*infrastructures*

MDPI

*Article*

# A Target-Free Vision-Based Method for Measuring Girder Rigid-Body Displacement Under Long-Distance Imaging Conditions

**Guangyu Li [1,2], Hai-Bin Huang [3], Shengzhi Ai [3,*], Yuan Cheng [2] and Dong Liang [3]**

[1] School of Economics and Management, Hebei University of Technology, Tianjin 300401, China
[2] Hebei Expressway Group Limited, Shijiazhuang 050081, China
[3] School of Civil and Transportation Engineering, Hebei University of Technology, Tianjin 300401, China; hbhuang@hebut.edu.cn (H.-B.H.); 1998049@hebut.edu.cn (D.L.)
* Correspondence: 202431603078@stu.hebut.edu.cn

**Abstract**

The rigid-body displacement of bridge girders, particularly the lateral displacement of curved girder bridges, is a critical indicator reflecting the structural safety reserve and durability of bridges. However, under long-distance imaging conditions, the inherent scale ambiguity and perspective distortion in monocular vision measurement, coupled with environmental interferences such as weakened natural edges and varying illumination, pose severe challenges to target-free, high-precision, and real-time displacement measurement. To this end, this paper proposes a target-free visual method for measuring rigid-body displacement of bridge girders under long-distance imaging. By fusing optical flow and Hough transform to extract seismic block edges and adopting hierarchical NCC matching for stable girder tracking, the method achieves millimeter-level accuracy, real-time performance, and strong illumination robustness. Model tests and field validation confirm its effectiveness for low-cost bridge health monitoring.

**Keywords:** bridge structural monitoring; girder rigid-body displacement; target-free measurement; seismic block; optical flow–Hough fusion; hierarchical template matching

## 1. Introduction

The rigid-body displacement of the girder, which refers to the translational and rotational components of the main girder as a rigid body in space, is a core mechanical indicator reflecting the stress state and safety reserve of bridge structures. Among these components, the lateral displacement component is directly related to the relative position changes between the girder and supporting members such as the bent cap, abutment back, and seismic block, and is of critical significance for evaluating structural stability and durability [1]. For instance, in curved girder bridges, the periodic lateral dynamic action induced by vehicle centrifugal force may cause the gap between the girder and the block to gradually narrow cumulatively, and even trigger overall lateral offset, which directly threatens driving safety and structural durability [2]. Therefore, achieving high-precision, real-time, and continuous monitoring of the rigid-body displacement of the girder (especially lateral displacement) possesses important engineering value for bridge structural health monitoring and condition assessment [3].

However, existing displacement monitoring technologies struggle to efficiently and cost-effectively meet the above requirements in practical engineering. Traditional methods

Check for updates

Academic Editors: Xiaodong Song, Qi Li and Chao Zou

Received: 8 April 2026
Revised: 30 April 2026
Accepted: 1 May 2026
Published: 6 May 2026

**Copyright:** © 2026 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.

mainly rely on instruments such as contact displacement meters [4] and total stations [5], but they suffer from significant limitations in engineering practice: contact-based methods require physical installation, involve complex construction, provide limited monitoring points, and are vulnerable to environmental interference [6]; although non-contact instruments avoid structural contact, most of them depend on artificial targets or aiming points, resulting in high deployment and maintenance costs and insufficient real-time performance [7]. More importantly, in many existing bridge monitoring systems, surveillance cameras are the only available on-site observation means. How to use these existing devices to achieve rigid-body displacement monitoring of the girder in a target-free, low-cost, and high-precision manner has become a critical issue that urgently needs to be addressed.

In recent years, machine vision technology has gradually become an important research direction in structural health monitoring owing to its advantages of non-contact, high precision and low cost [8]. C.Z. Dong et al. [9] proposed a machine vision-based method for dynamic response measurement and modal parameter identification of bridge structures, verifying the accuracy of the vision-based approach. Xie et al. [10] combined the YOLOv5 and DeepSORT algorithms for displacement monitoring during incremental launching construction of bridges. Hong et al. [11] proposed a feature matching algorithm integrating ROI image interpolation and gray projection to identify the vibration displacement of reinforced concrete beams. Zhang et al. [12] conducted structural displacement monitoring based on Mask R-CNN, with an average error of 2.05% for static displacement. Wu et al. [13] adopted a deep learning super-resolution algorithm to improve the performance of vision-based measurement. Although the above studies have promoted the development of vision measurement technology, their direct application to rigid-body displacement monitoring of bridge girders still faces the following core bottlenecks: (1) Contradiction between target dependence and practical applicability: To pursue high precision and robustness, most existing vision measurement schemes still cannot avoid reliance on artificial targets [14] (e.g., QR codes, circular markers). (2) Geometric distortion in long-distance imaging [15]: Bridge surveillance cameras are usually installed tens of meters away, resulting in reduced target feature scale, significant perspective distortion and lack of depth information, which affect the accurate mapping from pixels to physical quantities. (3) Detection instability caused by weakened natural edges and environmental variations: Under strong light, backlight, rain, fog and occlusion conditions, natural edges tend to disappear or break, leading to insufficient reliability of traditional Hough or Canny-based methods [16]. (4) Template matching vulnerable to weakened texture and illumination changes: The girder texture is weak in long-distance views, making traditional normalized cross-correlation (NCC) matching prone to drift, while matching based only on local features struggles to resist motion blur and drastic brightness changes [17].

To address the above bottlenecks, this paper takes seismic blocks as the fixed measurement reference and proposes a target-free vision-based method for rigid-body displacement monitoring of bridge girders under long-distance imaging conditions. The main contributions of this paper are as follows: (1) A dynamic edge line extraction algorithm based on optical flow–Hough collaborative fusion is proposed, which utilizes the temporal continuity of optical flow tracking and the geometric consistency constraint of Hough transform to achieve adaptive and robust extraction of seismic block edges under dynamic illumination. (2) A hierarchical normalized cross-correlation matching strategy is designed, which solves the problem of stable girder tracking under long-distance and weak-texture conditions through autonomous switching between “local fast tracking” and “global robust compensation”. (3) A complete target-free pixel-to-physical scale mapping and real-time visualization system is developed, which enables millimeter-level real-time monitoring of girder rigid-body displacement using existing surveillance cameras without relying on

artificial targets. Verified by test model tests and real bridge applications, the proposed method exhibits significant advantages in measurement accuracy, real-time performance and environmental robustness, providing a scalable technical approach for low-cost health monitoring of bridge structures.

## 2. Visual Measurement Method for Girder Lateral Displacement Based on Seismic Block Reference

This paper takes a high-pier curved beam bridge with a 150 m radius (Figure 1) as the engineering background. Taking seismic blocks as the fixed measurement reference, this paper proposes a target-free vision measurement method with engineering implementability, algorithmic robustness and real-time performance to address the difficulty in stably measuring the lateral rigid-body displacement of bridge girders under long-distance imaging conditions. As illustrated in Figure 2, there is a noticeable lateral gap between the main girder and the seismic block, which represents the lateral displacement of the girder relative to the block that is to be measured. By using the natural edges of seismic blocks and the natural texture of the girder as core features, the method establishes an integrated technical chain of "robust edge reference extraction—weak-texture target tracking—pixel-to-physical quantity conversion", thereby realizing reliable measurement of the girder's lateral displacement relative to the blocks. Compared with traditional schemes relying on artificial targets or single detection algorithms, the proposed method demonstrates higher stability and stronger engineering adaptability under complex conditions such as long distance, weak texture and illumination variations.



