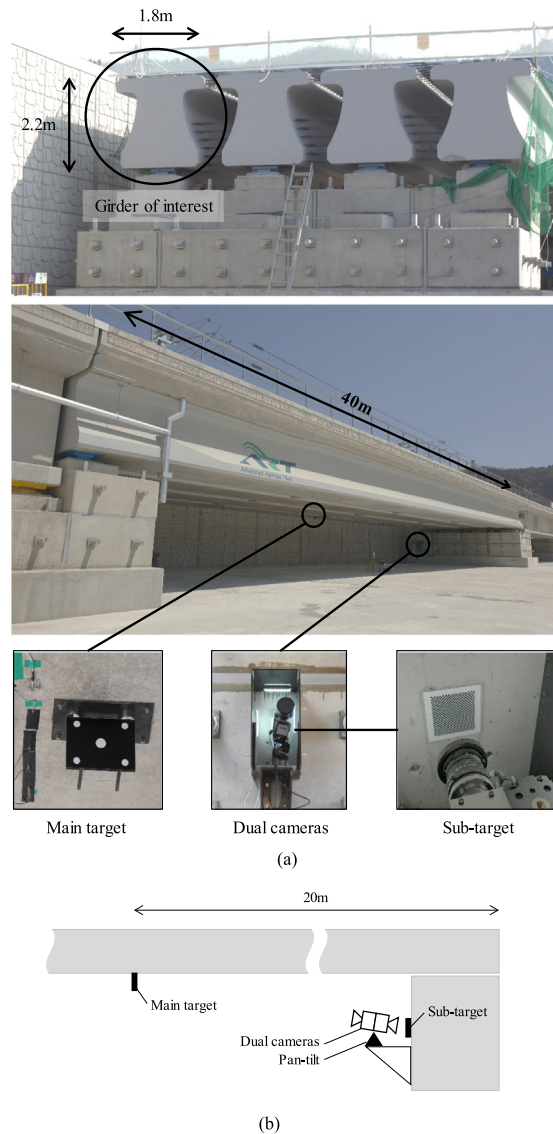


Fig. 12. Experimental setup: (a) installation of the dual-camera system on the test bed bridge; (b) schematic of the experimental setup.

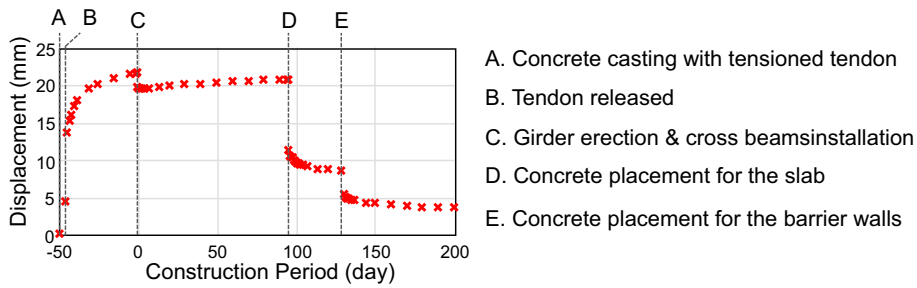


Fig. 13. Expected bridge displacement using the finite element model with the construction schedule.

[44], and loading histories relating to the construction process, such as the pre-stress transfer, girder erection, and slab concrete casting. The compressive strengths of the concrete used for the girders and slab were set at 49 MPa and 30 MPa, respectively. The calculated displacement for the construction process is shown in Fig. 13. The LiDAR-based method [28] was employed to provide a reference measurement. This method features a LiDAR position-independent algorithm that allows LiDAR equipment to be repositioned for each measurement, which makes permanent installation in the field unnecessary. Because LiDAR equipment needed to be brought to the bridge for each measurement, only weekly displacement measurements were possible.

To validate the proposed dual-camera system, the measured displacement was compared to the values from the finite element model and those measured by the LiDAR-based method [28]. Fig. 14 shows that the three long-term displacements shared similar overall trends that were consistent with the construction process. Note that the displacement from the dual-camera system was plotted with measurements every 10 min for a clear graphical demonstration. Detailed comparisons are provided below for each time period.

