



Robustness evaluation of phase-based displacement measurement under illumination variation conditions

Sida Ai^a, Zhenkun Li^{b,*} 

^a Department of Civil and Environmental Engineering, University of Alberta, Edmonton T6G 2W2, Alberta, Canada

^b Department of Civil Engineering, Aalto University, 02150 Espoo, Finland

ARTICLE INFO

Keywords:

Vision-based displacement measurement
Phase-based optical flow
Illumination variation
Low-light conditions
Time-varying illumination
Robustness

ABSTRACT

Vision-based structural displacement measurement has attracted increasing attention in structural health monitoring because of its non-contact nature, low cost, and high spatial resolution. However, the performance of many conventional intensity-based methods is highly sensitive to illumination conditions, which limits their long-term application in real engineering environments. Although phase-based methods have been suggested to be more robust to illumination variations, this view has largely been based on theoretical inference in previous studies, and systematic experimental validation under representative lighting conditions remains limited. Therefore, in this study, controlled laboratory experiments were conducted under different illumination intensities, time-varying illumination, and fog-induced dynamic spatial interference to investigate the performance of the phase-based method for vision-based structural displacement measurement. A laser displacement sensor was used as the reference measurement. For comparison, two additional optical flow methods, namely intensity-based optical flow and deep learning-based optical flow, were also implemented. The results show that, under the tested low-light and time-varying illumination conditions in laboratory settings, the phase-based method provides more robust displacement measurement, with better accuracy and signal stability, than the other two methods.

1. Introduction

Structural displacement response is a particularly important physical quantity in structural health monitoring (SHM) [1,2]. It can be used to identify changes in natural frequencies, stiffness, and mode shapes, thereby enabling a comprehensive assessment of the structural condition [3]. Therefore, accurate measurement of structural displacement response is a critical topic in SHM and has a direct impact on both monitoring cost and accuracy [3].

Among conventional approaches for measuring structural displacement, commonly used methods include accelerometers, linear variable displacement transducers (LVDTs), draw-wire sensors, and laser displacement meters [4–7]. Although these methods are mature and reliable, they suffer from several limitations, such as high installation cost, expensive equipment, limited service life, and a restricted number of measurement points, which leads to incomplete monitoring information [1]. Accordingly, integrating structural response monitoring with emerging technologies is a promising way to improve cost-effectiveness, efficiency, accuracy, and spatial coverage. In particular,

with advances in imaging sensors and computer vision, vision-based methods for measuring structural displacement response have attracted increasing attention in the SHM community because of their non-contact nature, low cost, and high spatial resolution [8–12].

In general, vision-based structural displacement measurement involves several basic steps. First, a vibrating structure is selected as the monitoring target. Then, cameras are deployed and calibrated to establish the transformation between physical coordinates and pixel coordinates, while the vibration process is recorded. Once a qualified video of structural vibration is obtained, computer vision algorithms are applied to extract the displacement signal from the recorded image sequence.

Camera deployment and calibration are key steps in vision-based structural displacement measurement, as they ensure data usability and enable quantitative conversion. Proper deployment and fixation of the camera determine a reasonable observation geometry and guarantee the stability of the camera pose, thereby preventing problems such as insufficient field of view, occlusion, defocus, or vibration of the camera support system, all of which may render the recorded video unusable

* Corresponding author.

E-mail addresses: sai6@ualberta.ca (S. Ai), zhenkun.li@aalto.fi (Z. Li).

<https://doi.org/10.1016/j.measurement.2026.122441>

Received 31 March 2026; Received in revised form 2 June 2026; Accepted 1 July 2026

Available online 7 July 2026

0263-2241/© 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

[1]. Calibration establishes the mapping between the world coordinate system and the pixel coordinate system, so that pixel-level measurements can be converted into real physical quantities. Calibration accuracy therefore has a direct influence on the final measurement accuracy [1,13]. Under common engineering conditions, where the intrinsic and extrinsic camera parameters and installation pose remain stable, the target can be approximately treated as planar, and lens distortion is sufficiently corrected, the conversion error from pixels to physical units is often systematic and repeatable, typically appearing as a scale-factor bias. Under these assumptions, such errors can often be identified and corrected through reference measurements or geometric verification [13,14].

Overall, once camera deployment and calibration satisfy the basic measurement requirements, the computer vision algorithm often becomes the dominant factor affecting the performance of vision-based methods. It largely determines the accuracy, efficiency, and robustness of vision-based structural displacement monitoring systems [3]. Since the 1980s and 1990s, various classes of computer vision algorithms have been developed and gradually applied to structural response monitoring. Representative early methods include template matching, digital image correlation (DIC), feature-point matching, and optical flow, all of which have been relatively well established in structural dynamics experiments [9,15–20]. Optical flow is one of the classical vision-based methods capable of achieving subpixel-level displacement measurement accuracy. Among optical flow approaches, the Lucas–Kanade (LK) method and its variants have been widely adopted because of their computational efficiency and ability to estimate subpixel motion from local intensity variations [21,22]. In structural displacement measurement, LK optical flow has shown good performance under controlled imaging conditions and has been used for tracking small motions of structural targets [22]. More recently, deep learning-based optical flow methods have emerged as a new direction in motion estimation. In particular, RAFT (Recurrent All-Pairs Field Transforms) has demonstrated strong performance in dense optical flow estimation by iteratively refining flow fields based on global correlation information, offering improved accuracy and robustness in complex motion scenarios [23]. These advances suggest the potential of modern optical flow algorithms for vision-based displacement measurement, especially in cases where dense motion estimation is desirable [24].

However, it should be noted that the successful applications of these methods have mostly been concentrated in laboratory environments with controllable imaging conditions. Their application in real-world engineering environments remains comparatively limited, particularly for the long-term monitoring of in-service structures [1,3]. One important reason is that these methods rely heavily on image intensity information for target localization and displacement estimation, making them sensitive to complex illumination conditions such as low lighting, temporal variations in light intensity, and shadows [17,25]. Under stable illumination, these methods usually perform well; however, under practical field conditions such as nighttime or day-to-night transitions, they are more likely to suffer from localization drift, matching failure, or tracking interruption [3].

Illumination conditions are particularly important in vision-based displacement measurement for two main reasons. First, vision-based methods fundamentally rely on pixel intensity and texture contrast to extract displacement information. Changes in illumination directly affect image brightness, contrast, and signal-to-noise ratio, making illumination one of the most important external factors influencing image quality and feature visibility [26,27]. Second, many other influencing factors, such as camera pose, calibration error, frame rate, and spatial resolution, can usually be kept relatively stable and controllable in engineering practice through proper installation, standardized calibration procedures, suitable hardware selection, and well-designed acquisition strategies. Likewise, issues such as motion blur and partial occlusion can often be mitigated through accumulated engineering experience and targeted mitigation strategies [10,19]. In contrast,

illumination is difficult to maintain stable over long-term field monitoring and directly determines image brightness, contrast, and the availability of useful texture information [1,3,27].

To address this challenge, several recent studies have attempted to improve the robustness of vision-based displacement measurement under challenging illumination conditions. For instance, deep learning techniques have been incorporated to adapt displacement tracking to ambient-light changes [27]. Robust feature tracking algorithms have also been developed and validated under changing illumination conditions [35]. Additionally, modified optical flow formulations have been proposed to relax the brightness constancy assumption under varying illumination [36]. While these approaches can partially mitigate the effects of illumination variation, they remain fundamentally dependent on intensity information and are therefore inherently limited when significant brightness changes occur. In contrast, phase-based methods offer a fundamentally different approach, as the phase information they rely on is inherently less sensitive to the magnitude of the source signal and its variations [28–34]. This enables phase-based methods to maintain higher stability under non-ideal illumination conditions and often achieve higher subpixel measurement accuracy [28,29,37–39]. Recent studies have further explored phase-based frameworks in the context of structural monitoring and motion magnification. Guo et al. [24] proposed an improved phase-based motion magnification method combined with deep learning for structural vibration measurement, demonstrating enhanced robustness against background noise and illumination changes. Yang et al. [40] developed a novel phase-based video motion magnification technique for non-contact measurement of micro-amplitude vibrations, advancing the capability of phase-based methods in capturing subtle structural motions. Ai et al. [41] developed a phase extraction method based on a complex Laplacian pyramid with condensed focus frequency to improve the computational efficiency of phase-based displacement measurement while maintaining measurement accuracy.

It should be noted, however, that although existing studies have suggested that phase-based methods are robust under unfavorable illumination conditions, systematic comparative validation under representative illumination variation conditions remains limited [29,33,37,42]. Systematic experimental investigation and evaluation of the illumination robustness of phase-based methods in specific vision-based displacement measurement tasks are still lacking. In particular, there remains insufficient comparative evidence showing whether phase-based methods truly offer a practical advantage in illumination robustness over other displacement measurement algorithms with sub-pixel accuracy.

Therefore, in this study, vision-based displacement measurement experiments were conducted under controlled indoor illumination conditions. Test scenarios were established to represent different illumination intensities, time-varying illumination intensities, and spatial dynamic disturbances (e.g., fog effects). The performance of phase-based methods and other displacement extraction algorithms was evaluated and compared under these conditions, followed by an analysis of their displacement measurement behavior. This study is expected to improve understanding of the performance of phase-based methods under complex illumination conditions and ultimately facilitate their stable application in real engineering environments.

The structure of this paper is organized as follows: Section 2 presents the methodology of phase-based structural displacement measurement, including the basic measurement process, phase extraction using the complex steerable pyramid, and phase-based optical flow. Section 3 details the illumination variation experiments designed to evaluate the performance of the phase-based method and presents the corresponding results and discussion.

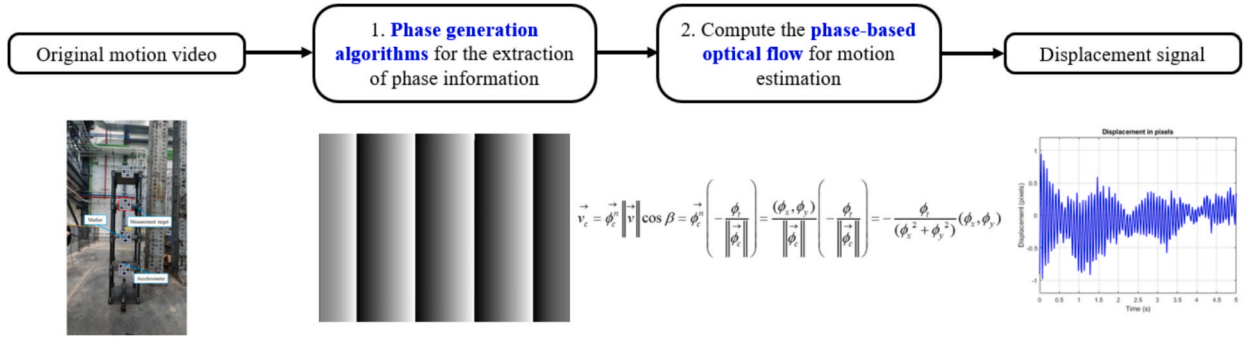


Fig. 1. The general process of phase-based displacement measurement.

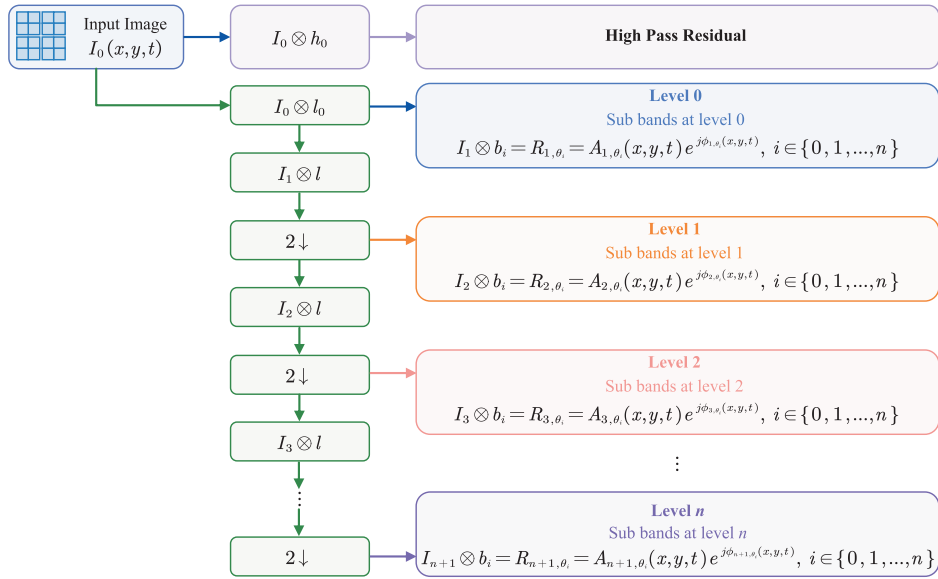


Fig. 2. The decomposition process of the CSP method.

2. Methodology

2.1. The process of phase-based displacement measurement

As shown in Fig. 1, the current phase-based displacement measurement method typically consists of two main steps: phase information extraction and the computation of phase-based optical flow. In the first step, phase extraction algorithms are used to extract phase information from frames of motion video. Two widely used approaches for phase extraction are the Complex Steerable Pyramid (CSP) and the Gabor wavelet methods [28,29,37,43,44]. In addition, several new phase extraction algorithms have recently been developed to simplify parameter selection in traditional methods and enhance computational efficiency. These include the Hilbert transform-enhanced method, the Riesz transform-enhanced method, and the derivative-enhanced method [32,42,45]. In our investigation, we selected a classic phase extraction method, the CSP method, which has been widely used in phase-based structural displacement measurement [28,33,39,46,47]. Once the phase information is extracted, the second step involves computing phase-based optical flow to obtain the displacement signal.

This phase-based measurement framework is particularly relevant for displacement estimation under varying illumination conditions. Unlike conventional intensity-based methods, which directly rely on image brightness and contrast for motion estimation, phase-based methods describe motion through phase variation in transformed image representations. Therefore, when illumination changes mainly

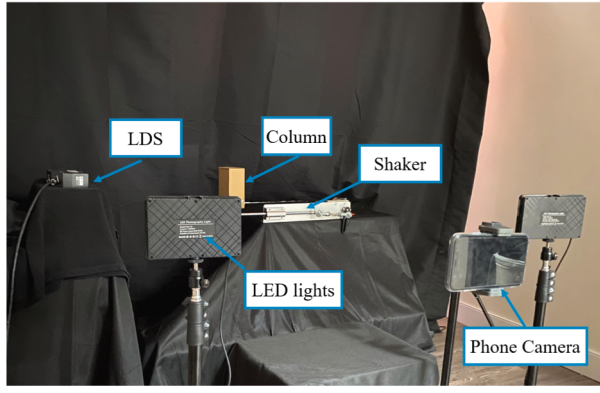
affect the intensity amplitude rather than the underlying local structural pattern, the extracted phase information may remain comparatively more stable. This characteristic provides the basic methodological motivation for employing phase-based displacement measurement in the present study.