**Figure 1.** A curved girder bridge with high piers and a radius of 150 m.




**Figure 2.** Main girder–stopper clearance in a high-pier curved girder bridge.

### *2.1. Overall Method Architecture*

The core idea of the target-free real-time measurement method proposed in this paper is to perform robust tracking and positioning on the seismic block (fixed reference) and the girder (moving target) respectively through monocular camera video streams, and then calculate the lateral distance between them to obtain the lateral rigid-body displacement component of the girder. The overall processing flow of the method is shown in Figure 3, which consists of four key stages: initialization, block edge detection, girder tracking, and distance calculation.

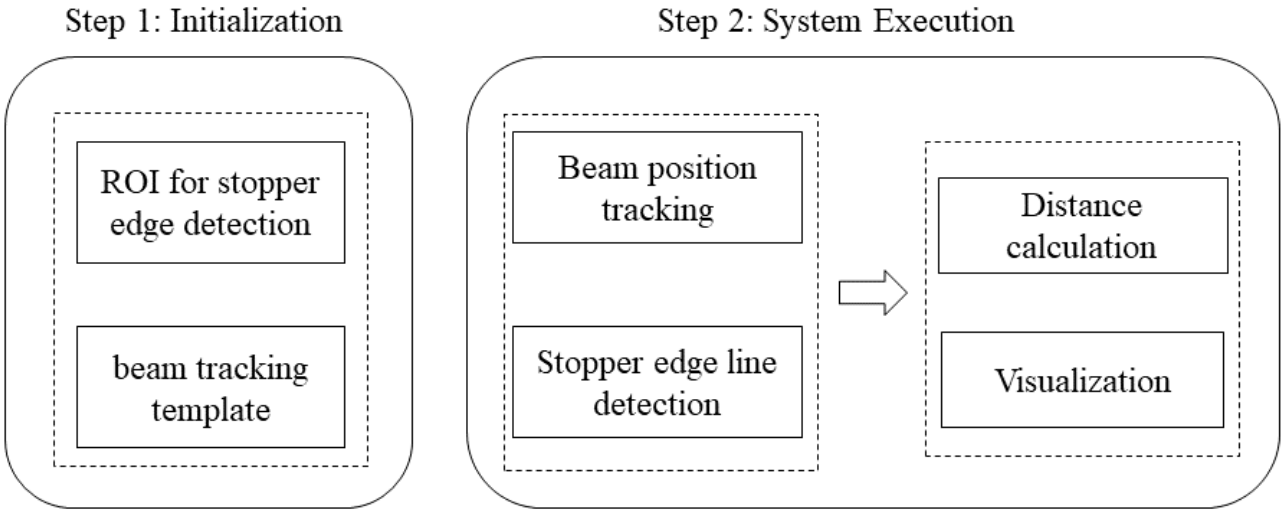



**Figure 3.** Overall method flowchart.

Initialization stage: In the first frame, the region of interest for seismic block edge detection and the template region for girder tracking are interactively defined respectively. This restricts computations to key structural components, significantly reducing background complexity and improving the real-time performance of subsequent detection and tracking.

Real-time block edge detection stage: To address the issues such as frequent breakage of natural edges and significant illumination interference in long-distance monitoring environments, an adaptive fusion strategy combining optical flow prediction and Hough geometric reconstruction is adopted. The optical flow module achieves rapid prediction by exploiting the temporal continuity of edges, while the Hough module compensates for the failure of optical flow under weak-texture or partial occlusion conditions through geometric constraints, thereby realizing robust detection of the linear position of the block and establishing a stable spatial reference for displacement measurement.

Real-time girder position tracking stage: A hierarchical normalized cross-correlation hybrid template matching algorithm is constructed to address the problems of weak texture at the girder end, sensitivity to illumination, and susceptibility to temporary occlusion. The local NCC is responsible for rapid position updating, while the global NCC is used to compensate for large displacement variations, ensuring stable and continuous tracking under various complex imaging conditions.

Distance calculation and visualization stage: Based on the linear equation of the block edge and the girder center position, the lateral distance between the girder and the block is calculated using the perpendicular distance from a point to a line, and smoothing filtering is applied to suppress accidental noise. The measurement results are overlaid on the video stream in real time to form a visual output that can be directly used for engineering monitoring.

### *2.2. Robust Extraction of Seismic Block Edge Lines Based on Optical Flow and Hough Transform*

Under long-distance imaging conditions, the seismic block serves as a fixed reference for displacement measurement. Its edge usually occupies only a few pixels in width and is susceptible to illumination variations, noise interference, and partial occlusion, resulting in broken, blurred, or even partially missing edge points [18]. The rigid edge of the block appears as a straight line in the image, and its accurate detection forms the basis for subsequent displacement measurement. As traditional single algorithms struggle to balance real-time performance and robustness, this paper proposes a straight-line extraction method based on adaptive fusion of optical flow prediction and Hough geometric reconstruction. Through the collaboration of continuity constraints and geometric consistency constraints, reliable edge localization under weak-feature scenarios is achieved, whose flow is shown in Figure 4.

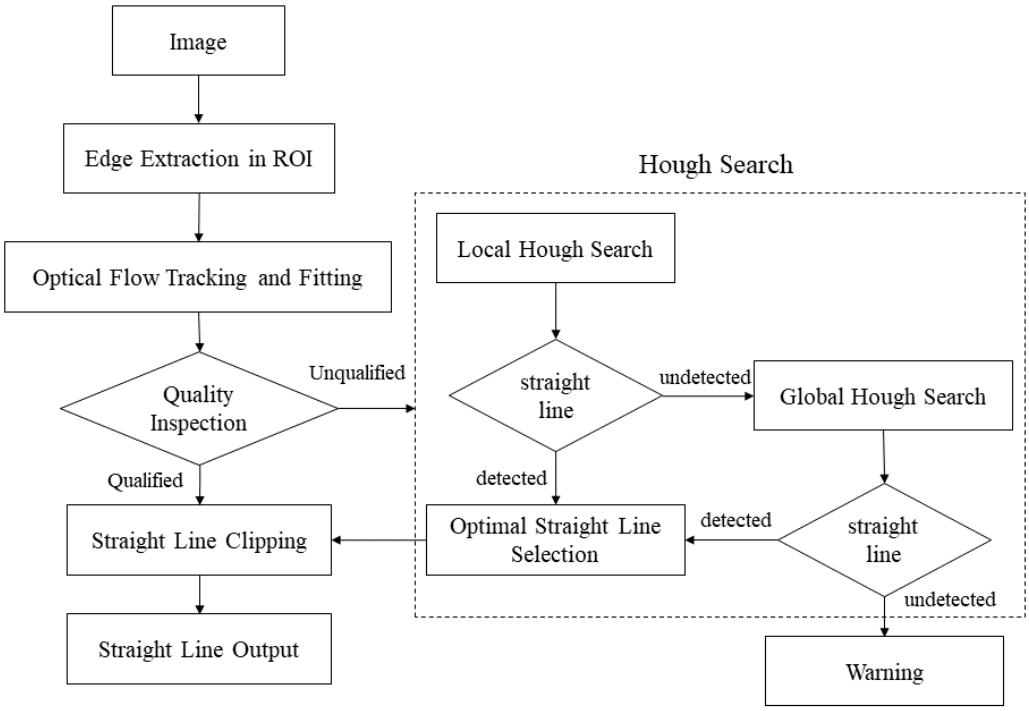



**Figure 4.** Flowchart of the stopper edge line detection algorithm.

#### 2.2.1. Motion Prediction Based on Sparse Optical Flow

To achieve efficient tracking in high-speed and low-noise scenarios, the algorithm preferentially employs prediction based on the temporal continuity of motion. Provided that the straight line $L_{t-1}$ is successfully detected in the previous frame, the algorithm takes it as prior information to predict its position in the current frame $t$.

