

Article

Continuous Structural Displacement Monitoring Using Accelerometer, Vision, and Infrared (IR) Cameras

Jaemook Choi, Zhanxiong Ma, Kiyoung Kim and Hoon Sohn *

Department of Civil and Environmental Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea

* Correspondence: hoonsohn@kaist.ac.kr

Abstract: With the rapid development of computer vision, vision cameras have been used as non-contact sensors for structural displacement measurements. However, vision-based techniques are limited to short-term displacement measurements because of their degraded performance under varying illumination and inability to operate at night. To overcome these limitations, this study developed a continuous structural displacement estimation technique by combining measurements from an accelerometer with vision and infrared (IR) cameras collocated at the displacement estimation point of a target structure. The proposed technique enables continuous displacement estimation for both day and night, automatic optimization of the temperature range of an infrared camera to ensure a region of interest (ROI) with good matching features, and adaptive updating of the reference frame to achieve robust illumination–displacement estimation from vision/IR measurements. The performance of the proposed method was verified through lab-scale tests on a single-story building model. The displacements were estimated with a root-mean-square error of less than 2 mm compared with the laser-based ground truth. In addition, the applicability of the IR camera for displacement estimation under field conditions was validated using a pedestrian bridge test. The proposed technique eliminates the need for a stationary sensor installation location by the on-site installation of sensors and is therefore attractive for long-term continuous monitoring. However, it only estimates displacement at the sensor installation location, and cannot simultaneously estimate multi-point displacements which can be achieved by installing cameras off-site.

Citation: Choi, J.; Ma, Z.; Kim, K.; Sohn, H. Continuous Structural Displacement Monitoring Using Accelerometer, Vision, and Infrared (IR) Cameras. *Sensors* **2023**, *23*, 5241. <https://doi.org/10.3390/s23115241>

Academic Editor: Gregorij Kurillo

Received: 4 May 2023

Revised: 27 May 2023

Accepted: 30 May 2023

Published: 31 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: displacement estimation; infrared camera; vision camera; accelerometer; multirate adaptive Kalman filter; continuous monitoring

1. Introduction

Displacement is a critical parameter that indicates the level of deformation or movement of civil infrastructure [1,2]. Measuring displacement helps to identify potential safety hazards and structural issues that could lead to failure or collapse. By monitoring the displacement, engineers can determine whether a structure is still within the acceptable limits of operation or if it requires repair or reinforcement. A linear variable displacement transducer (LVDT) is commonly used for bridge displacement measurement [3]. However, its usage requires the installation of a scaffold beneath the bridge, which may not be feasible for river crossings and overpass bridges where traffic flow interruptions are not permitted. Although real-time kinematic global navigation satellite systems (RTK-GNSS) have been widely applied for the continuous monitoring of structural displacement in large-scale bridges [4] and tall buildings [5], their precision and sampling rate are restricted, making them inadequate for monitoring small- or medium-scale structures. Though satellite-based interferometry techniques are advantageous for full-field displacement measurements for landslides [6] and bridges [7], they are limited to static displacement monitoring. Owing to the limitations of current displacement sensors,

accelerometers are commonly used for the continuous long-term monitoring of structures. However, the displacements estimated from acceleration measurements do not include the critical static and pseudo static components of structural displacement [8,9].

In addition to the abovementioned contact-type displacement measurement techniques, a range of noncontact displacement measurement techniques have been developed based on laser Doppler vibrometers (LDV) [10], radar systems [11,12], and vision cameras [13,14]. LDV and radar systems emit laser light and electromagnetic waves, respectively, toward the surface of a structure and subsequently receive reflected signals. These systems can accurately determine the displacement of a structure by measuring the time delay between the emission and reception of a signal. Although both LDV and radar systems can achieve high-precision measurements, their high cost limits their widespread use. On the other hand, vision cameras capture images of the structure and determine structural displacement by tracking changes in the structure position in these images. Although vision cameras are relatively inexpensive, they are sensitive to environmental conditions such as light and weather, and less accurate and efficient than LDV and radar systems. In addition, all sensors should be fixed at a stationary location, which makes them unsuitable for continuous long-term displacement monitoring rather than short-term measurements.

In recent years, combinations of different types of sensors have become increasingly popular for estimating the structural displacements [15,16]. Such a combination can provide complementary information and help improve the accuracy and efficiency of structural displacement estimation. Accelerometers are commonly fused with other types of sensors [17–21]. The authors previously explored the fusion of the vision camera and accelerometer for structural displacement estimation [22]. The vision camera and accelerometer were installed at a target structure, with the accelerometer measuring the structural acceleration at a high sampling rate, whereas the vision camera tracked a fixed target for the surroundings of the structure at a low sampling rate. Because these two sensors are installed at the same location, their data can be easily fused, resulting in highly accurate and highly efficient displacement estimation at a high sampling rate. In addition, the requirement for a stationary location was eliminated by the direct installation of these two sensors on the target structure, making these techniques more appropriate for long-term continuous displacement estimation. Nevertheless, vision cameras are incapable of working at night, which limits the practical application of the proposed technique to long-term continuous structural displacement monitoring.

In this study, a structural displacement estimation technique was developed by fusing accelerometers with vision and infrared (IR) cameras, particularly for long-term continuous displacement monitoring. Three sensors were installed at the displacement estimation point of the target structure, and their initial short-period measurements were first used to automatically estimate the scale factors required for unit conversion and to optimize the temperature range of the selected region of interest (ROI) for the IR camera. The proposed technique then continuously estimates the structural displacements. Specifically, it combines a vision camera and accelerometer to estimate the structural displacement during the day and an IR camera and accelerometer to estimate the structural displacement during the night. In addition, an adaptive reference frame updating algorithm was proposed and applied to enhance the robustness of the proposed technique against variations in illumination in the vision camera and temperature in the IR camera. The main contributions of this study are (1) day and night continuous displacement estimation by fusing the accelerometer, IR, and vision cameras; (2) automated optimization of the ROI temperature range of the IR camera for displacement estimation during the night; and (3) improved robustness of vision-based displacement estimation against variations in illumination in the vision camera and temperature in the IR camera by adaptively updating the reference frame.

The remainder of this paper is organized as follows: The proposed continuous displacement estimation technique is described in Section 2. The performance of the

proposed technique was experimentally validated using an indoor single-story building model test and outdoor pedestrian bridge test, as described in Section 3. Lastly, the concluding remarks are presented in Section 4.

2. Development of Structural Displacement Estimation Technique by Fusing Accelerometer, Vision, and IR Cameras

This study proposes a continuous displacement estimation method in which acceleration measurements are combined with collected vision and IR images for day and night, respectively. The accelerometer, vision, and IR cameras were mounted at the measurement point of the target structure, and the displacement was estimated, as shown in Figure 1a. The accelerometer measured the acceleration of the target structure at a high sampling rate. However, assuming that the natural targets in the surroundings of the target structures are stationary, the vision camera and IR camera track a natural target with rich features during the day and a natural target with a distinct temperature distribution during the night, both with a low sampling rate. A low sampling displacement was first estimated from the vision/IR images with adaptive reference frame updating, and the image-based displacement was then fused with the high-sampling acceleration using an adaptive multirate Kalman filter. Considering that the movements originally estimated from vision/IR images are in pixel units, the scale factors required to convert these pixel unit movements into structural displacements in a physical unit should be estimated in advance. Additionally, the displacement estimation performance of IR cameras depends on the selection of the temperature range, and it is necessary to optimize the temperature range for better displacement estimation performance. Therefore, the proposed technique is divided into two stages, as shown in Figure 1b: (1) automatic initial calibration for scale factor estimation and temperature range optimization (Section 2.1) and (2) continuous displacement estimation (Section 2.2).

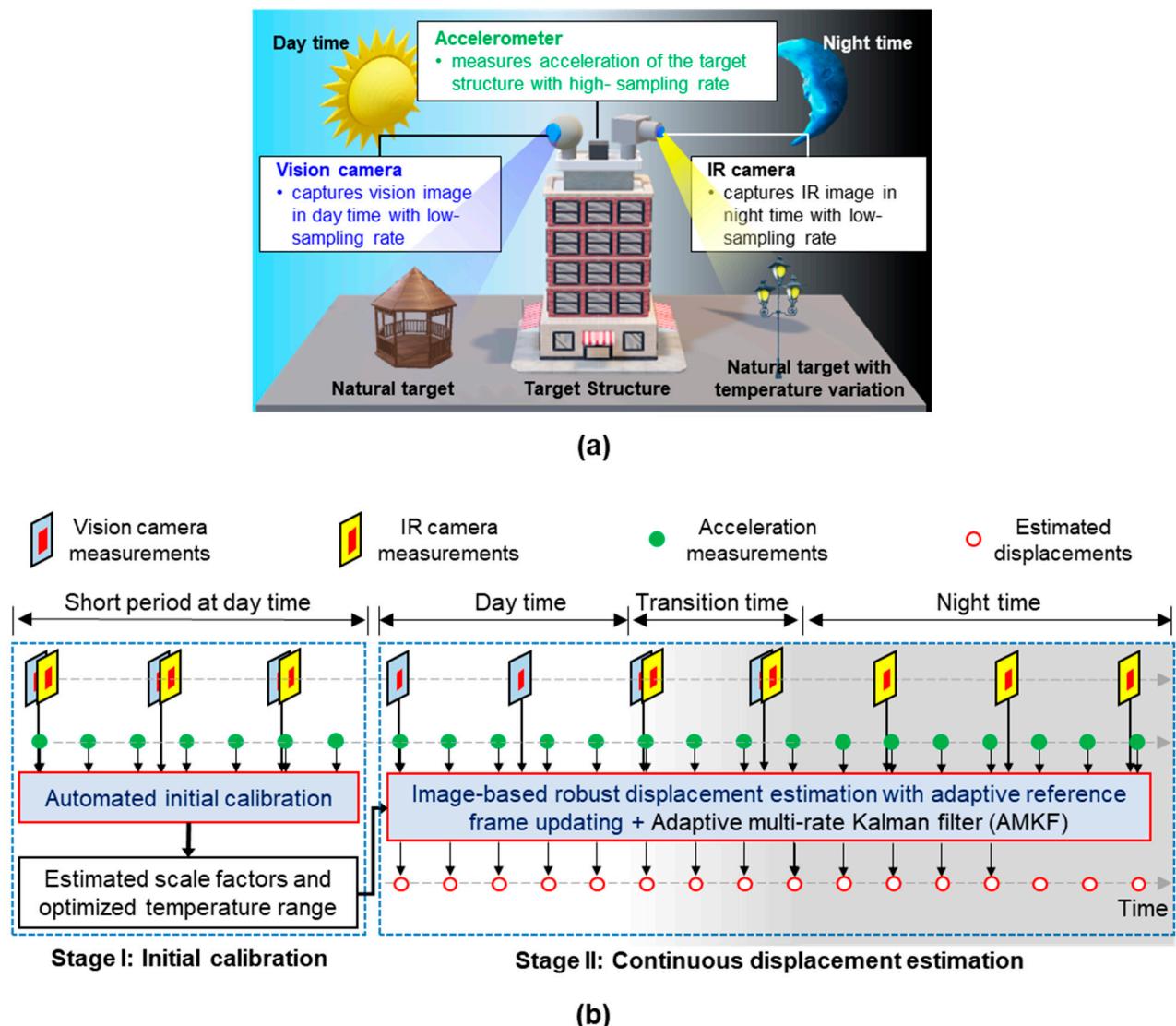


Figure 1. Overview of proposed displacement estimation technique: (a) sensor setup and (b) overall flowchart for continuous displacement estimation.

2.1. Stage I: Automated Initial Calibration

2.1.1. Scale Factor Estimation for Vision and IR Cameras

In this study, an acceleration-aided algorithm [22] was adopted to automatically estimate the scale factors required for image-based displacement estimation. As shown in Figure 2, translation d was first estimated from the collected short-term vision/IR images after ROI cropping and feature matching. In this study, the speeded-up robust features (SURF) [23] algorithm was used owing to its high accuracy and low computational cost. Subsequently, a bandpass filter was applied to d and the displacement u_a was estimated from the double integration of the acceleration measurement. The lower cutoff frequency of the filter was set to be sufficiently high to remove the low-frequency drift in u_a , and the upper cutoff frequency was set to 1/10 of the vision and IR camera sampling rate [24]. Finally, the scale factor α was estimated as the ratio of filtered translation d^f and filtered displacement u_a^f using a least-squares estimation (LSE) algorithm. Before applying the LSE algorithm, u_a^f was down-sampled to match the sampling rate with d^f .