2.2. Phase extraction method: complex steerable pyramid

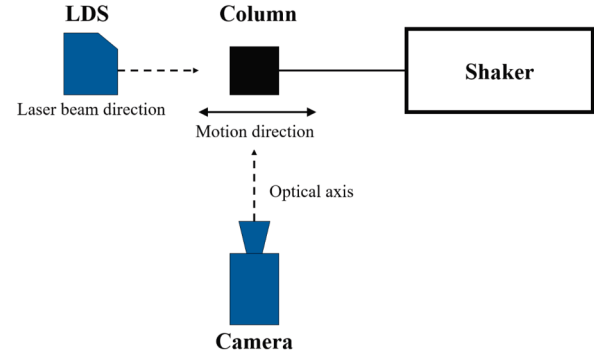
The steerable pyramid method was developed by Simoncelli et al. [43,48] to capture the directional information of images at various scales, and the imaginary part of CSP is the Hilbert transform version of the steerable pyramid [44]. In the CSP, the downsampling process is adopted to capture different scope details of the image at different scale levels, and the steerable filter is utilized to extract directional details from the image.

The framework of the decomposition process of CSP is shown in Fig. 2. The CSP utilizes low-pass filters l_0, l , high-pass filter h_0 , and steerable filters b_i for the decomposition process. The low-pass and high-pass filters facilitate the downsampling process, closely resembling the discrete wavelet transform [49]. The output from the low-pass filters at each scale level is convolved with the steerable filter to extract directional information R_{i,θ_i} . Generally, phase information derived from the directional information R_{i,θ_i} of CSP is employed to perform phase-based displacement measurement.

Among the available phase extraction methods, the CSP method was adopted in this study for several reasons. First, CSP is one of the most



(a) Experiment layout



(b) Schematic diagram

Fig. 3. Images of the experiment layout and schematic diagram.

established phase extraction approaches in vision-based structural vibration and displacement measurement, and its effectiveness has been demonstrated in a number of previous studies [28,33,39,46,47]. Second, its multi-scale and orientation-selective decomposition is well suited to motion analysis in structural targets with relatively clear vibration directions. Third, compared with some recently developed phase extraction methods, CSP provides a more classical and well-understood framework, which facilitates methodological interpretation and comparison with existing studies.

Physically, the scale level in CSP determines the spatial scale at which image patterns are represented and analyzed. Smaller scale levels are generally more sensitive to fine local details and subtle motion, but they are also more susceptible to image noise and local pattern instability. In contrast, larger scale levels tend to provide more stable phase information by emphasizing coarser structural patterns, although they may reduce sensitivity to small local variations. Therefore, the selection of an appropriate scale level is essential for balancing motion sensitivity and measurement stability in phase-based displacement estimation [33]. Once the direction of interest is set, the scale level is the only parameter of CSP that needs to be set for displacement measurement.

2.3. Phase-based optical flow

In 1990, Fleet et al. [37] proposed the phase constancy assumption for estimating displacement in phase-based optical flow, which plays a role analogous to the brightness constancy assumption used in conventional intensity-based optical flow. In contrast to intensity information, which reflects motion in two directions and is therefore subject to the aperture problem, phase information is intrinsically one-dimensional. As a result, estimating motion from phase along a selected direction has become a standard strategy in phase-based optical flow, since it can effectively avoid the aperture ambiguity encountered in intensity-based formulations.

Under the phase constancy assumption, taking the temporal derivative on both sides of Eq. (1) yields Eq. (2). In Eq. (2), ϕ_x and ϕ_y denote the spatial phase gradients along the horizontal and vertical axes, respectively, while ϕ_t represents the temporal phase gradient. In addition, v_x and v_y correspond to the pixel velocities in the horizontal and vertical directions.

$$\phi(x, y, t) = c \quad (1)$$

$$\phi_x v_x + \phi_y v_y + \phi_t = 0 \quad (2)$$

By projecting Eq. (2) onto the orientation of the Gabor wavelet, Eq. (3) can be obtained.

$$\left\| \vec{\phi}_c \right\| \left\| \vec{v} \right\| \cos \alpha + \phi_t = 0 \quad (3)$$

The complete velocity $\vec{v}$ is a vector that may point in any direction, whereas $\vec{\phi}_c$ denotes the phase gradient. The vectors $\vec{\phi}_c$ and $\vec{v}$ form an angle α .

Accordingly, the velocity component $\vec{v}_c$ along the Gabor wavelet direction can be derived from Eq. (4), where $\vec{\phi}_c^n$ is the unit vector of the phase gradient.

$$\begin{aligned} \vec{v}_c &= \vec{\phi}_c^n \left\| \vec{v} \right\| \cos \alpha = \vec{\phi}_c^n \left(- \frac{\phi_t}{\left\| \vec{\phi}_c \right\|} \right) = \frac{(\phi_x, \phi_y)}{\left\| \vec{\phi}_c \right\|} \left(- \frac{\phi_t}{\left\| \vec{\phi}_c \right\|} \right) \\ &= - \frac{\phi_t}{(\phi_x^2 + \phi_y^2)} (\phi_x, \phi_y) \end{aligned} \quad (4)$$

After estimating the pixel velocity, the pixel displacement can be calculated from the velocity and the time interval, as given in Eq. (5).

$$\vec{d}_c = \vec{v}_c \Delta t \quad (5)$$

Compared with conventional intensity-based optical flow, the essential difference of phase-based optical flow lies in the type of image information used for motion estimation. Conventional optical flow methods estimate motion primarily from the constancy of image intensity, and their performance is therefore more directly influenced by variations in brightness, contrast, and local texture visibility. By contrast, phase-based optical flow estimates motion from phase variation along a selected orientation, which is more closely related to the structural pattern of the image than to its absolute intensity magnitude. Therefore, when illumination changes mainly affect image amplitude rather than the local pattern itself, phase-based optical flow may provide comparatively more stable displacement estimation. This theoretical distinction forms the basis for the comparative experimental evaluation under illumination variation conditions in the present study.

3. Illumination variation experiments for evaluation of the performance of phase-based method

Fig. 3 (a) and (b) illustrate the experimental arrangement and corresponding schematic of the illumination variation experiments. In this setup, the moving target consists of a column mounted on the head of an electrodynamic shaker. The shaker generates reciprocating motion with controllable speed. Since the motion velocity can be adjusted to very low levels, an iPhone 11 was used to capture the target motion at a resolution of 720P and a frame rate of 30 fps. The camera was placed at a distance of 900 mm from the target, and its optical axis was aligned perpendicular to the motion direction. With the focal length set to 52 mm, the resulting imaging scale was approximately 0.402 mm/pixel.

To provide reference displacement data, a Keyence LK-G3000 laser

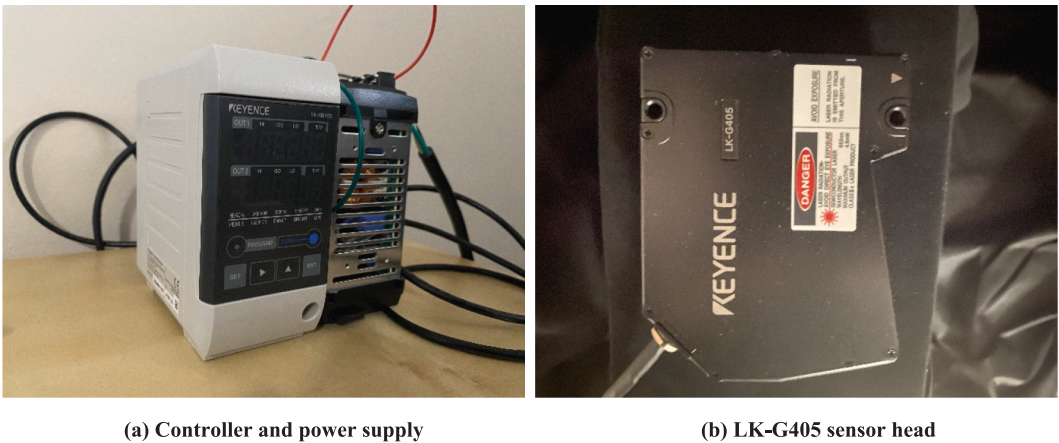


Fig. 4. Laser displacement sensor (Keyence LK-G3000 series).

Table 1
The general cases of the illumination variation experiments (check marks indicate the illumination conditions applied in each case).

Cases	Various fixed lightings	Temporal lighting changes	Under Fog
1	✓		
2	✓		✓
3		✓	
4		✓	✓

displacement sensor (LDS) fitted with an LK-G405 sensor head was used, as shown in Fig. 4. The sensor head has an effective operating range of $400\text{ mm} \pm 100\text{ mm}$, while the repeatability of the LDS system is specified as $2\text{ }\mu\text{m}$.

Based on the basic experimental setup, a series of experiments was conducted to investigate the effects of illumination variations. Specifically, the evaluated conditions include various lighting intensities and temporal lighting variations under both clear and foggy conditions, as shown in Table 1. The varying constant illumination cases are primarily designed to investigate the stability of each vision-based method under low-light conditions and the degree of accuracy degradation of these methods as illumination decreases. The corresponding real-world engineering scenarios include nighttime low-light environments or low-light conditions inside tunnels. The temporal illumination variation cases are designed to study the stability of each vision-based method when the illumination intensity changes during the measurement process. The corresponding engineering scenario is the day-night transition encountered in long-term structural health monitoring. The foggy cases are designed to investigate the effect of mild occlusion on each vision-based method. The corresponding engineering scenario is long-range vision-

based structural displacement measurement, where the influence of fog becomes more significant due to the extended optical path.

The lighting was provided by two adjustable LED lights with adjustable brightness (ALTSON LED Video Lighting Kit, USB-powered), as shown in Fig. 3 (a). The intensity of lighting is measured by a lux meter (TEKOPLUS Lux Meter with data logging capability) near the moving target, as shown in Fig. 5 (a). The lux meter is capable of time history recording of illuminance at a sampling rate of once per second.

To ensure consistency of the data acquisition process across all tested conditions, the exposure compensation value of the iPhone camera is set to 0 in all experimental cases. This means that the camera operates with the same automatic exposure strategy throughout all cases without any manual exposure compensation, ensuring that no artificial brightness bias is introduced across different illumination conditions. However, it should be noted that since the video is recorded using the built-in iPhone camera app, the underlying camera pipeline may still perform automatic exposure, auto gain, or noise suppression processing. Therefore, this experiment evaluates the adaptability of each vision-based method to different illumination conditions under a realistic smartphone imaging pipeline, rather than under fully controlled imaging conditions with locked physical exposure parameters. Importantly, under each illumination condition, all vision-based methods receive exactly the same video frames as input, and any automatic adjustments made by the internal camera pipeline affect all three methods equally, ensuring that the comparative results remain fair and valid.

The fog occlusion is simulated using a humidifier, which has been recognized in previous research as an effective and practical approach for generating artificial fog [10,50,51]. The humidifier (AquaOasis Cool Mist Humidifier, 2.2 L capacity) used to produce the fog is shown in Fig. 5 (b). The humidifier is placed beneath the target on the side closer

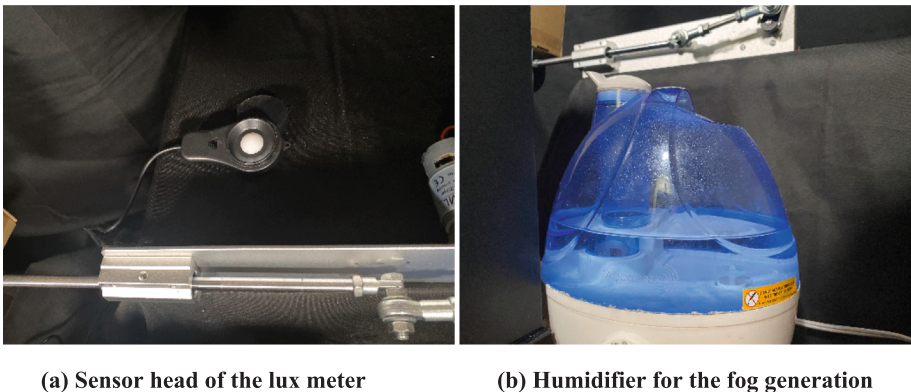


Fig. 5. The setups of lux meter and humidifier.

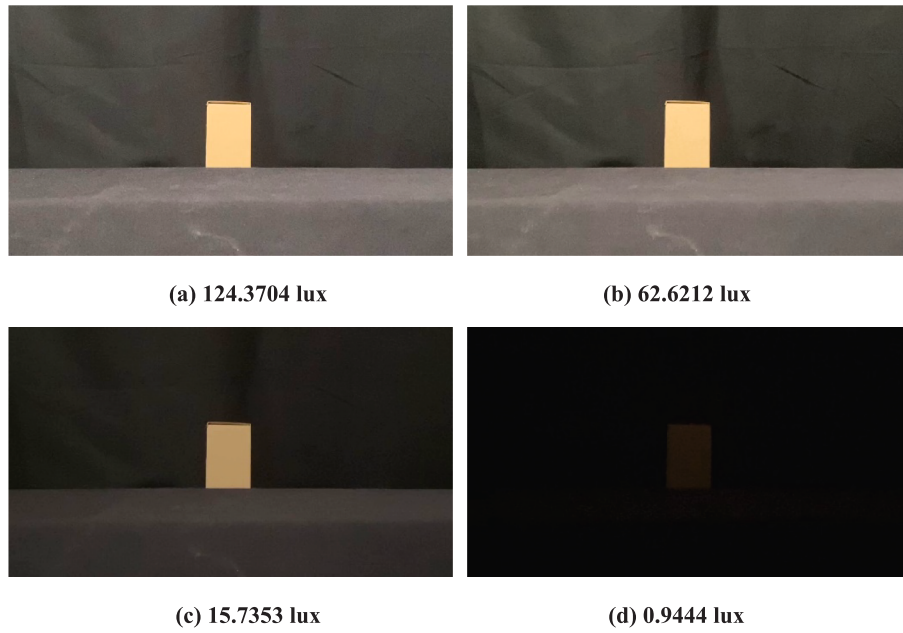


Fig. 6. Frames from experiment videos in Case 1 under clear conditions with varying illumination levels.