First, $N$ feature points are uniformly sampled on the straight line $L_{t-1}$ to form the point set $P_{t-1}$. Then, the pyramidal Lucas–Kanade optical flow method is adopted to track the corresponding positions $P_t$ of these feature points in the current frame. To eliminate mismatched points, forward–backward validation is introduced: for each point in $P_{t-1}$, its trajectory from frame $t-1$ to frame t (forward) and then back from frame $t$ to frame $t-1$ (backward) is calculated, and outliers with inconsistent end points between the forward and backward trajectories are removed. The remaining high-confidence inliers form the set $P_t^{inlier}$.

Based on the inlier set $P_t^{inlier}$, the candidate line $L_{cand}$ is obtained by fitting using the least squares method. To ensure the reliability of the prediction results, strict geometric consistency verification must be imposed on it.

If all the above conditions (in Table 1) are satisfied, the optical flow prediction is deemed successful, and the candidate line $L_{cand}$ is taken as the observed line $L_t$ of the current frame. This mode has the lowest computational cost and is the key to achieving real-time performance of the algorithm.

**Table 1.** Verification conditions for geometric consistency of optical flow prediction.

| Check Item | Description | Condition |
| --- | --- | --- |
| Angle | The included angle between $L_{cand}$ and $L_{t-1}$ | Not more than 1.0° |
| Displacement | Euclidean distance between the centers of the two lines | Not more than 1.0 pixel |
| Fit | Median of the fitting residuals | Not more than 2.5 pixel |
| Tracked Points | Number of successfully tracked feature points | No less than 8 |

#### 2.2.2. Hough Transform Correction Based on Hierarchical Search

When optical flow prediction fails due to severe motion, abrupt illumination changes, or occlusion (i.e., the geometric constraints in the previous section are not satisfied), the algorithm immediately switches from the "motion prediction" mode to the "observation reconstruction" mode, and relies instead on edge information from the underlying image for line searching. To improve search efficiency and noise immunity, this study proposes a hierarchical Hough transform detection strategy.

For the current frame image $I_t$ (where $t$ denotes the frame index), preprocessing is first performed to suppress background interference and highlight target edges. To eliminate interference from cluttered background textures, this study introduces an arbitrary polygonal region of interest (ROI), denoted as $R_{poly}$. All subsequent processing (edge detection and Hough transform) is confined within $R_{poly}$ to reduce computational complexity and improve anti-interference capability. The optimal line detected in the previous frame is defined as $L_{t-1}$ (if it exists), whose parametric form can be expressed in polar coordinates: $L : \rho = x \cos\theta + y \sin\theta$.

First, the image region corresponding to the polygonal ROI is cropped and converted into a grayscale image. Subsequently, the Canny operator is adopted to extract edge information of this region, yielding the initial edge map $E_{canny}$. To connect broken edges caused by noise or uneven illumination, morphological closing operation is applied to $E_{canny}$, generating the enhanced edge map $E_{closed}$. The edge map $E_{closed}$ serves as the basic input for subsequent line detection.

If the line $L_{t-1}$ from the previous frame exists, a local search is performed. First, a band-shaped local ROI with a width of $W_{band}$ pixels centered on $L_{t-1}$ is created, denoted as $R_{lcoal}$. The probabilistic Hough transform is implemented within this region to obtain the candidate line set $\mathcal{L}_{local} = \left\{ L_i^{local} \right\}$. This strategy leverages inter-frame continuity by focusing on the small area where the target is likely to appear, thus improving detection speed.

If the local search returns no candidate line segments, or the maximum support of the candidate line segments is lower than the preset threshold $T_{supp}$, a fallback mechanism is triggered. At this point, the search range of the Hough transform is expanded to the entire region of interest $R_{poly}$, and a global probabilistic Hough transform is performed to obtain the candidate line segment set $\mathcal{L}_{global} = \left\{ L_i^{global} \right\}$. This mechanism ensures that the line can still be captured when the local search fails.

To select the line with the most significant physical meaning from the candidate set $\mathcal{L}_{cand}$ (i.e., $\mathcal{L}_{local}$ or $\mathcal{L}_{global}$), this paper designs a comprehensive scoring function $S(L_i)$ that

integrates two indicators: edge support and geometric consistency. The optimal line $L_t$ is determined by the following equation:

$$L_t = \arg\max_{L_i \in \mathcal{L}_{\text{cand}}} S(L_i) \tag{1}$$

The comprehensive scoring function $S(L_i)$ is defined as

$$S(L_i) = Supp(L_i, E_{canny}) + \lambda Geom(L_i, L_{t-1}) \tag{2}$$

where $Supp(L_i, E_{canny})$ denotes the edge support of line $L_i$ on the edge map $E_{canny}$, which is specifically defined as the cumulative number of edge pixels lying on line $L_i$.

$$\text{Supp}(L_i, E_{\text{canny}}) = \sum_{(x,y)\in\mathcal{P}(L_i)} 1_{E_{\text{closed}}}(x, y) \tag{3}$$

where $\mathcal{P}(L_i)$ denotes the set of discrete pixel points of line $L_i$ within the ROI, and $1_{E_{closed}}(x, y)$ is an indicator function that takes 1 if the point $(x, y)$ is an edge pixel in $E_{canny}$ and 0 otherwise. A larger value indicates a higher coincidence degree between the line and the actual edge.

Geometric consistency $Geom(L_i, L_{t-1})$ measures the similarity between $L_i$ and the previous frame result $L_{t-1}$ in terms of angle and position. If $L_{t-1}$ does not exist, this term is set to 0. It is defined as

$$Geom(L_i, L_{t-1}) = \exp\left(-\frac{|\theta_i - \theta_{t-1}|}{\sigma_\theta}\right) + \exp\left(-\frac{|\rho_i - \rho_{t-1}|}{\sigma_\rho}\right) \tag{4}$$

where $\theta_i$ and $\rho_i$ are the parameters of line $L_i$, and $\sigma_\theta$ and $\sigma_\rho$ are the scale parameters that control the decay rate of the angle and distance (which can be preset according to actual scenarios). This form ensures that lines closer to the previous frame result receive higher scores, thereby enhancing tracking stability in the temporal domain. $\lambda$ is a balancing coefficient used to adjust the weight of the geometric consistency term, which is usually set experimentally ($\lambda = 0.5$).

Through the above scoring mechanism, the algorithm can select the optimal line from the candidate lines that not only conforms to the current edge features but also maintains temporal continuity, thereby achieving robust detection and tracking of the seismic block as a fixed reference.

### *2.3. Girder Displacement Tracking Based on Hierarchical NCC Matching*

Under long-distance imaging conditions, the girder region exhibits characteristics of weak texture and low contrast, and abrupt outdoor illumination changes occur frequently. A single matching method can hardly balance accuracy and robustness [19]. To achieve continuous and stable tracking of the girder relative to the block reference, this paper proposes a hierarchical adaptive template matching strategy driven by environmental stability, constructing a hybrid matching framework featuring "from coarse to fine, from fast to stable". The algorithm flow is shown in Figure 5.