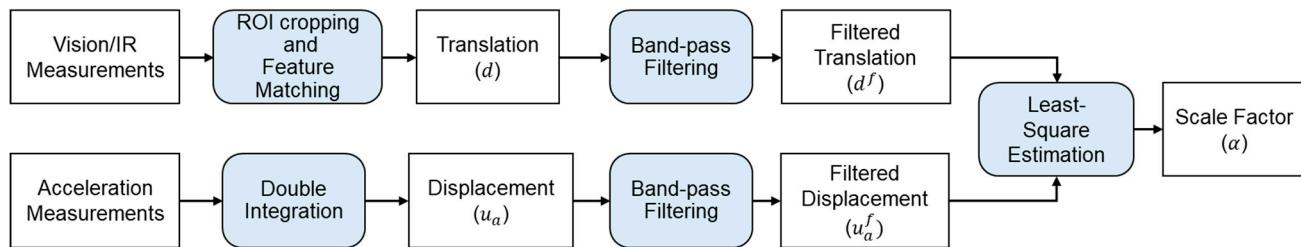


Figure 2. Flowchart of automated scale factor estimation algorithm [22].

2.1.2. Optimization of the Temperature Range for IR Camera

(a) Necessity of fixing and optimizing the temperature range

An IR image is essentially a temperature map, in which different colors represent different temperatures. If the target within the ROI has a stable and distinct temperature distribution, the IR-based displacement can be estimated to be the same as the vision-based displacement by applying a feature-matching algorithm between the reference and current ROIs. However, the difference from vision-based displacement estimation is that only temperature data are contained in an IR image. Therefore, when there is an external extreme heat source in the ROI, the color distribution within the ROI changes, as shown in Figure 3a, causing matching failure or no matching between the reference and the current ROIs. To reduce the above problem, the temperature range of the IR camera was fixed.

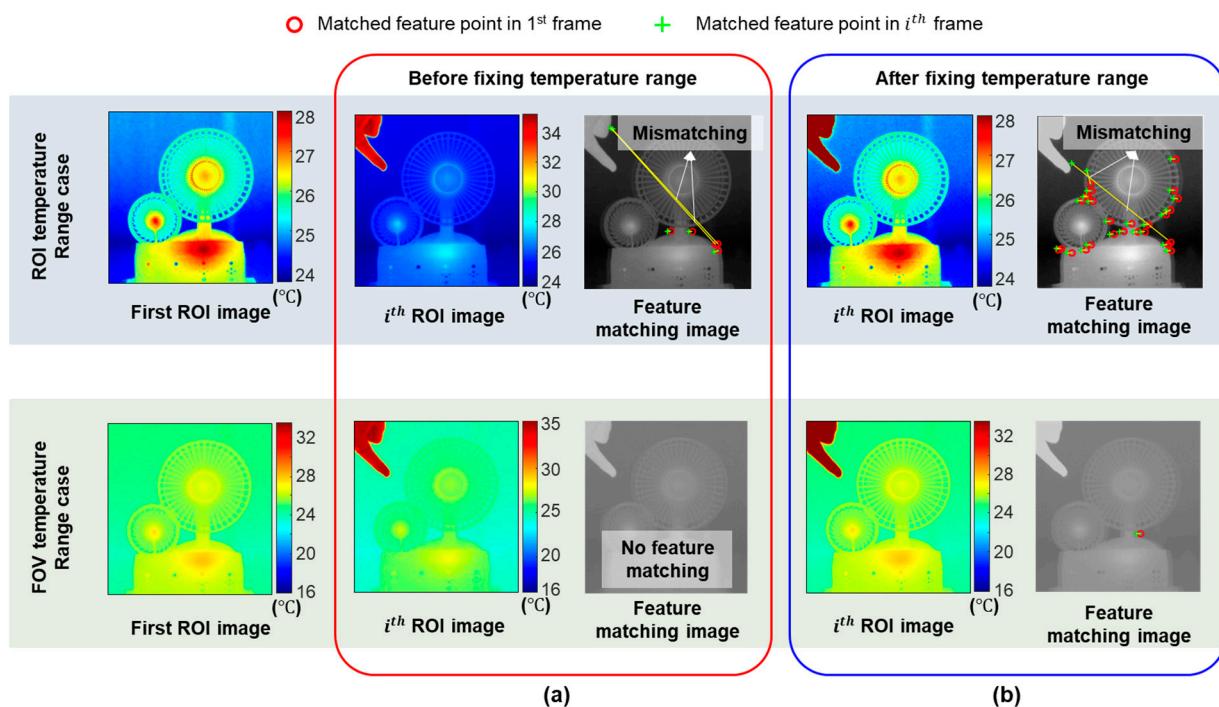


Figure 3. Example of region of interest (ROI) feature-matching results with heat source: (a) before fixing temperature range and (b) after fixing temperature range.

Figure 3b shows that the temperature range can be fixed using the maximum and minimum temperatures with the field of view (FOV) (T_{max}^F and T_{min}^F); however, relatively small temperature variations within the ROI cause less distinct features. On the other hand, the temperature range can also be fixed using the maximum and minimum temperatures with the ROI (T_{max}^R and T_{min}^R). Although more distinct features can be detected, they are not stable owing to the temperature measurement noise. Therefore, it is necessary to optimize the temperature range to ensure that sufficient and stable distinct features are available within the ROI.

(b) Working principle of automated optimization of the temperature range

The basic idea for temperature range optimization is to calculate the root-mean-square errors (RMSEs) between the acceleration- and IR-based displacements in various temperature ranges. The optimal temperature range was selected when the RMSE was the smallest. The detailed process of the proposed algorithm consists of three steps:

Step 1: First, the differences between T_{min}^R and T_{min}^F and between T_{max}^R and T_{max}^F were calculated and divided into M equal parts, as follows:

$$l = \frac{T_{min}^R - T_{min}^F}{M}, \quad h = \frac{T_{max}^F - T_{max}^R}{M} \quad (1)$$

where M is a constant value to divide the temperature difference between FOV and ROI into $M+1$ equal parts. Then, $(M+1)^2$ potential temperature ranges (δ) are generated as follows (Figure 4a):

$$\begin{aligned} \delta(m, n) &= [T_{min}^R - l(m-1), T_{max}^F + h(n-1)], \\ m &= (1, \dots, M+1), n = (1, \dots, M+1). \end{aligned} \quad (2)$$

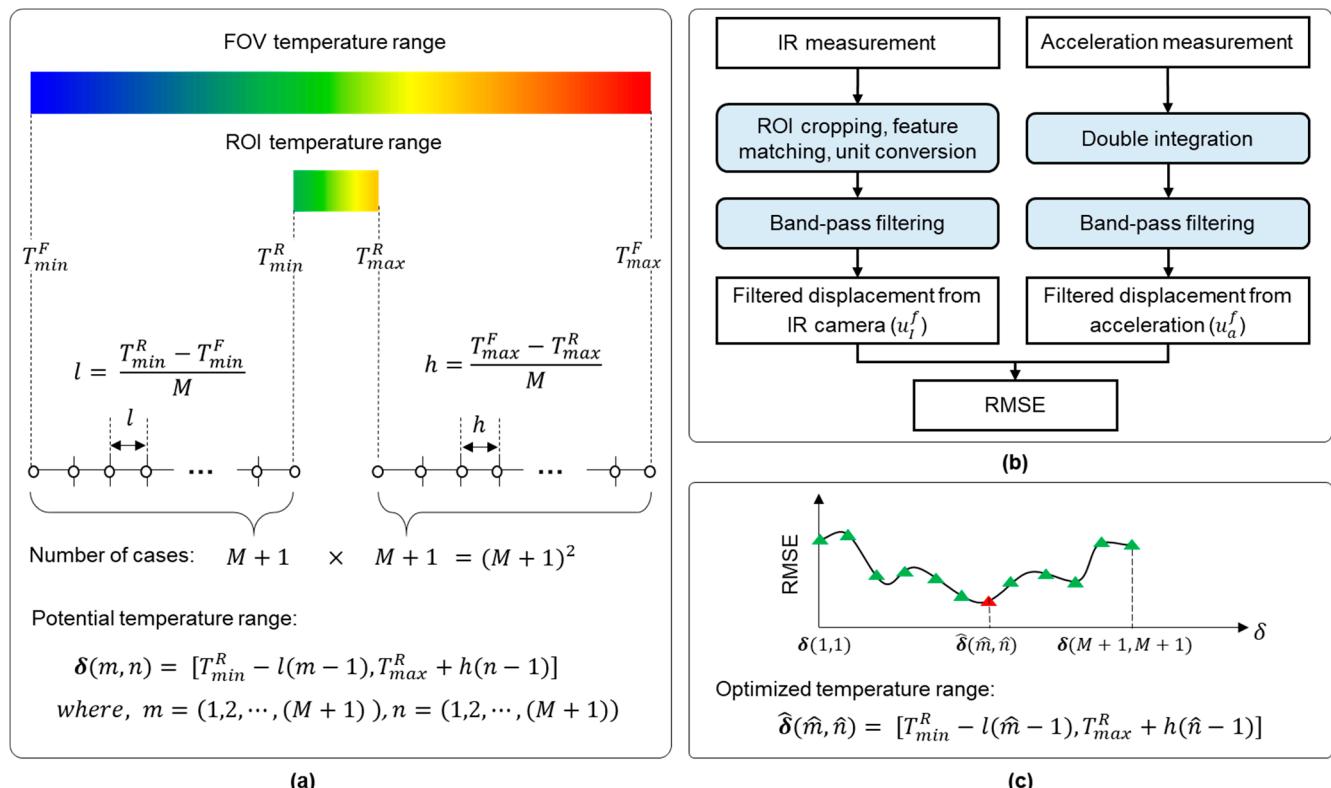


Figure 4. Automated optimization of the temperature range: (a) step 1: define potential temperature ranges (δ) with variables m and n , (b) step 2: calculate root-mean-square error (RMSE) using initial $\delta(1,1)$, and (c) step 3: select the optimized temperature range by repeating step 2 with different δ .

Step 2: The displacement was first estimated from the IR measurement using the first temperature range ($\delta(1,1)$) and the estimated scale factor and then bandpass-filtered to obtain the filtered IR-based displacement (u_I^f). The filtered displacement was estimated from the acceleration measurements using double integration and bandpass filtering. Subsequently, the RMSE between the filtered acceleration and IR-based displacements was calculated (Figure 4b).

Step 3: Step 2 was repeated for all $(M+1)^2$ potential temperature ranges. The temperature range with the smallest RMSE became the optimized temperature range ($\hat{\delta}(m, n)$). After that, $\hat{\delta}(m, n)$ was applied to IR-based displacement estimation in Stage II.

2.2. Stage II: Continuous Displacement Estimation Using Image-Based Robust Displacement Estimation Algorithm and Adaptive Multirate Kalman Filter (AMKF)

After the initial calibration, the displacement was continuously estimated by fusing the vision/IR-based displacement with asynchronous acceleration measurements using an adaptive multirate Kalman filter (AMKF) [22] developed by our group. The transition between vision and IR cameras is automatically achieved, and the reference frame is adaptively updated to improve the robustness of vision-based displacement estimation against illumination variations and IR-based displacement estimation against temperature variations, which are unavoidable in long-term continuous displacement estimation.

2.2.1. AMKF-Based Fusion of Asynchronous Image and Acceleration Measurement

Asynchronous accelerations and images were fused using the AMKF, which was formulated for three different time-step types, as shown in Figure 5. In a type-I time step, only acceleration is used, and the state vector (\hat{x}_z^+), which consists of displacement and velocity, is estimated using the previous time-step state vector (\hat{x}_{z-1}^+) and acceleration (a_{z-1}).