to the camera, with the fog outlet approximately 200 mm from the target

surface, allowing the fog to diffuse upward and cover both the target

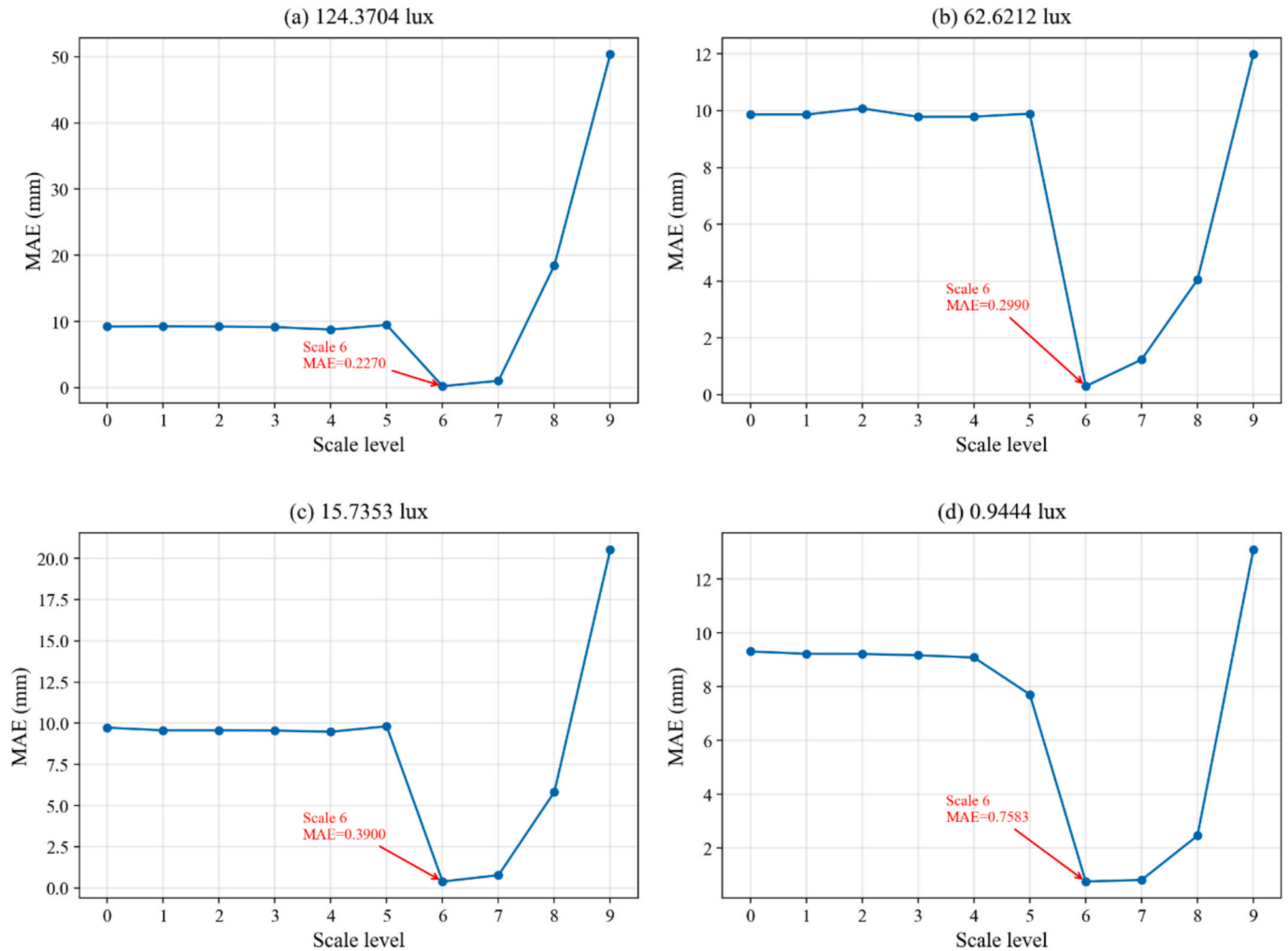


Fig. 7. Effect of scale level on MAE under varying illumination conditions in Case 1.

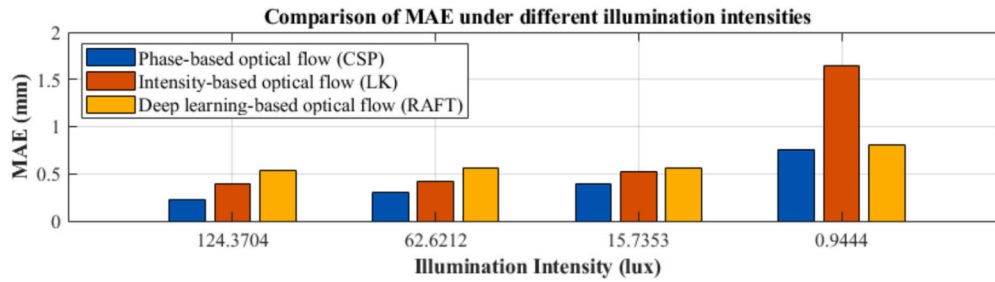


Fig. 8. Comparison of measurement performance under clear conditions with varying illumination levels.

region and the optical path between the target and camera. Before each foggy condition experiment, the humidifier is operated at maximum output for approximately 5 min to allow the fog concentration to reach a steady state. In the steady state, the humidity data obtained around the target area were as follows: the temperature and relative humidity at the center of the target were 16.9 °C and 90.13%, respectively, while the average temperature and relative humidity around the target area were 20.4 °C and 62.85%, respectively. These conditions are similar to those reported in previous laboratory studies [50], which were conducted to simulate the most unfavorable fog level. Video recording commences only after this stabilization period, and the humidifier remains at maximum output throughout the recording to maintain consistent fog concentration. Across all fog-related test conditions (Case 2 and Case 4), the humidifier position and output setting remain unchanged, ensuring a reasonably consistent fog environment across different cases.

It should be acknowledged that the current fog simulation has certain limitations compared with field deployment scenarios. The optical path in this laboratory setup is only 900 mm, whereas in real applications it may extend to tens or hundreds of meters with greater cumulative scattering effects. The indoor experiment uses a fixed-direction LED light source, whereas outdoor sunlight and diffuse skylight enter the fog from multiple directions, potentially producing more complex optical effects such as glare and non-uniform background brightness. Additionally, the humidifier generates pure water fog with uniform particle composition, whereas natural outdoor fog may coexist with haze, rain, and atmospheric turbulence, involving varying particle sizes and refractive indices that may produce more spatially heterogeneous optical distortions. Nevertheless, the laboratory experiment provides a controlled baseline to reveal the fundamental mechanisms by which spatial occlusion affects vision-based displacement measurement. Future work involving field deployment under natural fog and other adverse atmospheric conditions would be a worthwhile direction to further investigate the robustness of vision-based methods.

To ensure a well-controlled lighting environment, there are no other lighting sources around the moving target except for the two LED lights. In the general Case 1, the LED lights are adjusted to four brightness levels: maximum, medium, low, and near-zero (almost off), in order to evaluate the measurement performance of different vision-based methods. The corresponding light intensities are 124.3704 lx, 62.6212 lx, 15.7353 lx, and 0.9444 lx, respectively. The frames for these

cases are shown in Fig. 6.

Under the above conditions, three optical flow methods for vision-based displacement measurement are compared in this study, namely intensity-based optical flow (LK), deep learning-based optical flow (RAFT), and phase-based optical flow (CSP). To ensure a fair comparison, all three compared methods were implemented using reasonable parameter settings based on their methodological characteristics and previous studies.

To enable a direct comparison between the displacement results from the vision-based methods and the laser displacement sensor, the two signals were first synchronized in the time domain by aligning their initial response features to correct the time delay between the two acquisition systems. The laser displacement sensor was configured at a sampling rate of 1000 Hz during the measurement. To match the 30 Hz frame rate of the camera, the laser signal was resampled to the camera's time points using linear interpolation.

For the phase-based method, the CSP was adopted for phase extraction, and the scale level for displacement measurement was determined using the previously proposed strategy in Ref. [33]. Under all illumination intensities, the optimal scale level determined by this strategy is 6, based on the motion range and target size. To validate this selection, we extracted displacement measurements from all scale levels in the CSP and compared them with the ground truth. The point of interest (POI) was set at the center of each subband. The mean absolute error (MAE) at each scale level is plotted in Fig. 7 for different illumination intensities. As shown, scale level 6 achieves the minimum MAE across all illumination conditions, confirming the effectiveness and robustness of this strategy, even under low-light conditions.

For the intensity-based optical flow method, a tracking point was selected on the vertical edge of the column. This location provides a clear intensity gradient along the motion direction, which is favorable for LK optical flow tracking. The window size was set to 30×30 pixels, the pyramid level was set to 3, and the iteration termination criteria were set to a maximum of 10 iterations or a convergence threshold of 0.03. These parameters were kept consistent across all cases listed in Table 1.

For the deep learning-based optical flow method, the pretrained RAFT-Things model was adopted for optical flow estimation with its default inference settings, including 20 iterations for the recurrent update operator. Although RAFT was originally developed for general optical flow estimation, recent studies have demonstrated its effectiveness in structural displacement measurement. Specifically, Guo et al. [24] applied the same RAFT-Things model directly to structural vibration measurement and achieved sub-pixel accuracy, confirming its suitability for this task. It should be noted that no fine-tuning or domain adaptation was applied to the RAFT model in this study. This is because the ground truth available consists only of single-point displacement data, which does not satisfy the requirement of dense full-field displacement supervision for RAFT's training framework. Furthermore, the limited number of structural vibration videos in this study does not constitute a statistically representative dataset for proper partitioning into training and testing subsets. Performing fine-tuning on such a small dataset would essentially amount to ground-truth-guided optimization

Table 2

MAE results of different vision-based methods under clear conditions with varying illumination intensities.

Illumination intensity (lux)	MAE (mm)		
	Phase-based optical flow (CSP)	Intensity-based optical flow (LK)	Deep learning-based optical flow (RAFT)
124.3704	0.2270	0.3969	0.5301
62.6212	0.2990	0.4182	0.5655
15.7353	0.3900	0.5249	0.5669
0.9444	0.7583	1.6429	0.8013

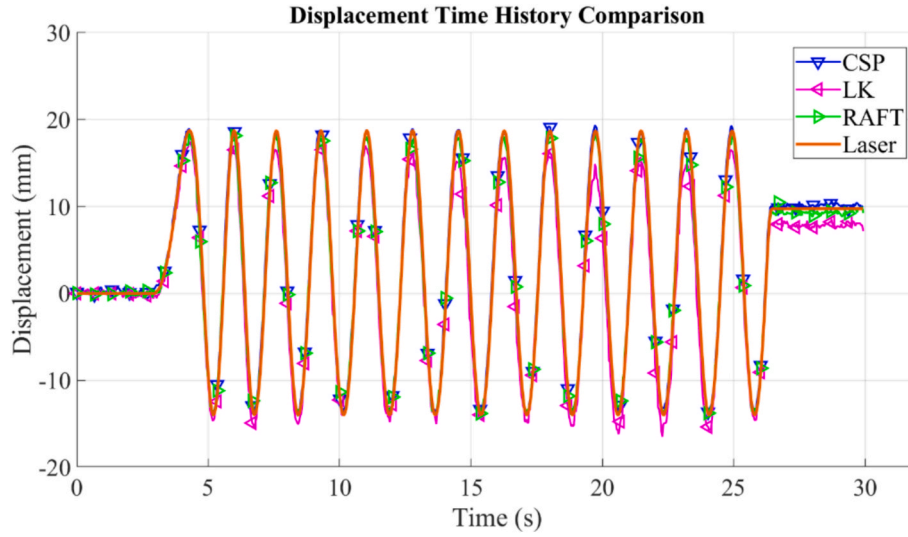


Fig. 9. Displacement time histories obtained by vision-based methods under 0.9444 lx illumination intensity.

on test data, rather than a meaningful training process that reflects the model's generalization capability. These settings were kept consistent across all cases listed in Table 1.

In this way, the compared methods were evaluated under reasonably optimized or representative settings, so that the comparison would reflect their actual measurement performance rather than differences caused by inappropriate parameter selection.

The measurement results of Case 1 based on different vision-based methods are presented in Fig. 8 and Table 2. As shown in Fig. 8 and Table 2, under ideal lighting conditions (e.g., 124.3704, 62.6212, and 15.7353 lx), all vision-based methods achieve good accuracy with low MAE values. Among them, the phase-based method consistently achieves better accuracy than the others, demonstrating the high measurement accuracy of the phase-based method.

However, under low illumination conditions (0.9444 lx), the phase-based method still provides good accuracy, while the performance of the intensity-based optical flow method degrades significantly. The time history displacement measurement results of the three methods in this case are presented in Fig. 9. As observed, the phase-based method maintains very close agreement with the laser reference over the entire vibration process. By comparison, the intensity-based optical flow method shows substantial errors, especially near local extrema, together with more pronounced instability in the extracted displacement signal. The deep learning-based optical flow method performs better than the conventional intensity-based optical flow method, but still shows slightly larger deviations than the phase-based method. These observations are further supported by the MAE results in Table 2. At 0.9444 lx, the phase-based method yields the smallest error of 0.7583 mm, whereas the corresponding errors of the intensity-based optical flow method and the deep learning-based optical flow method are 1.6429 mm and 0.8013 mm, respectively. Notably, in the present experiment, where the motion range was approximately 30 mm, the MAE of the phase-based method corresponds to only 2.53% of the displacement range, which is below the 5% criterion. This demonstrates the robustness and accuracy of the phase-based method under low-light conditions.

To provide a further mechanistic explanation of the measurement accuracy differences among the vision-based methods under the varying illumination conditions in Case 1, the one-dimensional intensity and phase signals along the target motion direction are extracted from the video frames for analysis. Since the target moves in the horizontal direction of the video, the pixel gradient and phase signal in the horizontal direction play a decisive role in target localization. The POI used in all cases and vision-based methods is located at the vertical center of the

video frame, and therefore the intensity and phase signals along row 360, which passes through the target region, are extracted and presented in Fig. 10. It should be noted that the phase signals presented here are extracted from the subband at scale level 6 of the CSP, which is the scale used for displacement measurement. Due to the successive down-sampling in the pyramid structure, the resolution of this subband is only 20×12 . For ease of comparison with the original image, the horizontal coordinate of each subband pixel is mapped back to the original image coordinate by $x = i \times 2^6$, where i is the pixel index in the subband.

As observed in Fig. 10, the intensity signals across all cases generally exhibit a step-like elevation within the target region, with a distinct contrast against the background. In the 124.3704 lx and 62.6212 lx cases, the overall intensity profiles show little difference, with the target-background intensity contrast both being approximately 150 and the target edges exhibiting a pronounced intensity gradient. This indicates that for the iPhone 11 camera, the imaging quality is stable and satisfactory above 62.6212 lx, and further increasing the illumination intensity yields limited improvement in the intensity contrast and gradient. In the cases below 62.6212 lx, the target-background intensity contrast decreases noticeably with reducing illumination, dropping to approximately 120 at 15.7353 lx and approximately 25 at 0.9444 lx.