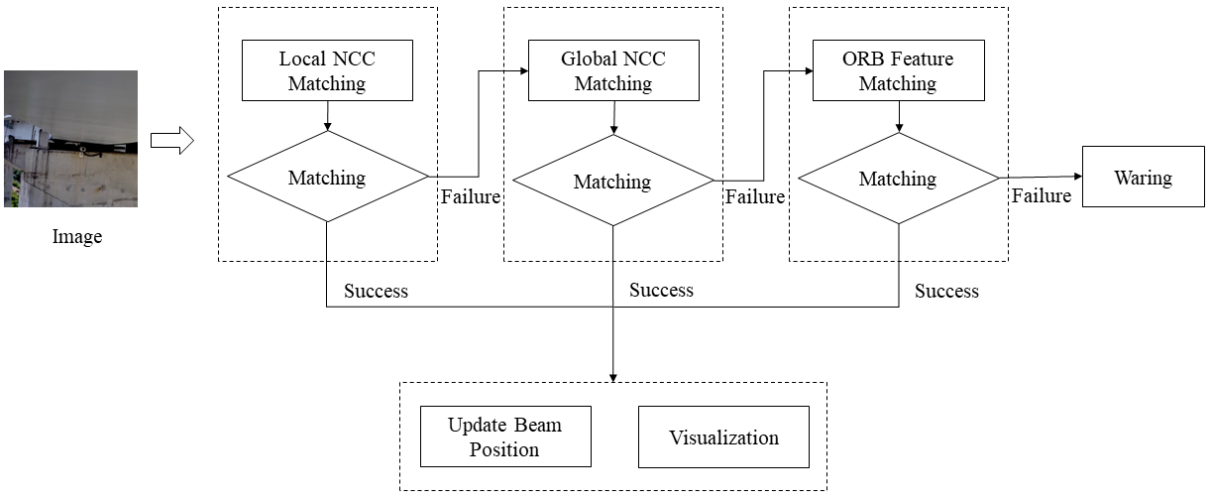



**Figure 5.** Flowchart of the hierarchical girder template tracking algorithm.

#### 2.3.1. Template Initialization and Preprocessing

During the algorithm initialization stage, a rectangular region $R_{templ}$ of the girder near the block is selected via human–computer interaction as the initial template, and the grayscale image of this region is extracted as the reference template $T$. To ensure the efficient operation of the subsequent fallback mechanism, precomputation and validity verification are performed once on the template $T$ during initialization: the ORB algorithm is used to extract stable feature points from the template $T$, with an upper limit of 800 feature points, and their corresponding descriptors $D_{templ}$ are calculated. If the number of valid feature points in the template is less than 50, the template is determined to be a weak-texture template, and the user is prompted to reselect the template region to guarantee the reliability of subsequent feature matching. This preprocessing procedure is executed only once in the initial frame, providing a unified texture reference and geometric feature basis for the subsequent full-sequence tracking.

#### 2.3.2. Hierarchical NCC Template Matching Strategy

For each new frame $I_t$, the system adaptively switches the matching strategy following the workflow shown in Figure 5.

Level 1: Local Fast Tracking. Taking the girder bounding box $B_{t-1}$ of the previous frame as the center, a local region $R_{local}$ is formed by expanding outward by 40 pixels, and NCC sliding-window matching is performed to obtain the position corresponding to the highest score. If $\rho_{\max} \geq \theta_{local}$ (where $\theta_{local} = 0.85$ in this paper), the matching is successful and $B_t$ is output.

Level 2: Global Robust Compensation. If local matching fails, the search range is expanded to a global search region $R_{global}$ formed by expanding outward by 120 pixels centered at $B_{t-1}$, and NCC matching is performed again. A lower threshold $\theta_{global}$ is adopted (where $\theta_{global} = 0.75$ in this paper). If the matching is successful, $B_t$ is output.

Level 3: Feature Fallback Assurance. If global matching still fails, ORB feature matching is activated: using the descriptor $D_{templ}$ extracted during the initialization stage, feature point matching is performed within $R_{global}$, and the current frame girder position is estimated via affine transformation.

This hierarchical matching structure dynamically adjusts the search strategy according to image quality and matching confidence, ensuring real-time tracking capability in long-distance imaging scenarios and significantly improving robustness against complex illumination and environmental disturbances.

### *2.4. Distance Calculation and Pixel–Physical Mapping Model*

To convert pixel distances to millimeter-level displacements, a reliable pixel-to-physical mapping model is established. To address the difficulty in directly determining the pixel ratio under long-distance oblique imaging, a lightweight mapping model based on on-site calibration is adopted, which implicitly compensates for the effects of camera pose, perspective distortion, and local geometric deformation.

#### 2.4.1. Pixel Distance Calculation and Filtering

Let the edge line equation of the block be $L_t : ax + by + c = 0$, and the center point of the rectangular curved girder region output by the girder tracking module be $C_t = (x_c, y_c)$. Then the pixel distance $d_p(t)$ from the center of the girder template to the block edge is calculated as

$$d_p(t) = \frac{|ax_c + by_c + c|}{\sqrt{a^2 + b^2}} \tag{5}$$

To suppress transient disturbances caused by image noise, edge extraction fluctuations, and matching errors, temporal smoothing is applied to the continuously output pixel distance sequence. This paper adopts median filtering (with a window length of 5 frames) to remove gross errors, and further introduces Exponential Moving Average (EMA) to enhance the continuity and stability of the data curve. The denoised pixel distance $\overline{d_p}(t)$ is finally obtained as the input for the subsequent physical conversion.

#### 2.4.2. Construction of Comprehensive Scale Factor and Physical Distance Mapping

Under long-distance oblique imaging conditions, the actual physical size represented by a single pixel in the image is not constant and usually exhibits a nonlinear variation along the vertical coordinate of the image. Therefore, instead of adopting a complex spatial transformation model, this paper constructs a practical mapping model suitable for engineering monitoring by combining on-site calibration and local geometric compensation.

First, the Zhang Zhengyou calibration method is used to obtain the camera intrinsic parameter matrix $K$ and distortion coefficients via a checkerboard calibration target, and undistortion preprocessing is performed on the original image to ensure the accuracy of the image geometric relationship, laying a foundation for subsequent distance measurement.

Second, the known physical dimensions of the block are obtained from the bridge design drawings. Combined with the block edge extracted from the image to obtain its pixel length $l_p$, the comprehensive scale factor $s$ is defined as

$$s = \frac{L_{real}}{l_p} \tag{6}$$

where $L_{real}$ is the design length of the block. This scale factor implicitly incorporates the combined effects of camera pose, perspective distortion, and local geometric correction, serving as a direct mapping parameter specific to the current monitoring scenario.

After obtaining the comprehensive scale factor $s$, the physical distance $d_m(t)$ of the $t$ frame in real-time monitoring is calculated as follows:

$$d_m(t) = s \times \overline{d_p}(t) \tag{7}$$

#### 2.4.3. Physical Distance Output and Visualization

The system maps the pixel distance to the millimeter-level lateral gap between the girder and the block in real time, which serves as a direct representation of the lateral rigid-body displacement component of the girder. Key data are superimposed on the

video frames in real time via a head-up display (HUD), providing visual output for engineering monitoring.