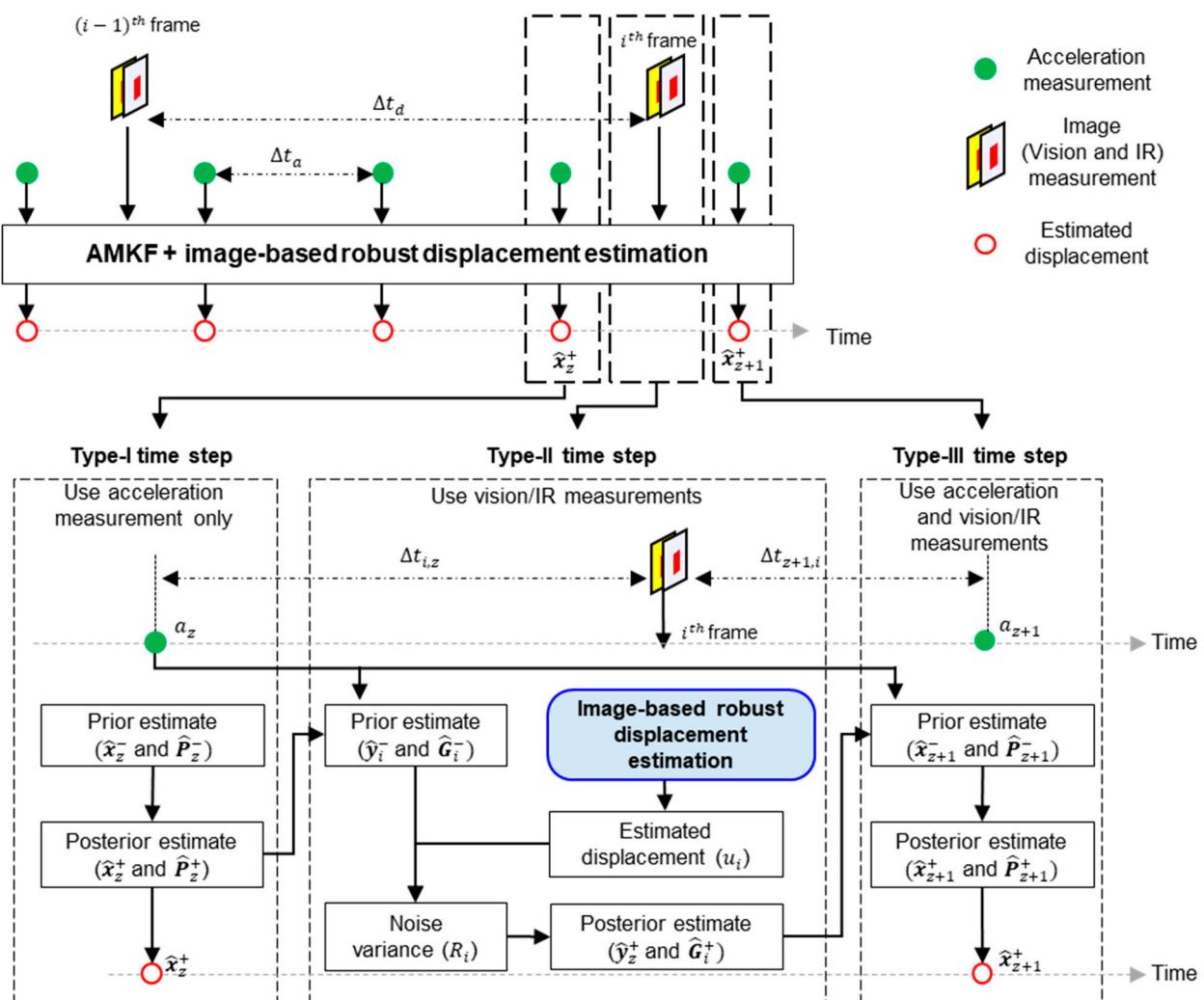


Figure 5. Overview of AMKF-based structural displacement estimation using accelerometer, vision, and infrared (IR) cameras.

$$\hat{x}_z^+ = \hat{x}_z^- = \mathbf{A}(\Delta t_a) \hat{x}_{z-1}^+ + \mathbf{B}(\Delta t_a) a_{z-1}; \mathbf{A}(\Delta t_a) = \begin{bmatrix} 1 & \Delta t_a \\ 0 & 1 \end{bmatrix}; \mathbf{B}(\Delta t_a) = \begin{bmatrix} \frac{\Delta t_a^2}{2} \\ \Delta t_a \end{bmatrix}, \quad (3)$$

where Δt_a denotes the time interval between the acceleration measurements. Next, the covariance ($\widehat{\mathbf{P}}_z^+$) of $\widehat{\mathbf{x}}_z^+$ is calculated as follows:

$$\widehat{\mathbf{P}}_z^+ = \widehat{\mathbf{P}}_z^- = \mathbf{A}(\Delta t_a) \widehat{\mathbf{P}}_{z-1}^+ \mathbf{A}^T(\Delta t_a) + q \mathbf{B}(\Delta t_a) \mathbf{B}^T(\Delta t_a), \quad (4)$$

where q denotes the noise variance in the acceleration measurements.

In a type-II time step, the prior state ($\widehat{\mathbf{y}}_i^-$) and its covariance ($\widehat{\mathbf{G}}_i^-$) are estimated as follows:

$$\begin{aligned} \widehat{\mathbf{y}}_i^- &= \mathbf{A}(\Delta t_{i,z}) \widehat{\mathbf{x}}_z^+ + \mathbf{B}(\Delta t_{i,z}) a_z, \\ \widehat{\mathbf{G}}_i^- &= \mathbf{A}(\Delta t_{i,z}) \widehat{\mathbf{P}}_z^+ \mathbf{A}^T(\Delta t_{i,z}) + q \mathbf{B}(\Delta t_{i,z}) \mathbf{B}^T(\Delta t_{i,z}); \Delta t_{i,z} = i \Delta t_d - z \Delta t_a, \end{aligned} \quad (5)$$

where Δt_d denotes the time interval between image measurements. Subsequently, the noise variance (R_i) in the i th image-based displacement (u_i) is expressed as follows [25]. The detailed process for estimating u_i is described in Section 2.2.2.

$$\begin{aligned} R_i &= \beta R_{i-1} + (1 - \beta)(\eta_i^2 - \mathbf{H} \widehat{\mathbf{G}}_i^- \mathbf{H}^T), 0 < \beta < 1, \\ \eta_i &= u_i - \mathbf{H} \widehat{\mathbf{y}}_i^-, \mathbf{H} = [1 \ 0]^T, \end{aligned} \quad (6)$$

where β and η denote the forgetting and innovation factors, respectively. The Kalman gain (\mathbf{K}) is calculated as

$$\mathbf{K} = \widehat{\mathbf{G}}_i^- \mathbf{H}^T (\mathbf{H} \widehat{\mathbf{G}}_i^- \mathbf{H}^T + R_i)^{-1}. \quad (7)$$

Finally, $\widehat{\mathbf{y}}_i^-$ and $\widehat{\mathbf{G}}_i^-$ are updated in a posterior process using K and u_i as follows:

$$\begin{aligned} \widehat{\mathbf{y}}_i^+ &= (\mathbf{I} - \mathbf{K} \mathbf{H}) \widehat{\mathbf{y}}_i^- + \mathbf{K} u_i, \\ \widehat{\mathbf{G}}_i^+ &= (\mathbf{I} - \mathbf{K} \mathbf{H}) \widehat{\mathbf{G}}_i^-. \end{aligned} \quad (8)$$

In a type-III time step, the state vector ($\widehat{\mathbf{x}}_{z+1}^+$) is estimated at the next acceleration time step (Δt_{z+1}). Details of the AMKF can be found in a study by Ma et al. [22]

$$\widehat{\mathbf{x}}_{z+1}^+ = \widehat{\mathbf{x}}_{z+1}^- = \mathbf{A}(\Delta t_{z+1,i}) \widehat{\mathbf{y}}_i^+ + \mathbf{B}(\Delta t_{z+1,i}) a_z; \Delta t_{z+1,i} = (z+1) \Delta t_a - i \Delta t_d \quad (9)$$

2.2.2. Image-Based Robust Displacement Estimation with Adaptive Reference Frame Updating

- (a) A brief review of the existing algorithm with a fixed reference frame and its limitations

When estimating displacement from visual measurements using a feature-matching algorithm, existing studies [22] set the first ROI as the reference ROI, and the displacement at each time step was estimated by matching the current and reference ROIs. Therefore, stable illumination conditions are required for a successful displacement estimation. This is not a problem for short-period displacement estimation, which was the focus of these existing studies. However, illumination variation is unavoidable in long-term continuous displacement estimation and may cause insufficient matches, as shown in Figure 6a, making continuous displacement estimation impossible. The IR-based displacement estimation suffers from the same issue. Unavoidable temperature variations may result in an insufficient match, as shown in Figure 6b. Therefore, an algorithm that improves the robustness of vision-based displacement estimation against illumination variations and IR-based displacement estimation against temperature variations is essential for a long-term continuous displacement estimation.

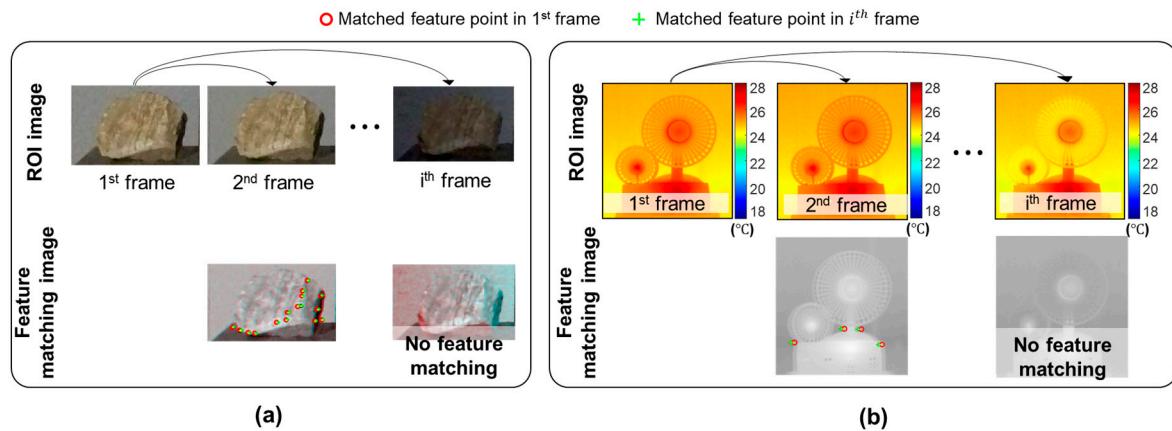


Figure 6. Overview of existing image-based displacement estimation algorithm [22] using a fixed reference frame and its limitation in long-term continuous displacement estimation: (a) vision measurement with illumination variation and (b) IR measurement with temperature variation.

(b) Working principle of adaptive reference frame updating

This study proposes an adaptive reference frame updating algorithm to improve the robustness of vision-based displacement estimation against illumination variations and IR-based displacement estimation against temperature variations. The basic principle of the proposed algorithm is to adaptively update the reference frame when the detected matches are insufficient, and update the reference frame back to the first frame if sufficient matches can be detected.

Figure 7 shows the working principle of the proposed algorithm. First, after obtaining the i^{th} image from the vision/IR cameras, the ROI was cropped from the FOV. Feature matching was then performed between the i^{th} ROI and the current reference ROI (i.e., the r^{th} ROI) and N_i^r matches were obtained. Note that the initial reference ROI is the first ROI. Owing to the relatively large variation in the number of matches, even with stable illumination, a moving average filter with an order of $(Q + 1)$ was applied to N_i^r to obtain an average value $(\bar{N}_{i-Q,i}^r)$.