The reduced intensity contrast resulting from lower illumination implies a significant decrease in the spatial intensity gradient at the target edges. For the LK optical flow method, displacement estimation is performed by solving the optical flow equation, the core of which involves inverting the structure tensor matrix constructed from the intensity gradients within the tracking window. The eigenvalues of this matrix are directly determined by the magnitude of the intensity gradients, decreasing as the intensity gradient diminishes [21]. In the 0.9444 lx case, the target-background intensity contrast is only approximately 25, resulting in small eigenvalues of the structure tensor matrix. The LK optical flow method is still able to output tracking results under this condition, indicating that the eigenvalues remain above the effective solving threshold of the algorithm. However, due to the small eigenvalues, the precision and stability of the solution are compromised, leading to significant degradation in displacement estimation accuracy.

For the RAFT method, although as a deep learning model it does not directly solve intensity gradient equations, its core mechanism still relies on pixel feature matching. RAFT estimates optical flow by constructing a 4D correlation volume between two consecutive frames. Under low intensity contrast conditions, the discriminability of extracted features is reduced and the matching signals become ambiguous, leading to a decline in optical flow estimation accuracy. Moreover, the pre-training datasets of RAFT predominantly consist of synthetic scenes with

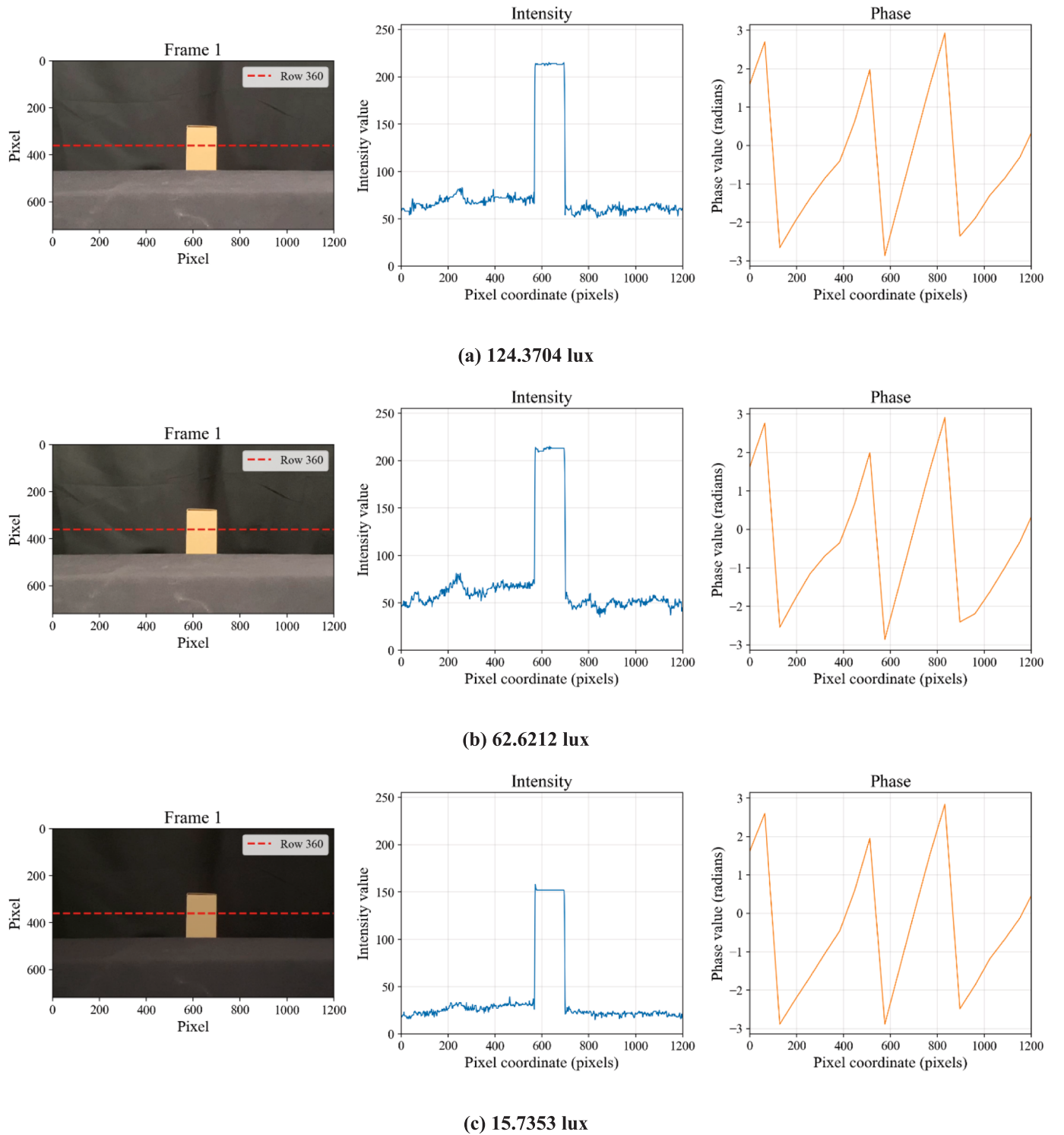


Fig. 10. Intensity and phase signals along row 360 under clear conditions with varying illumination levels.

adequate lighting and contrast conditions, limiting its generalization capability to extremely low illumination scenarios.

As shown in Fig. 10, the phase signal exhibits a fundamentally different morphology from the intensity signal, which displays a localized peak or plateau at the target region. Instead, the phase signal presents a more periodic sawtooth-like distribution, with values primarily confined within the range of approximately $-\pi$ to π and abrupt transitions from positive to negative values occurring at several locations due

to phase wrapping. Despite these spatial discontinuities, the phase signal maintains a relatively smooth variation trend between adjacent wrapping points, reflecting the local structural changes in the image along the horizontal direction. For phase-based displacement measurement methods, such phase characteristics constitute a critical signal source for target localization. It should be noted that since the CSP filter possesses a certain frequency bandwidth, the extracted signal represents the superposition of multiple frequency components within a frequency band.

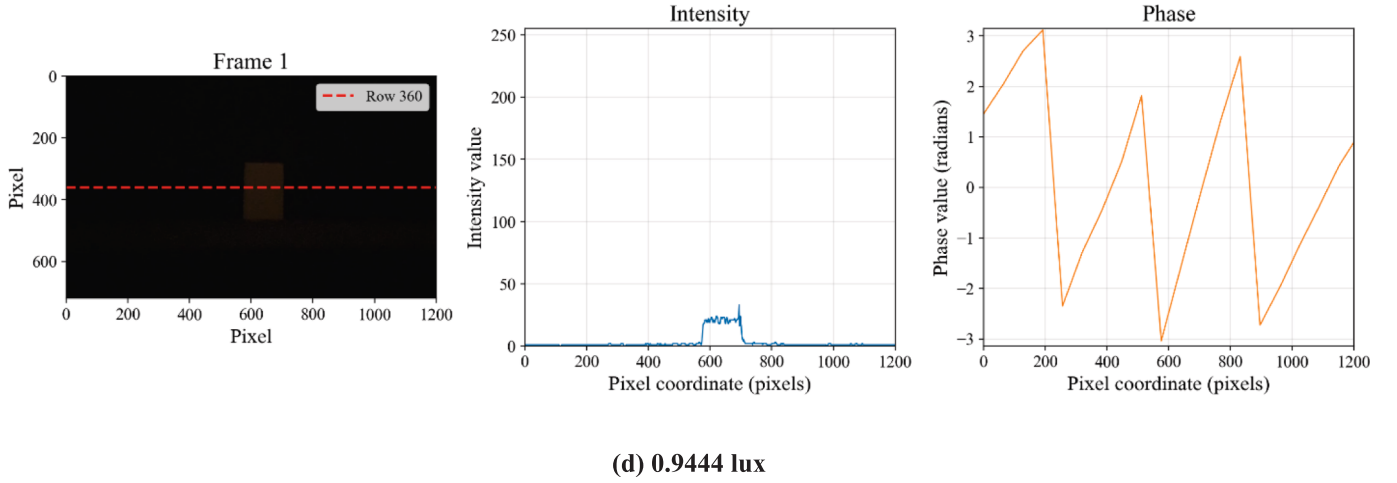


Fig. 10. (continued).

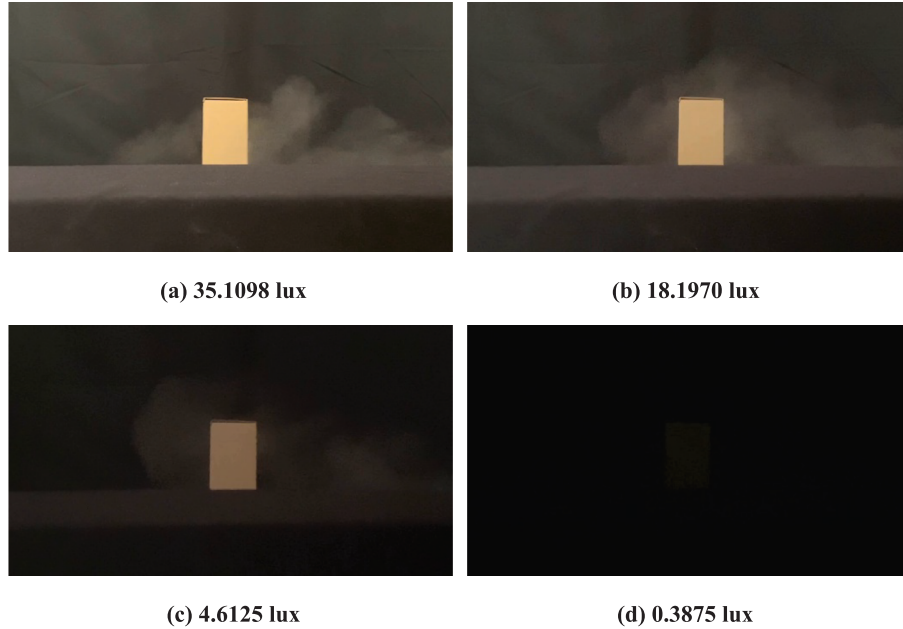


Fig. 11. Frames from experiment videos in Case 2 under foggy conditions with varying illumination levels.

As the phases of these individual components are not perfectly aligned, the synthesized phase values do not necessarily reach $\pm \pi$ at every extremum.

In contrast to the intensity signal, the phase signal exhibits a high degree of consistency across the different illumination conditions in Case 1, with all four cases presenting three complete sawtooth-like cycles of similar amplitude range near $\pm \pi$. In particular, in the vicinity of the target region, the phase signal maintains a relatively smooth and locally near-linear variation, which is a critical foundation for accurate displacement estimation and target localization [29]. Therefore, the results in Fig. 10 demonstrate from the perspective of phase signal morphology that the phase signal possesses strong robustness against illumination reduction, which explains why the phase-based method is able to maintain reliable displacement measurement accuracy under low illumination conditions.

From the perspective of measurement applicability, the result under 0.9444 lx is particularly noteworthy because such an illumination level represents an extremely unfavorable imaging condition for conventional vision-based methods. Even in this case, the phase-based method

maintains a displacement error that is small relative to the motion range of the present experiment. This indicates that phase-based displacement measurement can still provide practically meaningful results when the available illumination is very limited, whereas conventional intensity-based tracking becomes much less reliable.

In Case 2, four foggy lighting intensity levels are presented, corresponding to maximum, medium, low, and near-zero (almost off) LED lighting levels, with measured values of 35.1098 lx, 18.1970 lx, 4.6125 lx, and 0.3875 lx. The lighting intensity is reduced due to the avoidance of direct illumination, which can cause the Tyndall effect. This effect may make the fog appear too prominent, potentially leading to its misidentification as the moving target [52,53]. Frames from these cases are shown in Fig. 11.

The optimal scale level of CSP for all subcases in Case 2 is also determined as 6 via the aforementioned measurement strategy, based on the target size and motion range. The MAE at each scale level is presented in Fig. 12. As shown, even under the combined effects of varying illumination intensities and fog, scale level 6 consistently achieves the minimum MAE, indicating that the optimal scale level remains

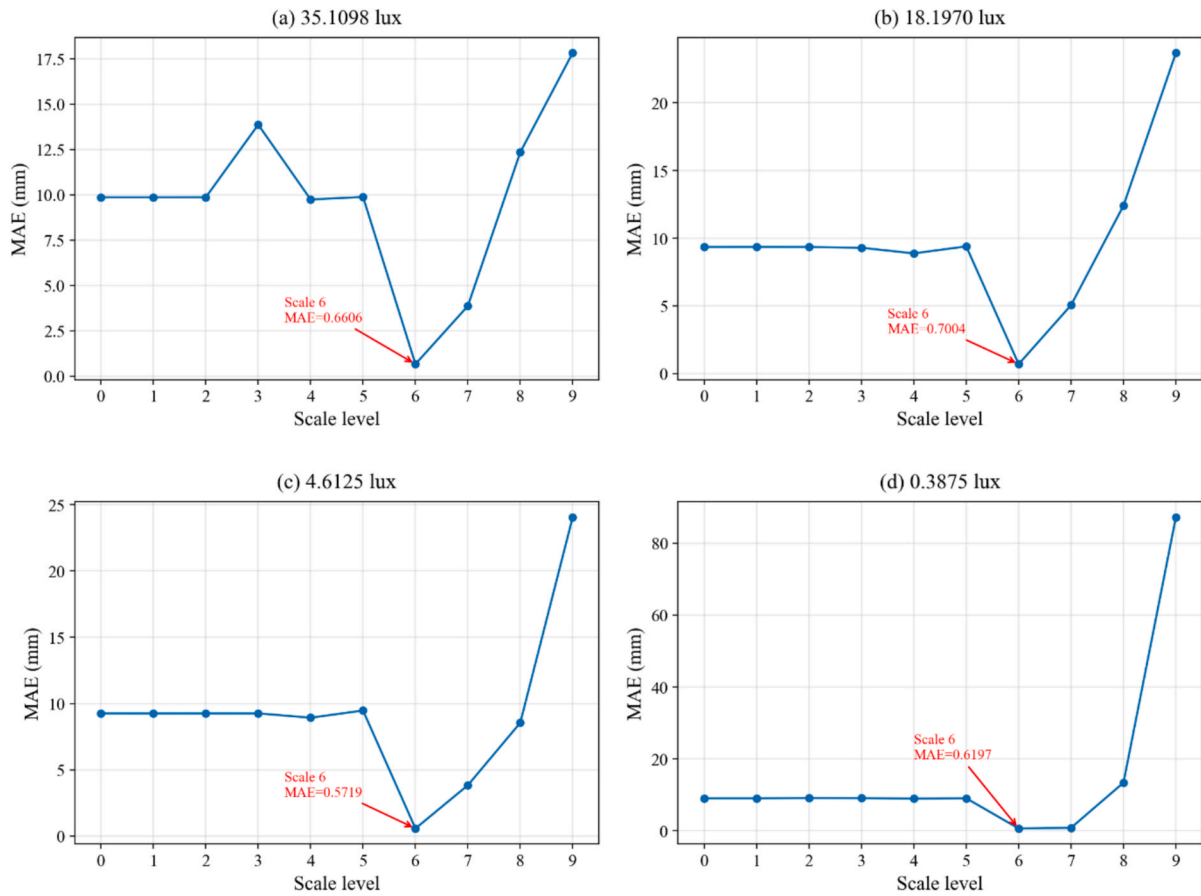


Fig. 12. Effect of scale level on MAE under varying illumination conditions in Case 2.