### *2.5. Method Summary*

This section proposes a target-free visual measurement method for the rigid-body displacement of the girder with the seismic block as the reference. It mainly includes: ① a robust block edge extraction algorithm based on optical flow–Hough fusion, which establishes a stable spatial reference; ② a girder displacement tracking strategy using hierarchical NCC matching, which achieves continuous and stable tracking of weak-texture targets; ③ a pixel-to-physical mapping model based on a comprehensive scale factor, which completes the engineering magnitude conversion.

The three modules work collaboratively to form a complete technical chain from image feature extraction to displacement output.

## 3. Experimental Design and Performance Evaluation

To verify the reliability, accuracy, and engineering applicability of the proposed method in long-distance imaging environments, a multi-group experimental system consisting of indoor test model tests and field condition simulations is designed, and quantitative evaluations are conducted from three aspects: algorithm performance, tracking stability, and physical measurement accuracy.

### *3.1. Experimental Design*

A girder–block test model was established (as shown in Figure 6a). The model consists of two rectangular blocks: the fixed rectangular block (length × width × height = 200 mm × 10 mm × 200 mm) is used to simulate the seismic retaining block and serves as the measurement reference; the laterally movable rectangular block (length × width × height = 2000 mm × 300 mm × 250 mm) is adopted to simulate the beam body and act as the moving target. As a small-scale laboratory specimen, this model is not geometrically scaled for a specific bridge. Its designed dimensions can sufficiently reproduce the expected relative motion and mechanical behavior between the beam body and the retaining block.

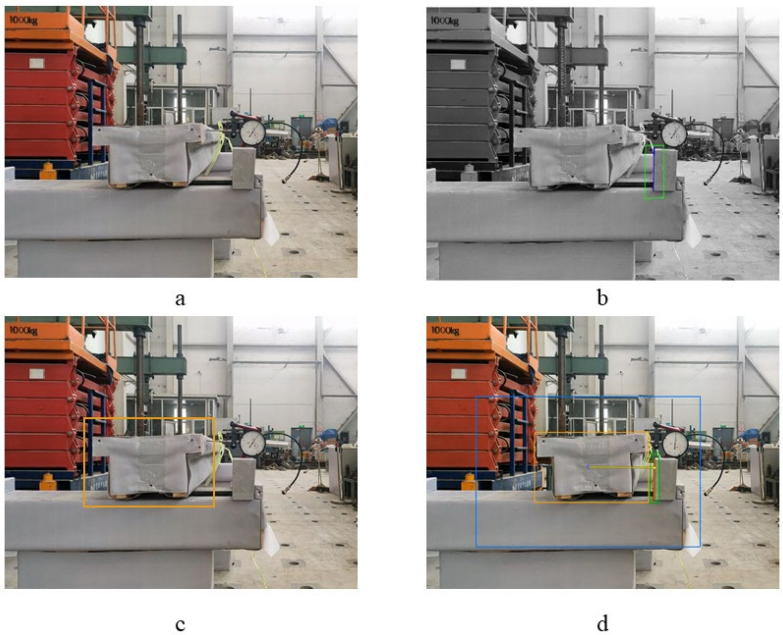



**Figure 6.** Simulation experiment results ((**a**). Overall experimental setup. (**b**). Simulated block line detection. (**c**). Girder formwork selection. (**d**). Result visualization).

The displacement of the girder was controlled and measured by a dial indicator with a measuring range of 100 mm and an accuracy of 0.01 mm, which served as the ground truth. A fixed camera with a resolution of 1920 × 1080 was used to capture the movement of the girder at a viewing distance of 5 m. The camera intrinsic parameters and distortion coefficients were obtained using the Zhang Zhengyou calibration method, with a pixel scale of 0.861 mm/pixel.

### *3.2. Evaluation Metrics and Comparison Methods*

Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are adopted to evaluate the measurement accuracy, and the average processing frame rate is used to assess real-time performance. Three baseline methods are set up for comparison (Table 2) to verify the effectiveness of the optical flow–Hough fusion strategy and the hierarchical NCC matching strategy.

**Table 2.** Comparative experimental methods.

| Name | Stopper Edge Detection | Girder Tracking |
|---|---|---|
| Proposed Method | Optical flow–Hough fusion | Hierarchical NCC matching |
| Baseline Method A | Hough Transform only | Hierarchical NCC matching |
| Baseline Method B | Optical flow–Hough fusion | Global NCC matching only |
| Baseline Method C | Hough Transform only | Global NCC matching only |

Baseline Method A is used to verify the effectiveness of the optical flow–Hough fusion module, and Baseline Method B is used to verify the effectiveness of the hierarchical NCC matching module. For Baseline Method C, only the Hough transform is applied for edge detection, and only global NCC matching is employed for beam tracking. That is, neither of the two proposed innovative modules is enabled. It serves as the most fundamental traditional process benchmark to comprehensively evaluate the overall improvement of the dual strategy proposed in this paper.

### *3.3. Accuracy and Real-Time Performance Analysis*

The proposed algorithm and the two baseline methods are tested on the laboratory video sequences. The overall performance comparison is shown in Table 3.

**Table 3.** Performance in the laboratory simulation environment.

| Method | MAE (mm) | RMSE (mm) | FPS |
|---|---|---|---|
| Proposed Method | 1.092 | 1.460 | 23.9 |
| Baseline Method A | 1.131 | 1.456 | 20.2 |
| Baseline Method B | 1.114 | 1.467 | 17.9 |
| Baseline Method C | 1.821 | 2.099 | 10.9 |

As shown in Table 3, the proposed method achieves the MAE of 1.092 mm, which is lower than those of Baseline A (1.131 mm), Baseline B (1.114 mm) and Baseline C (1.821 mm). The corresponding RMSE is 1.460 mm—slightly above that of Baseline A (1.456 mm) but below those of Baseline B (1.467 mm) and notably Baseline C (2.099 mm). Compared with Baseline C, the proposed method reduces MAE and RMSE by approximately 40.0% and 30.4%, respectively, further verifying the accuracy improvement contributed by the optical flow–Hough fusion and hierarchical NCC matching strategies.

The error characteristics are further examined to identify the dominant sources. First, environmental factors—subtle ambient light fluctuations and thermal-induced camera vibration—introduce low-frequency drift that slightly elevates the baseline error. Second, hardware limitations, including finite sensor resolution and residual calibration errors, set

a theoretical lower bound on the attainable MAE. Third, algorithmic factors contribute to sporadic larger residuals. Baseline A exhibits a higher MAE (1.131 mm) due to inherent Hough transform jitter, while Baseline B suffers from matching drift in low-texture regions (RMSE 1.467 mm). The proposed method alleviates these issues through optical flow–Hough fusion and hierarchical NCC matching, achieving the lowest MAE of 1.092 mm. However, its RMSE (1.460 mm) is marginally higher than that of Baseline A (1.456 mm). This slight performance degradation can be attributed to a small number of frames with remarkably large residuals, which may stem from the optical flow–Hough fusion under local blurring. Specifically, in regions with weak textures or indistinct edges in a very few frames, the structural displacements recovered by optical flow estimation and the Hough transform are mutually inconsistent. Such ambiguity introduces considerable random bias in individual frames. Although it has a limited impact on MAE, it is significantly amplified due to the sensitivity of RMSE to outliers.

In terms of operational efficiency, the proposed method achieves an average frame rate of 23.9 FPS, satisfying the requirements of real-time monitoring. Compared with Baseline B (17.9 FPS), the processing speed is increased by about 33.5% while simultaneously improving accuracy, demonstrating a favorable balance between precision and efficiency. Even against the considerably slower Baseline C (10.9 FPS), the speed advantage is more than twofold, highlighting the practicality of the proposed approach.