The 145 days of monitoring, starting from Day 41, helped to understand the long-term behavior of the railway bridge during the construction stage. Each subfigure in Fig. 15 shows a detailed view of Fig. 14 for a more comprehensive understanding. For the first 5 days, an upward displacement of about 5 mm was observed, which was attributable to sunlight. Only the daytime displacement was measured because the main target was invisible at night. A long discontinuity was observed between days 45 and 60 that was due to a power failure. Snow was stacked on the girder on days 64–78, which can be indirectly identified in Fig. 15(a). This blocked the sunlight, so less daily fluctuation was observed. A lighting system was installed on day 64 so that the displacement could be measured during the nighttime. On day 95, an instantaneous displacement of 6.5 mm was measured because the concrete for the slab was being put into place, as shown in Fig. 15(b). Then, the bridge was entirely covered by a tent for steam curing, which resulted in an upward displacement of about 4 mm during days 95–97. Subsequently, a displacement of about 8 mm gradually developed during days 95–185, as shown in Fig. 15 (c); this was mostly caused by concrete creep. The creep coefficient was estimated to be 1.23. Heavy rain during days 130–137, 150–153, and 168–171 (see Fig. 16) blocked direct sunlight, which prevented the bridge from deflecting upward. Abrupt changes in displacement were observed on days 173, 178, 183, and 184 because of disturbances from bugs drawn by the lighting system.

The dual-camera system was designed to compensate for the false displacement induced by camera motions. Fig. 17 shows the 6-DOF motions of the sub-camera with respect to CS(1). Even though the cameras were tightly fixed to the pier, the camera system was prone to movement because of its self-weight and the harsh weather conditions. The camera motions resulted in the drift error shown in Fig. 18, where the single camera (i.e., main camera only) measured a displacement 30 mm larger than that measured by the dual cameras. The false displacement was even larger than the true structural displacement, which clearly demonstrates why the camera motion-induced error must be removed for reliable long-term measurement. The dual-camera system was demonstrated to accurately measure the long-term structural displacement by compensating for the error due to camera motion.

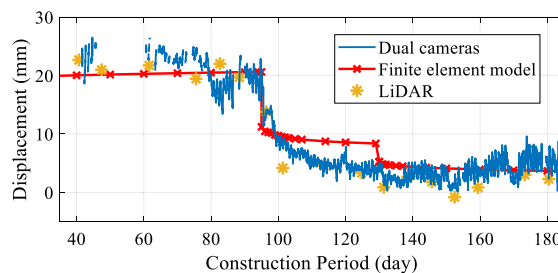


Fig. 14. Comparison of the displacements with the dual cameras, LiDAR, and finite element analysis.

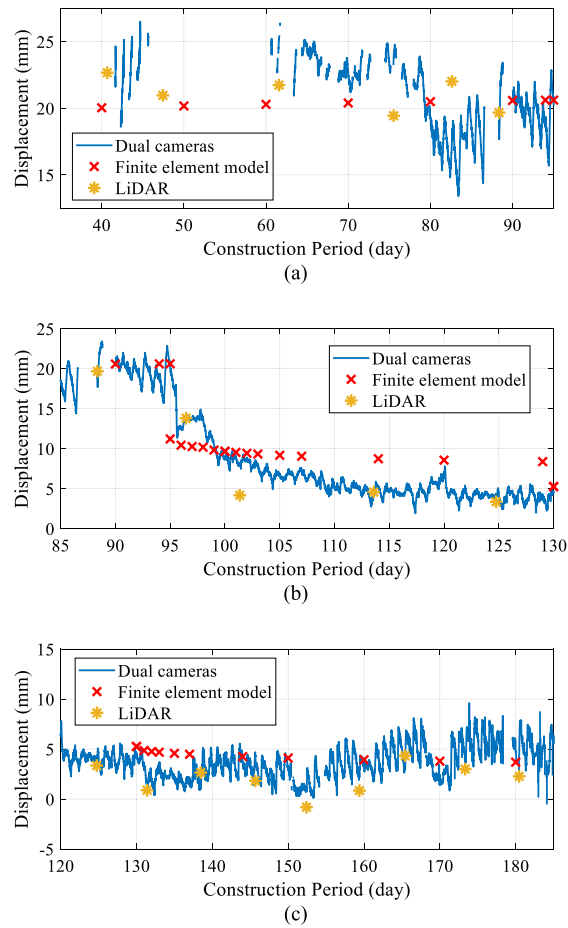


Fig. 15. Field test results: (a) days 35–95, (b) days 85–130, and (c) days 120–185.