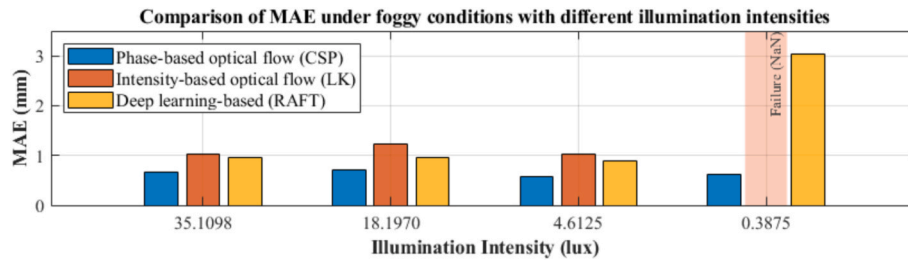


Fig. 13. Comparison of measurement performance under foggy conditions with varying illumination levels.

Table 3

MAE results of different vision-based methods under foggy conditions with varying illumination intensities.

Illumination intensity (lux)	MAE (mm)		
	Phase-based optical flow (CSP)	Intensity-based optical flow (LK)	Deep learning-based optical flow (RAFT)
35.1098	0.6606	1.0380	0.9631
18.1970	0.7004	1.2291	0.9653
4.6125	0.5719	1.0370	0.8885
0.3875	0.6197	N/A (tracking failure)	3.0272

dependent on the target size and motion range rather than environmental conditions.

The measurement results from different vision-based methods for Case 2 are presented in Fig. 13 and Table 3. As observed from Table 3,

the phase-based method is affected by the foggy condition, with its accuracy under the more ideal illumination levels of 35.1098 lx and 18.1970 lx dropping significantly compared to that at 15.7353 lx in Case 1. This may be attributed to the phase-based method is sensitive to any pattern changes, even when the pattern intensity is not significantly high, due to the nature of phase information. Other methods are also affected by the fog, showing reduced accuracy compared to Case 1. However, the phase-based method still achieves higher accuracy than the others, which may be attributed to its intrinsically high precision rather than robustness against fog.

To further investigate the mechanism by which fog degrades the measurement accuracy of all vision-based methods, the intensity and phase signals along row 360 are extracted from the fog case at 18.1970 lx and presented in Fig. 14 for analysis.

As shown in Fig. 14, for the intensity signal, the signal morphology under foggy conditions differs significantly from that under the clear scenario in Case 1 (Fig. 10). Under clear conditions, the intensity in the target region presents a rectangular pulse with steep edges, with the

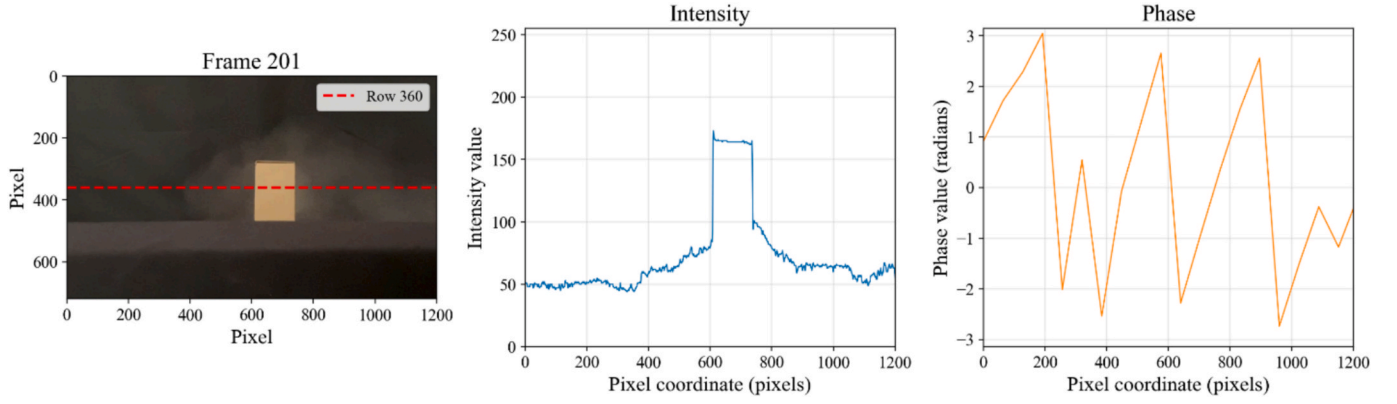


Fig. 14. Intensity and phase signals along row 360 under foggy conditions at 18.1970 lx.

transition between the target and background being nearly step-like and the background intensity remaining relatively stable. Under foggy conditions, however, the scattering effect of fog elevates the overall background intensity, reducing the contrast between the target and background. Meanwhile, the steepness of the target edges decreases noticeably, with the right edge in particular exhibiting a prolonged gradual transition, transforming the original step-like boundary into a slope-like profile and significantly blurring the intensity gradient information at the edges.

These changes in the intensity signal have direct implications for the displacement measurement methods that rely on intensity information. For the LK optical flow method, the reduction and blurring of the intensity gradient at the target edges undermine the solving stability of the optical flow equation. For the RAFT method, the edge blurring and reduced target-background intensity contrast diminish the discriminability of feature matching. Moreover, since the pre-training datasets of RAFT predominantly consist of scenes under normal visual conditions, the generalization capability of the model to fog-degraded images may be limited.

As shown in Fig. 14, the phase signal morphology under foggy conditions also differs notably from that under the clear scenario (Fig. 10). Under clear conditions, the phase signal presents three complete sawtooth-like cycles within the field of view, with regular waveforms and smooth linear segments. Under foggy conditions, however, approximately five sawtooth-like cycles appear within the same spatial range. These additional sawtooth-like cycles indicate that more non-target frequency components have been introduced into the frequency band. This is because fog introduces additional spatial intensity variations and texture structures, which are captured by the bandpass filters of the CSP and manifest as additional fluctuations in the phase signal. It

should be noted that if these additional textures were stationary, they could be regarded as a complex but fixed background, and the phase-based method would still be able to largely eliminate their influence through inter-frame phase differencing without significantly affecting the displacement measurement accuracy. However, fog varies dynamically during video recording. The phase fluctuations induced by fog are therefore inconsistent between frames. As a result, the CSP cannot distinguish these fluctuations from the actual motion of the target, effectively incorporating the dynamic scattering of fog into the displacement estimation and thereby interfering with the true measurement results.

The above analysis reveals that fog affects the measurement process in a fundamentally different manner from simple illumination reduction. Rather than merely reducing intensity contrast, fog introduces dynamic spatial interference that disrupts the local structural patterns relied upon by the phase-based method for displacement estimation. This explains why the phase-based method remains robust under low-light conditions in Case 1 but exhibits noticeable accuracy degradation in the presence of fog in Case 2.

Nevertheless, the comparative results in Table 3 still show that the phase-based method outperforms the other two methods across all tested foggy cases. This indicates that, although the phase-based method is not immune to spatially varying disturbances caused by fog, it still retains a comparatively higher level of measurement reliability.

Moreover, in the extremely low illumination case of 0.3875 lx, the impact of fog is significantly reduced, and the main limiting factor becomes the low light intensity itself. In this scenario, the deep learning-based optical flow method produces inaccurate measurement results, as shown in Table 3 and Fig. 15. Moreover, the intensity-based optical flow fails to provide results due to the extremely low intensity gradients

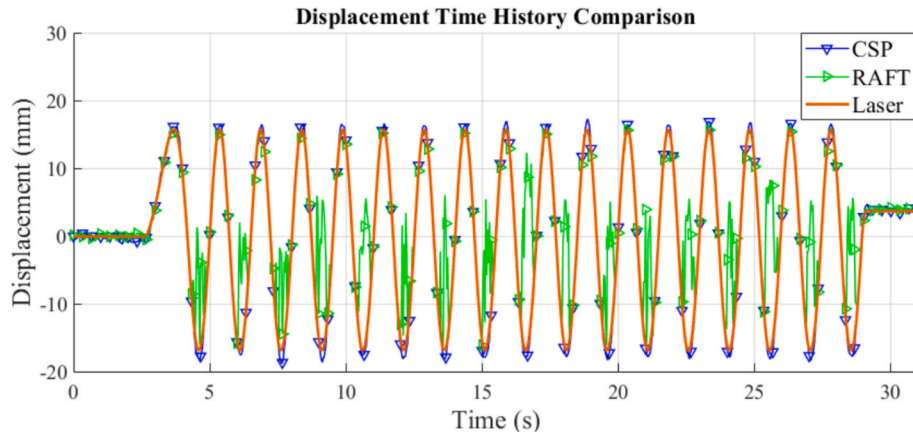


Fig. 15. Displacement time histories obtained by vision-based methods under 0.3875 lx illumination intensity and fog.

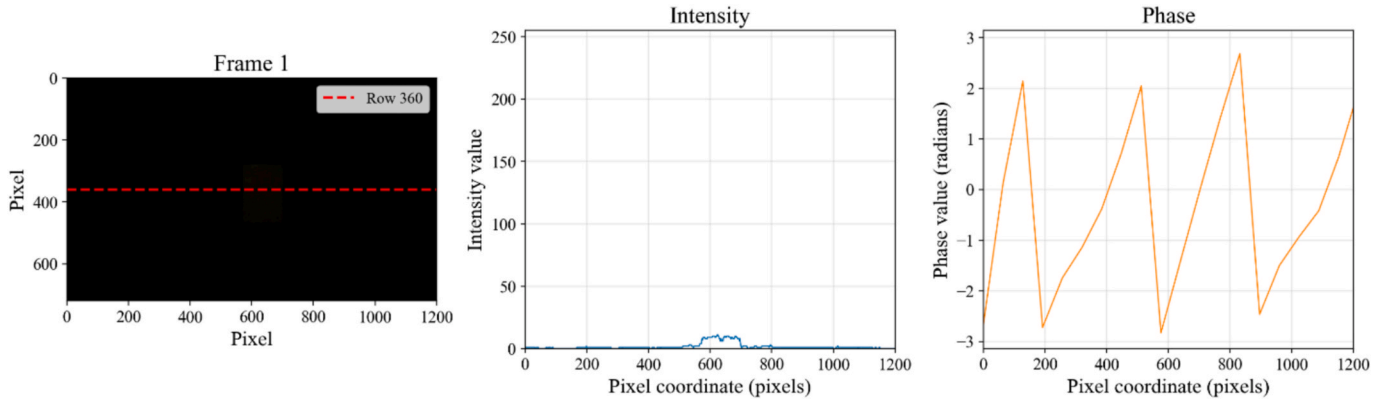


Fig. 16. Intensity and phase signals along row 360 under foggy conditions at 0.3875 lx.

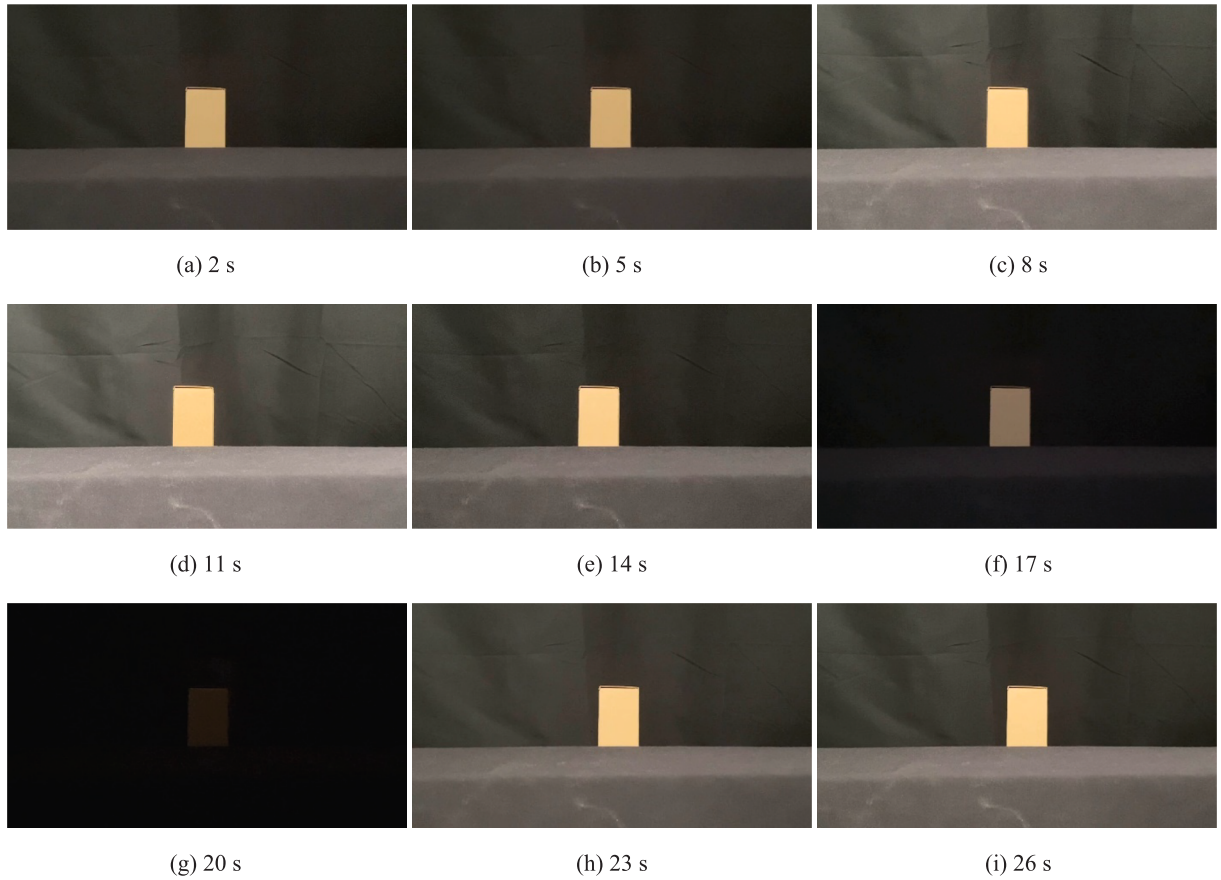


Fig. 17. Representative frames extracted from the video in Case 3 at different time instants under time-varying illumination conditions.

around the target, which makes the optical flow equation ill-posed. However, with the influence of fog diminished, the phase-based method achieves even higher accuracy than at higher illumination levels, which can be attributed to its robustness to low-light conditions.