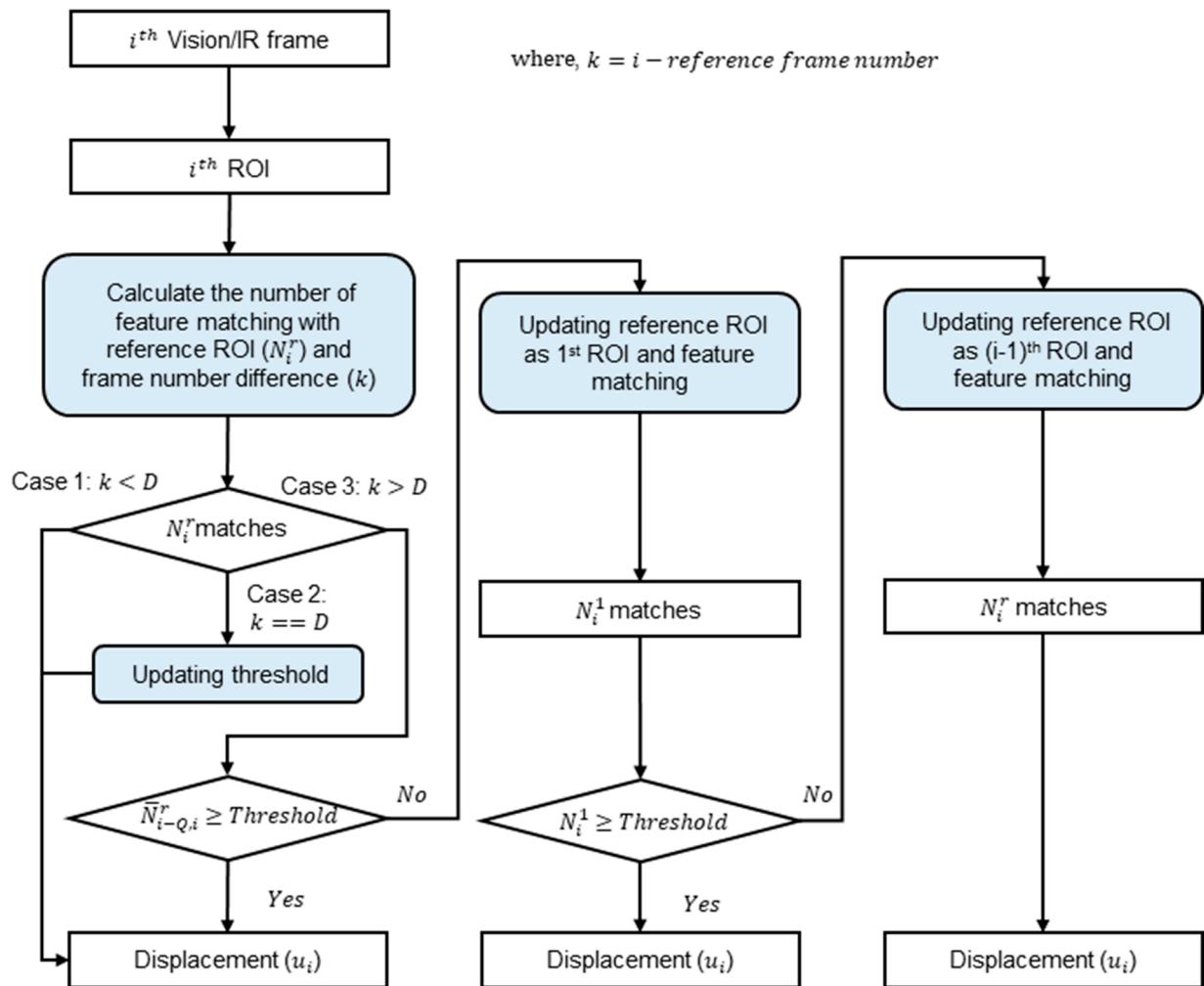


Figure 7. Flowchart of proposed displacement estimation algorithm for estimating displacement from vision/IR images with adaptive reference ROI updating.

$$\bar{N}_{l-Q,i}^r = \left(\frac{1}{Q+1} \sum_{j=i-Q}^i N_j^r \right) \quad (10)$$

If $\bar{N}_{i-Q,i}^r$ is larger than the threshold (N_T^r), it is not necessary to update the reference frame, and the i th image-based displacement (u_i^1) is calculated as

$$u_i^1 = \alpha d_i^r + u_r^1, \quad (11)$$

where d_i^r is the relative translation between the i th and reference ROIs and u_r^1 is the relative displacement between the 1st and reference ROIs. α denotes the scale factor for vision/IR measurements. Note that N_T was determined as follows:

$$N_T^r = \left(\frac{1}{D} \sum_{j=r+1}^{r+D} N_j^r \right) - 3\sigma [N_j^r | j = (r+1), \dots, (r+D)]. \quad (12)$$

Otherwise, feature matching was performed between the 1st and i th ROIs, and N_i^1 matches were obtained. If N_i^1 is larger than a threshold (N_T^1), the reference frame is updated to the first frame, and u_i^1 is estimated as

$$u_i^1 = \alpha d_i^1, \quad (13)$$

where d_i^1 is the relative translation between the i th and 1st ROIs, and N_T^1 is determined as

$$N_T^1 = \left(\frac{1}{D} \sum_{j=2}^{D+1} N_j^1 \right) - 3\sigma[N_j^1 | j = 2, \dots, (D + 1)]. \quad (14)$$

Denoting k as the difference in the number of frames between the i th frame and the reference frame, reference frame updating is executed differently in the three cases by comparing the rules of k and D as follows:

Case 1: k is less than D . The reference frame changes only within the prior D timesteps. The i th ROI and reference ROI are directly matched, and i th image-based displacement (u_i) is calculated using Equation (11).

Case 2: $k = D$. Here, the threshold value is calculated by applying Equation (12) using D frames sequentially after the reference frame. The subsequent process was the same as that for Case 1.

Case 3: $k > D$. If $\bar{N}_{i-D,i}^r$ is larger than the threshold, u_i can be estimated using (11). However, if $\bar{N}_{i-D,i}^r$ is smaller, then the reference frame is updated to the first frame. The number of matches (N_i^1) compared to the first frame was compared to the threshold value (N_T^1). If N_i^1 was larger than N_T^1 , u_i was calculated using Equation (13); however, if N_i^1 was less than N_T^1 , the reference frame was updated to the $(i - 1)$ th frame. Finally, u_i was estimated using (11) with the updated reference frame.

Figure 8 shows an example of the threshold updating process. The threshold value (N_T^1) was initially determined using D frames after the first frame. In the i th frame, the moving-averaged feature-matching number ($\bar{N}_{i-Q-1,i-1}^1$) is less than N_T^1 . Therefore, the reference frame is updated to the $(i - 1)$ th frame and the threshold is updated (N_T^r) using D frames after the $(i - 1)$ th frame.

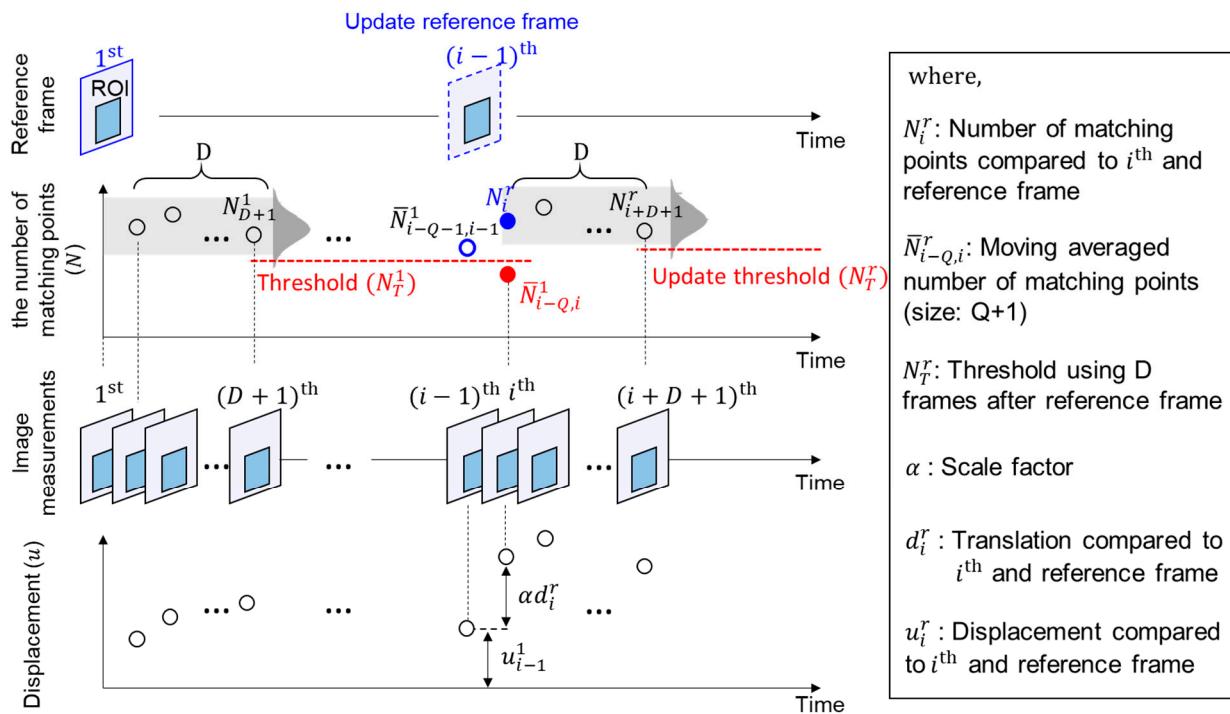


Figure 8. Example of reference frame and threshold updating process.

The proposed algorithm was applied separately to vision and IR images. A vision camera was used during the day and an IR camera was used during the night. However, vision and IR images were acquired simultaneously during the transition time (e.g., day-to-night or night-to-day) and the numbers of matches were calculated respectively. Note that the transition time can be set to approximately 1 h before and after the beginning of

the morning nautical twilight (BMNT) and the end of evening nautical twilight (EENT). Therefore, at the transition time, the robust number of features matching from the vision and IR images can be calculated as \bar{N}_V and \bar{N}_I , respectively, and then compared. Displacement was estimated by selecting an application (vision or IR) with a large matching number.

3. Experimental Validation

To validate the performance of the proposed technique, a laboratory-scale test was conducted on a single-story building model with various excitation signals considering illumination and temperature variations, as described in Section 3.1, and a field experiment was conducted on a pedestrian bridge. Note that long-term measurement was impossible for the pedestrian bridge owing to safety issues, and it was difficult to consider the illumination and temperature variations in the field test. Therefore, only the applicability of the IR camera for displacement estimation under field conditions was verified in Section 3.2 by estimating the bridge displacement for a short period under stable illumination and temperature. In both tests, the displacement estimation performance of the proposed technique was compared to that of an existing technique [22] to highlight its superiority.

3.1. Lab-Scale Test Using Single-Story Building Model Test

The proposed technique was first validated using a single-story building model test, the setup of which is illustrated in Figure 9a. A single-story building model composed of stainless steel was firmly attached onto an Electro-Seis APS 400 vibration shaker, which produced a horizontal movement of the building model. A Kinemetric EpiSensor ES-U2 uniaxial force-balanced accelerometer, Insta360 Pro2 camera, and FLIR A655sc IR cameras were mounted on the top of the building model. A PSV-400-M4 LDV was used to measure the reference displacement of the building model with a resolution of 0.5 pm (Figure 9b). The sampling frequency for the vision and IR cameras was set to 30 Hz, and the acceleration and LDV measurements were sampled at 100 Hz. For ideal displacement estimation, a sufficient number of feature points should be detected within the ROI and the translations of all these features should be close to each other. In this study, a stone and two fans placed approximately 2 m from the building model were used as targets for the vision and IR cameras, respectively, as shown in Figure 9b,c. A cup of hot water and a cup of cold water were placed approximately 3 m away from the targets to simulate objects with high and low temperatures, and were included in the FOV of the IR camera, as shown in Figure 9d. Seven cases were considered to validate the proposed technique fully, as listed in Table 1. Note that the spatial resolutions of the vision and IR camera were 2880 by 3840 and 640 by 240, respectively, and all recorded images by the fisheye camera (insta360 pro 2) were calibrated by MATLAB built-in toolbox [26] for distortion correction.

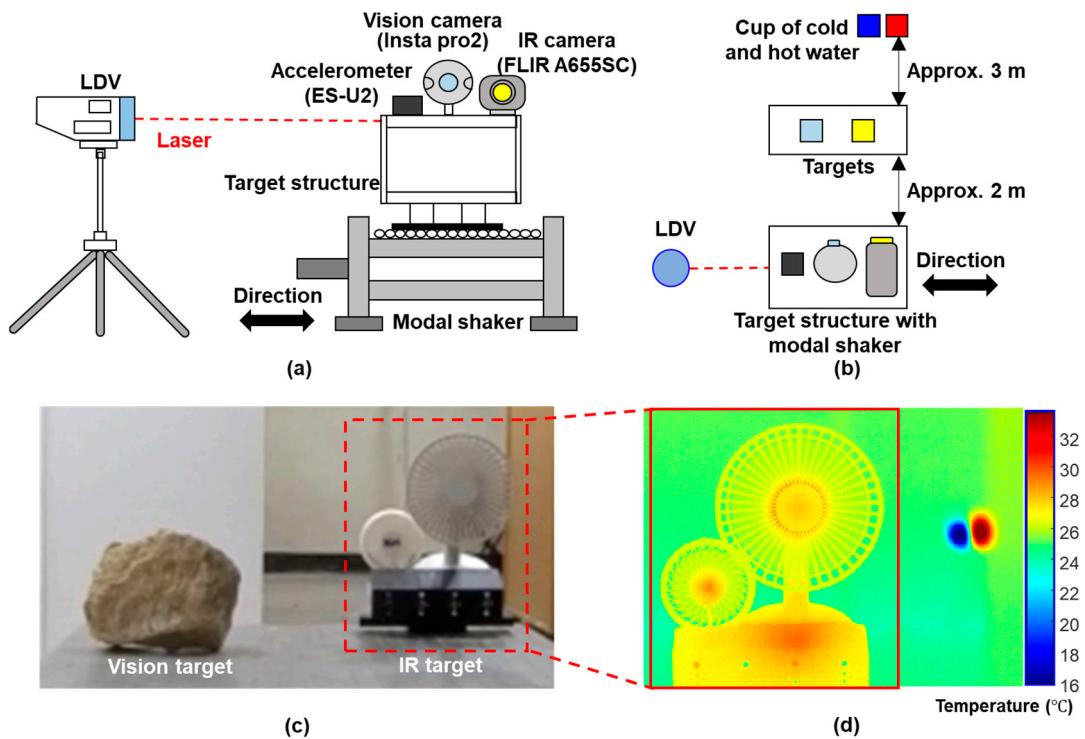


Figure 9. Overview of lab-scale test on a single-story building model: (a) front view of experiment setup, (b) top view of the experiment setup, (c) targets for vision and IR cameras, and (d) cropped IR image.