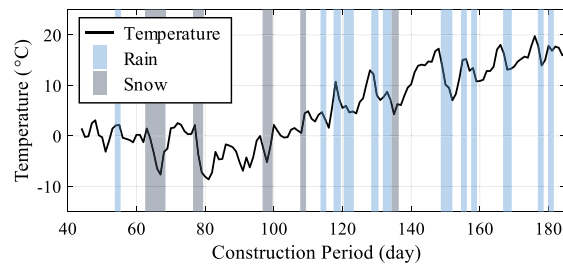


Fig. 16. Weather conditions during the displacement measurement.

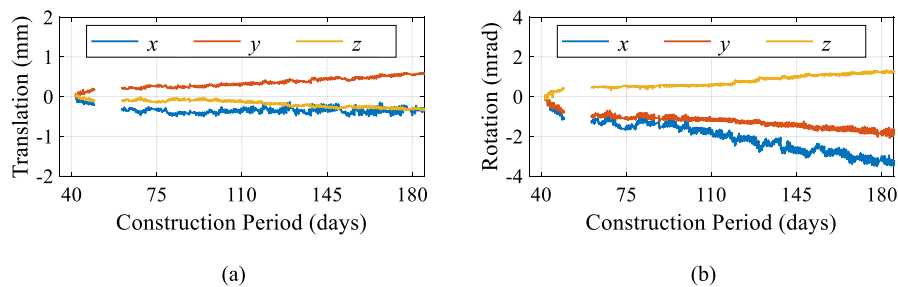


Fig. 17. 6-DOF camera motions for 145 days: (a) translation and (b) rotation.

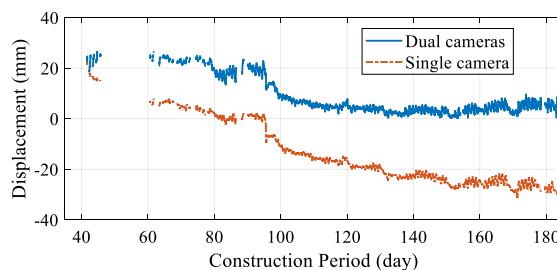


Fig. 18. Comparison between the dual- and single-camera methods.

4.4. Discussion

The 145 days of long-term measurement clearly revealed that methods of compensating for the error induced by camera motions is important for reliable displacement measurements. Complete removal of the camera motions was difficult in the field test due to the harsh field environment and self-weight of the camera system. The movements of the dual-camera body were clearly observed in Fig. 17, despite the efforts made to minimize the camera motions through bolting the system tightly to the bridge pier. These motions can cause a significant error unless properly handled, because the camera is generally located at a distance from the target. Indeed, a false displacement of 30 mm was observed as shown in Fig. 18, which was even larger than the true displacement of 20 mm. Thus, the error induced by the camera motions, which is difficult to control in practice, must be compensated for.

The error sources other than the camera motions could be suppressed by means of an appropriate design of the dual-camera system. The error sources that are not considered in the proposed method include: (1) relative motion between the cameras and (2) movement of the sub-target. As discussed in Section 4.2, two cameras were combined using the steel plates, which constrained the relative motion between the cameras under the loadings, owing to its self-weight and temperature changes. As for the sub-target, the light-weight material foamex pvc was adopted so that the self-weight of the target would not induce displacement. Another concern is the movement of the stationary reference point where the sub-target is attached. Because the deformation of the bridge girder is of primary interest in the field testing of railway bridges, the displacement at the midspan relative to the bridge piers was the desired measurement. Indeed, the dual camera system provides an appropriate displacement measurement in this field test, even when the piers are sinking or shortened. As such, these two error sources were controlled practically by carefully designing and positioning the dual-camera system, and only the movements of the dual-camera body were considered in the error compensation. As can be seen in Fig. 14, the resulting displacement showed good agreement with that of the LiDAR-based approach. Hence, the steel plates used to combine the cameras and the light-weight material used for the sub-target were seen to be effective in handling the error sources.