### *3.4. Illumination Robustness Analysis*

Actual illumination changes are simulated by adjusting the brightness in software, with brightness coefficients ranging from 0.300 to 1.700 applied to the original video (1.000 corresponds to the original brightness). The imaging results under different brightness conditions are shown in Figure 7.




**Figure 7.** Examples of different luminance conditions ((**a**). Normal. (**b**). Bright. (**c**). Dark).

The proposed algorithm is tested under different illumination conditions, and the measurement results of the girder–stopper spacing are shown in Figure 8.

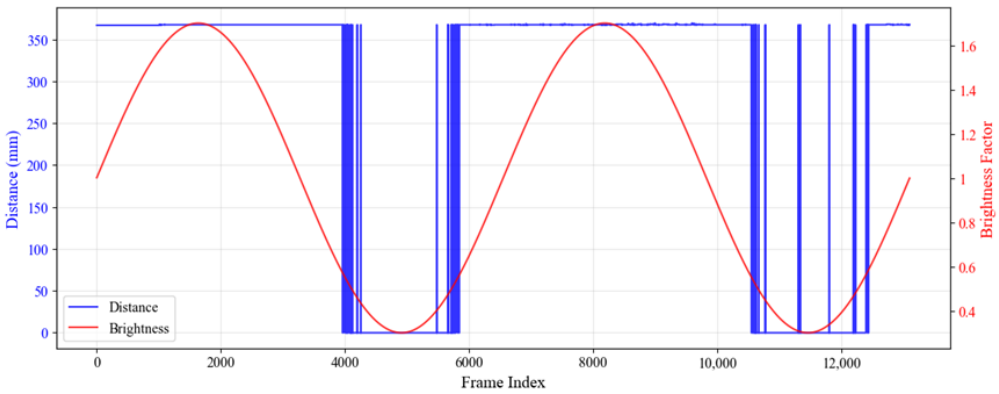



**Figure 8.** Component spacing as a function of luminance.

Within the brightness coefficient range of 0.650 to 1.700, the measurement system performs excellently: statistics based on 8730 data points show that the average distance is 368.214 mm, the standard deviation is 0.272 mm, and the range is 2.519 mm, indicating high consistency of the measurement results and good adaptability of the method to conventional illumination variations. Under extremely low-brightness conditions (0.300 to 0.650), measurement failure occurs, mainly due to the failure of block edge detection caused by excessively dark images.

### *3.5. Discussion*

Experimental results show that the proposed method outperforms the baseline methods in both measurement accuracy and operational efficiency. The performance advantages stem from the collaboration and complementarity of the algorithm: in the block reference detection stage, optical flow ensures continuity and high efficiency, while Hough transform provides geometric reconstruction when tracking fails; in the girder tracking stage, hierarchical matching achieves a dynamic balance between computational efficiency and tracking robustness.

Within the conventional illumination range (0.650–1.700), the measurement results are stable with a range less than 2.519 mm. Combined with the processing speed of 23.9 FPS, the method demonstrates the potential for capturing structural dynamic responses and field deployment on real bridges. However, under extremely low-illumination conditions, block reference detection fails, limiting direct nighttime application. Future improvements can be made by incorporating infrared supplementary lighting or low-light enhancement algorithms.

## 4. Real Bridge Application Validation

To further verify the implementability and engineering value of the proposed method in real bridge monitoring scenarios, it was applied to the field observation of an in-service bridge to obtain the lateral rigid-body displacement variation in the curved girder under long-term traffic loads.

### *4.1. Field Deployment and Monitoring System Configuration*

The video data were collected from a fixed surveillance camera mounted on the cap beam of an in-service expressway bridge, used to monitor the girder and its seismic block at the adjacent cap beam. The horizontal distance between the two cap beams is 30.0 m, representing typical long-distance imaging conditions. The video has a resolution of 1920 $\times$ 1080 and a frame rate of 25.0 FPS. The on-site monitoring image is shown in Figure 9a.

Camera parameters were uniformly set in advance: focal length locked, exposure parameters fixed, and frame rate locked, eliminating image inconsistencies caused by automatic adjustment. Through user interaction, the region of interest (ROI) of the seismic block and the girder template region were respectively selected in the initial frame to provide initialization conditions for the algorithm. The entire process requires no installation of artificial targets and fully utilizes existing monitoring equipment and natural structural features.




**Figure 9.** Actual bridge monitoring ((**a**). Surveillance footage of the actual bridge. (**b**). System monitoring interface).

### *4.2. Monitoring Process and Algorithm Execution*

During the four consecutive days of monitoring, the system operated automatically under unattended conditions, and fully recorded the lateral displacement responses of the girder under various working conditions such as day–night alternation, illumination variation, and vehicle passage.

For block reference detection, the optical flow–Hough fusion algorithm maintained stable extraction of the seismic block edge under complex conditions including illumination changes and partial occlusion. Optical flow prediction achieved efficient tracking, and the system automatically switched to Hough transform reconstruction when tracking failure occurred due to sudden illumination changes, ensuring the continuous availability of the measurement reference. Figure 10 illustrates the typical illumination variations in the real bridge during the monitoring period. It can be observed that the bridge deck and retaining block areas are affected by the solar altitude angle and cloud cover at different times, leading to changes in surface illuminance, shadow direction and contrast, which verifies the illumination non-stationarity of the monitoring scenario.

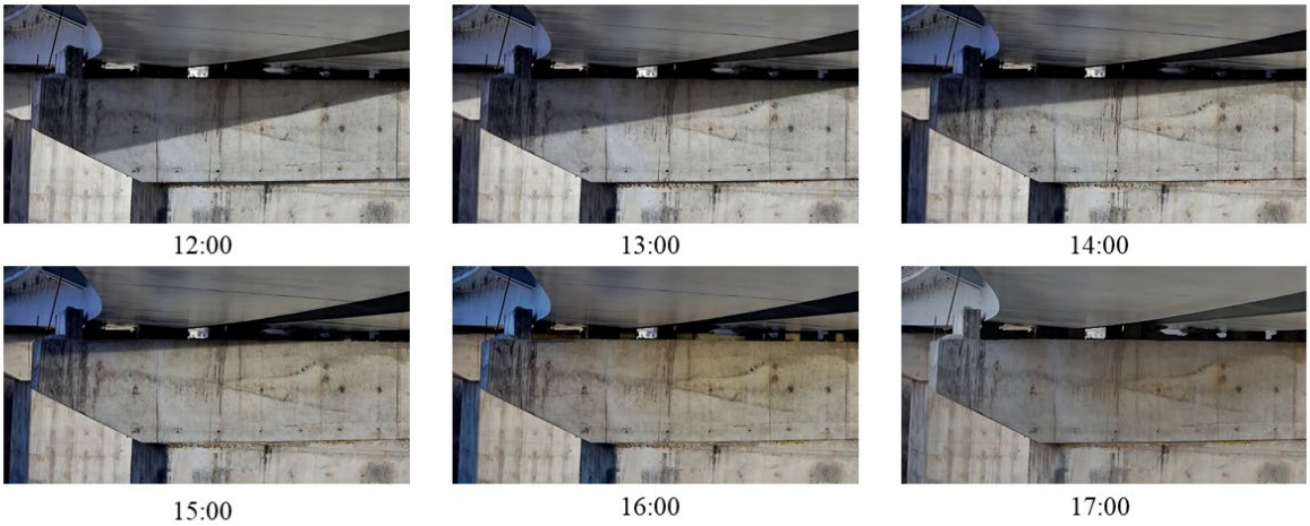



**Figure 10.** Light variation at different times on the real bridge.

For girder displacement tracking, the hierarchical NCC matching strategy continuously provided the girder center position under conditions of weak texture and dynamic interference. Local fast tracking accomplished matching at low computational cost in most

frames; when local failure resulted from sudden illumination changes or vehicle occlusion, the system automatically switched to global robust compensation. No continuous tracking failure or obvious drift occurred throughout the monitoring process.