To further investigate the mechanism behind the measurement performance of each method under the 0.3875 lx foggy condition, the intensity and phase signals for this case are presented in Fig. 16. From the intensity plot, it can be observed that the background intensity is nearly zero, with the target region exhibiting only a minor step-like elevation of approximately 10. Notably, no significant influence of fog on the intensity signal morphology is observed, as the scattering effect of fog is substantially weakened under extremely low illumination, making it difficult to form any discernible intensity pattern in the image.

However, an intensity contrast of only approximately 10 is insufficient to support the normal solving process of the LK optical flow method. As discussed earlier, the LK algorithm estimates displacement by inverting the structure tensor matrix constructed from the intensity gradients within the tracking window, and the eigenvalues of this matrix directly reflect the magnitude of the intensity gradients. Under such a low intensity contrast, the eigenvalues may have fallen below the effective solving threshold of the algorithm, rendering the matrix non-invertible and causing the algorithm to directly flag the tracking as failed. For the RAFT method, its optical flow estimation relies on pixel feature matching between two consecutive frames. Under an intensity contrast of only approximately 10, the target and background become nearly indistinguishable in the feature space, which is fundamentally

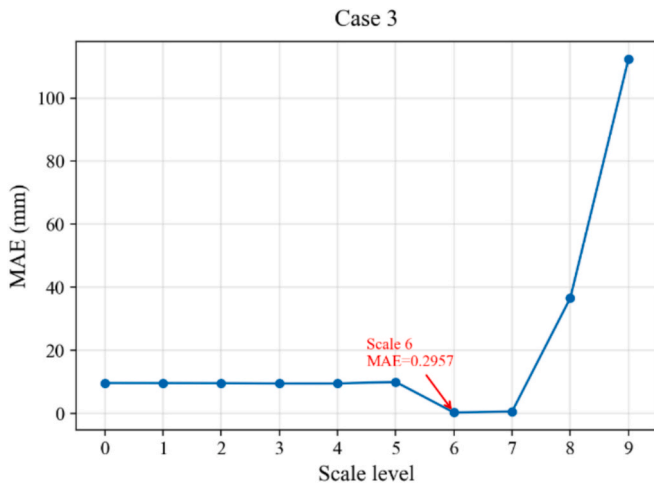


Fig. 18. Effect of scale level on MAE in Case 3.

consistent with the cause of LK failure due to insufficient intensity gradients. Regardless of whether a traditional or deep learning-based approach is employed, when the image itself no longer contains sufficient visual information to differentiate the target from the background, any displacement measurement method based on intensity information will inevitably fail.

However, for the phase signal, even under such extremely low illumination, its morphology remains similar to that in Case 1, presenting three complete sawtooth-like cycles within the field of view with regular waveforms and smooth linear segments. This indicates that the phase signal maintains good linearity and high single-frequency purity, which is favorable for accurate phase extraction and avoids interference from non-target frequency components, thereby providing a reliable signal foundation for accurate displacement measurement.

In Case 3, the lighting conditions were varied during camera recording. Representative video frames at different time instants are shown in Fig. 17. As shown in Fig. 17, the recorded image sequence exhibits substantial temporal variation in scene brightness, including a pronounced transition from relatively bright conditions to extremely low-light conditions. These frames provide direct visual evidence of the temporal illumination variation introduced in this case.

The optimal scale level in Case 3 is also determined as 6 via the aforementioned strategy, based on the target size and motion range. The

MAE at each scale level is presented in Fig. 18. As shown, scale level 6 consistently achieves the minimum MAE among all scale levels, confirming the effectiveness of this strategy even under temporal illumination variation conditions.

The time history displacement measurement results of the three methods and the illuminance time history are shown in Fig. 19. As shown in Fig. 19, the illuminance undergoes a sharp transition between approximately 17 s and 20 s. During this same period, both the intensity-based optical flow method and the deep learning-based optical flow method exhibit evident deviations from the laser reference, indicating that their displacement estimation is strongly affected by the abrupt lighting change. By contrast, the phase-based method maintains close agreement with the reference measurement throughout the entire recording, including the period of rapid illumination variation. This observation is further supported by the MAE results in Table 4, where the phase-based method achieves the smallest error of 0.2957 mm, compared with 0.9583 mm for the intensity-based optical flow method and 0.5951 mm for the deep learning-based optical flow method.

This is because intensity-based and deep learning-based methods rely heavily on raw intensity values or intensity features learned from training data. Abrupt changes in brightness can disrupt their ability to accurately identify pixel correspondences across frames. In contrast, phase information is derived from intensity and represents the trend of the intensity. As such, uniform changes in intensity amplitude have minimal impact on phase, allowing the phase-based method to remain stable under temporal illumination changes.

To provide a further mechanistic explanation of the measurement performance of each method, the intensity and phase signals at frame 510 (17 s) and frame 590 (19.67 s), during a period of significant illumination variation, are presented in Fig. 20. Based on the displacement measured by the phase-based method in pixels, a tracking point is defined at horizontal coordinate 640 as the origin, and its corresponding positions at the two frames are marked on the intensity and phase plots along with the respective intensity and phase values.

Table 4

The measurement performances of different vision-based methods under temporal illumination changes.

	Phase-based optical flow (CSP)	Intensity-based optical flow (LK)	Deep learning-based optical flow (RAFT)
MAE (mm)	0.2957	0.9583	0.5951

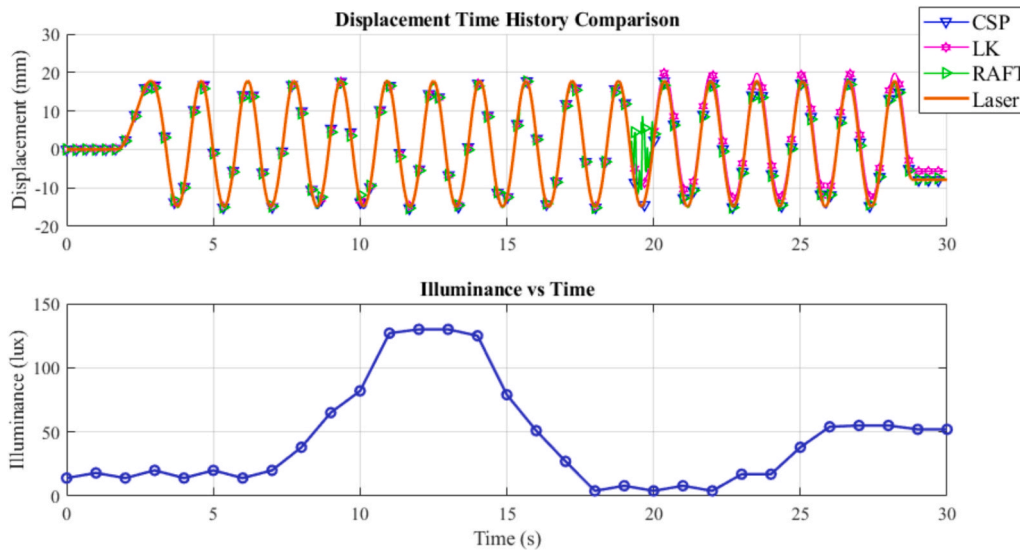


Fig. 19. Measured displacement time histories of vision-based methods under dynamic illumination with corresponding illuminance time history.

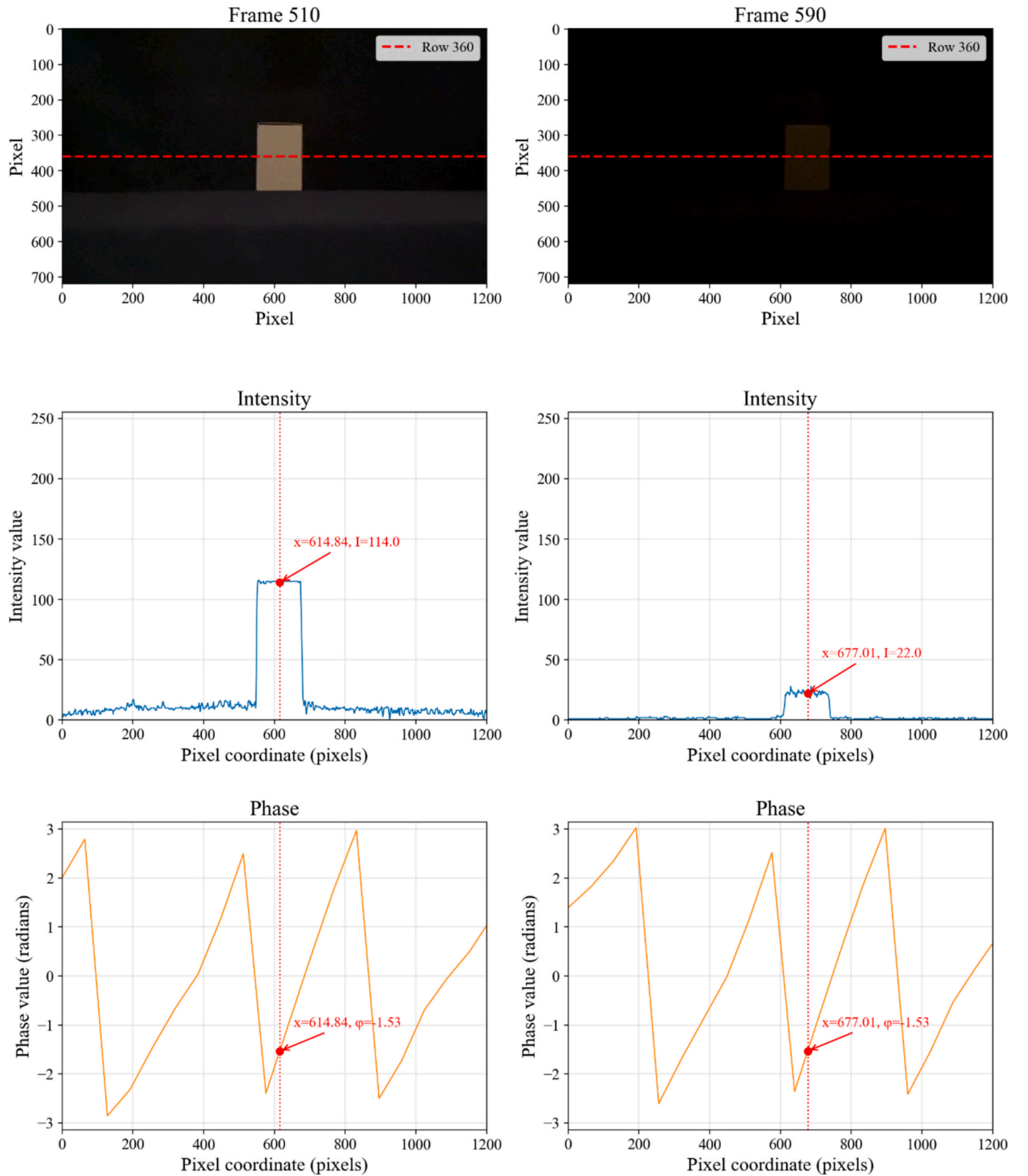


Fig. 20. Intensity and phase signals along row 360 at frame 510 and frame 590 during temporal illumination variation.

From the intensity plots, it can be observed that the intensity at the tracking point is 114 at frame 510 and drops to 22 at frame 590, indicating a significant change in the intensity values within the target region during the illumination decay.

The LK optical flow equation is built upon the brightness constancy

assumption, which requires that the intensity of the same pixel remains unchanged between consecutive frames. However, during the continuous decay of illumination, even if the target remains stationary, the pixel intensity decreases progressively with the declining illumination. The LK algorithm misinterprets such inter-frame intensity differences

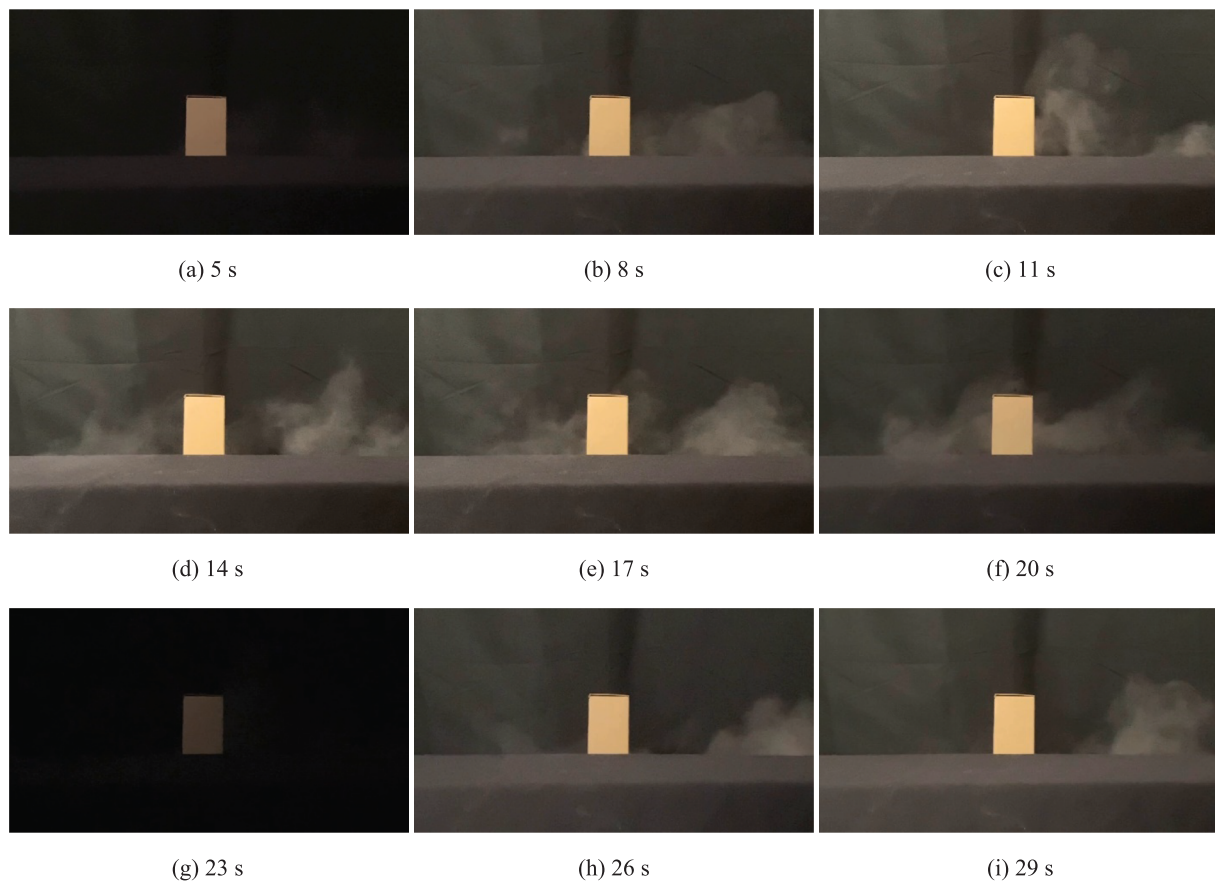


Fig. 21. Representative frames extracted from the video in Case 4 at different time instants under foggy and time-varying illumination conditions.

caused by illumination variation as target motion, thereby introducing systematic errors into the displacement estimation. Therefore, under conditions of dynamically varying illumination, the accuracy degradation of the LK method is primarily attributable to the persistent violation of the brightness constancy assumption caused by the illumination change.