Table 1. Descriptions of seven lab-scale test cases.

# of Cases	Test Duration (s)	Illumination Variation	Temperature Variation	Purposes
1	30	No	No	Initial calibration
2	30	No	Yes	Robustness of the proposed technique to external heat source
3				
4	30	No	No	Performance of displacement estimation using optimized temperature range
5				
6	210	Yes (Extreme)	No	Robustness of the proposed technique to extreme variations in illumination
7	120	Yes	Yes	Application selection of transition time for continuous displacement estimation

3.1.1. Initial Calibration Results (Case 1)

The scale factors for vision and IR measurements were estimated using the algorithm proposed in Section 2.1.1, exciting the building model with a 1 Hz sinusoidal signal. A bandpass filter was used before estimating the ratio between the acceleration-based displacement and the camera-based translation. The lower and upper cutoff frequencies of the bandpass filter were set to 0.3 Hz and 3 Hz, respectively, considering the effective frequency range of the accelerometer and the sampling rate of vision and IR cameras [24]. Figure 10 shows the estimated scale factors (α_v and α_l) for the vision and IR cameras, which are 0.945 mm/pixel and 1.094 mm/pixel, respectively.

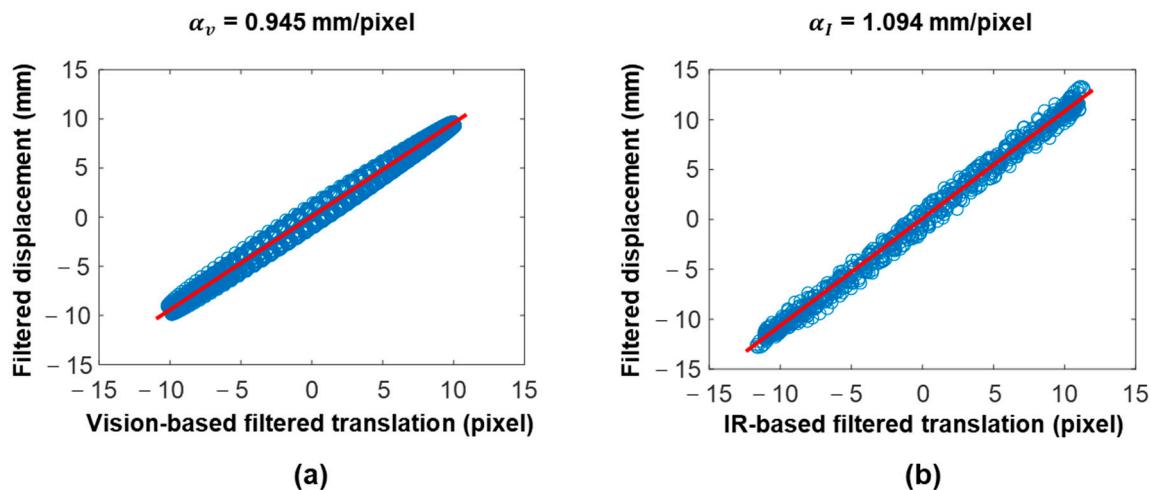


Figure 10. Scale factor estimations in the lab-scale test: (a) vision camera and (b) IR camera.

Next, the temperature range was optimized for the IR measurements. Figure 11a shows the FOV of the IR camera and cropped ROI. The maximum and minimum temperatures were 33.52 °C and 15.94 °C, respectively, for the FOV due to the existence of two cups of water, while they were 28.08 °C and 23.65 °C, respectively, for the ROI. The differences between the maximum and minimum temperatures of FOV and ROI were equally divided into nine parts, and the values of l and h were set to 0.857 °C and 0.604 °C, respectively, as shown in Figure 11a. Potential temperature ranges were generated for different combinations of l and h , and the corresponding RMSEs were calculated, as shown in Figure 11b. Finally, the temperature range [16.79 °C, 28.08 °C] corresponding to the smallest RMSE was selected as the optimized temperature range.

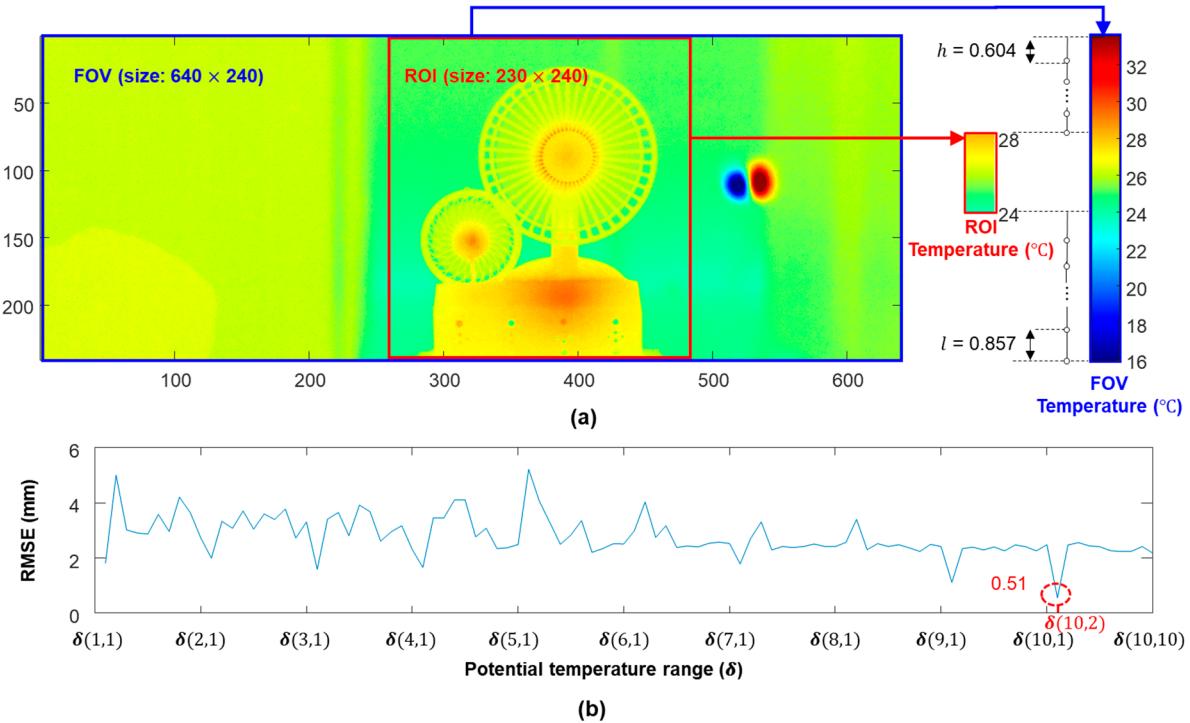


Figure 11. Temperature range optimization results: (a) temperature ranges of the ROI and field of view (FOV) and (b) RMSE under different temperature ranges.

3.1.2. Displacement Estimation Results

(a) IR-based displacement estimation using optimized temperature range (Cases 2–5)

The superiority of using the optimized temperature range was first verified when the building model was subjected to a 1 Hz sinusoidal signal excitation (Case 2). Note that the researcher's finger appeared in the ROI at 23.5 s and disappeared from the ROI at 27 s to simulate an external heat source that may appear in the ROI in practice. Figure 12a compares the displacements estimated using the FOV and ROI temperature ranges, and the temperature range optimized by the proposed algorithm. The best displacement estimation performance was obtained using the optimized temperature range, indicating that the performance of IR-based displacement is sensitive to the temperature range. Note that when using the FOV temperature range, displacement cannot be continuously estimated and then its RMSE cannot be calculated. Figure 12b compares the number of matched feature points using the FOV, ROI, and optimized temperature ranges. Using the FOV temperature range generated many more matched feature points than using the ROI temperature range; however, the number of matched feature points decreased in both cases with the appearance of a finger (e.g., from 26 to 27 s). Therefore, during this period, the displacement was estimated with extremely large errors using the FOV temperature range, whereas the displacement could not be estimated using the ROI temperature range. However, the external heat source had less of an effect on the number of matched feature points when using the optimized temperature range, and sufficient feature points were stably matched to ensure a reliable displacement estimation.

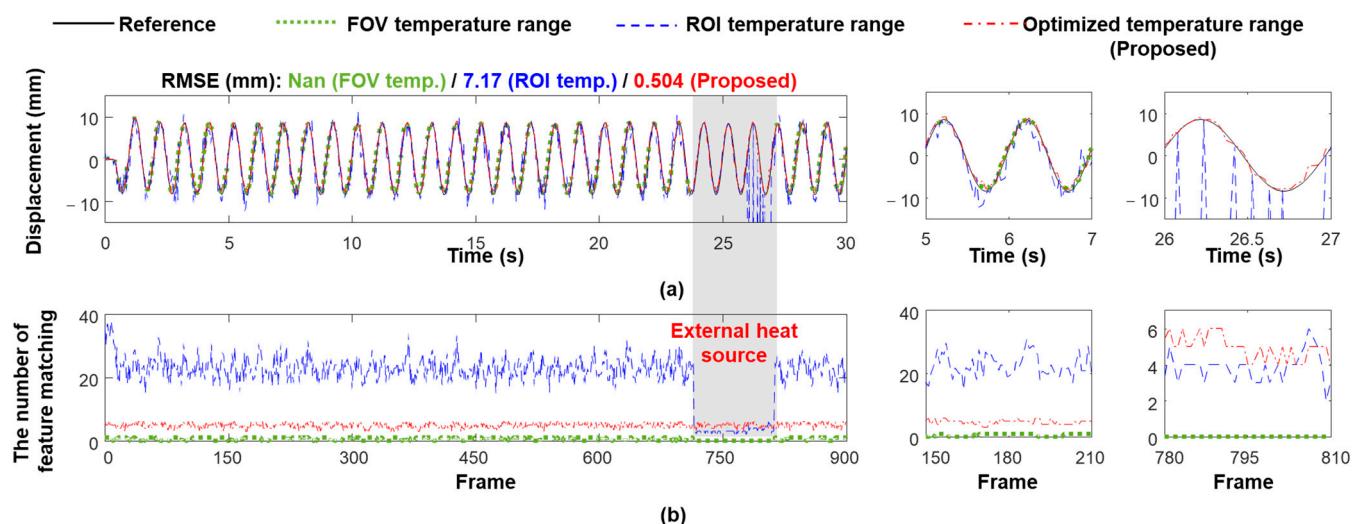


Figure 12. Comparison of (a) IR-based displacements and (b) number of matched feature points with an external heat source suddenly appearing in the ROI (Case 2).

Figure 13a compares the ROI images and corresponding matching results at the 340th frame without an external heat source (i.e., the finger). Only one feature point was matched when the FOV temperature range was used, because of the relatively small temperature variations within the ROI. Although a relatively large number of feature points (27) matched when the ROI temperature range was used, many were mismatched. Therefore, the displacement estimation accuracy was low in both cases. Figure 13b compares the ROI images and corresponding matching results at the 768th frame with the external heat source (i.e., the finger). No feature point was matched when the FOV temperature range was used, which caused failure in the continuous displacement estimation. The external heat source significantly reduced the number of matched feature points, and decreased the displacement estimation accuracy when using the ROI temperature range. However, the external heat source had less effect on the number of matched feature points

when using the optimized temperature range, and four sufficient feature points were stably matched in both the 340th and 768th frames.