Field operation results show that the algorithm performs excellently in continuous monitoring. Even under complex conditions such as drastic illumination changes (cloud occlusion, shadow sweeping), the system can automatically resume stable detection within a very short time.

### *4.3. Engineering Evaluation Results*

Automated analysis of the four consecutive days of surveillance video demonstrates the stable performance of the system:

(1) System Performance and Robustness. The average single-frame processing time is 36.0 ± 3.5 ms, corresponding to a real-time frame rate of 28.0 FPS, which meets the real-time requirements of engineering monitoring. The girder template matching confidence is 0.976 ± 0.014, indicating that the hierarchical matching strategy is highly robust to illumination variations and slight vibration disturbances. Neither the block edge detection nor the girder tracking module experienced continuous tracking failure or cumulative drift.

(2) Measurement Accuracy and Physical Rationality. As shown in Figure 11, the girder–block gap remained generally stable during the four-day monitoring period, while exhibiting periodic fluctuations associated with traffic loads. The measured values range from 90.000 to 108.000 mm (pixel conversion rate: 5.340 mm/pixel), with a mean value of 100.450 ± 6.000 mm.

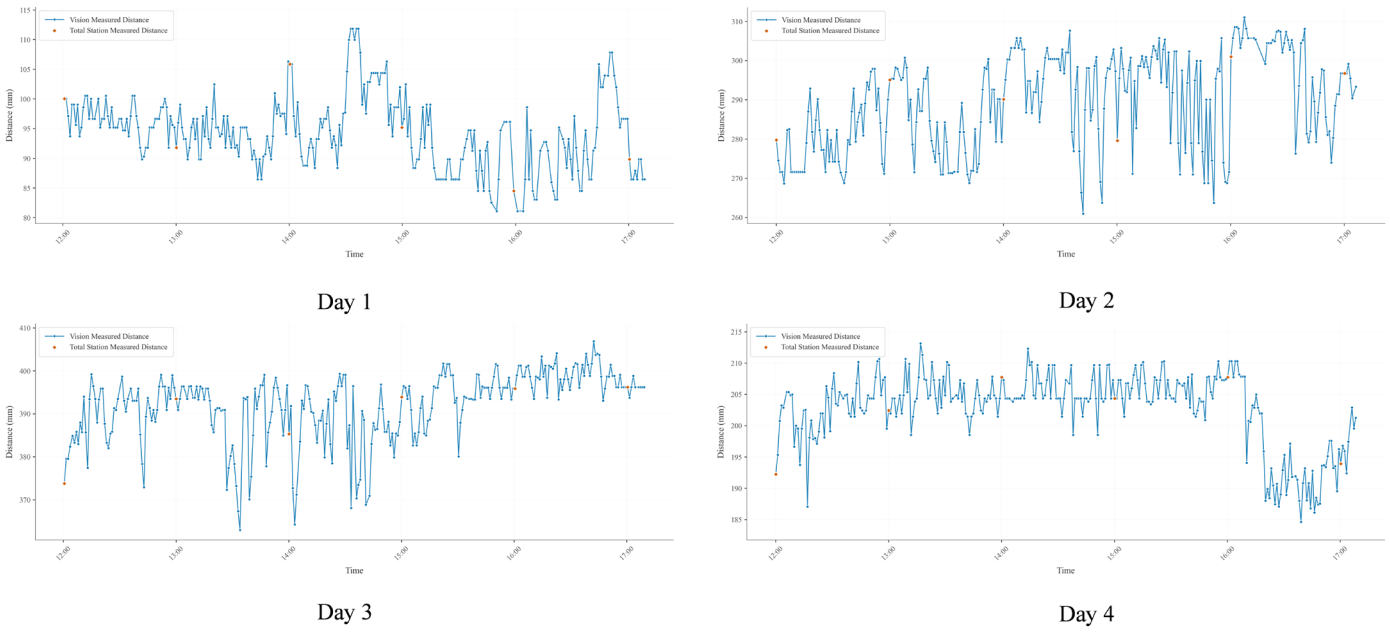



**Figure 11.** Variation in girder–stopper spacing over time in the practical engineering application.

During the monitoring period, manual measurements were carried out using a high-precision total station at 1 h intervals. A comparison with the proposed method showed that the absolute differences were all below 2.500 mm. The detailed results are listed in Table 4.

**Table 4.** Absolute differences between total station and proposed method (mm).

| Time | 12:00 | 13:00 | 14:00 | 15:00 | 16:00 | 17:00 |
|---|---|---|---|---|---|---|
| Day1 | 0.286 | 1.791 | 0.598 | 0.781 | 1.439 | 1.349 |
| Day2 | 1.646 | 0.892 | 0.792 | 0.515 | 1.314 | 2.127 |
| Day3 | 0.498 | 1.230 | 1.592 | 0.982 | 1.439 | 1.307 |
| Day4 | 0.294 | 1.910 | 2.491 | 1.037 | 0.404 | 0.325 |

Table 4 summarizes the absolute differences between the total station measurements and the proposed method over four consecutive days. Measurements were taken every hour from 12:00 to 17:00. The differences range from 0.286 mm to 2.491 mm, all remaining below the 2.500 mm threshold. The maximum deviation, 2.491 mm, occurred on Day 4 at 14:00.

When heavy vehicles pass, the girder produces an instantaneous lateral displacement response under centrifugal force, manifested as a brief change in the gap; under temperature or wind loads, a slower periodic drift is observed. These patterns are consistent with the mechanical mechanism of bridges, verifying the physical rationality of the measurement results. No cumulative drift occurred in the displacement data over the four-day monitoring period.

(3) Engineering Application Value. The system can perform real-time over-limit early warning and visual indication for lateral displacement based on preset safety thresholds (Figure 9b). Key data are superimposed on video frames in real time, providing intuitive monitoring information for operation and maintenance personnel. The scheme requires no artificial targets, no modification to the bridge structure, and can directly utilize existing monitoring equipment.

### *4.4. Discussion and Limitations*

The real-bridge validation demonstrates that the proposed method exhibits favorable feasibility, stability, and accuracy in practical engineering environments. Under the 30 m long-distance imaging condition, the system operated continuously for four days without drift, the measurement deviation from the total station was less than 2.500 mm, and the lateral displacement responses induced by vehicle passage were successfully captured.

The difference between the two measurement methods is less than 2.500 mm. This difference may mainly result from image measurement uncertainty, camera calibration error, manual total station measurement error, and the inconsistency in measurement time and measurement points between the two methods. The pixel conversion rate used in this study is 5.340 mm/pixel. Therefore, a difference of 2.500 mm corresponds to approximately 0.468 pixels. For a long-distance video monitoring condition of about 30 m, this error level is considered reasonable.