Advanced structural health monitoring could be achieved by considering the long-term displacement. First, the concrete creep and shrinkage of the bridge could be identified. In the field test using the railway bridge, a displacement of 8 mm due to concrete creep was observed as it developed over 50 days. Second, the temperature gradient due to direct sunlight was the main influence on the daily displacement. The exposure of the bridge surface to sunlight induced an upward displacement during the daytime, while the bridge returned to its original position in the nighttime. Similarly, the sunblock effect by clouds limited the upward movement of the bridge. Lastly, the instantaneous displacement from the overlying substructures could provide immediate feedback regarding the structural integrity; this can be used to check the construction quality and assure worker safety. As such, the long-term displacement can be employed in various civil engineering structures to assure their safety, from construction to in-service stages.

5. Conclusions

This paper proposes a long-term displacement measurement system based on computer vision. Because conventional vision-based approaches are tailored to the short-term measurement of dynamic displacements with a fixed camera, an unexpected camera motion during measurement may result in significant errors. The proposed system introduces main and sub-cameras to compensate for the motion-induced error during displacement measurement. The main camera measures the structural displacement like conventional methods, so the camera motion-induced error is included. The sub-camera is attached to the main camera and measures its 6-DOF camera motions. The physical relationship between the 6-DOF camera motions and the motion-induced error was identified, which requires nine unknown parameters to be determined before measurement. The unknown parameters are determined through a proposed calibration process, which allows the motion-induced error to be estimated for each frame. Removing the measurement error from the main camera allows the true structural displacement to be directly obtained. The proposed dual-camera system is designed to measure long-term displacements.

The proposed system was validated through a numerical simulation and both laboratory and field experiments. The numerical simulation was conducted by modeling the vision-based displacement measurements with moving cameras. The simulation validated the error compensation capability of the proposed method, while its performance depended on the camera pixel resolution. A dual-camera system was fabricated and tested in a laboratory environment by measuring the displacement of the fixed target with the moving dual cameras. The motion-induced error was over 44.1 mm, which was suppressed to 1.1 mm with the proposed method. The proposed method was applied to a full-scale railway bridge under construction. The displacement measured by the dual-camera system showed good agreement with the reference displacements obtained with LiDAR and estimated by a finite element model over 145 days of measurement. The proposed system successfully measured the long-term bridge displacement due to construction, which coincided well with the numerical analysis results and LiDAR measurements.

CRedit authorship contribution statement

Junhwa Lee: Methodology, Software, Validation, Formal analysis, Writing - original draft. **Kyoung-Chan Lee:** Validation, Investigation, Writing - review & editing. **Seunghoo Jeong:** Validation, Formal analysis. **Young-Joo Lee:** Methodology, Investigation, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Sung-Han Sim:** Conceptualization, Methodology, Formal analysis, Investigation, Writing - review & editing, Supervision.

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.

Acknowledgments

This research was supported by a grant from R&D Program of the Korea Railroad Research Institute, Republic of Korea.