For the RAFT method, the continuous decay of illumination likewise affects its optical flow estimation. RAFT extracts pixel features from two consecutive frames through a feature encoder and constructs a correlation volume for matching. When the illumination varies between frames, the feature representations of the same target differ across the two frames, causing the matching signals in the correlation volume to deviate from the true motion correspondence. Particularly near frame 590, where the target intensity has dropped to only 22 and the image approaches near-total darkness, the feature encoder of RAFT can hardly extract meaningful feature information, resulting in a complete failure of the optical flow estimation.

As for the phase signal, the phase signal morphology at frame 510 and frame 590 shows no fundamental difference, both presenting three complete sawtooth-like cycles with regular waveforms and smooth linear segments, with only a horizontal shift occurring during the process of illumination decay and target motion. The phase values at the tracking point remain highly consistent between the two frames, indicating that the drastic illumination change has not significantly affected the phase information. This implies that the phase consistency condition upon which the phase-based optical flow relies can still be satisfied under rapid illumination variation, enabling the phase-based displacement measurement to maintain stable measurement accuracy under temporal illumination variation conditions.

In Case 4, the lighting was adjusted during camera recording under foggy conditions. Representative video frames at different time instants

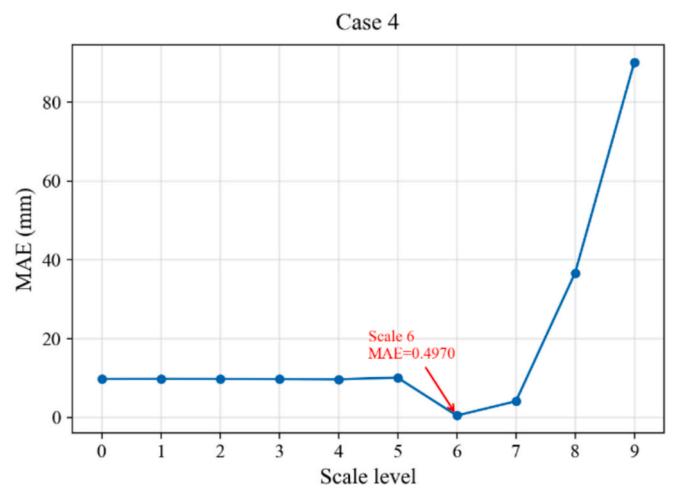


Fig. 22. Effect of scale level on MAE in Case 4.

are shown in Fig. 21. As shown in Fig. 21, both fog-induced visual interference and substantial temporal variation in scene brightness are present in the recorded image sequence. In particular, the brightness of the scene changes noticeably over time, while the fog remains distributed around the target region, making this case visually more challenging than the clear-condition case with temporal illumination variation. These frames provide direct visual evidence of the combined adverse conditions introduced in this case.

The optimal scale level in Case 4 is also determined as 6 via the aforementioned strategy. The MAE at each scale level is presented in Fig. 22. As shown, scale level 6 achieves the minimum MAE among all

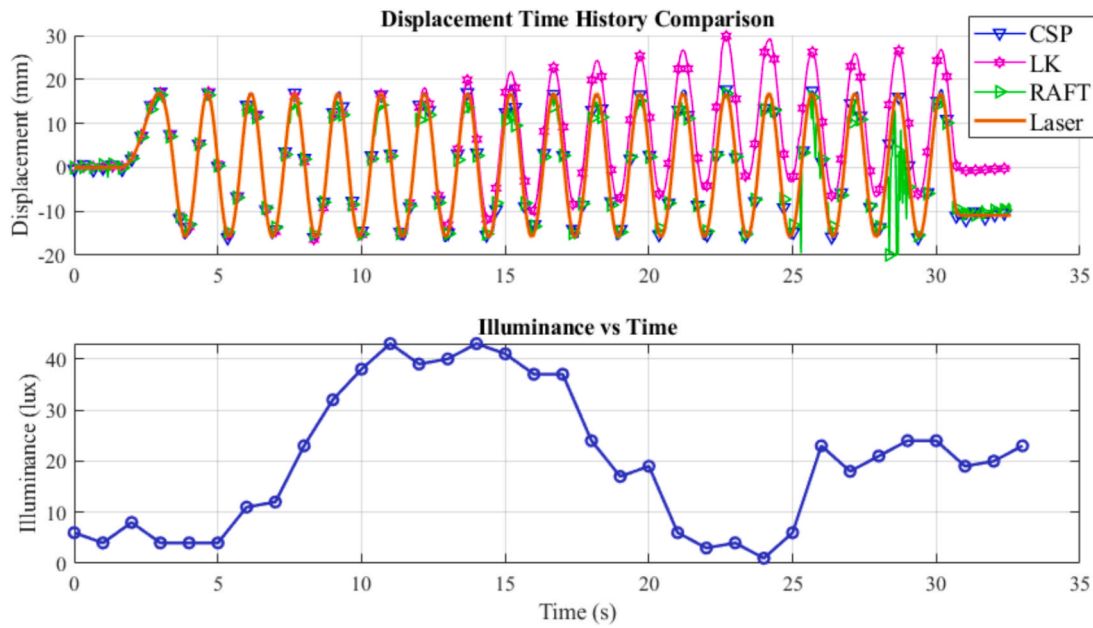


Fig. 23. Measured displacement time histories of vision-based methods under fog and dynamic illumination with corresponding illuminance time history.

Table 5

The measurement performances of different vision-based methods under fog and temporal illumination changes.

	Phase-based optical flow (CSP)	Intensity-based optical flow (LK)	Deep learning-based optical flow (RAFT)
MAE (mm)	0.4970	5.4262	1.0763

scale levels, further confirming the effectiveness of this strategy even under unfavorable conditions involving both fog and temporal illumination variations.

The corresponding measured displacement time history results, together with the illuminance time history, are presented in Fig. 23, and the quantitative comparison is summarized in Table 5. As shown in Fig. 23, the phase-based method achieves relatively higher accuracy in this case than under the foggy conditions in Case 2, while still maintaining closer agreement with the laser reference than the other two methods throughout the measurement process. By contrast, both the intensity-based optical flow method and the deep learning-based optical flow method exhibit more pronounced deviations, especially during the period of illumination transition. This observation is consistent with the MAE results in Table 5, where the phase-based method achieves the smallest error of 0.4970 mm, compared with 5.4262 mm for the intensity-based optical flow method and 1.0763 mm for the deep learning-based optical flow method. These results indicate that the phase-based method remains the most reliable among the three methods under the combined effects of fog and temporal illumination variation.

To further investigate the mechanism behind the measurement performance of each method in this case, the intensity and phase signals at frame 715 (23.83 s) and frame 810 (27 s) during the illumination variation are presented, with the tracking point marked in the same manner as described previously, as shown in Fig. 24.

From the intensity plots, it can be observed that at the low illumination moment (frame 715), the influence of fog on the intensity signal is relatively minor, and the target region still retains a step-like intensity profile. At frame 810, however, as the illumination intensity increases, the overall intensity values rise significantly, but the scattering effect of fog is also enhanced accordingly, introducing noticeable fluctuations in the background intensity and weakening both the target-background

intensity contrast and the edge gradient. The intensity value at the tracking point changes from 53 to 192, also exhibiting a substantial variation. Due to the combined effects of the dynamic temporal variation in intensity values and the blurring of target edges caused by fog, the accuracy of the LK and RAFT methods further declines compared with Case 3.

As for the phase signal, under the low illumination conditions at frame 715, its morphology is similar to that in Case 3, presenting three complete sawtooth-like cycles with regular waveforms and smooth linear segments. However, at frame 810, as the illumination intensity increases, the scattering effect of fog becomes more pronounced, and the phase signal morphology changes considerably, with approximately five sawtooth-like cycles appearing within the same spatial range, indicating that more non-target frequency components have been introduced into the frequency band. Although the phase values at the tracking point remain relatively close between the two frames (-1.47 vs -1.50), the dynamic variation of fog introduces unstable phase fluctuations between frames, interfering with the frame-by-frame phase difference estimation and leading to a decline in the displacement measurement accuracy of the phase-based method in this case compared with Case 3.

A possible explanation for the relatively better performance of the phase-based method in Case 4, compared with Case 2, is that the influence of fog is partly reduced during the low-illumination periods in the overall measurement process. As a result, the dominant disturbance may shift from strong fog-related pattern interference to temporal variation in brightness. Since the phase-based method is comparatively robust to low-light conditions and temporal illumination variation, it can still maintain stable displacement estimation in this combined-condition case. Nevertheless, some loss of accuracy is still observed relative to the clear-condition cases, indicating that the coexistence of fog-induced interference and temporal illumination change remains challenging. Overall, Case 4 further demonstrates that the phase-based method is comparatively more robust than the other two methods under the combined unfavorable visual conditions considered in the present experimental setting.

In conclusion, this experiment proves that the phase-based method achieves high measurement accuracy and is robust against low illumination and temporal lighting variations. However, the phase-based method can also be affected by the dynamic spatial interference (such as fog), since phase information is sensitive to any pattern changes, even

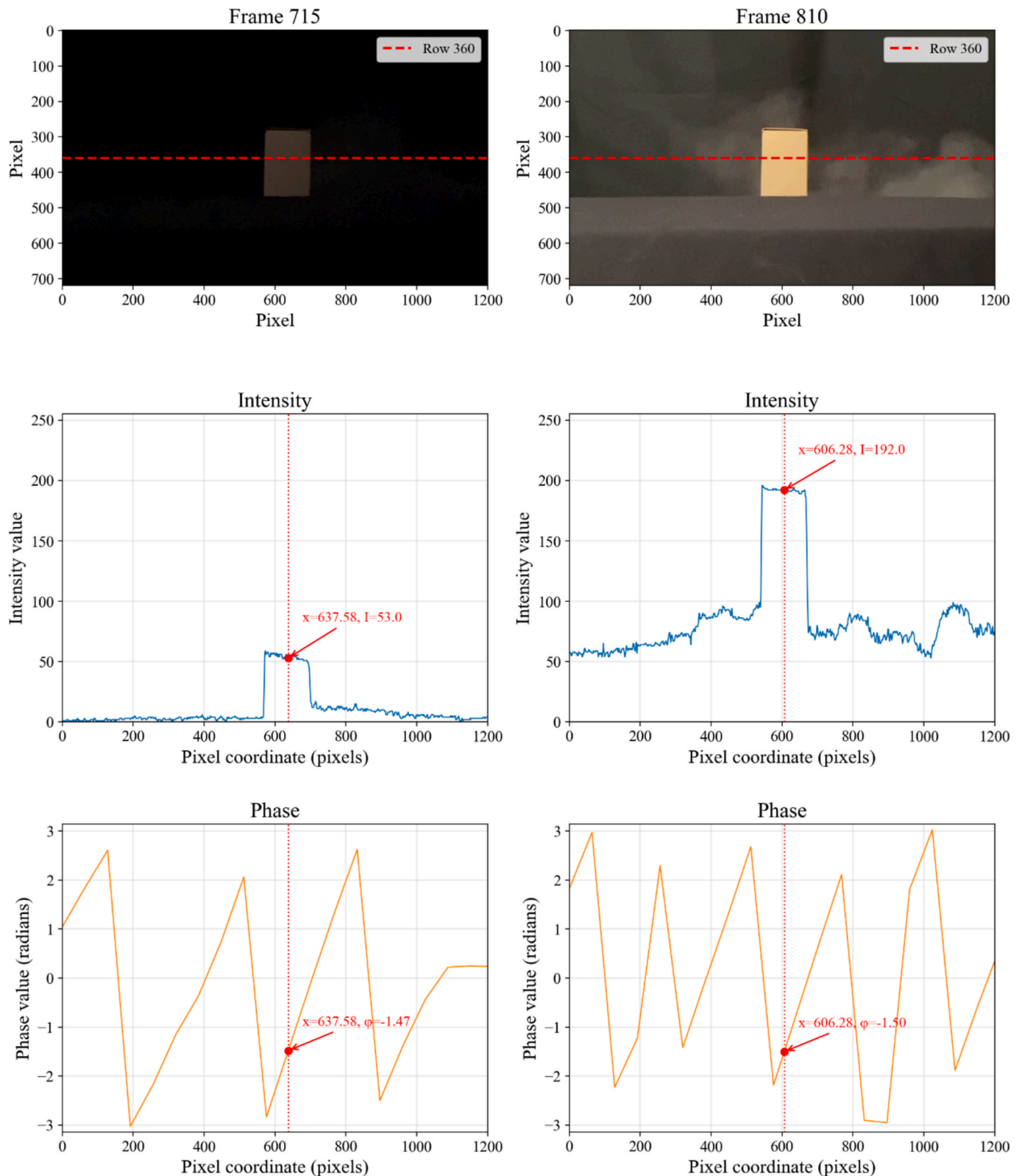


Fig. 24. Intensity and phase signals along row 360 at frame 715 and frame 810 during temporal illumination variation with fog.

when the pattern intensity is not significantly high. The superior performance of the phase-based method under low illumination conditions (below 1 lx) indicates its potential for structural displacement monitoring at nighttime or poorly illuminated environments. Its robustness against temporal illumination variation could also benefit long-term structural displacement monitoring involving day-night cycles.

4. Conclusion

This study conducted vision-based displacement measurement experiments under controlled illumination conditions and spatial disturbances. The investigated scenarios included different illumination intensities, time-varying illumination, and dynamic spatial interference.

Using laser displacement sensor measurements as the ground truth, the performance of a phase-based method was evaluated under different conditions and compared with that of commonly used intensity-based methods.