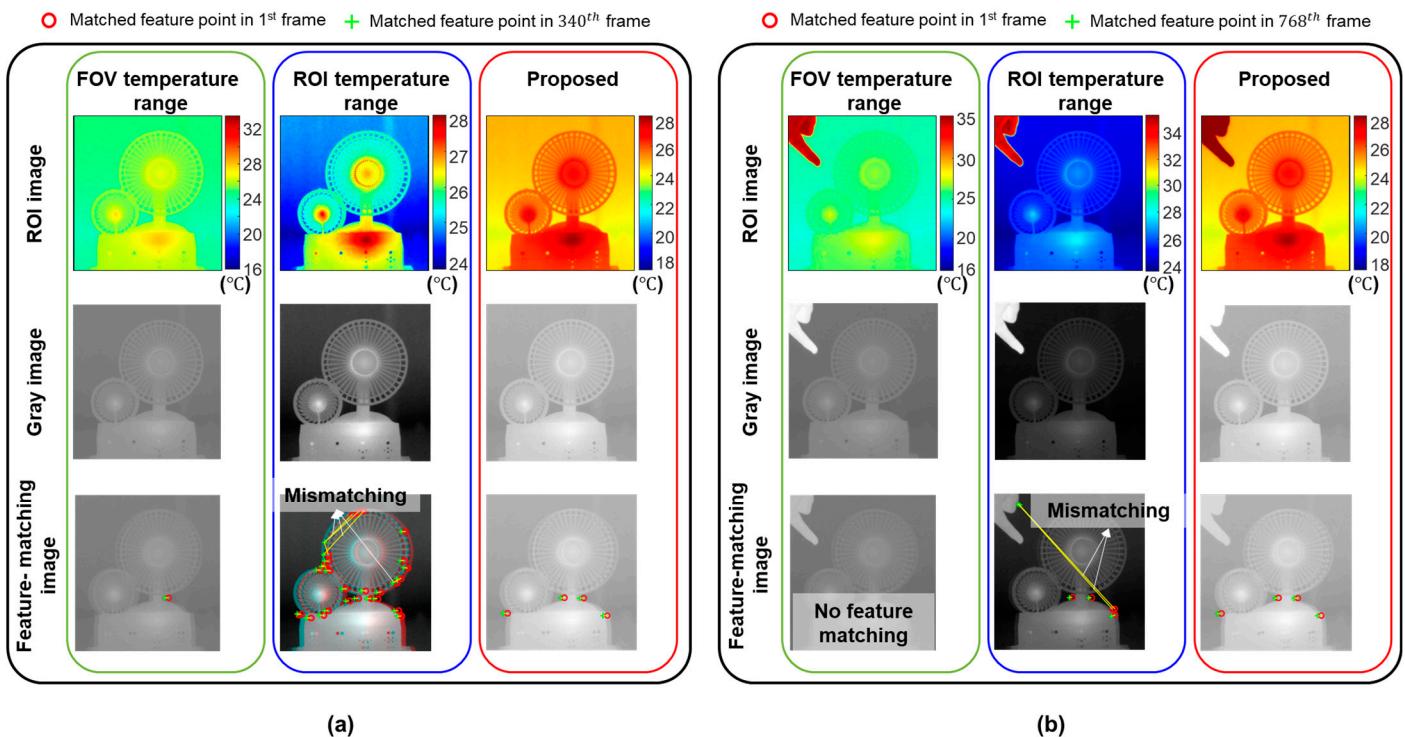


Figure 13. ROI, gray, and feature-matching images: (a) with and (b) without the external heat source (i.e., the finger).

The superiority of using the optimized temperature range was further validated under three additional excitation signals, that is, a 2 Hz sinusoidal signal, 0–3 Hz sweep signal, and a recorded real bridge vibration signal. The displacements estimated using the different temperature ranges are compared in Figure 14. Using the optimized temperature range reduced the average RMSEs by 72.57% compared to using the ROI temperature range, and displacements were estimated accurately with RMSEs below 0.8 mm. Note that displacements were not continuously estimated using the FOV temperature range. Therefore, the RMSEs could not be calculated.

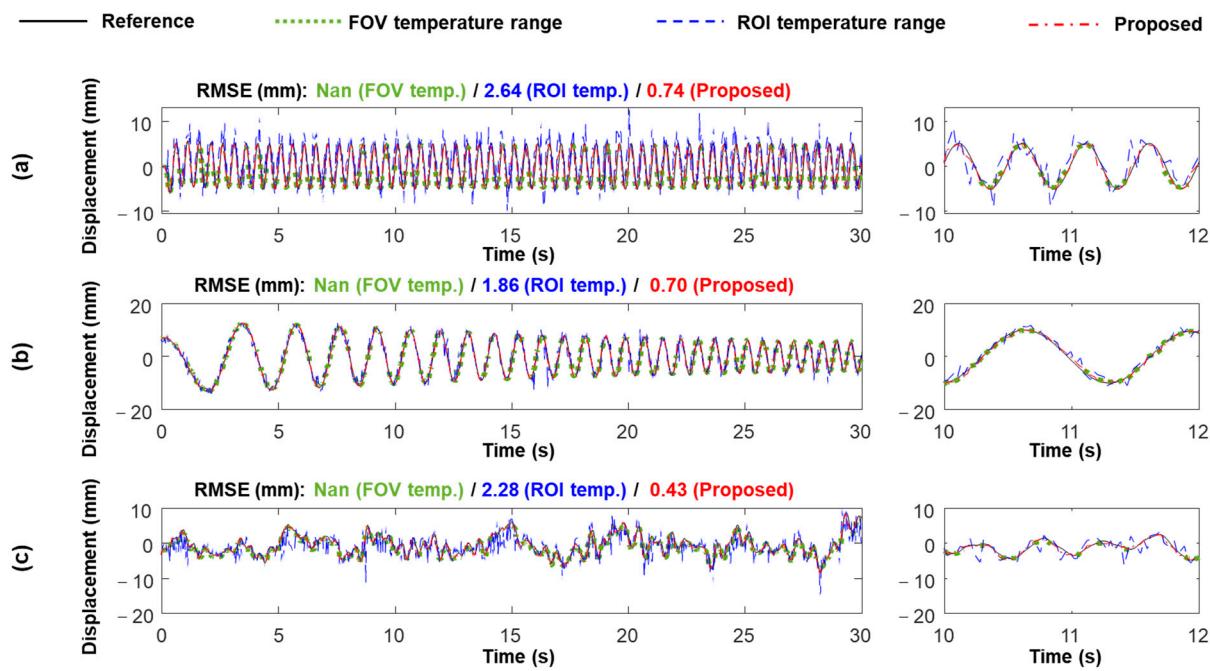


Figure 14. IR-based displacements estimated using FOV, ROI, and optimized temperature ranges under (a) 2 Hz sinusoidal (Case 3), (b) 0~3 Hz sweep (Case 4), and (c) recorded real bridge vibration signal (Case 5) inputs.

(b) Vision-based displacement estimation using adaptive reference frame updating (Case 6)

The superiority of the adaptive reference frame updating was verified under varying illumination conditions when the building model was simultaneously subjected to a 1 Hz sinusoidal signal and pseudo-static signal excitation (Case 6). A flashlight approximately 40 cm from the target of the vision camera was used as the light source and was moved, as shown in Figure 15, to simulate varying illumination conditions.

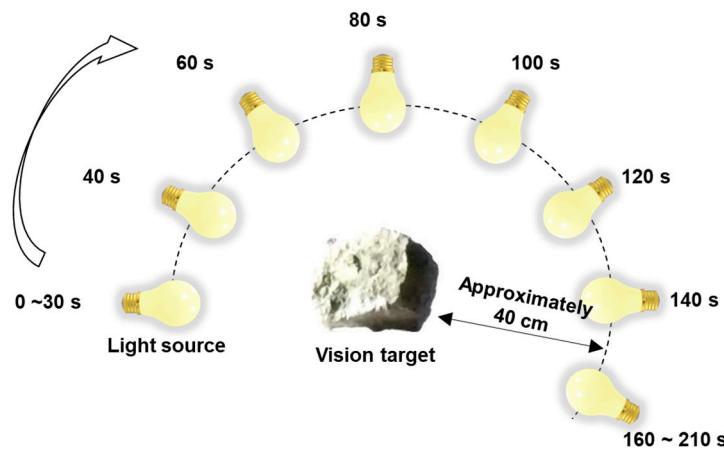


Figure 15. Simulation of varying illumination conditions for Case 6 of the lab-scale test using a moving light source (i.e., flashlight).

Figure 16a compares the captured target images at different times, and illumination variation is clearly observed. Figure 16b,c compare the number of matched feature points and estimated displacements using the existing and proposed techniques. The existing technique fixes the reference frame to the first frame, and enough feature points are matched in the first 30 s without illumination variations. Subsequently, the movement of

the light source causes illumination variations in the ROIs. Therefore, the number of matched feature points decreased significantly until it reached zero at 80 s (Figure 16b), and the displacements were only estimated at 80 s (Figure 16c). However, the proposed technique adaptively updated the reference frame, and matched sufficient feature points under the dramatical illumination variation to continuously estimate displacement for 210 s with 0.89 mm RMSE (Figure 16b,c). However, the reference frame was updated 28 times using the proposed technique (Figure 16d), where dramatically varying illumination conditions were considered and the reference frame was updated frequently. However, considering slowly varying illumination under field conditions, less frequent frame updates are required.

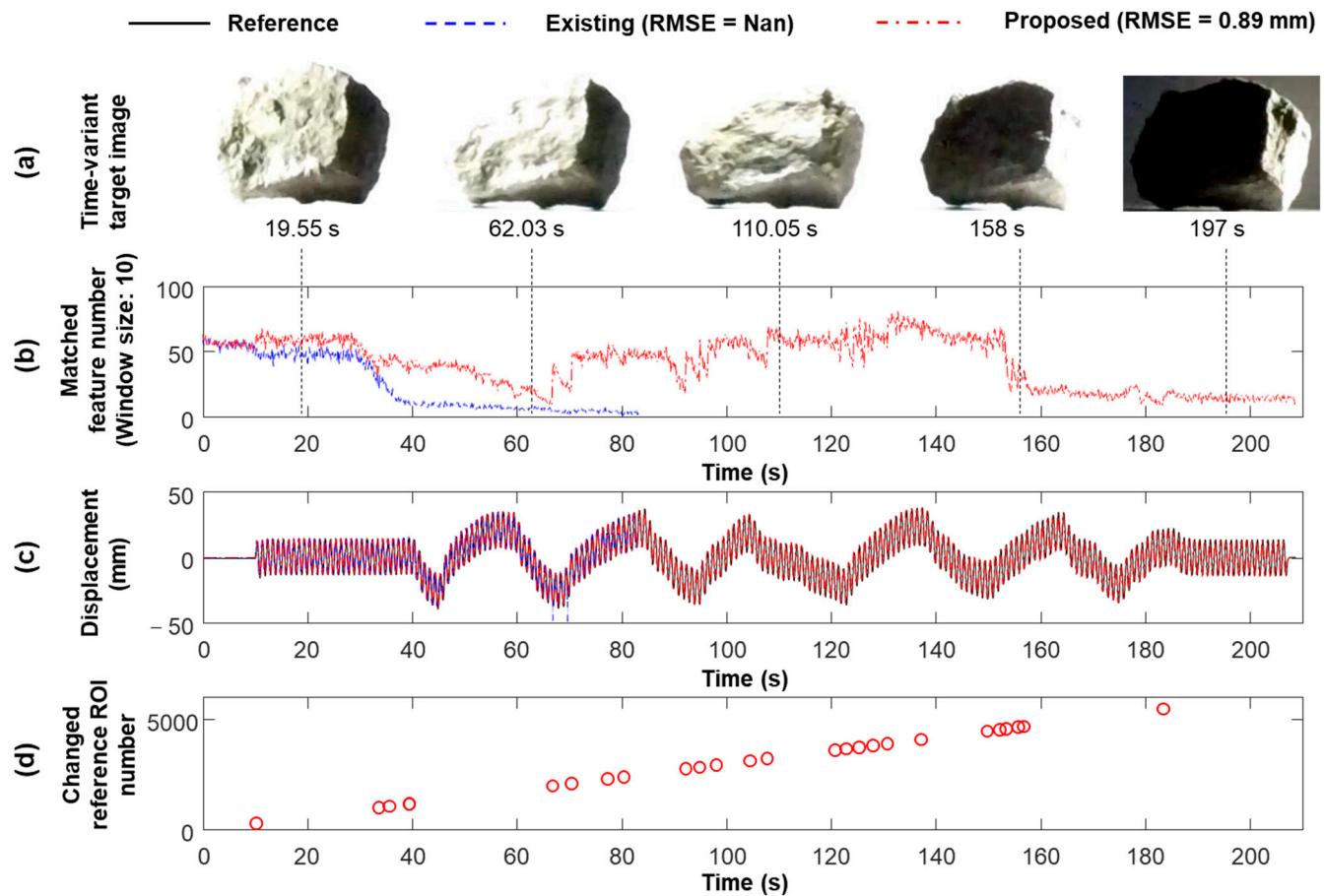


Figure 16. Estimation results Case 6 of the lab-scale test: (a) vision ROIs at different time steps, (b) vision-based displacements estimated with and without the proposed adaptive reference frame updating algorithm, (c) number of matched feature points with and without the proposed adaptive reference frame updating algorithm, and (d) time steps when updating the reference frame.