References

- [1] J.-W. Park, K.-C. Lee, S.-H. Sim, H.-J. Jung, B.F. Spencer Jr., Traffic safety evaluation for railway bridges using expanded multisensor data fusion, *Comput.-Aided Civ. Infrastruct. Eng.* 31 (10) (2016) 749–760.
- [2] Y.-J. Lee, R.E. Kim, W. Suh, K. Park, Probabilistic fatigue life updating for railway bridges based on local inspection and repair, *Sensors* 17 (4) (2017) 936.
- [3] F. Huseynov, C. Kim, E.J. O'Brien, J.M.W. Brownjohn, D. Hester, K.C. Chang, Bridge damage detection using rotation measurements—Experimental validation, *Mech. Syst. Sig. Process.* 135 (2020) 106380.
- [4] E.S. Tomé, M. Pimentel, J. Figueiras, Damage detection under environmental and operational effects using cointegration analysis—Application to experimental data from a cable-stayed bridge, *Mech. Syst. Sig. Process.* 135 (2020) 106386.
- [5] Z. Yi, C.-W. Kim, K.F. Tee, A. Garg, A. Garg, Long-term health monitoring for deteriorated bridge structures based on copula theory, *Smart Struct. Syst.* 21 (2) (2018) 171–185.
- [6] M. Kim, D. Sung, Experimental investigation on effects of track configurations on long-term behavior of ballasted track, *J. Struct. Integrity Maint.* 4 (2) (2019) 76–85.
- [7] C. Pearson, N. Delatte, Collapse of the Quebec bridge, 1907, *J. Perform. Constr. Facil.* 20 (1) (2006) 84–91.
- [8] Z.P. Bažant, Q. Yu, G.H. Li, Excessive long-time deflections of prestressed box girders. I: Record-span bridge in Palau and other paradigms, *J. Struct. Eng.* 138 (6) (2012) 676–686.
- [9] N. Bagge, J. Nilimaa, T. Blanksvärd, L. Elfgren, Instrumentation and full-scale test of a post-tensioned concrete bridge, *Nordic Concrete Research* 51 (2014) 63–83.
- [10] S. Cho, J. Lee, S.-H. Sim, Comparative study on displacement measurement sensors for high-speed railroad bridge, *Smart Struct. Syst.* 21 (5) (2018) 637–652.
- [11] H. Jo, S.-H. Sim, A. Tatkowski, B.F. Spencer Jr., M.E. Nelson, Feasibility of displacement monitoring using low-cost GPS receivers, *Struct. Control. Health Monit.* 20 (9) (2013) 1240–1254.
- [12] P.A. Psimoulis, S.C. Stiros, Measuring deflections of a short-span railway bridge using a robotic total station, *J. Bridge Eng.* 18 (2) (2013) 182–185.
- [13] H.H. Nassif, M. Gindy, J. Davis, Comparison of laser Doppler vibrometer with contact sensors for monitoring bridge deflection and vibration, *NDT and E Int.* 38 (3) (2005) 213–218.
- [14] K. Kim, H. Sohn, Dynamic displacement estimation by fusing LDV and LiDAR measurements via smoothing based Kalman filtering, *Mech. Syst. Sig. Process.* 82 (2017) 339–355.
- [15] K.T. Park, S.H. Kim, H.S. Park, K.W. Lee, The determination of bridge displacement using measured acceleration, *Eng. Struct.* 27 (3) (2005) 371–378.
- [16] M. Gindy, R. Vaccaro, H. Nassif, J. Velde, A state-space approach for deriving bridge displacement from acceleration, *Comput.-Aided Civ. Infrastruct. Eng.* 23 (4) (2008) 281–290.
- [17] J.W. Park, S.-H. Sim, H.J. Jung, Development of a wireless displacement measurement system using acceleration responses, *Sensors* 13 (7) (2013) 8377–8392.
- [18] F. Moreu, J. Li, H. Jo, R.E. Kim, S. Scola, B.F. Spencer Jr., J.M. LaFave, Reference-free displacements for condition assessment of timber railroad bridges, *J. Bridge Eng.* 21 (2) (2016) 04015052.
- [19] L.H. Kang, D.K. Kim, J.H. Han, Estimation of dynamic structural displacements using fiber Bragg grating strain sensors, *J. Sound Vib.* 305 (3) (2007) 534–542.
- [20] S. Shin, S.U. Lee, Y. Kim, N.S. Kim, Estimation of bridge displacement responses using FBG sensors and theoretical mode shapes, *Struct. Eng. and Mech.* 42 (2) (2012) 229–245.
- [21] J.-W. Park, S.-H. Sim, H.-J. Jung, Displacement estimation using multimetric data fusion, *IEEE/ASME Trans. Mechatron.* 18 (6) (2013) 1675–1682.
- [22] 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, *Mech. Syst. Sig. Process.* 38 (2) (2013) 460–481.
- [23] S. Cho, S.-H. Sim, J.-W. Park, J. Lee, Extension of indirect displacement estimation method using acceleration and strain to various types of beam structures, *Smart Struct. Syst.* 14 (4) (2014) 699–718.