The results show that the phase-based method provides clear advantages in displacement measurement under low-illumination and time-varying illumination conditions when compared with intensity-based optical flow and deep learning-based optical flow. These findings confirm the superior robustness of the phase-based method under adverse lighting conditions, particularly in extremely low-light environments and in the presence of temporal illumination changes.

From an engineering perspective, these results indicate that phase-based displacement measurement has the potential to serve as a promising technical approach for vision-based structural displacement monitoring under adverse lighting conditions. However, it should be noted that the experiments in this study were conducted under controlled indoor laboratory conditions, which do not fully represent the complex factors involved in long-term outdoor monitoring. Future work should further investigate the deployment stability and practical performance of phase-based methods in real-world engineering environments with challenging lighting conditions, such as day-to-night transitions and nighttime scenes with little or no ambient illumination.

CRediT authorship contribution statement

Sida Ai: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Zhenkun Li:** Writing – review & editing, Visualization, Validation, Investigation, Formal analysis.

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.

Acknowledgement

The authors have no acknowledgements to declare.

Data availability

Data will be made available on request.

References

- [1] C.-Z. Dong, F.N. Catbas, A review of computer vision-based structural health monitoring at local and global levels, *Struct. Health Monit.* 20 (2021) 692–743, <https://doi.org/10.1177/1475921720935585>.
- [2] Y. Xu, J.M.W. Brownjohn, Review of machine-vision based methodologies for displacement measurement in civil structures, *J. Civil Struct. Health Monit.* 8 (2018) 91–110, <https://doi.org/10.1007/s13349-017-0261-4>.
- [3] D. Feng, M.Q. Feng, Computer vision for SHM of civil infrastructure: from dynamic response measurement to damage detection – a review, *Eng. Struct.* 156 (2018) 105–117, <https://doi.org/10.1016/j.engstruct.2017.11.018>.
- [4] F. Gomez, J.-W. Park, B.F. Spencer, Reference-free structural dynamic displacement estimation method, *Struct. Control Health Monit.* 25 (2018) e2209, <https://doi.org/10.1016/j.ymsp.2024.111398>.
- [5] S.C. Saxena, S.B.L. Seksen, A self-compensated smart LVDT transducer, *IEEE Trans. Instrum. Meas.* 38 (1989) 748–753, <https://doi.org/10.1109/19.32186>.
- [6] Y. Chen, D. Joffre, P. Avitabile, Underwater dynamic response at limited points expanded to full-field strain response, *J. Vib. Acoust.* 140 (2018) 051016, <https://doi.org/10.1115/1.4039800>.
- [7] E. Cai, Y. Zhang, X. Lu, X. Ji, X. Gao, J. Hou, J. Shi, W. Guo, Estimating structural motions in extreme environmental conditions—A dynamic correlation filter based computer vision approach, *Mech. Syst. Sig. Process.* 215 (2024) 111398, <https://doi.org/10.1016/j.ymsp.2024.111398>.
- [8] W. Zhu, Z. Cui, L. Chen, Z. Zhou, X. Chu, S. Zhu, Robust monocular vision-based monitoring system for multi-target displacement measurement of bridges under complex backgrounds, *Mech. Syst. Sig. Process.* 225 (2025) 112242, <https://doi.org/10.1016/j.ymsp.2024.112242>.
- [9] Y. Weng, S.-T. Quek, J. Ker-Wei Yeoh, Robust vision-based sub-pixel level displacement measurement using a complementary strategy, *Mech. Syst. Signal Proc.* 223 (2025) 111898, <https://doi.org/10.1016/j.ymsp.2024.111898>.
- [10] W. Guo, J. Li, Y. Hu, Vision-based displacement measurement method of large-scale bridges using tilt shift camera and fast spatio-temporal context learning, *Mech. Syst. Sig. Process.* 224 (2025) 112165, <https://doi.org/10.1016/j.ymsp.2024.112165>.
- [11] Y. Zhang, G. Li, Z. Xu, C. Xin, M. Qin, A novel hybrid method for vision-based target-free dynamic displacement measurement of civil structures under seismic excitation, *Earthq. Eng. Struct. Dyn.* 55 (2026) 199–215.
- [12] C. Lin, J. Li, H. Hao, Deep learning-based motion magnification and frames matching for structural displacement measurement using computer vision, *Eng. Struct.* 346 (2026) 121647.
- [13] Z. Zhang, A flexible new technique for camera calibration, *IEEE Trans. Pattern Anal. Machine Intell.* 22 (2000) 1330–1334, <https://doi.org/10.1109/34.888718>.
- [14] C.-Z. Dong, O. Celik, F.N. Catbas, E.J. O'Brien, S. Taylor, Structural displacement monitoring using deep learning-based full field optical flow methods, *Struct. Infrastruct. Eng.* 16 (2020) 51–71, <https://doi.org/10.1080/15732479.2019.1650078>.
- [15] X.T. Nguyen, G. Jeon, V. Vy, G. Lee, P.T. Lam, H. Yoon, A hybrid approach for vision-based structural displacement measurement using transforming model prediction and KLT, *Mech. Syst. Sig. Process.* 223 (2025) 111866, <https://doi.org/10.1016/j.ymsp.2024.111866>.
- [16] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, Speeded-up robust features (SURF), *Comput. Vis. Image Underst.* 110 (2008) 346–359, <https://doi.org/10.1016/j.cviu.2007.09.014>.
- [17] R. Brunelli, T. Poggio, Template matching: matched spatial filters and beyond, *Pattern Recogn.* 30 (1997) 751–768, [https://doi.org/10.1016/S0031-3203\(96\)00104-5](https://doi.org/10.1016/S0031-3203(96)00104-5).
- [18] P. Etchepareborda, M.-H. Moulet, M. Melon, Random laser speckle pattern projection for non-contact vibration measurements using a single high-speed camera, *Mech. Syst. Sig. Process.* 158 (2021) 107719, <https://doi.org/10.1016/j.ymsp.2021.107719>.
- [19] C. Zhang, Z. Lu, X. Li, Y. Zhang, X. Guo, A two-stage correction method for UAV movement-induced errors in non-target computer vision-based displacement measurement, *Mech. Syst. Sig. Process.* 224 (2025) 112131, <https://doi.org/10.1016/j.ymsp.2024.112131>.
- [20] H. Zhang, S. Wu, X. Luo, Y. Huang, H. Li, Efficient matching of Transformer-enhanced features for accurate vision-based displacement measurement, *Autom. Constr.* 171 (2025) 105962, <https://doi.org/10.1016/j.autcon.2025.105962>.
- [21] B.D. Lucas, T. Kanade, An iterative image registration technique with an application to stereo vision, in: *IJCAI'81: 7th International Joint Conference on Artificial Intelligence*, Vancouver, Canada, 1981, pp. 674–679.
- [22] X. Pan, T.Y. Yang, Y. Xiao, H. Yao, H. Adeli, Vision-based real-time structural vibration measurement through deep-learning-based detection and tracking methods, *Eng. Struct.* 281 (2023) 115676, <https://doi.org/10.1016/j.engstruct.2023.115676>.
- [23] Z. Teed, J. Deng, RAFT: recurrent all-pairs field transforms for optical flow, in: A. Vedaldi, H. Bischof, T. Brox, J.-M. Frahm (Eds.), *Computer Vision – ECCV 2020*, Springer International Publishing, Cham, 2020, pp. 402–419, 10.1007/978-3-030-58536-5_24.
- [24] L. Guo, W. Guo, D. Chen, B. Duan, Z. Shi, Structural vibration measurement based on improved phase-based motion magnification and deep learning, *Mech. Syst. Sig. Process.* 224 (2025) 111945, <https://doi.org/10.1016/j.ymsp.2024.111945>.
- [25] S.S. Beauchemin, J.L. Barron, The computation of optical flow, *ACM Comput. Surv.* 27 (1995) 433–466, <https://doi.org/10.1145/212094.212141>.
- [26] R. Szeliski, *Computer vision: algorithms and applications*, Springer, Nature (2022).
- [27] B. Lu, B. Bai, X. Zhao, Vision-based structural displacement measurement under ambient-light changes via deep learning and digital image processing, *Measurement* 208 (2023) 112480.
- [28] D.H. Diamond, P.S. Heyns, A.J. Oberholster, Accuracy evaluation of sub-pixel structural vibration measurements through optical flow analysis of a video sequence, *Measurement* 95 (2017) 166–172, <https://doi.org/10.1016/j.measurement.2016.10.021>.
- [29] Y. Miao, J.Y. Jeon, Y. Kong, G. Park, Phase-based displacement measurement on a straight edge using an optimal complex Gabor filter, *Mech. Syst. Sig. Process.* 164 (2022) 108224, <https://doi.org/10.1016/j.ymsp.2021.108224>.
- [30] M. Li, G. Liu, Z. Mao, Phased-based motion estimation through short-distance Hilbert transform, *Mech. Syst. Sig. Process.* 211 (2024) 111219, <https://doi.org/10.1016/j.ymsp.2024.111219>.
- [31] Y. Miao, Y. Kong, H. Nam, S. Lee, G. Park, Phase-based vibration imaging for structural dynamics applications: Marker-free full-field displacement measurements with confidence measures, *Mech. Syst. Sig. Process.* 198 (2023) 110418, <https://doi.org/10.1016/j.ymsp.2023.110418>.
- [32] S. Liu, L. Yu, W. Niu, J. Wang, Z. Zhong, J. Huang, M. Shan, Fast and accurate visual vibration measurement via derivative-enhanced phase-based optical flow, *Mech. Syst. Sig. Process.* 209 (2024) 111089, <https://doi.org/10.1016/j.ymsp.2023.111089>.
- [33] S. Ai, C. Dong, Q. Mei, The optimal scale level of complex steerable pyramid for phase-based motion estimation under different motion ranges and target sizes, *Measurement* 253 (2025) 117440, <https://doi.org/10.1016/j.measurement.2025.117440>.
- [34] S. Collier, T. Dare, Informed pixel pushing: a new method of large-motion handling for phase-based optical flow, *Measurement* 213 (2023) 112711, <https://doi.org/10.1016/j.measurement.2023.112711>.

- [35] S. Zhang, Q. Han, K. Jiang, X. Lu, G. Wang, Robust and versatile vision-based dynamic displacement monitoring of natural feature targets in large-scale structures, *J. Civ. Struct. Heal. Monit.* 14 (2024) 1819–1837.
- [36] X. Liao, Z. Cai, J. Chen, T. Liu, J. Lai, Physics-based optical flow estimation under varying illumination conditions, *Signal Process. Image Commun.* 117 (2023) 117007.
- [37] D.J. Fleet, A.D. Jepson, Computation of component image velocity from local phase information, *Int. J. Comput. Vision* 5 (1990) 77–104, <https://doi.org/10.1007/BF00056772>.
- [38] J.L. Barron, D.J. Fleet, S.S. Beauchemin, Performance of optical flow techniques, *Int. J. Comput. Vision* 12 (1994) 43–77, <https://doi.org/10.1007/BF01420984>.
- [39] S. Collier, T. Dare, Accuracy of phase-based optical flow for vibration extraction, *J. Sound Vib.* 535 (2022) 117112, <https://doi.org/10.1016/j.jsv.2022.117112>.
- [40] Y. Yang, Q. Jiang, A novel phase-based video motion magnification method for non-contact measurement of micro-amplitude vibration, *Mech. Syst. Sig. Process.* 215 (2024) 111429.
- [41] S. Ai, C. Dong, Q. Mei, A fast and accurate phase-based displacement measurement method based on complex laplacian pyramid with condensed focus frequency, *Comput. Aided Civ. Inf. Eng.* 42 (2026) 100023.
- [42] G. Liu, M.Z. Li, Z. Mao, Q.S. Yang, Structural motion estimation via Hilbert transform enhanced phase-based video processing, *Mech. Syst. Sig. Process.* 166 (2022) 108418, <https://doi.org/10.1016/j.ymssp.2021.108418>.
- [43] E.P. Simoncelli, W.T. Freeman, E.H. Adelson, D.J. Heeger, Shiftable multiscale transforms, *IEEE Trans. Inform. Theory* 38 (1992) 587–607, <https://doi.org/10.1109/18.119725>.
- [44] J. Portilla, E.P. Simoncelli, A parametric texture model based on joint statistics of complex wavelet coefficients, *Int. J. Comput. Vision* 40 (2000) 49–70.
- [45] M.Z. Li, G. Liu, Z. Mao, Q.S. Yang, J.W. Gu, Two-dimensional motion estimation using phase-based image processing with Riesz transform, *Mech. Syst. Sig. Process.* 188 (2023) 110044, <https://doi.org/10.1016/j.ymssp.2022.110044>.
- [46] Z. Shang, Z. Shen, Multi-point vibration measurement and mode magnification of civil structures using video-based motion processing, *Autom. Constr.* 93 (2018) 231–240, <https://doi.org/10.1016/j.autcon.2018.05.025>.
- [47] Y.-J. Cha, J.G. Chen, O. Büyüköztürk, Output-only computer vision based damage detection using phase-based optical flow and unscented Kalman filters, *Eng. Struct.* 132 (2017) 300–313, <https://doi.org/10.1016/j.engstruct.2016.11.038>.
- [48] E.P. Simoncelli, W.T. Freeman, The steerable pyramid: a flexible architecture for multi-scale derivative computation, in: *Proceedings., International Conference on Image Processing*, IEEE Comput. Soc. Press, Washington, DC, USA, 1995, pp. 444–447. Doi: 10.1109/ICIP.1995.537667.
- [49] D.M. Simpson, An introduction to the discrete orthogonal wavelet transform, *Revista Brasileira De Engenharia* 9 (1993) 57–81.
- [50] C.-Z. Dong, O. Celik, F.N. Catbas, E. O'Brien, S. Taylor, A robust vision-based method for displacement measurement under adverse environmental factors using spatio-temporal context learning and taylor approximation, *Sensors* 19 (2019) 3197, <https://doi.org/10.3390/s19143197>.
- [51] M. Li, X. Chen, C. Peng, S. Du, Y. Li, Modeling the occlusion problem in thermal imaging to allow seeing through mist and foliage, *J. Opt. Soc. Am. A* 36 (2019) A67, <https://doi.org/10.1364/JOSAA.36.000A67>.
- [52] M. Bender, The use of light scattering for determining particle size and molecular weight and shape, *J. Chem. Educ.* 29 (1952) 15, <https://doi.org/10.1021/ed029p15>.
- [53] X.W. Ye, T.-H. Yi, C.Z. Dong, T. Liu, Vision-based structural displacement measurement: System performance evaluation and influence factor analysis, *Measurement* 88 (2016) 372–384, <https://doi.org/10.1016/j.measurement.2016.01.024>.