(c) Continuous displacement estimation (Case 7).

The proposed technique was validated by considering both illumination and temperature variations when the building model was subjected to a real bridge vibration signal excitation (Case 7). As shown in Figure 17a, to simulate the 24 h illumination variation in practice, the light source (i.e., the flashlight), approximately 2 m away from the target of the vision, slowly moved from the left side to the right side and then turned off at 60 s. Subsequently, it was moved to its original location and turned on for 100 s. Here, [0, 38 s], [38 s, 52 s], [52 s, 92 s], [92 s, 108 s], and [108 s, 120 s] simulate the day, the transition time from day to night, the night, the transition time from night to day, and during the day, respectively. When the light source was turned off, the air conditioner was turned on to simulate low-temperature conditions at night (Figure 17b).

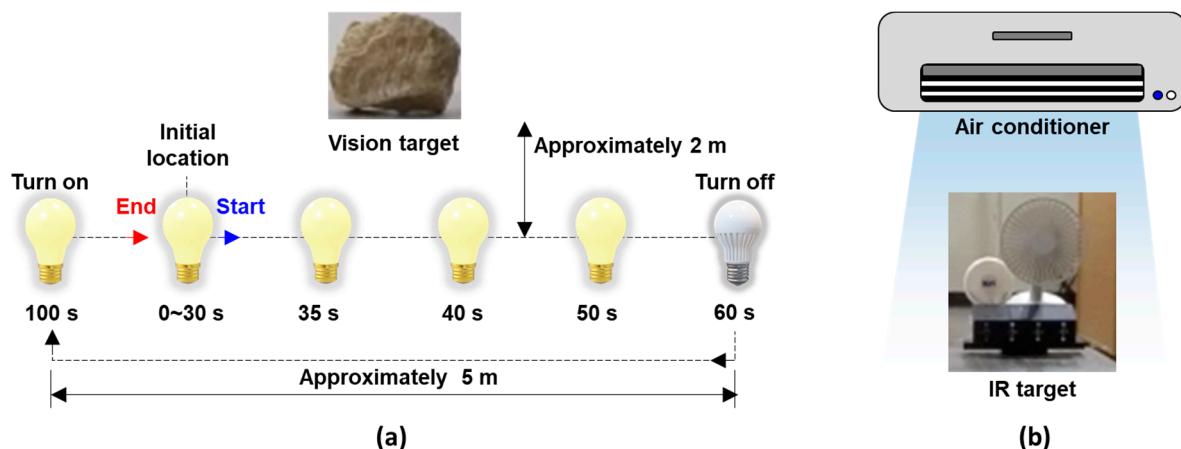


Figure 17. Simulation of (a) varying illumination and (b) varying temperature conditions in 24 h continuous displacement using the light source and air conditioner, respectively.

As shown in Figure 18a, in the first 38 s, the illumination was relatively stable, and sufficient (more than 30) feature points were matched from the vision measurement. Therefore, the proposed technique estimates the displacement using vision and acceleration measurements. Subsequently, the illumination dramatically decreased, causing a decrease in the number of matched feature points in the vision measurement (Figure 18d). IR measurements were also obtained from the beginning of the transition from day to night (i.e., 41 s). Because the IR measurements were not sensitive to illumination variations, the number of matched feature points in the IR measurements was constant. When more feature points were matched from IR measurements rather than vision measurements (i.e., 45 s), the proposed technique switched to using IR and acceleration measurements to estimate displacement. At 82 s, the temperature variation induced by the air conditioner caused the ROI image of the IR measurement to be significantly different from that of the reference IR frame (i.e., the first IR frame), even when the optimized temperature range was used (Figure 18b). Therefore, the matched feature points were insufficient for displacement estimation, and the reference IR frame was updated to increase the number of matched feature points (Figure 18d). Vision measurements were obtained again from the beginning of the transition from night to day (92 s). When more feature points were matched from vision measurements than IR measurements (i.e., 100 s), the proposed technique switched back to using vision and acceleration measurements to estimate displacement. Through automated switching between the vision and IR cameras, the proposed technique continuously estimated the displacement with 0.41 mm RMSE (Figure 18c). However, displacements could only be estimated for the first 42 s using a vision camera and accelerometer.

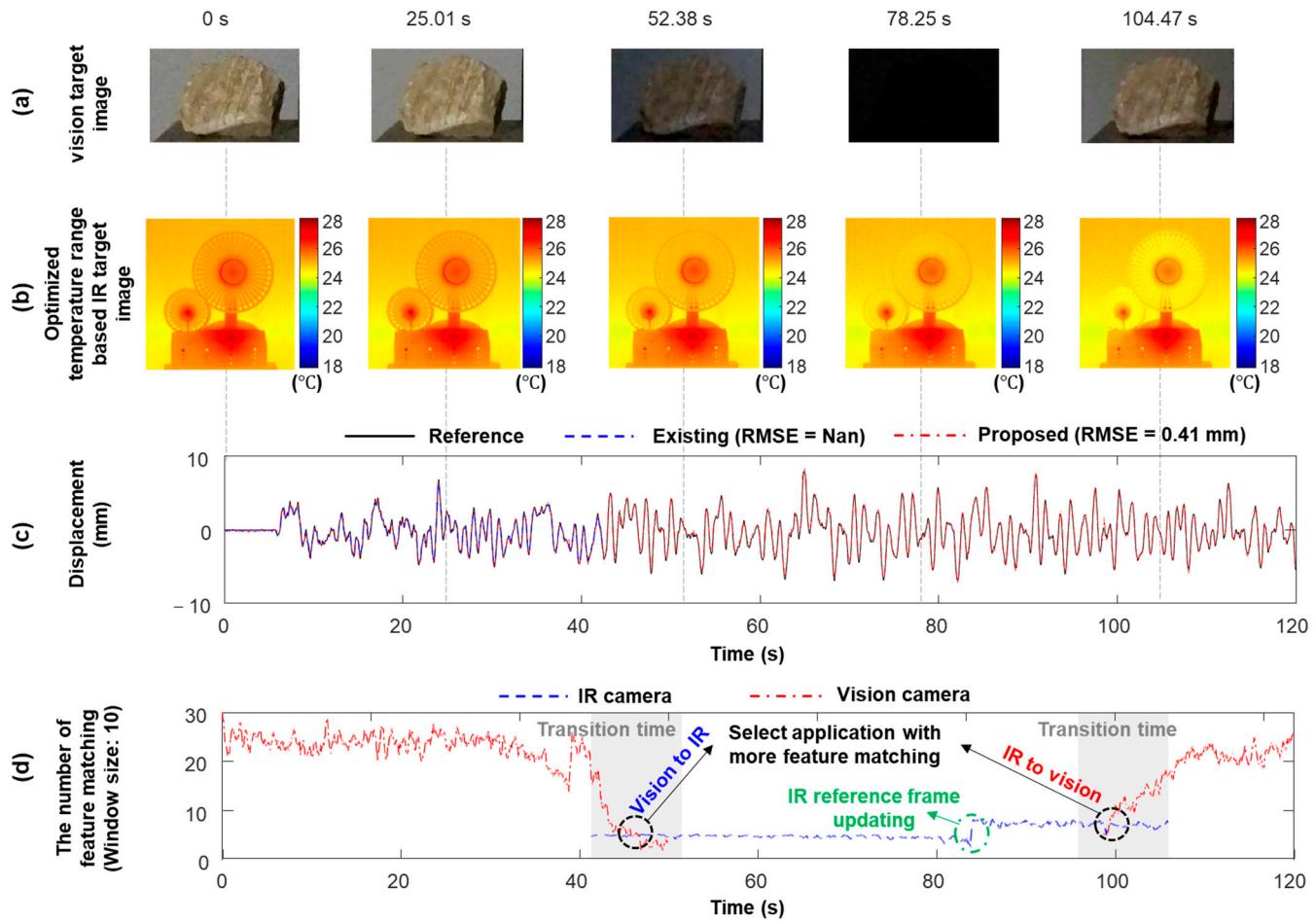


Figure 18. Estimation results for Case 9 of the lab-scale test: (a) vision ROIs at different time steps, (b) IR ROIs at different time steps using the optimized temperature range, (c) estimated displacements, and (d) the number of matched feature points.

3.2. Field Test

3.2.1. Experimental Setup

Figure 19 presents an overview of the field test on a pedestrian bridge. The bridge shown in Figure 19a is located in Daejeon, Korea, and has a length of 45 m and width of 8 m. An IR camera and uniaxial force balance accelerometer identical to those used in the lab-scale test were installed at approximately 1/4 of the span length of the bridge, as shown in Figure 19b. The main purpose of this test was to verify the applicability of the IR camera for displacement estimation under field conditions. A Polytec RSV-150 LDV was installed at a stationary location under the bridge to measure the reference displacement. Figure 19c shows the first frame of IR measurement. The pedestrian bridge was excited by four people jumping near the measurement point. Note that long-term measurement is required to consider temperature and illumination variation, but this is not possible owing to safety reasons. Therefore, the bridge displacements were estimated for a few minutes with stable illumination and temperature to verify the applicability of the IR camera for displacement estimation under field conditions.

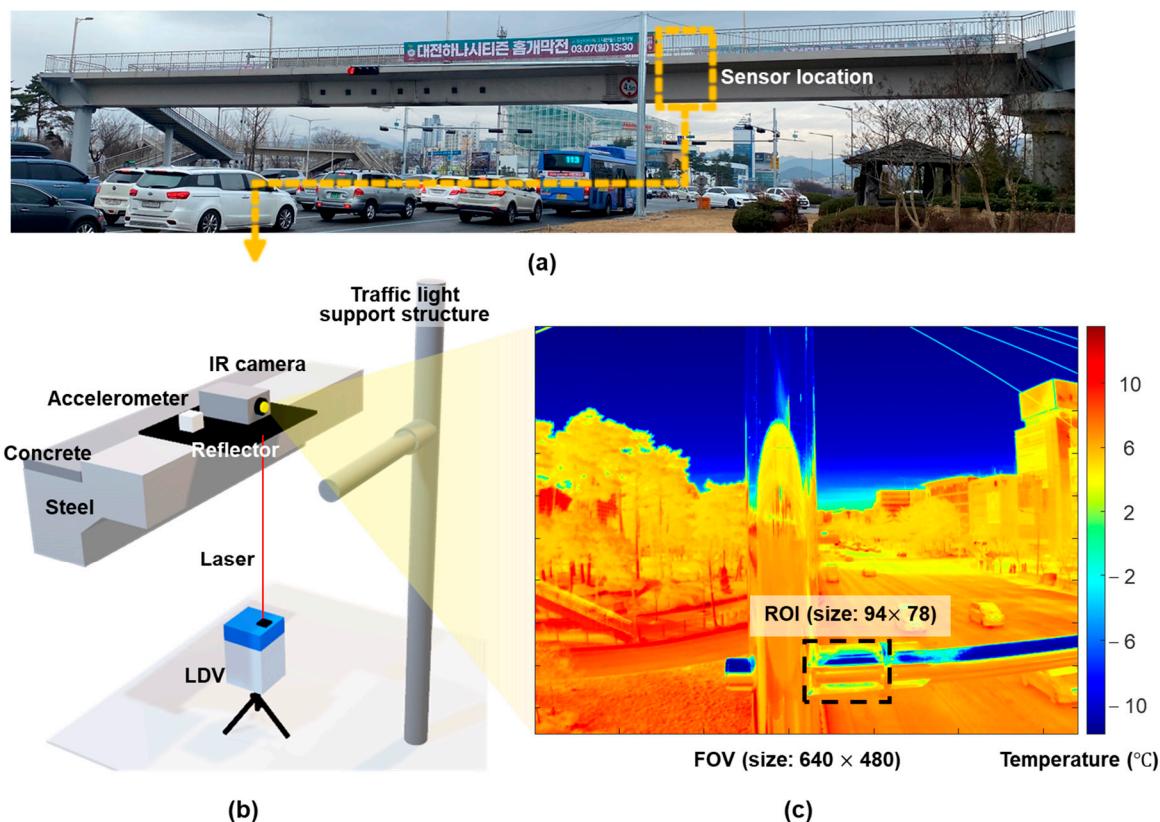


Figure 19. Overview of field test: (a) pedestrian steel box girder bridge, (b) sensor setup on bridge, and (c) view from the IR camera.