- [24] S. Cho, C.B. Yun, S.-H. Sim, Displacement estimation of bridge structures using data fusion of acceleration and strain measurement incorporating finite element model, *Smart Struct. Syst.* 15 (3) (2015) 645–663.
- [25] G.H. Cao, S. Zhang, W. Zhang, X.R. Peng, Long-term deflection test and theoretical analysis on cracked prestressed concrete box beams, *KSCE J. Civ. Eng.* 22 (2) (2018) 688–695.
- [26] O. Burdet, Experience in the long-term monitoring of bridges, In: 3rd fib International Congress, 108–113, 2010.
- [27] A. Beltempo, O.S. Bursi, C. Cappello, D. Zonta, M. Zingales, A viscoelastic model for the long-term deflection of segmental prestressed box girders, *Comput.-Aided Civ. Infrastruct. Eng.* 33 (1) (2018) 64–78.
- [28] J. Lee, K.-C. Lee, S. Lee, Y. Lee, S.-H. Sim, Long-term displacement measurement by using a LiDAR system, *Struct. Control. Health Monit.* 26 (10) (2019).
- [29] B.F. Spencer Jr., V. Hoskere, Y. Narazaki, Advances in computer vision-based civil infrastructure inspection and monitoring, *Eng.* 5 (2) (2019) 199–222.
- [30] J.J. Lee, M. Shinozuka, A vision-based system for remote sensing of bridge displacement, *NDT and E Int.* 39 (5) (2006) 425–431.
- [31] C.C. Chang, Y.F. Ji, Flexible videogrammetric technique for three-dimensional structural vibration measurement, *J. Eng. Mech.* 133 (6) (2007) 656–664.
- [32] Y. Fukuda, M.Q. Feng, Y. Narita, S.I. Kaneko, T. Tanaka, Vision-based displacement sensor for monitoring dynamic response using robust object search algorithm, *IEEE Sens. J.* 13 (12) (2013) 4725–4732.
- [33] H. Yoon, H. Elanwar, H. Choi, M. Golparvar-Fard, B.F. Spencer Jr., Target-free approach for vision-based structural system identification using consumer-grade cameras, *Struct. Control. Health Monit.* 23 (12) (2016) 1405–1416.
- [34] J. Lee, K.-C. Lee, S. Cho, S.-H. Sim, Computer vision-based structural displacement measurement robust to light-induced image degradation for in-service bridges, *Sensors* 17 (10) (2017) 2317.
- [35] C.A. Santos, C.O. Costa, J. Batista, A vision-based system for measuring the displacements of large structures: Simultaneous adaptive calibration and full motion estimation, *Mech. Syst. Sig. Process.* 72 (2016) 678–694.
- [36] D. Lydon, M. Lydon, S. Taylor, J.M. Del Rincon, D. Hester, J. Brownjohn, Development and field testing of a vision-based displacement system using a low cost wireless action camera, *Mech. Syst. Sig. Process.* 121 (2019) 343–358.
- [37] S.W. Kim, B.G. Jeon, N.S. Kim, J.C. Park, Vision-based monitoring system for evaluating cable tensile forces on a cable-stayed bridge, *Struct. Health Monit.* 12 (5–6) (2013) 440–456.
- [38] J.G. Chen, A. Davis, N. Wadhwa, F. Durand, W.T. Freeman, O. Büyüköztürk, Video camera-based vibration measurement for civil infrastructure applications, *J. Infrastruct. Syst.* 23 (3) (2016) B4016013.
- [39] H. Yoon, J. Shin, B.F. Spencer Jr., Structural displacement measurement using an unmanned aerial system, *Comput.-Aided Civ. Infrastruct. Eng.* 33 (3) (2018) 183–192.
- [40] R. Hartley, A. Zisserman, Multiple view geometry in computer vision, second ed., Cambridge University Press, 2000.
- [41] A. Ben-Israel, T.N. Greville, Generalized inverses: theory and applications, Springer Science & Business Media, 2003.
- [42] Z. Zhang, Flexible camera calibration by viewing a plane from unknown orientations, in: Proceedings of the Seventh IEEE International Conference on Computer Vision (ICCV), 1999, pp. 666–673.
- [43] Midas Information Technology Co., Ltd. Midas Civil Available at: https://en.midasuser.com/product/civil_overview.asp [Accessed 05 July 2019].
- [44] Korea Concrete Institute (KCI), Structural concrete design code, 2012.