More specifically, the edge localization of the seismic block, template matching of the main girder, illumination variation, local image blur, and slight camera or bridge vibration may all introduce small measurement errors. In addition, the transverse displacement of the main girder may be affected by vehicle loads, temperature changes, and wind loads. Therefore, slight differences in measurement time and measurement position between the visual monitoring method and the manual total station measurement may also lead to small deviations. Moreover, the total station measurement may be influenced by manual point selection, aiming accuracy, and field operation conditions.

Limitations: In the absence of auxiliary lighting at night, the image signal-to-noise ratio drops sharply, resulting in the failure of block edge detection and the inability to establish a stable measurement reference.

## 5. Conclusions

This paper proposes a target-free visual method using seismic blocks as fixed references to measure rigid-body girder displacements (particularly lateral) under long-distance imaging. The main conclusions are:

(1) An optical flow–Hough fusion edge detector with a quality control module is developed to adaptively and robustly extract block edges under weak-edge conditions, providing a stable spatial reference despite illumination changes, noise, and partial occlusion.

(2) A hierarchical NCC-based hybrid template matching strategy is constructed, which autonomously switches between local fast tracking and global robust compensation. It balances efficiency and robustness, achieves 23.9 FPS (vs. 17.9 FPS with global NCC), and maintains high tracking success under sudden illumination changes, partial occlusion, and blur.

(3) Indoor tests at 5 m yield an MAE of 1.092 mm, superior to the compared baselines; measurement fluctuation remains below 2.519 mm for brightness coefficients of 0.65–1.70. In a 30 m real curved girder bridge application using existing surveillance cameras, continuous 4-day monitoring achieves 28.0 FPS, deviations from total station data less than 2.5 mm, no time drift, and captures vehicle-induced lateral displacements.

The method requires no artificial targets or structural modification, can directly utilize existing monitoring equipment, and offers clear advantages in deployment cost, environmental adaptability, and accuracy for operational bridge displacement monitoring. Future work will extend nighttime capability, full rigid-body displacement via multi-view and sparse 3D reconstruction, and end-to-end depth estimation for simplified calibration.

**Author Contributions:** Conceptualization, G.L. and D.L.; Methodology, G.L. and H.-B.H.; Software, G.L., S.A. and Y.C.; Validation, G.L., H.-B.H., S.A. and Y.C.; Investigation, Y.C.; Data curation, H.-B.H.; Writing—original draft, G.L.; Writing—review & editing, G.L., H.-B.H., S.A., Y.C. and D.L.; Visualization, G.L., H.-B.H. and S.A.; Supervision, G.L. and D.L.; Project administration, G.L. and D.L.; Funding acquisition, D.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study was funded by the National Natural Science Foundation of China (Grant No. 52478310).

**Data Availability Statement:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Acknowledgments:** The authors would like to thank the members of the research group for their valuable discussions and support during the preparation of this manuscript.

**Conflicts of Interest:** Authors Guangyu Li and Yuan Cheng were employed by the company Hebei Expressway Group Limited. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Brownjohn, J.M.W.; Koo, K.Y.; Scullion, A.; List, D. Operational deformations in long-span bridges. *Struct. Infrastruct. Eng.* **2015**, *11*, 556–574. [CrossRef]
2. Cao, H.T.; Lu, Y.; Chen, D.H. Analysis of Vehicle-Bridge Coupling Vibration Characteristics of Curved Girder Bridges. *Appl. Sci.* **2024**, *14*, 2021. [CrossRef]
3. Wang, X.P.; Zhao, Q.Z.; Xi, R.J.; Li, C.F.; Li, G.Q.; Li, L.A. Review of Bridge Structural Health Monitoring Based on GNSS: From Displacement Monitoring to Dynamic Characteristic Identification. *IEEE Access* **2021**, *9*, 80043–80065. [CrossRef]
4. Ko, J.M.; Ni, Y.Q. Technology developments in structural health monitoring of large-scale bridges. *Eng. Struct.* **2005**, *27*, 1715–1725. [CrossRef]
5. Pramudita, A.A.; Lin, D.B.; Dhiyani, A.A.; Ryanu, H.H.; Adiprabowo, T.; Yudha, E.A. FMCW Radar for Noncontact Bridge Structure Displacement Estimation. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 14. [CrossRef]

6. Farrar, C.R.; Worden, K. An introduction to structural health monitoring. *Philos. Trans. R. Soc. A-Math. Phys. Eng. Sci.* **2007**, *365*, 303–315. [CrossRef] [PubMed]
7. Ye, X.W.; Dong, C.Z.; Liu, T. A Review of Machine Vision-Based Structural Health Monitoring: Methodologies and Applications. *J. Sens.* **2016**, *2016*, 10. [CrossRef]
8. Bao, Y.Q.; Tang, Z.Y.; Li, H.; Zhang, Y.F. Computer vision and deep learning-based data anomaly detection method for structural health monitoring. *Struct. Health Monit.* **2019**, *18*, 401–421. [CrossRef]
9. Dong, C.Z.; Ye, X.W.; Jin, T. Identification of structural dynamic characteristics based on machine vision technology. *Measurement* **2018**, *126*, 405–416. [CrossRef]
10. Xie, H.B.; Liao, Q.Y.; Liao, L.; Qiu, Y.H. Machine Vision-Based Real-Time Monitoring of Bridge Incremental Launching Method. *Sensors* **2024**, *24*, 7385. [CrossRef] [PubMed]
11. Hong, Y.; Zhao, Z.K.; Tian, Q.W.; Li, S.Y. Vibration measurement in machine vision based on ROI gray-scale projection feature matching algorithm. *Adv. Struct. Eng.* **2025**, *28*, 2488–2509. [CrossRef]
12. Zhang, Y.; Liu, P.; Zhao, X.F. Structural displacement monitoring based on mask regions with convolutional neural network. *Constr. Build. Mater.* **2021**, *267*, 120923. [CrossRef]
13. Wu, L.J.; Cai, Z.W.; Lin, C.H.; Chen, Z.C.; Cheng, S.Y.; Lin, P.J. Investigation of the super-resolution methods for vision based structural measurement. *Smart Struct. Syst.* **2022**, *30*, 287–301. [CrossRef]
14. Jo, B.W.; Lee, Y.S.; Jo, J.H.; Khan, R.M.A. Computer Vision-Based Bridge Displacement Measurements Using Rotation-Invariant Image Processing Technique. *Sustainability* **2018**, *10*, 1785. [CrossRef]
15. Feng, D.M.; Feng, M.Q.; Ozer, E.; Fukuda, Y. A Vision-Based Sensor for Noncontact Structural Displacement Measurement. *Sensors* **2015**, *15*, 16557–16575. [CrossRef] [PubMed]
16. Spencer, B.F., Jr.; Hoskere, V.; Narazaki, Y. Advances in Computer Vision-Based Civil Infrastructure Inspection and Monitoring. *Engineering* **2019**, *5*, 199–222. [CrossRef]
17. Dong, C.Z.; Catbas, F.N. A review of computer vision-based structural health monitoring at local and global levels. *Struct. Health Monit.* **2021**, *20*, 692–743. [CrossRef]
18. Luo, L.X.; Feng, M.Q.; Wu, Z.Y. Robust vision sensor for multi-point displacement monitoring of bridges in the field. *Eng. Struct.* **2018**, *163*, 255–266. [CrossRef]
19. Khuc, T.; Catbas, F.N. Structural Identification Using Computer Vision-Based Bridge Health Monitoring. *J. Struct. Eng.* **2018**, *144*, 13. [CrossRef]


**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.