3.2.2. IR-Displacement Estimation Results

For the IR camera, a scale factor was estimated as 3.827 mm/pixel and the temperature range was optimized as $[-10.2\text{ }^{\circ}\text{C}, 6.4\text{ }^{\circ}\text{C}]$. The displacements estimated using the ROI, FOV, and optimized temperature ranges are compared in Figure 20a. The optimized temperature range exhibited the best displacement estimation performance with an RMSE of only 0.189 mm. Figure 20a shows a comparison of the number of matched feature points. Although the ROI temperature range had the most matched feature points, many of these feature points were mismatched and the displacement estimation accuracy was worse than that obtained using the optimized temperature range. Only a few feature points were matched when using the FOV temperature range, and a matching failure occurred when using an external heat source (e.g., bus exhaust). Therefore, the displacement could not be continuously estimated. However, the use of the optimized temperature ensured sufficient correctly matched feature points with and without a heat source (Figure 20b,c).

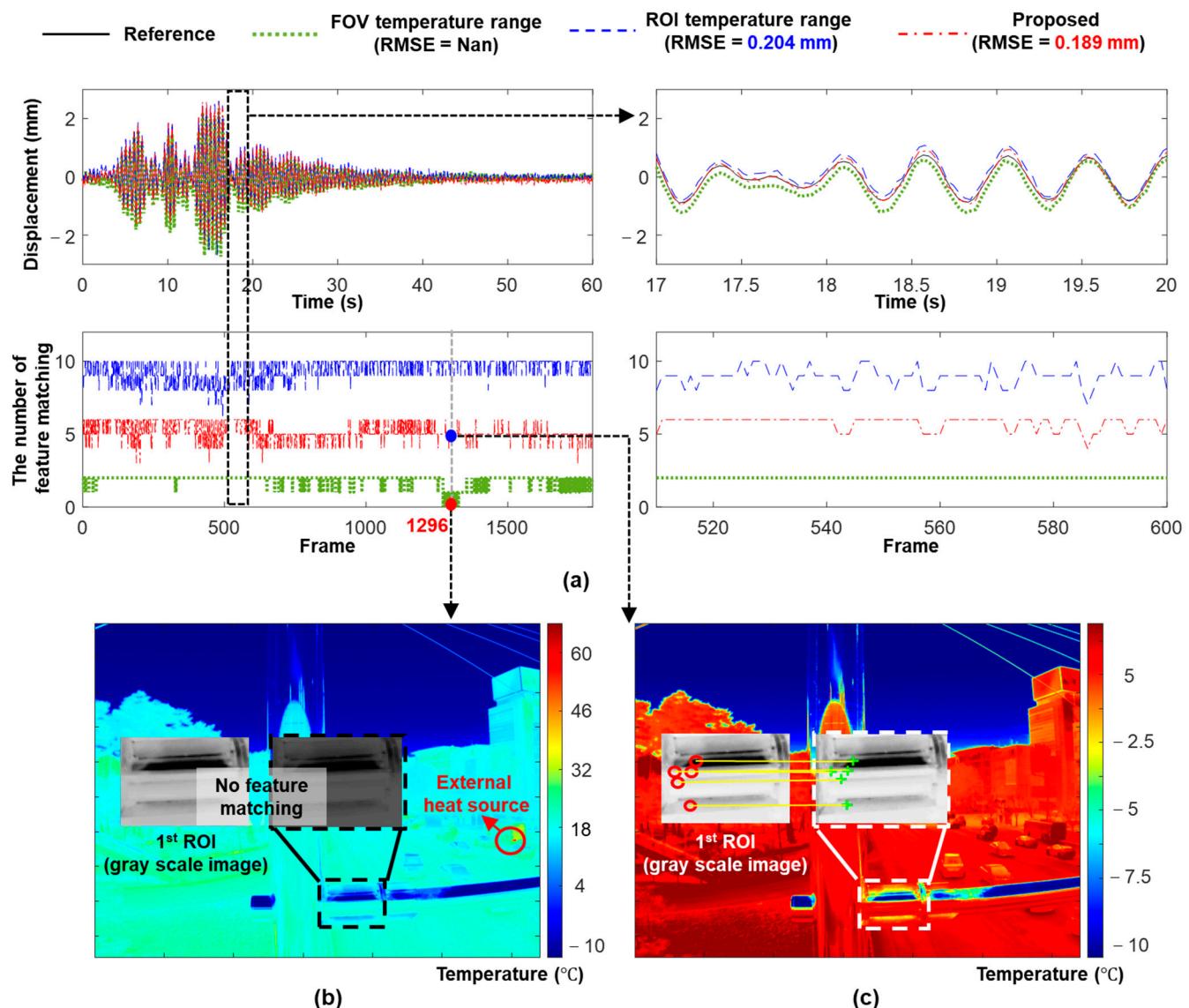


Figure 20. IR-based displacement estimation results: (a) IR-based displacement and matching results, (b) 1296th IR image (based on FOV temperature) before temperature optimization, and (c) after temperature optimization.

4. Conclusions

This study proposes a continuous structural displacement estimation technique using a collocated accelerometer, vision, and IR cameras. The proposed technique first estimates two scale factors for converting translation in a pixel unit to displacement in a length unit for the vision and IR cameras and then optimizes the temperature range for the IR camera. Subsequently, the displacement was continuously estimated by adaptively updating the reference frame and automated switching between the vision and IR cameras during the day and night. The main contributions of this study are (1) day and night continuous displacement estimation using an accelerometer, IR, and vision cameras; (2) automated temperature range optimization for the IR camera; and (3) adaptive reference frame updating for improved robustness against illumination and temperature variations. The proposed technique was validated through a laboratory-scale test that considered illumination and temperature variations, and the displacements were estimated using RMSEs below 1 mm. The applicability of the IR camera for displacement estimation under field conditions was validated using a pedestrian bridge test. The proposed reference frame updating improves the robustness against illumination and temperature variations.

but also causes error accumulation until the reference frame is updated back to the first frame. Further studies are required to address this issue. In addition, the proposed technique was validated only for a short time period, and further validation of its long-term performance under field conditions with illumination and temperature variations is required.

Author Contributions: J.C.: Conceptualization, Methodology, Software Validation, Writing—original draft. Z.M.: Validation, Writing—Review and Editing. K.K.: Writing—Review and Editing. H.S.: Supervision, Writing—Review and Editing, Funding Acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by National Research Foundation of Korea (NRF) grants funded by the Korean government (MSIT) (grant numbers 2017R1A5A1014883 and 2022R1C1C2008186).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Catbas, F.N.; Aktan, A.E. Condition and damage assessment: Issues and some promising indices. *J. Struct. Eng.* **2002**, *128*, 1026–1036.
2. Mau, S. *Introduction to Structural Analysis: Displacement and Force Methods*; CRC Press: Boca Raton, FL, USA, 2012.
3. Nassif, H.H.; Gindy, M.; Davis, J. Comparison of laser Doppler vibrometer with contact sensors for monitoring bridge deflection and vibration. *NDT E Int.* **2005**, *38*, 213–218.
4. Nakamura, S.-I. GPS measurement of wind-induced suspension bridge girder displacements. *J. Struct. Eng.* **2000**, *126*, 1413–1419.
5. Xiong, C.; Wang, M.; Chen, W. Data analysis and dynamic characteristic investigation of large-scale civil structures monitored by RTK-GNSS based on a hybrid filtering algorithm. *J. Civ. Struct. Health Monit.* **2022**, *12*, 857–874.
6. Infante, D.; Di Martire, D.; Confuorto, P.; Tessitore, S.; Tomas, R.; Calcaterra, D.; Ramondini, M. Assessment of building behavior in slow-moving landslide-affected areas through DInSAR data and structural analysis. *Eng. Struct.* **2019**, *199*, 109638.
7. Nettis, A.; Massimi, V.; Nutricato, R.; Nitti, D.O.; Samarelli, S.; Uva, G. Satellite-based interferometry for monitoring structural deformations of bridge portfolios. *Autom. Constr.* **2023**, *147*, 104707.
8. Gomez, F.; Park, J.W.; Spencer, B.F., Jr. Reference-free structural dynamic displacement estimation method. *Struct. Control Health Monit.* **2018**, *25*, e2209.
9. Park, J.-W.; Sim, S.-H.; Jung, H.-J.; Spencer, B.F., Jr. Development of a wireless displacement measurement system using acceleration responses. *Sensors* **2013**, *13*, 8377–8392.
10. Reu, P.L.; Rohe, D.P.; Jacobs, L.D. Comparison of DIC and LDV for practical vibration and modal measurements. *Mech. Syst. Signal Process.* **2017**, *86*, 2–16.
11. Huang, Q.; Wang, Y.; Luzi, G.; Crosetto, M.; Monserrat, O.; Jiang, J.; Zhao, H.; Ding, Y. Ground-based radar interferometry for monitoring the dynamic performance of a multitrack steel truss high-speed railway bridge. *Remote Sens.* **2020**, *12*, 2594.
12. Zhang, G.; Wu, Y.; Zhao, W.; Zhang, J. Radar-based multipoint displacement measurements of a 1200-m-long suspension bridge. *ISPRS J. Photogramm. Remote Sens.* **2020**, *167*, 71–84.
13. Jeong, J.H.; Jo, H. Real-time generic target tracking for structural displacement monitoring under environmental uncertainties via deep learning. *Struct. Control Health Monit.* **2022**, *29*, e2902.
14. Khuc, T.; Catbas, F.N. Computer vision-based displacement and vibration monitoring without using physical target on structures. *Struct. Infrastruct. Eng.* **2017**, *13*, 505–516.
15. Lee, S.; Kim, H.; Sim, S.-H. Equation Chapter 1 Section 1 nontarget-based displacement measurement using LiDAR and camera. *Autom. Constr.* **2022**, *142*, 104493.
16. Nasimi, R.; Moreu, F. A methodology for measuring the total displacements of structures using a laser–camera system. *Comput.-Aided Civ. Infrastruct. Eng.* **2021**, *36*, 421–437.
17. Moschas, F.; Stiros, S. Measurement of the dynamic displacements and of the modal frequencies of a short-span pedestrian bridge using GPS and an accelerometer. *Eng. Struct.* **2011**, *33*, 10–17.
18. Ozdagli, A.; Gomez, J.; Moreu, F. Real-time reference-free displacement of railroad bridges during train-crossing events. *J. Bridge Eng.* **2017**, *22*, 04017073.
19. Park, J.W.; Moon, D.S.; Yoon, H.; Gomez, F.; Spencer, B.F., Jr.; Kim, J.R. Visual–inertial displacement sensing using data fusion of vision-based displacement with acceleration. *Struct. Control Health Monit.* **2018**, *25*, e2122.

20. Zhou, Q.; Li, Q.-S.; Han, X.-L.; Wan, J.-W.; Xu, K. Horizontal displacement estimation of high-rise structures by fusing strain and acceleration measurements. *J. Build. Eng.* **2022**, *57*, 104964.
21. Xu, Y.; Brownjohn, J.M.; Huseynov, F. Accurate deformation monitoring on bridge structures using a cost-effective sensing system combined with a camera and accelerometers: Case study. *J. Bridge Eng.* **2019**, *24*, 05018014.
22. Ma, Z.; Choi, J.; Sohn, H. Real-time structural displacement estimation by fusing asynchronous acceleration and computer vision measurements. *Comput.-Aided Civ. Infrastruct. Eng.* **2022**, *37*, 688–703.
23. Bay, H.; Tuytelaars, T.; Gool, L.V. Surf: Speeded up Robust Features. In Proceedings of the 9th European Conference on Computer Vision (ECCV 2006), Graz, Austria, 7–13 May 2006; Springer: Berlin/Heidelberg, Germany, 2006; pp. 404–417.
24. Brandt, A. *Noise and Vibration Analysis: Signal Analysis and Experimental Procedures*; John Wiley & Sons: Hoboken, NJ, USA, 2011.
25. Akhlaghi, S.; Zhou, N.; Huang, Z. Adaptive Adjustment of Noise Covariance in Kalman Filter for Dynamic State Estimation. In Proceedings of the 2017 IEEE Power & Energy Society General Meeting, Chicago, IL, USA, 16–20 July 2017; pp. 1–5.
26. Scaramuzza, D.; Martinelli, A.; Siegwart, R. A Toolbox for Easily Calibrating Omnidirectional Cameras. In Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 9–15 October 2006; pp. 5695–5701.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.